NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

THE USE OF EPI-SPLINES TO MODEL EMPIRICAL SEMIVARIOGRAMS FOR OPTIMAL SPATIAL ESTIMATION

by

Peter M. P. Tydingco II

September 2016

Thesis Advisor: Douglas P. Horner
Co-Advisor Johannes O. Royset

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704–0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2016	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE THE USE OF EPI-SPLINES TO MODEL EMPIRICAL SEMIVARIOGRAMS FOR OPTIMAL SPATIAL ESTIMATION	**5. FUNDING NUMBERS**
6. AUTHOR(S) Peter M. P. Tydingco II	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB number ____N/A____.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.	**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

This research investigates the ability of epi-splines to improve upon current methods of creating empirical semivariograms for use in optimal spatial estimation (OSE). Models utilizing traditional methods of curve fitting for semivariograms (spherical, exponential, and Matérn) used in the spatial estimation process are compared to a proposed model that employs an epi-spline for curve fitting. The resulting semivariograms are then used for kriging to produce spatial estimation using a sparse number of measurements. The epi-spline model improves upon the mean absolute error, mean standard error, and range of errors when compared to traditional methods. However, the comparisons indicate that goodness of fit does not drastically improve the resultant spatial estimation. The benefit of epi-splines, in addition to their ability to more accurately depict the relationship between data points, is their ability to incorporate soft information in the form of constraints and the tighter variance of estimates resulting from their use.

14. SUBJECT TERMS epi-splines, terrain aided navigation, semivariograms, kriging, optimal spatial estimation	**15. NUMBER OF PAGES** 81
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540–01-280-5500 Standard Form 298 (Rev. 2–89)
Prescribed by ANSI Std. 239–18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

THE USE OF EPI-SPLINES TO MODEL EMPIRICAL SEMIVARIOGRAMS FOR OPTIMAL SPATIAL ESTIMATION

Peter M. P. Tydingco II
Lieutenant, United States Navy
B.S., University of Arizona, 2007

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 2016

Approved by: Douglas P. Horner
Thesis Advisor

Johannes O. Royset
Co-Advisor

Patricia A. Jacobs
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This research investigates the ability of epi-splines to improve upon current methods of creating empirical semivariograms for use in optimal spatial estimation (OSE). Models utilizing traditional methods of curve fitting for semivariograms (spherical, exponential, and Matérn) used in the spatial estimation process are compared to a proposed model that employs an epi-spline for curve fitting. The resulting semivariograms are then used for kriging to produce spatial estimation using a sparse number of measurements. The epi-spline model improves upon the mean absolute error, mean standard error, and range of errors when compared to traditional methods. However, the comparisons indicate that goodness of fit does not drastically improve the resultant spatial estimation. The benefit of epi-splines, in addition to their ability to more accurately depict the relationship between data points, is their ability to incorporate soft information in the form of constraints and the tighter variance of estimates resulting from their use.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

A2AD	Anti-access/Area Denial
AUV	Autonomous Underwater Vehicle
DOD	Department of Defense
E-fit	Process of Fitting an Epi-Spline to Data
GP	Gaussian Process
GPS	Global Positioning System
GRF	Gaussian Random Field
IDW	Inverse Distance Weighting
INS	Inertial Navigation System
L2 normal	Lebesgue Space squared
lsc	Lower semicontinuous
MAE	Mean Absolute Error
MSE	Mean Squared Error
NPS	Naval Postgraduate School
OSE	Optimal Spatial Estimation
RF	Random Field
RV	Random Variable
SLAM	Simultaneous Localization and Mapping
TERCOM	Terrain Contour Mapping
TAN	Terrain Aided Navigation
UUV	Unmanned Underwater Vehicle

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The proliferation of unmanned systems in military and civilian sectors has occurred at lightning speed. In the case of Autonomous Underwater Vehicles or Unmanned Underwater Vehicles (AUV/UUVs), the expansion of their use and the potential that remains has brought to light areas in need of improvement. One of those areas is their reliance on Global Positioning System (GPS) for accurate long-term navigation. With the stated goals of operations in areas of Anti-Access/Area Deniability (A2AD) by the Department of Defense (DOD) [1], it is imperative that the ability for accurate navigation during long-term underwater operations by unmanned systems be improved.

Terrain Aided Navigation (TAN) and Simultaneous Localization and Mapping (SLAM) are two of the leading methodologies that have been demonstrated to be successful at solving this problem. TAN is a process of vehicle position estimation [2]. As a vehicle transits an environment, measurements are taken. Those measurements are compared to a known map of the operating area. The vehicle's location is most likely to be in areas where the profile of the measurement most closely matches the profile of the known map data. Like TAN, SLAM is a method of position estimation that relies on map data [3]. In this process, the creation of the map occurs as the vehicle is navigating the environment. It can be used to create a sparse map of the environment or to revise an a priori map.

This thesis focuses on TAN. Better maps will permit more accurate estimates of the unmanned vehicle's position. This in turn creates more accurate sensor measurements taken from the vehicle. One technique to build the map is the use of Optimal Spatial Estimation (OSE). A benefit of this approach is that it creates an estimate of the terrain that ensures minimal errors.

This technique is accomplished in two steps. The first step is creation of the semivariogram. The semivariogram is a means of depicting the spatial autocorrelation of the known data points. The second step is to use the information gained from the

semivariogram for spatial estimation. This is an interpolation procedure that produces a mean estimate at points of interest by scaling measurements based on distance and semivariogram function. The geostatistical community calls this procedure Kriging. It produces a mean and variance estimate of a surface. The use of the autocorrelation values produced by the semivariogram allows kriging to produce predictions while also providing a measure of the error or accuracy of those predictions.

The key for OSE is the accuracy of the semivariogram. It determines the accuracy of the mean and variance of the estimates. Epi-splines present a curve fitting methodology which may be applied to semivariograms in order to fit a curve that is better able to model the data. In addition, epi-splines present a way of incorporating soft data or a priori information into the semivariogram. This thesis investigates the use of epi-splines for modeling the semi-variance relation and its influence on the resultant spatial estimation.

These concepts are tested and validated using bathymetry data collected from Pavilion Lake, British Columbia. Two sections of 200-meter square areas are selected. One is a region of rough terrain; the second is a region of smooth terrain. The data is then divided in order to utilize *k*-fold cross validation to verify the model's accuracy. Upon setting aside validation data, each region is divided into 20-meter sub-regions where each sub-region is subjected to the kriging process. This process will produce the estimate of the square's midpoint using six randomly chosen data points within the 20-meter region. That estimation is then compared to the points within the 20-meter square that have been set aside as validation data. This process will be performed using traditional methods, with semivariograms produced using the spherical, exponential and Matérn formulas, and then again with semivariograms employing an epi-spline.

The results of this thesis indicate that epi-splines are capable of improving OSE visually and quantitatively. The epi-spline, with its better fit of the data in the region of rough terrain, demonstrates improvement in semi-variance for data points at farther distances than the range of the sill. The Lebesgue space squared (L2 normal) comparison of the curves produced by the Matérn and epi-spline models, after 16,000 estimations, averaged 2056.8 in a region of rough terrain and 59.7326 in the region of smooth terrain.

Despite the improvements in goodness of fit by the epi-spline, the resulting estimations are not as drastic as one might expect. The improvement in mean squared error (MSE) is 2.366 and mean absolute error (MAE) is 1.724 when applied to a region of rough terrain. Applied to smooth terrain, MSE is 0.396 and MAE is 0.168. Furthermore, the errors demonstrate Gaussian behavior and, especially in the case of the region of smooth terrain, are close to zero.

References

[1] "Unmanned systems integrated roadmap FY2013-2038," Dept. of Defense. [Online]. Available: http://www.defense.gov/Portals/1/Documents/pubs/DOD-USRM-2013.pdf.

[2] L. Zhou, X. Chang, Y. Zhu, and Y. Lu, "Terrain aided navigation for long-range AUVs using a new bathymetric contour matching method," presented at the IEEE International Conference on Advanced Intelligent Mechatronics, Busan, Korea, 2015.

[3] S. Riisgaard and M. R. Blas, (n.d.). SLAM for Dummies: A tutorial approach to simultaneous localization and mapping [Online]. Available: http://ocw.mit.edu/courses/aeronautics-and-astronautics/16-412j-cognitive-robotics-spring-2005/projects/1aslam_blas_repo.pdf

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to thank Dr. Horner and Dr. Royset for the tremendous opportunity investigation of this thesis afforded me. The education and experience I have gained is much appreciated and will last a lifetime. Professor Mike Cook from the NPS Oceanography Department, your MATLAB expertise and eagerness to help were invaluable.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

In its Unmanned System Integrated Roadmap, FY2013-2038 [1], the Department of Defense (DOD) lists the ability to operate in areas of anti-access/area denial (A2AD) as a requirement. It explicitly states unmanned systems will be critical to U.S. operations. One challenge for unmanned systems is position estimation when GPS is unavailable. Without an accurate position estimate, navigation and data collection becomes difficult.

This is especially true for Autonomous Underwater Vehicles/Unmanned Underwater Vehicles (AUVs/UUVs) since GPS is not available underwater. Coming to the surface may be problematic since in congested areas, like a harbor, the AUV may collide with other vehicles or obstacles. In deeper water, any time and power used to surface for GPS access, reduces the possible duration of the mission. Finally, for clandestine military missions, staying underwater may be required.

Many of the military's more sophisticated UUV/AUVs employ expensive, complex inertial navigation systems (INS) used in combination with GPS and onboard sensors to estimate their position. When operating underwater, in the absence of GPS, these systems begin to develop errors in their positional estimate. These errors grow over time. Terrain Aided Navigation (TAN) and Simultaneous Localization and Mapping (SLAM) are two methodologies for positional estimation where GPS is unreliable or unavailable.

TAN is a process of vehicle position estimation. As a vehicle transits an environment, measurements are taken. Those measurements are compared to a known map of the operating area. The vehicle's location is most likely to be in areas where the profile of the measurement most closely matches the profile of the known map data. Like TAN, SLAM is a method of position estimation that relies on map data. In this process, the creation of the map occurs as the vehicle is navigating the environment. It can be used to create a sparse map of the environment or to revise an a priori map.

Whereas TAN traditionally relies on a priori maps, this thesis is part of a broader interest seeking to employ Optimal Spatial Estimation (OSE) to build a map suitable for

use in TAN. An important part of the OSE process is the creation of semivariograms. This is a measurement of the dissimilarity of data over increasing distance. These measurements are aggregated into a series of bins and this empirical data can be fit to a curve. This is currently accomplished using one of several parametric functions. Epi-splines employ piecewise polynomial functions, subject to predefined constraints, to more closely fit a curve to a given dataset. It is this resulting improvement that this thesis seeks to exploit in order to improve OSE for use in AUV/UUVs employing TAN.

A. PROBLEM DESCRIPTION

The goal of this thesis is to produce a terrain map based on the sensor measurements of an AUV/UUV. The technique that will be used to accomplish that goal is OSE. It is a process commonly termed "geostatistics," which finds its origins in the mining and mineral industry [2]. OSE is a technique used to map a surface using a limited number of data points.

Kriging is a two-step process. The first step in this process is the production of the semivariogram. The second step is to produce an estimate, or prediction, at points of interest. This is accomplished through a system of linear equations. Normally this is accomplished as part of a batch operation where the data is first collected and the kriging process is run afterwards. While not a specific goal for this thesis, the eventual goal is the near-real time estimate of the terrain model as measurements are gathered, enabling an AUV to perform TAN in-stride.

The thesis hypothesis is that incorporating epi-splines for OSE can improve terrain estimation since it can model the semi-variance relation with greater fidelity and permits soft or hard data, constraints, or model assumptions when fitting a line of regression to the semivariogram. The ability to more accurately model the terrain will directly benefit procedures of terrain estimation that do not rely on GPS.

The goal of this thesis is to evaluate epi-splines as an improved nonparametric representation to form empirical semivariograms for kriging. The improvement realized in the kriging process can then be translated into an improved terrain estimate for use in TAN.

1. Relation to Terrain Aided Navigation

TAN is a process of vehicle position estimation. A vehicle takes measurements as it transits over terrain. The measurements are then compared to a previously known map to form the estimation. The vehicle's position is likely to be over terrain where the profile of the measurements most closely matches the data from the known map [3]. Figure 1 demonstrates the general idea of TAN. While traversing an area, the vehicle's sensors take measurements of the terrain below as depicted in the image to the left. The right image demonstrates the comparison of the data to the a priori map of the terrain. The yellow line at the top of the image represents the correlation between the measurements and the map data, the vehicle is more likely to be over the region where the largest peak occurs.

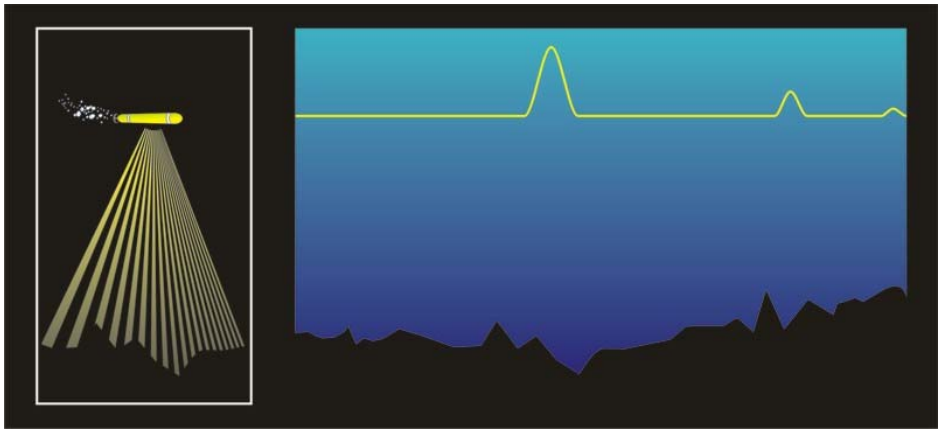

Figure 1. One Dimensional Example of TAN, Demonstrates the Correlation Between Measurement Data and Map Data. Source: [3].

a. Brief History of TAN

TAN was originally developed for use in TERCOM (Terrain Contour Matching); a guidance system for cruise missiles. Advances in computational power, data storage, and sensor accuracy have allowed it to become an alternative method of navigation for AUVs/UUVs that currently require GPS or resurfacing for GPS, and acoustic navigation [4]. Current implementations of TAN coupled with a high accuracy INS have achieved accuracy rates within meters and drift rates of less than 1%, when used with high quality sonars and/or terrain maps [4].

2. How Improved Optimal Spatial Estimation Relates to TAN

Typically TAN utilizes a previously known map to compare against sensor measurements for position estimation. Often, these maps are made with sparse information. OSE is capable of producing mean and uncertainty estimates using a limited number of measurements. This thesis posits that OSE can be utilized to create the map of comparison or to improve upon the map in use. The ability to accurately predict the operational environment accurately without having previously mapped the entire domain will not only reduce time requirements, it will also increase the accuracy of the navigational solution produced using the improved maps. This will result in increased chances of success for all missions.

B. LITERATURE REVIEW

Kriging is an advanced method of interpolation that produces an optimal estimate of an unknown point using a scattered set of known data in the surrounding area. Here optimality is measured with respect to minimal variance. It is the method used to produce the estimates for investigation of this thesis. The literature devoted to kriging is numerous. The most complete sources of information for kriging are by Cressie [5] and Chilés et al [6]. Cressie's work is the standard reference to which other works regarding kriging typically refer. The books by Cressie and Chilés are replete with mathematical proofs, key terms, definitions, and examples of kriging, suitable not only as an introduction but also for in depth questions and understanding of the topic. Bohling's paper, [7], provides a survey of kriging in an abbreviated form.

Semivariograms are a major part of this thesis and therefore a thorough understanding of their meaning qualitative and quantitatively is critical. While some mention regarding semivariograms can be found in [5] and [6], Bohling [8] and Gandhi [9] provide introductory surveys. Stein compares differing methods of approaching semivariograms [10]. Specifically addressed is the difference between parametric forms of semivariograms and non-parametric forms of semivariograms and the advantages of each.

As the subject of this thesis, an understanding of epi-splines is paramount. Literature regarding epi-splines is sparse compared to kriging. Two papers by Royset et al., [11] and [12], combine to provide an in-depth survey of epi-splines. The benefits epi-splines offer and examples of their use are discussed thoroughly. Examples of how to formulate a proposed epi-spline's objective function and formulation of some common constraints one may wish to impose on the optimization problem can be found in [12]. Additionally, examples of the real world applications of epi-splines are presented and discussed.

C. THESIS ORGANIZATION

This thesis is organized as follows: Chapter II serves as an overview of epi-splines. Chapter III provides an introduction to semivariograms, covering terms, definitions and examples. Chapter IV is a survey of the approach and terminology of the kriging process. Chapter V is an explanation of the implementation, results, and analysis. Finally, Chapter VI presents conclusions and recommendations.

THIS PAGE INTENTIONALLY LEFT BLANK

II. EPI-SPLINES

An epi-spline is a piecewise lower-semicontinuous (lsc) function formed by a discrete number of polynomials, subject to a finite number of constraints [12]. "E-fit" refers to the practice of fitting epi-splines to data. Information used in simulation or analysis is never perfect. There may be physical laws or information from sensors that require the curve to be increasing or at least non-negative. An analyst may have insight about the upper and lower bounds or range of the particular function. E-fit is a means of incorporating such information into the curve fitting process, in the form of constraints. The benefits of an E-fit is that it is flexible; allows the inclusion of outside information; can be quickly implemented by a variety of programming software; requires little calculation time; and because of epi-convergence, can converge to any function.

A. KEY TERMS AND GUIDELINES

As stated in [11], an epi-spline is defined in terms of *order*, which is a nonnegative integer, p, the number of partitions S, and its *mesh* $m = \{m_k\}_{k=0}^{S}$, where $m_{k-1} < m_k$, k = 1,2,…S. When these quantities are defined, an epi-spline is a real valued function defined on the closed interval $[m_0, m_S]$ that is a polynomial of degree p for each segment (m_{k-1}, m_k), where k=1, 2…S. Each segment of the epi-spline requires (p+1) parameters to define a polynomial of degree p.

By increasing the number of partitions S, an epi-spline of any order p can approximate any lsc function to an arbitrary level of accuracy in the sense of epi-convergence [11]. Thus, optimal points of a function are approximated by corresponding points of the approximating epi-spline. Extensive experience has shown low-order splines to be preferred over higher order ones because they do not exhibit the oscillatory behavior of high-degree polynomials [12]. An order of two or three is recommended and is what will be used for the investigation of this thesis. Reasons for selection of higher order polynomials are discussed in [12].

Choosing the *mesh* is straightforward. When not facing computational constraints, it is recommended to select as fine a *mesh* as possible [12]. In cases where the actual function closely resembles polynomial behavior or when computing time is a *concern*, coarse meshes are acceptable [12]. The spacing of mesh-points may also be dictated by external information, such as a point of discontinuity. In the absence of such information, evenly spaced mesh-points are the natural choice [12]. Though not addressed here, situations of greater than one dimension exist and are more complex; solutions to such cases are discussed in [12].

Figure 2 depicts an epi-spline with multiple discontinuities; however it is presented here to highlight the mesh-points. Explanation of the mathematical considerations for formulations that contain discontinuities can be found in [12]. As stated in section 3.1 of [12], when a continuous or continuously differentiable function suffices, as is the case for the E-fit proposed for this thesis, such applications do not require the full generality of lower- and upper-semicontinuous functions.

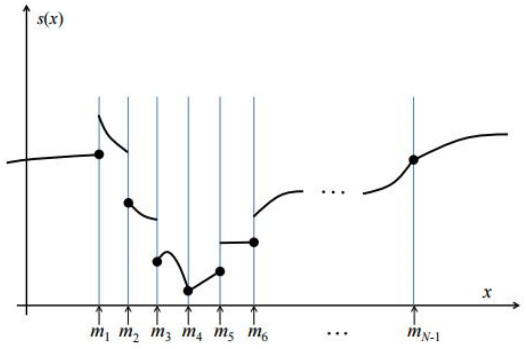

Figure 2. Depiction of Mesh-Points. Source: [12].

B. GENERAL EXAMPLE

A general E-fit in standard NPS format [13] is

Index

$i \in I = \{1, 2, ..., n\}$, n = the number measurements used for curve fitting

$k \in K = \{1, 2, ..., S\}$, S = the number of segments in the epi-spline

Data

d is the x coordinate of the data point

m_k, are the mesh points $\forall k = 0,1,\ldots,S$

$f(d)$ is the y coordinate for the data point.

Decision Variables

$a_0^k,\ a_1^k,\ a_2^k$

where a_0^k is the constant term, a_1^k is the first degree term, and a_2^k is the leading term/second degree term for the second degree polynomial of the *k*th segment.

Formulation

$$\underset{x}{\text{Min}} \quad \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$
\begin{aligned}
s.t. \quad & f(d_i) - \left[a_0^{k_i} + a_1^{k_i} d_i + a_2^{k_i} d_i^2\right] \le x_i \quad \forall i = 1,\ 2,...,n \ \text{(first two constraints relate to the absolute error)} \\
& \left[a_0^{k_i} + a_1^{k_i} d_i + a_2^{k_i} d_i^2\right] - f(d_i) \le x_i \quad \forall i = 1,\ 2,...,n \\
& a_2^k \le 0 \quad \forall k = 1,2\ldots,S \ \text{(concavity)} \\
& a_0^k + a_1^k m_k + a_2^k m_k^2 = a_0^{k+1} + a_1^{k+1} m_k + a_2^{k+1} m_k^2, \forall k = 1,2,\ldots,S-1 \ \text{(continuous)} \\
& a_1^k + 2a_2^k m_k = a_1^{k+1} + 2a_2^{k+1} m_k \quad \forall k = 1,2,\ldots,S-1 \ \text{(derivative is continuous)} \\
& a_0^1 = 0 \ \text{(epi-spline begins at 0)}
\end{aligned}
$$

1. Objective Function

The objective function seeks to minimize the distance (error) between the data points and the curve that is fit to the data. In mathematical terms, the objective can be to minimize the absolute error, mean squared error, least squares error, or any other user defined error. The purpose of this function is to minimize the average of x_i, over all values of *a* and *x*. Recall that *a* represents the coefficients for each of the piecewise

polynomials, $\left[a_0^{k_i} + a_1^{k_i} d_i + a_2^{k_i} d_i^2 \right] = \hat{f}(d_i)$, and $x_i = \left| f(d_i) - \hat{f}(d_i) \right|$ which is the absolute value of the difference between $f(d_i)$ and the prediction, $\hat{f}(d_i)$, or the absolute error.

2. Constraints

The constraints are a means of incorporating soft information into the model. Soft information is contextual information about the data, such as shape, slope, continuity, smoothness, symmetry, etc. The constraints for this E-fit and the reasons for choosing them are discussed in the following text. Additional information relating to the constraints discussed for this model as well as a discussion and examples of additional constraints can be found in [11] and [12].

The first two constraints stem from the absolute value of the error desired in the objective function. Here, $f(d_i)$ is the actual value of a function at the i^{th} data point and $\left[a_0^{k_i} + a_1^{k_i} d_i + a_2^{k_i} d_i^2 \right] = \hat{f}(d_i)$, is the value of the epi-spline at the point.

The third constraint requires the second derivative of the polynomials to be non-positive and, coupled with the continuity constraints, ensure the resulting epi-spline is concave [12].

The fourth constraint requires that the epi-spline begin at zero. In the case where information dictates that the epi-spline begin somewhere other than the origin, this constraint could be adjusted to reflect such information.

The fifth constraint requires that the epi-spline is continuous. This is accomplished by imposing a constraint of equality at the mesh-points. That is to say, it requires the intersection between segment *k* and segment *k+1* to be equal for the resulting piecewise polynomials of each segment.

The final constraint ensures that the first derivative of the epi-spline is continuous by requiring the derivatives of the polynomials to be equal at the mesh-points when moving from left to right,

C. OPTIMIZATION IMPLEMENTATION

With the formulation of the problem complete, it is necessary to solve the system of equations. This can be accomplished through an optimization library. This section is included as an elementary tutorial. What follows is a discussion of the formulation of the minimization problem and how the required vectors and matrices are defined. The general format of the objective function for the minimization problem is

$$\begin{aligned} \min \quad & \boldsymbol{f}^T \boldsymbol{x} \\ s.t. \quad & \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \\ & \boldsymbol{C}\boldsymbol{x} = \boldsymbol{d} \\ & \boldsymbol{l} \le \boldsymbol{x} \le \boldsymbol{u}. \end{aligned}$$

Prior to implementing the minimization function, it is necessary to form the $\boldsymbol{f}$, $\boldsymbol{m}$, $\boldsymbol{\tau}$, $\boldsymbol{A}$, $\boldsymbol{b}$, $\boldsymbol{C}$, $\boldsymbol{d}$, $\boldsymbol{l}$, and $\boldsymbol{u}$ matrices. Note that $\boldsymbol{m}$ and $\boldsymbol{\tau}$ are vectors that represent the mesh-points and the y-coordinates of the data points, respectively, and are used in the formation of the matrices necessary for the implementation of the optimization problem.

Recall S is the number of segments chosen for the epi-spline and n is the number of data points. In this case, S=20 and n=15. The output of the function will be the $\boldsymbol{x}$ matrix, the size of which will be (3S+n)-by-1. For the intended E-fit formulation, $\boldsymbol{x}$ will be a 75-by-1 vector, the contents of which will be the variables that the minimization program will vary in order to meet the imposed constraints while satisfying the objective function. The first 60 columns of the $\boldsymbol{x}$ matrix represent the coefficients of the 2nd order polynomial that will be formed to represent each segment of the epi-spline, for this example, 3 coefficients times 20 segments = 60. The remaining 15 columns of each row represent the x_i variable for each segment. The $\boldsymbol{f}$ vector is a t-by-1 vector, where t=(3S+n)=75. The first 3S=60 rows are zero, the remaining n rows are 1/n. The $\boldsymbol{f}$ and $\boldsymbol{x}$ vectors are represented as

$$\boldsymbol{f}^T = \left[0 \quad \cdots \quad 0 \quad \frac{1}{n} \quad \cdots \quad \frac{1}{n} \right] \text{ and } \boldsymbol{x}^T = [a_0^1 \;\; a_1^1 \;\; a_2^1 \;\; \cdots \; a_0^{20} \;\; a_1^{20} \;\; a_2^{20} \;\; x_1 \;\; \cdots \;\; x_{15}].$$

The transposition of ***f*** and ***x*** are shown in the interest of space.

Before moving on to the constraints, it is necessary to define a vector representing the values of the mesh-points for the epi-spline, this size of which will be (S+1)-by-1. For this example, this vector is defined as m. For one iteration of the model implemented in this thesis, m is a 21-by-1 vector, the first row is zero, and the subsequent rows increase by 6.8, resulting in an endpoint of 136. In this example, the first mesh-point will occur at 6.8, the second at 13.6, etc. The end points of m represent the range over which the proposed epi-spline is defined. It is also necessary to define a vector representing the y coordinates of the data points. This symbol chosen for this example is τ, which is an n-by-1 vector.

1. Constraints

The inequality equation is $\boldsymbol{Ax} \le \boldsymbol{b}$. For the proposed E-fit, the $\boldsymbol{A}$ matrix will be a t-by-v matrix, where $t = (2n+S)=50$ and $v = (3S+n)=75$. The variables of a single row correspond to the variables in $\boldsymbol{x}$ and their values will be assigned prior to executing the minimization. It is necessary to apply some simple algebra to arrange the inequalities into a more usable format. Recall the first three constraints from the E-fit formulation were the inequalities

$$\gamma(d_i) - \left[a_0^{k_i} + a_1^{k_i} d_i + a_2^{k_i} d_i^2\right] \le x_i \quad \forall i = 1,\ 2,\ldots,n$$
$$\left[a_0^{k_i} + a_1^{k_i} d_i + a_2^{k_i} d_i^2\right] - \gamma(d_i) \le x_i \quad \forall i = 1,\ 2,\ldots,n$$
$$a_2^k \le 0 \quad \forall k = 1,2\ldots,S.$$

After manipulation, their format resembles $Ax \le b$,

$$-a_0^{k_i} - a_1^{k_i} d_i - a_2^{k_i} d_i^2 - x_i \le -\gamma(d_i)$$
$$a_0^{k_i} + a_1^{k_i} d_i + a_2^{k_i} d_i^2 - x_i \le \gamma(d_i)$$
$$a_2^k \le 0.$$

Now the $\boldsymbol{A}$ and $\boldsymbol{b}$ matrices are easier to visualize.

The $\boldsymbol{A}$ matrix for the first constraint will be n=15 rows. Given the case that S=20, n=15, and the first two data points are in segments 1 and 2, respectively, the first two rows of $\boldsymbol{A}$ corresponding to the first constraint are

$$\begin{bmatrix} -1 & -d_1 & -d_1^2 & 0 & 0 & 0 & a_0^3 & a_1^3 & a_2^3 & \cdots & a_0^{20} & a_1^{20} & a_2^{20} & -1 & 0 & x_3 & \cdots & x_{15} \\ 0 & 0 & 0 & -1 & -d_2 & -d_2^2 & a_0^3 & a_1^3 & a_2^3 & \cdots & a_0^{20} & a_1^{20} & a_2^{20} & 0 & -1 & x_3 & \cdots & x_{15} \end{bmatrix}.$$

Note that a_0^3, a_1^3, a_2^3, a_0^{20}, a_1^{20}, a_2^{20}, x_3 and x_{15} are shown to illustrate the size of the matrix. Their actual values in this example would be zero. This highlights a key concept in the formation of the $\boldsymbol{A}$ matrix, which is that the location of a data point d matters. The segment of the epi-spline within which the data point falls will determine the set of polynomial coefficients (a_0^k, a_1^k, a_2^k) assigned a value. For instance, if x_1 falls in segment 1, then a_0^1, a_1^1, a_2^1 will be assigned the values of $-1, -d_1, -d_1^2$ and x_1 will be assigned the value of -1. The values of d are the actual x-coordinate for the i^{th} data point. This process is the same for constraints 1 and 2.

The second inequality is implemented in the same way as the first. Given the case that S=20, n=15, and the first two data points were in segments 1 and 2, respectively, rows 16 and 17 of $\boldsymbol{A}$, which correspond to the second constraint are

$$\begin{bmatrix} 1 & h_1 & h_1^2 & 0 & 0 & 0 & a_0^3 & a_1^3 & a_2^3 & \cdots & a_0^{20} & a_1^{20} & a_2^{20} & -1 & 0 & x_3 & \cdots & x_{15} \\ 0 & 0 & 0 & 1 & h_2 & h_2^2 & a_0^3 & a_1^3 & a_2^3 & \cdots & a_0^{20} & a_1^{20} & a_2^{20} & 0 & -1 & x_3 & \cdots & x_{15} \end{bmatrix}.$$

As before, a_0^3, a_1^3, a_2^3, a_0^{20}, a_1^{20}, a_2^{20}, x_3 and x_{15} are shown to illustrate the size of the matrix. Their actual values in this example would be zero.

The final 20 rows of the $\boldsymbol{A}$ matrix correspond to the third constraint, which requires that $a_2^k \le 0$ for all segments. This is as simple as the next S=20 rows having a 1 placed at the a_2^k placeholder for every segment and is formed as

$$\begin{bmatrix} 0 & 0 & 1 & \cdots & a_0^{20} & a_1^{20} & a_2^{20} & x_1 & \cdots & x_{15} \\ \vdots & & & & & & & & & \vdots \\ a_0^1 & a_1^1 & a_2^1 & \cdots & 0 & 0 & 1 & x_1 & \cdots & x_{15} \end{bmatrix}.$$

Note that $a_0^1, a_1^1, a_2^1, a_0^{20}, a_1^{20}, a_2^{20}, x_1$ and x_{15} are shown to illustrate the size of the matrix. Their actual values in this example would be zero.

The ***b*** matrix is an *t*-by-1 vector, where $t=(2n+S)=50$. As stated earlier, the first 15 rows of ***b*** are simply the negative of **τ**. The second 15 rows of ***b*** are equal to **τ**, and the final 20 rows are zeros, corresponding to the third constraint.

The equality constraints are now used to form $\boldsymbol{Cx} = \boldsymbol{d}$. Here, ***C*** is a *t*-by-*v* matrix and ***d*** is a *t*-by-1 vector, with $t=(2S-1)=39$ and $v=(3S+n)=75$. As was the case for the inequalities, the variables of a single row correspond to the variable in ***x***, and will also have values prior to executing the minimization. The first row, corresponds to constraint 4 and requires that $a_0^1 = 0$, which is

$$\begin{bmatrix} 1 & 0 & 0 & \ldots & a_0^{20} & a_1^{20} & a_2^{20} & x_1 & \cdots & x_{15} \end{bmatrix}.$$

Again, a_0^{20}, a_1^{20}, a_2^{20}, x_1, and x_{15} are shown to illustrate the size of the matrix. Their actual values in this example would be zero. This constraint could also be implemented in the lower and upper bound vectors, ***l*** and ***u***.

The fifth and sixth constraints require manipulation to ease the formation of ***C***. Constraint five, which imposes continuity of the epi-spline, is

$$a_0^k + a_1^k m_k + a_2^k m_k^2 - a_0^{k+1} - a_1^{k+1} m_k - a_2^{k+1} m_k^2 = 0, \quad \forall k = 1,2,\ldots,S-1 \ .$$

Written in a form that facilitates implementation, the sixth constraint is reformulated as

$$a_1^k + 2a_2^k m_k - a_1^{k+1} - 2a_2^{k+1} m_k = 0, \quad \forall k = 1,2,\ldots,S-1 .$$

The first two rows of $\boldsymbol{C}$, defined by constraint 5 appear as the following,

$$\begin{bmatrix} 1 & m_1 & m_1^2 & -1 & -m_1 & -m_1^2 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & m_2 & m_2^2 & -1 & -m_2 & -m_2^2 & \cdots & 0 & 0 & 0 \end{bmatrix},$$

and this pattern would repeat until row 19. Here, m_i corresponds to the i^{th} value of $\boldsymbol{m}$. Rows 20 through 39 correspond to constraint 6. The first two rows corresponding to constraint 6 for this example are 20 and 21. They appear as

$$\begin{bmatrix} 0 & 1 & 2m_1 & 0 & -1 & -2m_1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2m_2 & 0 & -1 & -2m_2 & \cdots & 0 & 0 & 0 \end{bmatrix},$$

and this pattern would repeat to row 39.

The $\boldsymbol{d}$ matrix will be a t-by-1 vector of zeros, where $t = (2S - 1) = 39$.

2. Lower and Upper Bounds

The lower and upper bounds will serve to set the limits of the variables in the $\boldsymbol{x}$ vector as they are varied during the minimization process. Knowing that $\boldsymbol{x}$ will be a 75-by-1 vector requires the $\boldsymbol{l}$ and $\boldsymbol{u}$ vectors to be the same size. As required by the proposed constraints, a_0^1 will have a lower and upper bound of 0. The slope terms, a_1^k , will have a lower bound of zero and an upper bound of infinity, to ensure they are non-negative which in turn ensures the epi-spline is non-decreasing. The a_2^k (second or leading coefficient) for the polynomial associated with each segment of the epi-spline will be less than or equal to zero to ensure concavity. The lower bound for these variables will be minus infinity and their upper bound will be zero. Any remaining variables will have a lower bound of negative infinity and an upper bound of infinity.

Having defined these matrices, one need only implement the minimization problem in capable software. The method chosen for this thesis is MATLAB and its linprog function.

THIS PAGE INTENTIONALLY LEFT BLANK

III. SEMIVARIOGRAMS

The broader interest of this thesis is terrain estimation with sparse information containing measurement errors. The goal is to use the resulting estimation to form a grid map. A natural way to think about this is to use an n-dimensional Random Field (RF). RFs are a stochastic process indexed by a spatial variable [14].

When the stochastic processes of a RF involve Gaussian (normal) probability density functions, such a RF is a Gaussian Random Field (GRF). These models are specified by a vector of means and matrix of covariances at any discrete location [15].

GRFs of one and two dimensions are normally called a Gaussian Process (GP). Similar to a GRF, a GP is a collection of random variables, any finite number of which have a joint Gaussian distribution [16]. Typically, GPs are defined over time and space. For this thesis, the variables will represent a two-dimensional location. Two reasons for the importance of GPs are that, upon assumption of a GP, virtually all prediction, estimation, and distribution theory are much easier [5]. The second reason stems from asymptotic considerations where the net result of many small order effects is approximately Gaussian [5].

Predictions based on GPs date back at least as far as 1949 and are well known in geostatistics where the process is known as kriging [16]. They are essentially the same, GP being more commonly used in the field of robotics and kriging being the common term in the field of statistics. This thesis will use vernacular associated with kriging.

A. KERNELS

Covariance functions, or kernels, are the key to GPs. Kernels encode all assumptions about the function being modeled and are a representation of similarity between measurements. A kernel can represent covariance (similarity) or semi-variance (dissimilarity). Important terms associated with kernels relate to the degree of stationarity of the GP.

If the covariance and semivariogram functions depend on two points, x_1 and x_2, many realizations of the pair of random variables (RV) would be necessary for any statistical inference to be possible [17]. However, if these functions depend only on the distance between the two points, statistical inference becomes possible. A RF is stationary, in the strict sense, if its spatial law is invariant under translation. Specifically the covariance between two locations a fixed distance apart is the same for all estimations.

A function is weakly or second-order stationary if its mean function exists and is not dependent on the point; e.g., the mean is constant. Additionally, for each pair of RVs, the covariance must exist and depend on the separation between the two points [17]. Cressie states that when a random process is Gaussian, second order-stationarity and strong stationarity coincide because a Gaussian process is defined by its mean and its covariance function [5].

In practice, the structural function, be it covariance or semivariogram, is used for limited distances, $|h| \leq b$, where $|h|$ is the distance between two points and b represents the limit of the neighborhood of estimation [17]. In such cases, the limitation of the hypothesis of second-order stationarity corresponds to a hypothesis of quasi-stationarity [17].

Additionally, when $\boldsymbol{x}$ and $\boldsymbol{x}'$ are two points, the covariance function is a function of

$$\left|\boldsymbol{x} - \boldsymbol{x}'\right|,$$

the RF is called isotropic [16]. Isotropy amounts to assuming there is no reason to distinguish one direction from another for the RF under consideration [10]. When a RF under consideration shows a different autocorrelation structure in different directions, it is called anisotropic.

The data used for investigation of this thesis does not have a time dimension. Instead there is a space dimension. Here too, the concept of stationarity is equally important and applicable. An important assumption about the underlying statistical process of the data is made. As it relates to this thesis, the assumption of quasi-stationarity is important with regard to the way in which the data is divided. A 200-meter region of bathymetry data will be divided into 20-by-20 meter sub-regions. The assumption of quasi-stationarity is applied to each of these sub-regions. Furthermore, it is assumed that each of these regions is isotropic.

B. SEMI-VARIANCE

Semi-variance is a type of kernel and is used to express the degree of dissimilarity between two points on a surface. Semi-variance is simply half of the variance of the differences between all points in a data set spaced a constant distance apart. A plot of semi-variance versus distance between pairs of data is called a semivariogram. When covariance is used, the plot is called a variogram, and represents the degree of similarity between two points. Figure 3 depicts a semivariogram and covariogram to demonstrate the difference between the two.

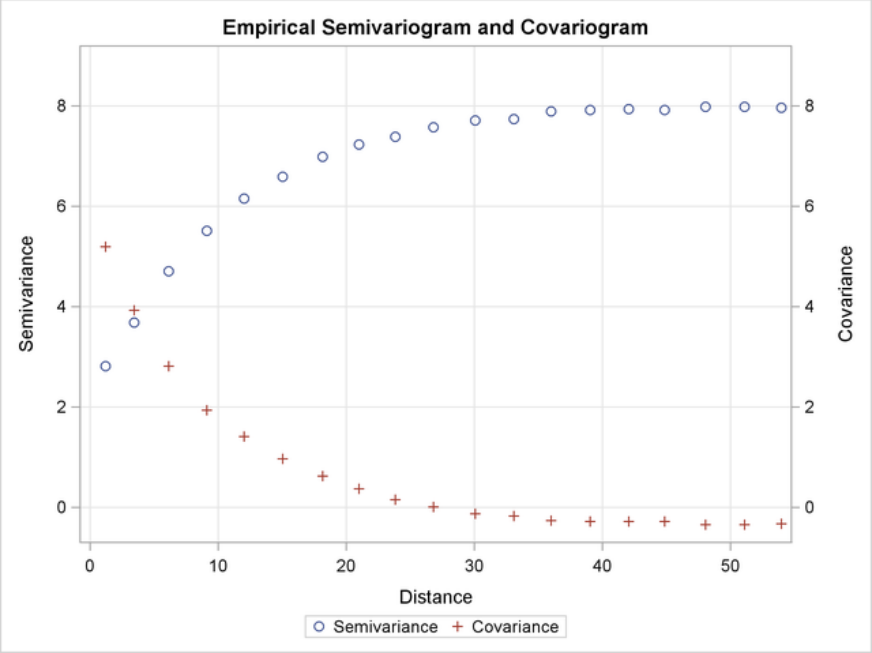

Figure 3. Depiction of a Covariogram and Semivariogram.

Kriging typically utilizes semivariograms, although either can be used, to characterize the spatial correlation between sample points on a surface. In the case of the semivariogram, the vertical axis represents a measure of semi-variance, γ. The formula for γ is

$$\gamma(h) = \sum_{i=1}^{N} \frac{\left[z(x_i + h) - z(x_i) \right]^2}{2N},$$

where N is the number of pairs of sample points of observations of the value of attribute z, separated by distance h [17]. The horizontal axis represents the separation distance between the points, or lag distance. Semivariograms used for kriging must be *non-negative definite* to ensure that the kriging equations are *non-singular* [8]. There are three main characteristics of a semivariogram: the nugget, sill, and range. Figure 4 is a diagram of a semivariogram that points out the characteristics.

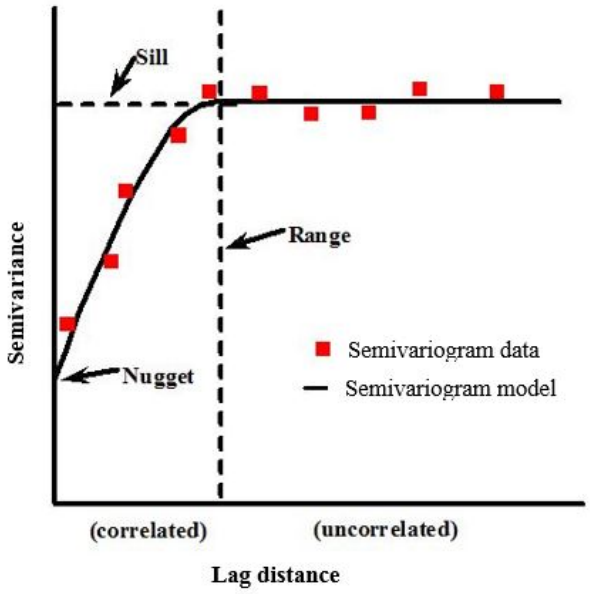

Figure 4. Semivariogram with Nugget, Sill and Range Depicted.

Theoretically, the value of the semivariogram at the origin should be zero. When it is not, it is referred to as the nugget [8]. The nugget is a measure of variability in the data at small distances and includes measurement error [8]. The sill refers to the value at which the slope of the semivariogram levels off. In cases when the nugget is not zero, there can be a partial sill, which is simply the sill minus the nugget. The range is related to the sill, it is the distance at which the semivariogram achieves the sill value. Correlation between points in the data set whose separation distance is greater than the range is zero. In his explanation of semivariograms, Bohling states that the process of fitting a model to empirical semivariograms is "more of an art than a science," [8] noting that the proper method and protocol for doing so varies with authority.

C. STANDARD MODELS

Standard models of semivariograms include spherical, penta-spherical, circular, exponential, Gaussian, Whittle and Matérn. For this thesis, a model based on an E-fit will be compared to three of the most popular models, spherical, exponential and Matérn. The data used for the semivariograms is from a 5-meter resolution data set of bathymetry measurements taken in Pavilion Lake, British Columbia.

1. The Spherical Model

The simplest and most widely used model for semivariograms in geological and hydrological fields is the spherical model [10]. The spherical model is smooth in regard to continuous differentiability, making the assumption that correlations are exactly zero at large enough distances. Stein finds no basis for its popularity however and cites specific prediction problems associated with it [10]. Figure 5 depicts a semivariogram based on the spherical model. The formula for the spherical model is

$$K(h) = \begin{cases} c\left(1.5\left(\frac{h}{a}\right) - 0.5\left(\frac{h}{a}\right)^3\right), & \text{if } h \le a \\ c, \quad \text{otherwise,} \end{cases}$$

where c represents the sill, a represents the range, and h represents the lag distance between two points.

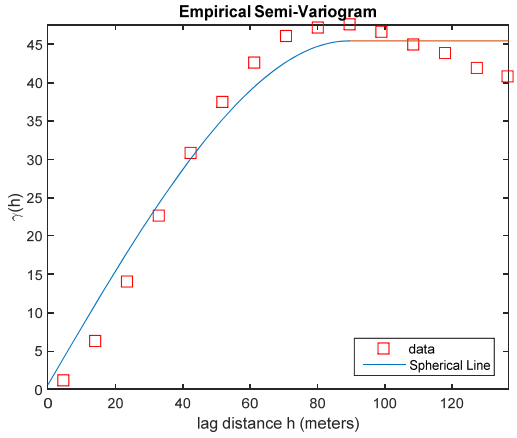

Figure 5. Best Fitting Spherical Model (in blue) Based on an Empirical Semi-Variogram (in red).

2. The Exponential Model

The exponential model makes the assumption that correlations may become arbitrarily small at large distances but they are never zero. Stein considers the exponential model a better substitute for the spherical model because of its improved performance in

certain three dimensional problems [10]. Figure 6 is a depiction of a semivariogram based on the exponential model. The formula for the exponential model is

$$K(h)=\begin{cases}0, & h=0\\ c\left(1-\exp\left(\dfrac{-3h}{a}\right)\right), & h>0,\end{cases}$$

where c represents the sill, a represents the range, and h represents the lag distance between two points.

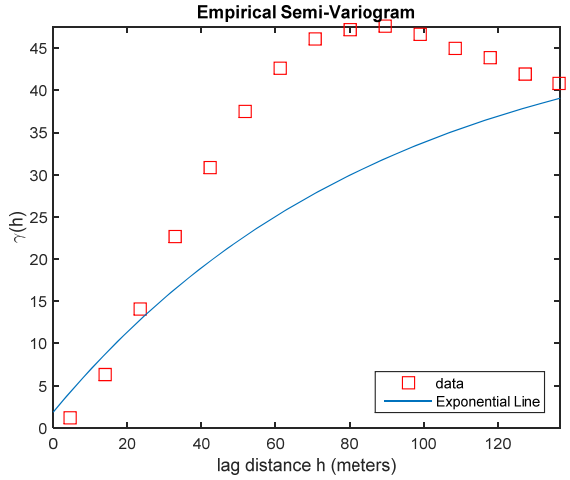

Figure 6. Best Fitting Exponential Model (Blue Line) Based on an Empirical Semivariogram (Red Squares).

3. The Matérn Model

The Matérn class of semivariograms provides much greater range for the possible local behavior of the RF GP [10]. Figure 7 depicts semivariogram based on the Matérn model, where ν=0.25, 0.5, 1, and 2. The formula for the isotropic auto-covariance function for the Matérn model is

$$K(h)=\frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}\left(\frac{2\nu^{1/2}h}{\rho}\right)\kappa_\nu\left(\frac{2\nu^{1/2}h}{\rho}\right),$$

where Γ is the gamma distribution function, κ_ν is the Bessel function, ρ and ν are non-negative parameters of the covariance, and *h* is the lag distance between two points. The fact that the function necessitates the calculation of a Bessel function does not create a serious obstacle to its use as programs capable of making such calculations are readily available.

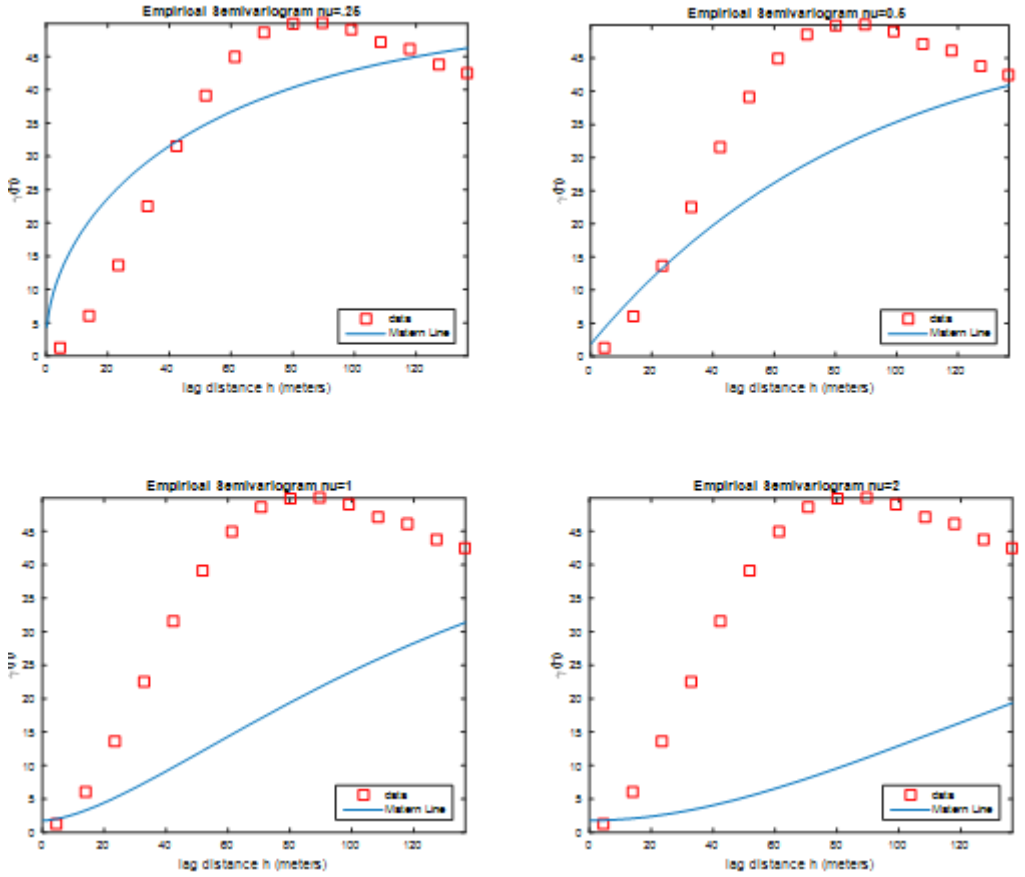

Figure 7. Best Fitting Matérn Model, ν=0.25, 0.5, 1 and 2 (Blue Lines) Based on an Empirical Semivariogram (Red Squares).

As depicted in Figure 7, the Matérn model's semivariogram never quite reaches a clear sill and range, which implies that there is never a distance where two measurements are unrelated. With that in mind, it is not surprising that the Matérn model, with ν=0.5, is equivalent to the exponential model. Stein suggests the most important reason for choosing the Matérn model over the others is the inclusion of the parameter ν into the

model [10]. ν controls the smoothness of the RF [10]. Unless there is some theoretical or empirical basis for fixing the degree of smoothness of a RF a priori, Stein sees no justification for selecting semivariogram models such as the spherical or exponential as they provide no flexibility in the degree of smoothness [10].

D. E-FIT FOR SEMI-VARIANCE

With its great flexibility, incorporation of soft information, and speed, E-fit represents a method that can improve current methods of curve fitting employed in the spatial estimation process. The formulation of the E-fit proposed for this thesis is

Index

$i \in I = \{1,2,...,n\}$, n = the number of data points on the semivariogram

$k \in K = \{1,2,...,\mathrm{S}\}$, S = the number of segments in the epi-spline

Data

h_i is the lag distance of data point i

m_k, are the mesh points $\forall k = 0,1,\ldots,S$

$\gamma(h_i)$ is the value of the γ function at a distance h_i from the semivariogram

Decision Variables

$a_0^k,\ a_1^k,\ a_2^k$

where a_0^k is the constant term, a_1^k is the first degree term, and a_2^k is the leading term/second degree term for the second degree polynomial of the kth mesh point.

Formulation

$$\min_x \ \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$
\begin{aligned}
s.t. \quad & \gamma(h_i) - \left[a_0^{k_i} + a_1^{k_i} h_i + a_2^{k_i} h_i^2 \right] \le x_i \quad \forall i = 1,\ 2,..., n \\
& \left[a_0^{k_i} + a_1^{k_i} h_i + a_2^{k_i} h_i^2 \right] - \gamma(h_i) \le x_i \quad \forall i = 1,\ 2,..., n \\
& a_2^k \le 0 \quad \forall k = 1, 2 \ldots, S \\
& a_0^k + a_1^k m_k + a_2^k m_k^2 = a_0^{k+1} + a_1^{k+1} m_k + a_2^{k+1} m_k^2, \forall k = 1, 2, \ldots, S-1 \\
& a_1^k + 2 a_2^k m_k = a_1^{k+1} + 2 a_2^{k+1} m_k \quad \forall k = 1, 2, \ldots, S-1 \\
& a_0^1 = 0
\end{aligned}
$$

Figure 8 depicts the proposed E-fit being applied to the same data as the standard models shown previously. Here the curve fits much closer to the data while simultaneously allowing for the clear depiction of the range, sill, and nugget. One big distinction in the case of the E-fit model is that there appears to be increasing value from data points at distances greater than the range.

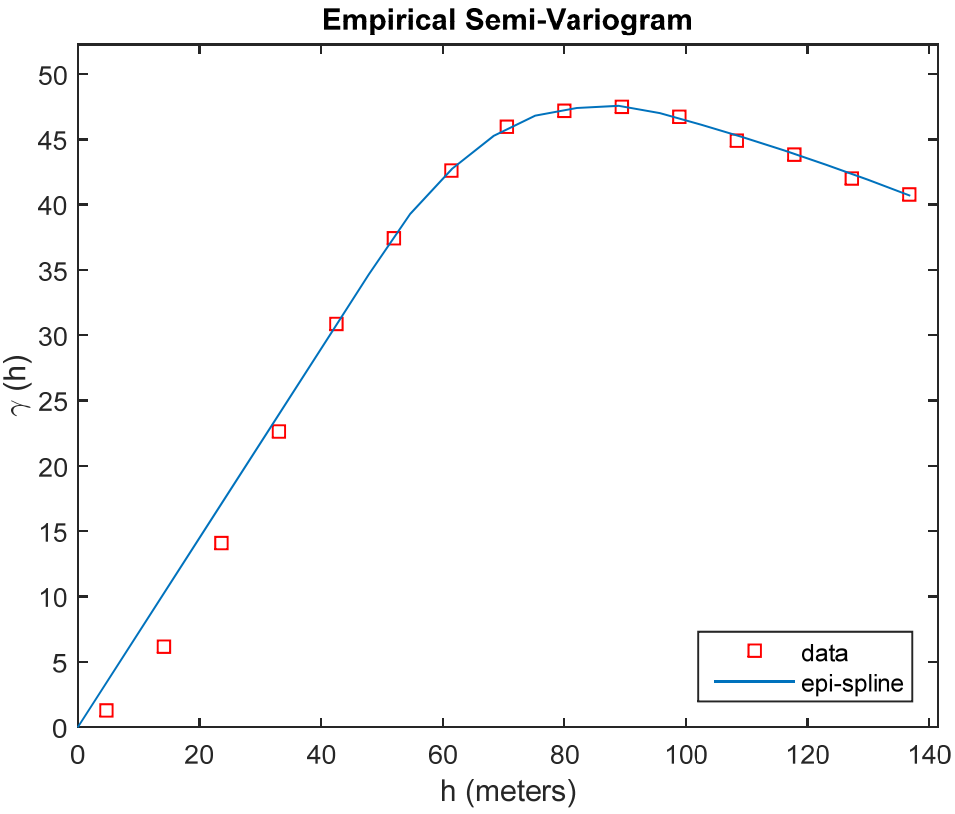

Figure 8. Best Fitting E-fit Model (Blue Line) Based on an Empirical Semivariogram (Red Squares).

As evidenced by the information presented in this chapter, E-fit as a means of curve fitting presents a better alternative to modeling empirical semivariograms. Their ability to model data that is either linear or oscillatory in nature provides a more accurate representation of the information found within the data set. Furthermore, the incorporation of soft and hard constraints increases the accuracy to which the resulting curve models the data.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. KRIGING

There are times when it is necessary to estimate unknown values of a regionalized variable. For this thesis that application is AUV TAN. A surveying AUV collects a sparse number bathymetry measurements used to build the reference map. Desired attributes of a spatial estimation process should include the following:

1. The reference map needs to be complete. At each point of interest an estimate is calculated.
2. It needs to have error estimates at all points of interest in the survey region. This uncertainty map is useful for route planning for the AUV.
3. Estimates should be optimal or minimize the error between actual values and the estimation.
4. Estimates should be unbiased, meaning that the mean error is close to zero.

Kriging is a geostatistical interpolation technique that provides all of these qualities. The general approach of kriging is that it uses the weighted sum of known, surrounding data values. Like other techniques, kriging assigns weights to values. These weights are determined by the semi-variance function. This typically means that the weights decrease as the distance increases between the point of interest and the selected measurements. Used correctly, kriging produces estimates that are optimal and unbiased [18]. The ability to provide the estimation error at each of the estimated points, which serves as a measure of confidence in the model, is a key feature of kriging and the reason it is a statistical technique versus a deterministic method [18].

Kriging comes in different "flavors," the most common of which are: simple kriging, ordinary kriging, and universal kriging or kriging with a trend. While it is not the goal of this paper to provide an in-depth presentation of kriging in its various forms, a brief introduction to simple kriging, ordinary kriging, kriging with a trend, and their formulas follows. A comprehensive explanation of kriging and its methods can be found in several sources, most notably, [5] and [6].

A. APPROACH AND TERMINOLOGY

The approach and terminology that follow are from [18]. Since kriging is a technique similar to that of Inverse Distance Weighting (IDW) interpolation, a comparison of the two reveals the benefits of kriging. In IDW, sample points within a radius are used to estimate each grid node. The weighted distance of a sample point from the grid node being estimated determines its influence on the estimate. In this manner, the closer a point is to the node, the greater its influence will be on the estimated value. IDW's disadvantage is that it treats all sample points within the search radius the same [18]. In contrast, kriging is capable of using different weighting functions, dependent upon distance and location of the data points relative to the estimation point, and the pattern in which the data points lie [18].

The prediction is

$$Z^* = \sum_{i=1}^{N} \lambda_i Z(s_i),$$

where N is the number of observed or measured values, λ_i is the weighted value for the observed or measured value at the i^{th} location, and $Z(s_i)$ is the measured value at the i^{th} location.

As previously stated, kriging assigns weights based on distance and orientation of the known data points to the estimation point. To do this, the spatial autocorrelation must be calculated [18]. Essentially, kriging requires two tasks, accomplished via a two-step process. The first task, uncovering the dependency information from the data, is accomplished by creating a variogram or semivariogram of the data. The semivariogram serves to reveal the spatial autocorrelation values of the data [18]. The second task, calculate an estimate or prediction, is accomplished by coupling the weighting values and the spatial autocorrelation data via a function, to compute the estimation.

1. Kriging Process

The formulas that are used come from [7]. In its most basic form, the kriging estimator is

$$Z^*(u) = \sum_{i=1}^{n} \lambda_i \left[Z(u_i) - m(u_i) \right] + m(u),$$

where $Z^*(u)$ is the basic linear regression estimator and *u* is the location vector for the estimation point. $Z(u_i)$ is a known measurement at location *i,* u_i is the location vector for the known data point (indexed by *i*), and λ_i is the kriging weight assigned to data point *i*. Finally, *n* is the number of known data points used for estimation, $m(u)$ is the mean of $Z(u)$, and $m(u_i)$ is the mean of $Z(u_i)$.

Consider $Z(u)$ as a RF consisting of a trend component, $m(u)$, and a residual component $R(u) = Z(u) - m(u)$. The residual at a point is estimated as a weighted sum of residuals at the neighboring data points [7]. Derivation of kriging weights, λ_i, can be from a semivariogram function as previously discussed, or from a covariance function.

With the kriging estimator in mind, it is necessary to determine the kriging weights, λ_i, that will result in the minimum variance of the estimator while adhering to the unbiased constraint [7]. The variance and estimation error are

$$\sigma^2(u) = \mathrm{Var}\left\{ Z^*(u) - Z(u) \right\}$$
$$\mathrm{E}\left\{ Z^*(u) - Z(u) \right\} = 0 .$$

Solving $R(u) = Z(u) - m(u)$, for $Z(u)$ allows the decomposition of the RF $Z(u)$ into a residual and trend component, $Z(u) = R(u) + m(u)$. Here, the residual component is a RF with stationary mean of 0 and stationary covariance equal to

$$\mathrm{E}\left\{ R(u) \right\} = 0$$
$$\mathrm{Cov}\left\{ R(u), R(u+h) \right\} = \mathrm{E}\left\{ R(u) \cdot R(u+h) \right\} = C_R(h),$$

where h is the lag distance.

Typically, the residual covariance function is derived from the semivariogram model as $C_R(h) = \delta - \gamma(h)$, where δ represents the sill and h is the lag distance between two points. In this regard, the semivariogram used for input into the kriging model is a representation of a variable's residual component [7]. As stated earlier, there are several varieties of kriging. The three main types are simple, ordinary, and kriging with a trend. The difference between the three is in the way they treat the trend component or the mean, $m(u)$ [7].

a. Simple Kriging

Once a semivariogram model is selected the kriging process can begin. Sometimes termed "kriging with a mean," this is the simplest form of kriging [6] and relies on the assumption that the trend component is a constant and known over the entire domain. The assumption for simple kriging is that the data set has a stationary variance and a stationary mean. User input of the mean is required [6], [18]. In this case,

$$m(u) = m \text{ and } Z^*(u) = m + \sum_{i=1}^{n} \lambda_i \left[Z(u_i) - m \right].$$

Since $\mathrm{E}\left[Z(u_i) - m \right] = 0$, the estimate is unbiased and

$$\mathrm{E}\left[Z^*(u) \right] = m = \mathrm{E}\left[Z(u) \right].$$

Additionally, the estimation error, $Z^*(u) - Z(u)$, is a linear combination of random variables representing the residuals at the known data points, u_i, and the estimation point, u [7]. Applying the rules for the variance of a linear combination of random variables, the error variance is

$$\sigma^2(u) = \mathrm{Var}\{R^*(u)\} + \mathrm{Var}\{R(u)\} - 2\,\mathrm{Cov}\{R^*(u), R(u)\}$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i(u)\lambda_j(u) C_R(u_i - u_j) + C_R(0) - 2\sum_{i=1}^{n} \lambda_i(u) \mid C_R(u_i - u).$$

By taking the derivative of the error variance formula with respect to each of the kriging weights and setting them to zero, the error variance is minimized, such that

$$\sum_{j=1}^{n} \lambda_j(u) C_R(u_i - u_j) = C_R(u_i - u), \qquad i = 1,...,n\,.$$

Due to the constant mean assumption in simple kriging, the covariance function for $Z(u)$ is equal to that of the residual component, thus $C(h) = C_R(h)$. The simple kriging equation can therefore be written in terms of $C(h)$ as

$$\sum_{j=1}^{n} \lambda_j(u) C_R(u_i - u_j) = C(u_i - u), \qquad i = 1,...,n\ .$$

Expressed in matrix form, K$\boldsymbol{\lambda}$=k, where K is the matrix of covariance between data points, k is the vector of covariance between the data points and the estimation point, and λ is the vector of kriging weights for the known data points. The structure of K and k are

$$\boldsymbol{K} = \begin{pmatrix} \gamma(d(z_1, z_1)) & \cdots & \gamma(d(z_1, z_n)) \\ \vdots & \ddots & \vdots \\ \gamma(d(z_n, z_1)) & \cdots & \gamma(d(z_n, z_n)) \end{pmatrix}, \quad \boldsymbol{k} = \begin{pmatrix} \gamma(d(z_1, z^*)) \\ \vdots \\ \gamma(d(z_n, z^*)) \end{pmatrix}.$$

$\boldsymbol{K}$ is a square *n*-by-*n* matrix, where *n* is equal to the number of known points used in the estimation. $\boldsymbol{\gamma}$ is derived from the semivariogram, z^* is the estimation point, and $d(z_i, z_j)$ is the distance between two points where $i = 1,...,n$ and $j = 1,...,n$.

Assuming no two data points are collocated, the covariance matrix is positive definite and the kriging weights can be determined by taking the inverse of $\boldsymbol{K}$, resulting

in $\boldsymbol{\lambda}=\boldsymbol{K}^{-1}\boldsymbol{k}$. The dimensions of $\boldsymbol{\lambda}$ are 1-by-*n*. It will contain the kriging weights and will resemble $\boldsymbol{\lambda}^T=\begin{pmatrix}\lambda_1 & \cdots & \lambda_n\end{pmatrix}$.

Having the kriging weights allows calculation of the kriging estimate and the kriging variance. The estimate or prediction is

$$z^*(u)=m+\sum_{i=1}^{n}\lambda_i(u)\left[z(u_i)-m\right],$$

where m is the mean of the RF and is known, λ_i is the kriging weight for each measurement, and $z(u_i)$ is the known measurement. The kriging variance is given by substituting the kriging weights into the equation for error variance and is equal to

$$\sigma^2(u)=C(0)-\lambda^T k=C(0)-\sum_{i=1}^{n}\lambda_i C(u_i-u).$$

b. Ordinary Kriging

Ordinary kriging assumes that the mean of a RF is constant in the local neighborhood of the estimation point. That is to say that $m(u_i)=m(u)$ for each data point z_i being used to estimate z^* [7] and the formula is

$$\begin{aligned}Z^*(u)&=m(u)+\sum_{i=1}^{n}\lambda_i(u)\left[Z(u_i)-m(u)\right]\\&=\sum_{i=1}^{n}\lambda_i(u)Z(u_i)+\left[1-\sum_{i=1}^{n}\lambda_i(u)\right]m(u).\end{aligned}$$

Requiring the kriging weights to sum to one filters the unknown local mean and results in the following equation for the estimator

$$Z^*(u) = \sum_{i=1}^{n} \lambda_i(u) Z(u_i) \quad \text{with} \sum_{i=1}^{n} \lambda_i(u) = 1.$$

The unit sum constraint imposed on the weights requires an additional step in order to minimize the error variance. The system must be set to minimize the error variance plus an additional term, the Lagrange parameter, $\mu(u)$ [7]. This results in a new equation for the error variance,

$$\sigma^2(u) + 2\mu(u)\left[1 - \sum_{i=1}^{n} \lambda_i(u)\right].$$

Minimizing the formula with respect to $\mu(u)$ forces adherence to the constraint requiring summation of the kriging weights to equal 1 [7]. Now, the equations for the kriging weights are

$$\begin{cases} \sum_{j=1}^{n} \lambda_j(u) C_R(u_i - u_j) + \mu(u) = C_R(u_i - u_j), & i = 1,...,n \\ \sum_{j=1}^{n} \lambda_j(u) = 1. \end{cases}$$

As with simple kriging, $C_R(h)$ is the covariance function for the variable's residual component. However, since a constant mean is no longer assumed over the entire domain, $C_R(h) \neq C(h)$. In practice, that substitution is often made based on the assumption that the semivariogram effectively filters the influence of large scale trends in the mean [7]. The unit sum constraint placed on the kriging weights allows the ordinary kriging model to be stated in terms of the semivariogram. As an interpolation approach, ordinary kriging follows naturally from semivariogram in that both tools filter the trends in the mean [7].

The matrix manipulation process is much the same process except for the addition of Lagrange multipliers into the ***K*** and ***k*** matrices, which allows us to solve for μ. In this case, the matrices appear as

$$\boldsymbol{K}=\begin{pmatrix} \gamma\left(d\left(z_1,z_1\right)\right) & \cdots & \gamma\left(d\left(z_1,z_n\right)\right) & 1 \\ \vdots & \ddots & & 1 \\ \gamma\left(d\left(z_n,z_1\right)\right) & \cdots & \gamma\left(d\left(z_n,z_n\right)\right) & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}, \boldsymbol{k}=\begin{pmatrix} \gamma\left(d\left(z_1,z^*\right)\right) \\ \vdots \\ \gamma\left(d\left(z_n,z^*\right)\right) \\ 1 \end{pmatrix}, \boldsymbol{\lambda}=\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu \end{pmatrix}.$$

As with simple kriging, it is now possible to calculate the estimate,

$$z^*\left(u\right)=\mu+\sum_{i=1}^{n}\lambda_i\left(u\right)\left[z\left(u_i\right)-\mu\right],$$

where μ is the calculated Lagrange multiplier, λ_i is the kriging weight for the i^{th} measurement, and z_i is the known measurement at location i. Upon calculation of the kriging weights and the Lagrange parameter, the error variance is

$$\sigma^2\left(u\right)=C\left(0\right)-\sum_{i=1}^{n}\lambda_i C\left(u_i-u\right)-\mu\left(u\right).$$

c. Kriging with a Trend

Also known as universal kriging, kriging with a trend is handled in a manner similar to ordinary kriging. The difference in this case being, instead of fitting a local mean in the local neighborhood of the estimation point, a linear or higher order trend in the (*x, y*) plane of the coordinates corresponding to the data points [7] is used.

Using this model for the trend in the kriging process requires the same adaptation of the Lagrange parameters as used in ordinary kriging. In this case, there are two additional Lagrange parameters and the two associated rows and columns in the ***K*** matrix.

While higher order trends, such as quadratic or cubic, can be implemented in the same way, anything higher than a first order trend is rare [7].

B. SUMMARY

All interpolation algorithms (inverse distance squared, radial basis functions, triangulation, etc.) estimate the value at an unknown location as a weighted sum of surrounding data values [7]. Most assign weights according to functions that assign a decreasing weight with increasing distance. Kriging assigns weights according to a data-driven weighting function but it is still just an interpolation algorithm. Some of the advantages of kriging are

- It helps compensate for the effects of data clustering by treating clusters more like single points.
- Kriging variance provides an estimate of estimation error.
- The availability of the estimation error provides a basis for stochastic simulation.

Like all methods of interpolation, kriging has the following weaknesses

- When data locations fall in clusters and there are large gaps in between, the estimates will be unreliable. This is true for any interpolation algorithm [7].
- Almost all interpolation algorithms underestimate the highs and overestimate the lows [7]. This is an inherent characteristic to averaging, which is found in all interpolation algorithms.

The semivariogram is the key function in the kriging process because it is used to model temporal or spatial correlation of the observed phenomenon. For the reasons proposed by Stein, the Matérn model was chosen as one model of comparison for this thesis. The spherical and exponential models were chosen in order to provide additional comparisons and provide information to highlight the subtleties associated with the formation of a semivariogram.

Some difficulty was encountered when implementing these models. There were instances when negative kriging weights were calculated. Negative weights arise in ordinary kriging when data close to the location being estimated occlude outlying data [19].

These negative kriging weights, when used for prediction, may lead to negative and non-physical estimates and therefore must be corrected. In his paper, Deutsch discusses other consequences of negative kriging weights and possible ways to deal with them [19]. For this thesis, the process used to correct negative kriging weights amounts to resetting the negative weights to zero and re-standardizing the remaining weights to sum to one [19].

V. RESULTS AND ANALYSIS

The hypothesis of this thesis is that E-fit can be used to improve semi-variance representation. This has a direct impact on OSE and in turn has a direct impact on the ability to build a reference bathymetry map for undersea TAN and ultimately, the end state ability of position estimation in GPS denied environments. This chapter provides the results and analysis for building an undersea reference map using the above-described methodology. It begins with an investigation into the effect that the number of data points will have on the formation of a semivariogram. Comparison is then made between semivariograms with curves fit using standard models and a semivariogram with an E-fit. Finally, kriging is carried out using the four models and the results and comparison are presented.

This thesis employed a dataset of underwater topography collected in Pavilion Lake, British Columbia. The 5-meter resolution data set was created by the University of Delaware as part of testing conducted by NASA. It was augmented in the middle and mid-southern basin areas by the REMUS. Figure 9 is a satellite image of Pavilion Lake with a topographical representation of the measurements overlaid. The red box over the northern portion highlights the region of smooth terrain and the box over the southern portion highlights the region of rough terrain. Figure 10 is a picture of the REMUS vehicle (yellow, white, and black vehicle lower left half of image) on the ice at Pavilion Lake.

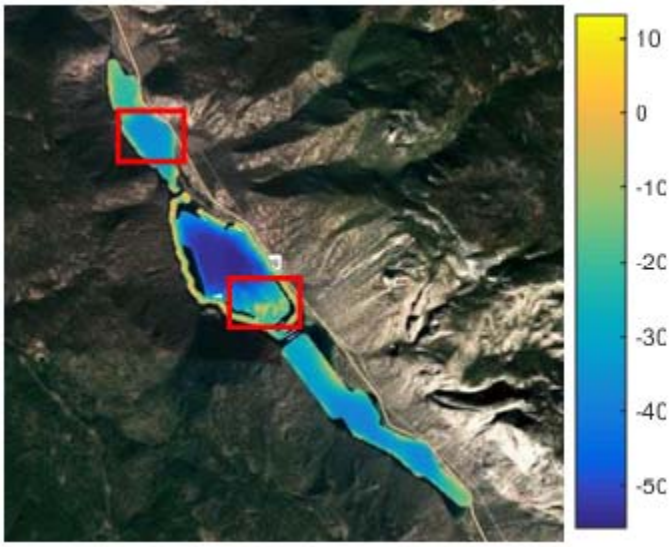

Figure 9. Pavilion Lake, BC, with Topographical Overlay and Red Boxes to Highlight Regions of Interest.

Figure 10. REMUS Vehicle on Surface Ice at Pavilion Lake, BC, February 2015.

A. ACCURACY OF SEMIVARIOGRAMS

As a precursor to the implementation of the proposed model, an investigation into the formation of the semivariograms was undertaken. The purpose was to determine the accuracy of semivariograms as a function of the number of measurements to determine

the impact the number of data points used in the formulation of the semivariogram will have on the resulting curve. This information is important because it provides insight as to how much information is needed in order to perform OSE in real time with confidence.

Using the previously described 200-meter regions of rough and smooth terrain in the Pavilion Lake data set, a series of semivariograms is created, beginning with 5% of the total data points and gradually increasing to 100%. At 100% the maximum number of data points for each region is 1,681. Beginning with the region of rough terrain, the data points used for each percentage group were randomly chosen from the 200-meter subset and a semivariogram is estimated. This process was replicated 10 times for each percentage. Based on information about semivariograms, it is anticipated that the precision of the semivariogram will increase as the number of data points increases. More specifically, the semivariogram represented by smaller percentages will have larger variation than those containing larger percentages. Upon completion of this preliminary investigation, model implementation and comparison begin.

The software used for investigation of this thesis was MATLAB, and its open source *variogram* and *variogramfit* functions. Figure 11 shows the results of 5%, 10%, 20%, 50%, 80% and 100% of the data points being used for the semivariogram in the region of rough terrain.

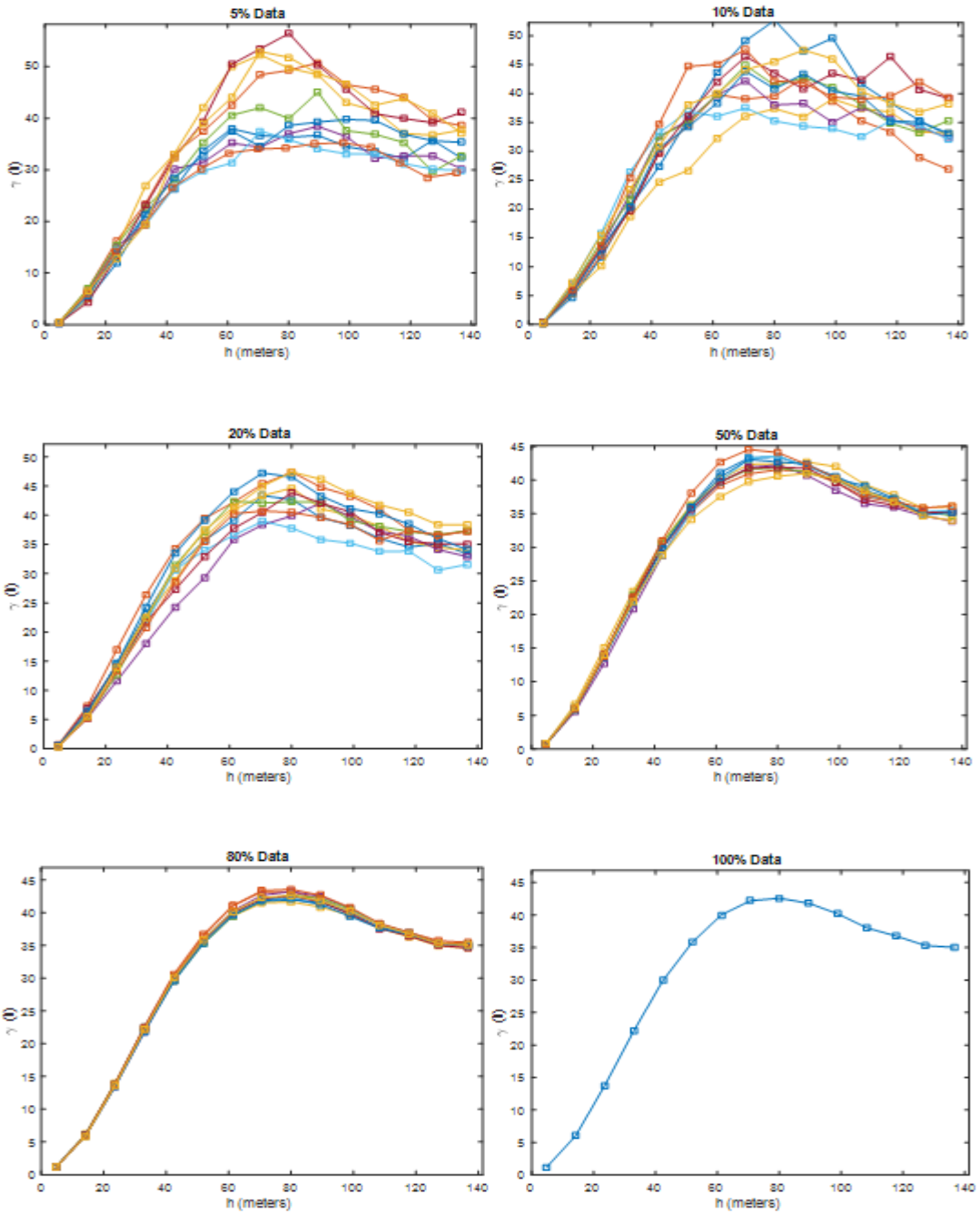

Figure 11. Comparison of Semivariograms Composed of 5%, 10%, 20%, 50%, 80%, and 100% of the Data in the Region of Rough Terrain.

As evidenced by the semivariograms, the number of data points used for the creation of the semivariogram does reduce the variation in semivariogram because there are fewer points available for sampling. The general shape of the semivariogram appears

to be consistent. As it relates to the range and sill of the semivariograms, it appears that the range remains consistent when the percentage of data points used is 20% or greater.

Figure 12 is a comparison of the semivariograms for the same percentages of data in a region of smooth terrain.

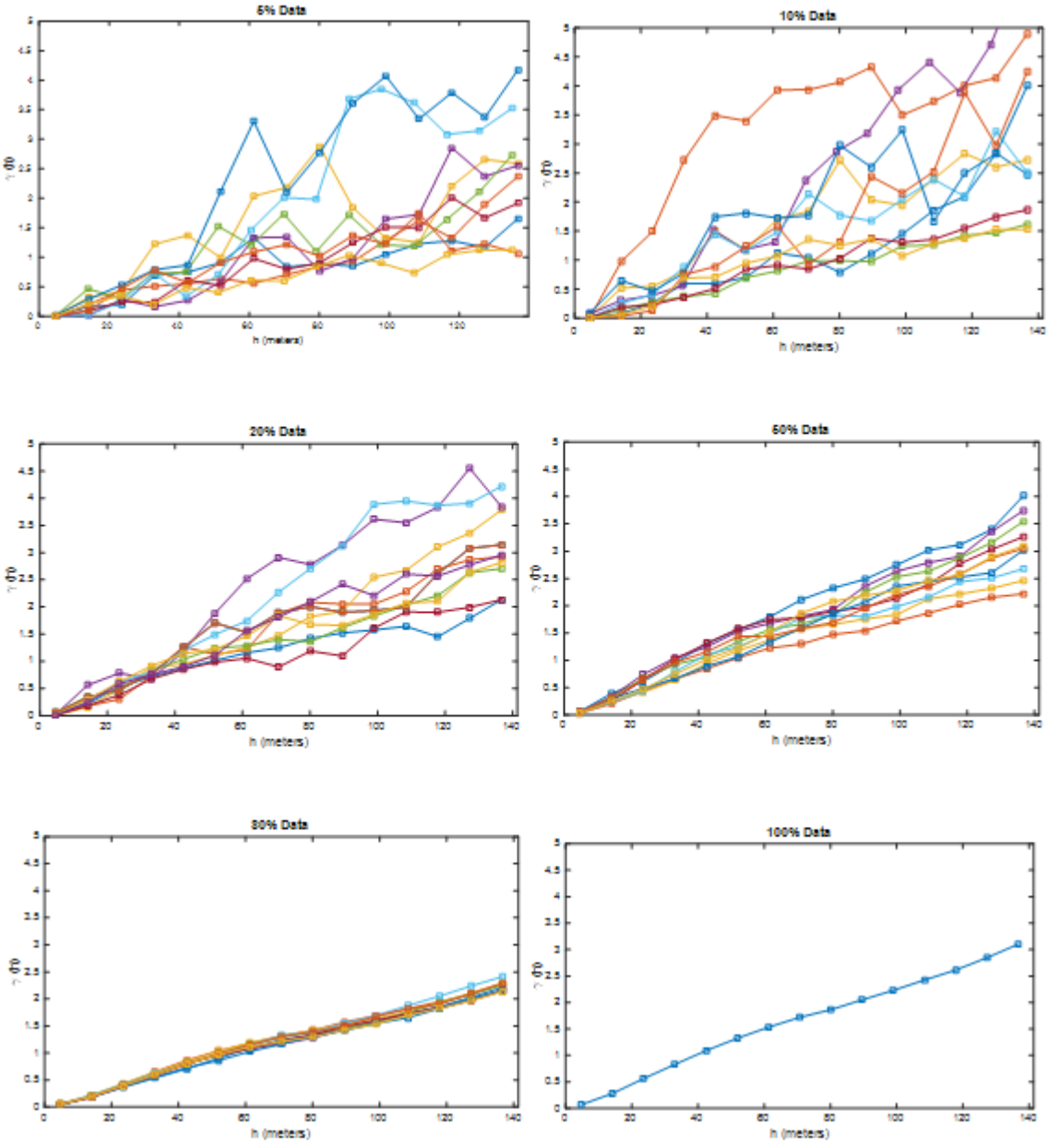

Figure 12. Comparison of Semivariograms Composed of 5%, 10%, 20%, 50%, 80%, and 100% of Data in Region of Smooth Terrain.

As observed in the region of rough terrain, the variation between semivariograms produced decreases significantly with the number of data points. At values less than 20%, the shape varies significantly. The range and sill are not as obvious in the semivariograms representing the region of smooth terrain, as the data exhibits a more linear behavior. Another difference observed in Figure 12 is that the peak value and the range at which it occurs (the range and sill) do not become consistent until above 50%.

B. COMPARISON OF CURVE FITTING

This section begins with a description of the process used to form semivariograms and the subsequent curve fitting for each of the models. A comparison of E-fit and traditional methods is conducted in two different terrain areas in Pavilion Lake. It concludes with the results for each model, an analysis of the results and how those results relate to OSE.

First, the five-meter resolution data from Pavilion Lake was imported into MATLAB. Using the contour plot, two areas were selected for application of the process to ensure its capability over different types of terrain. The first area, located in the mid-southern portion of the lake, consisted of rough terrain. The second area, in the northern portion of the lake, is a region of smooth terrain. Investigation began with the rough terrain.

Upon selection of the 200-by-200 meter area, the data is divided into k=10 sub-groups as necessary for the k-fold cross validation process. For every k^{th} iteration, the *variogram* function, an open source MATLAB function, was used to produce the empirical semivariogram with fifteen bins. For the spherical, exponential, and Matérn models, the *variogramfit* function, another open source MATLAB function, was used to fit a curve to the data points.

The first step toward testing each model is creating the semivariogram of the 200-meter grid in the region of rough terrain, dividing the 1,681 data points into fifteen bins with the *variogram* function. Bins are simply the number of segments the distance between points will be grouped into. So if the maximum distance between points was 100 meters and five bins were desired, the first bin would be 0 to 20 meters, the next 20 to 40 meters, etc. The default for the *variogram* function is 20. Fifteen bins were used for all

semivariograms created for this thesis. The second step was to fit a curve to the semivariograms. MATLAB's v*ariogramfit* function, for the Matérn, spherical and exponential based models. The E-fit model was presented at the end of Chapter III.

1. Semivariograms in a Region of Rough Terrain

The top-left image in Figure 13 depicts the 15 bins from the *variogram* function as red squares and the output of v*ariogramfit,* using the spherical formula, as a blue line that turns red at the range of the semivariogram. The image on the upper-right shows the same process except that the blue line is a result of the exponential formula, here no range is shown. This is a characteristic of the exponential model. The lower-left and right images depict the curve fit using the Matérn and E-fit models, respectively. It is easy to see here that the E-fit is able to more accurately fit a line to the data points.

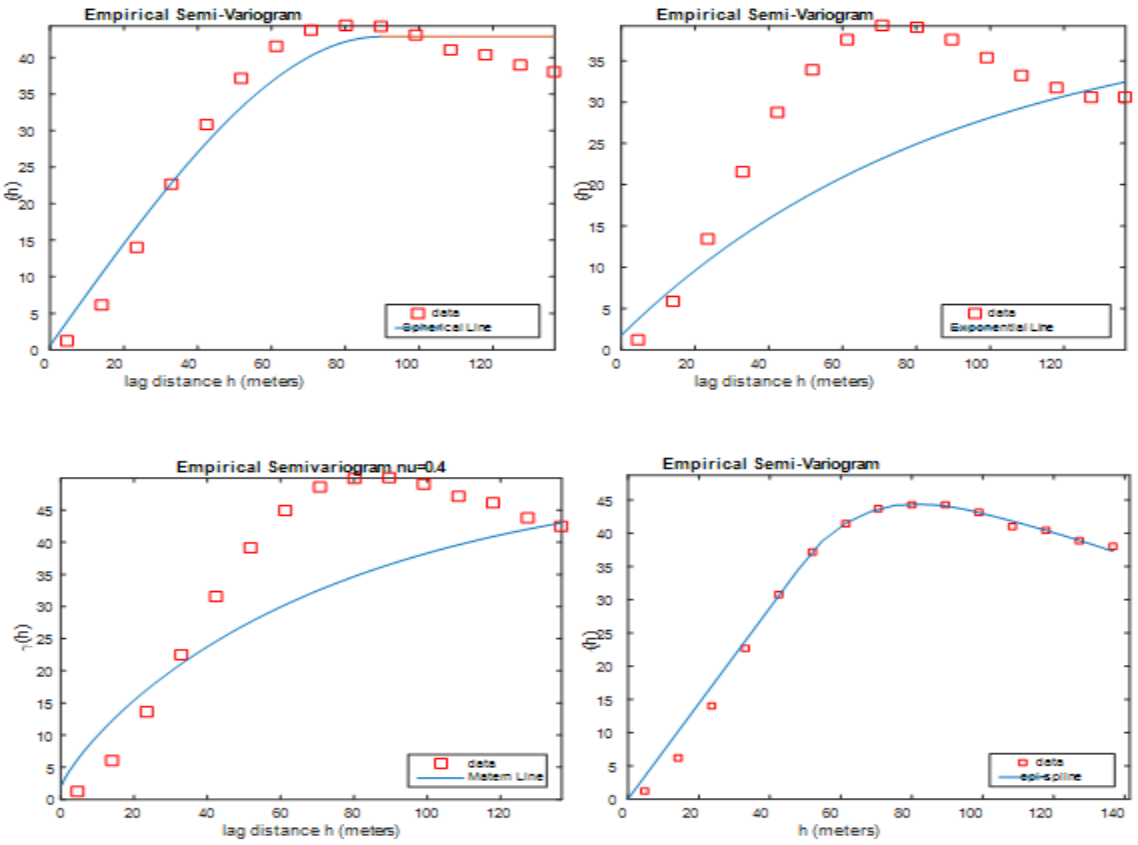

Figure 13. Semivariograms Produced Using the Spherical, Exponential, Matérn with ν=0.4, and E-fit Based Semivariograms for Rough Terrain.

Looking at the spherical model, the sill and range for the semivariogram are much more obvious. The exponential and Matérn models result in the least accurate curve fitting to the data. Specifically, the exponential and Matérn models do not depict an asymptotic sill because those models assume that there is no distance at which two points are unrelated. Thus, the sill for the exponential and Matérn models appears to occur at the maximum value. On the E-fit based semivariogram, the range and sill are clearly depicted and the curve fits the data much more closely. Somewhat counter-intuitively, the decline that occurs subsequent to the sill suggests that data points at those distances become more useful or their similarity increases. This information is important because it shows that data points beyond the range can be used in the map building process.

To provide a measure of the improvement realized by the E-fit over the Matérn model, the commonly used L2 norm was calculated for every iteration of the simulation as well. The L2 norm is a method of evaluating what the error or difference between the approximation of the Matérn model and the approximation of the E-fit model is. The L2 norm (also known as two-norm, mean-square norm or least squares norm) is basically minimizing the sum of the square of the differences between the target values (epi-spline values) and the estimated values (Matérn values) [20]. The spherical, exponential, Matérn, and E-fit models were applied to a random subset of the 200-meter sub-region of rough terrain. Each subset consisted of 90% of the 1,681 data points and each model was tested 100 times. Upon simulation of the semivariogram, the L2 norm comparison was calculated for each of the 100 simulated semivariograms created by the Matérn and E-fit models. The average over the 100 simulations was 2056.8 in the region of rough terrain. In terms of the difference in total area between the two curves, 2056.8 indicates a big improvement by the E-fit model.

2. Semivariograms in a Region of Smooth Terrain

Figure 14 depicts the spherical, exponential, Matérn and E-fit based semivariograms in the region of smooth terrain. These images represent one iteration of the model however they are indicative of the performance of each model.

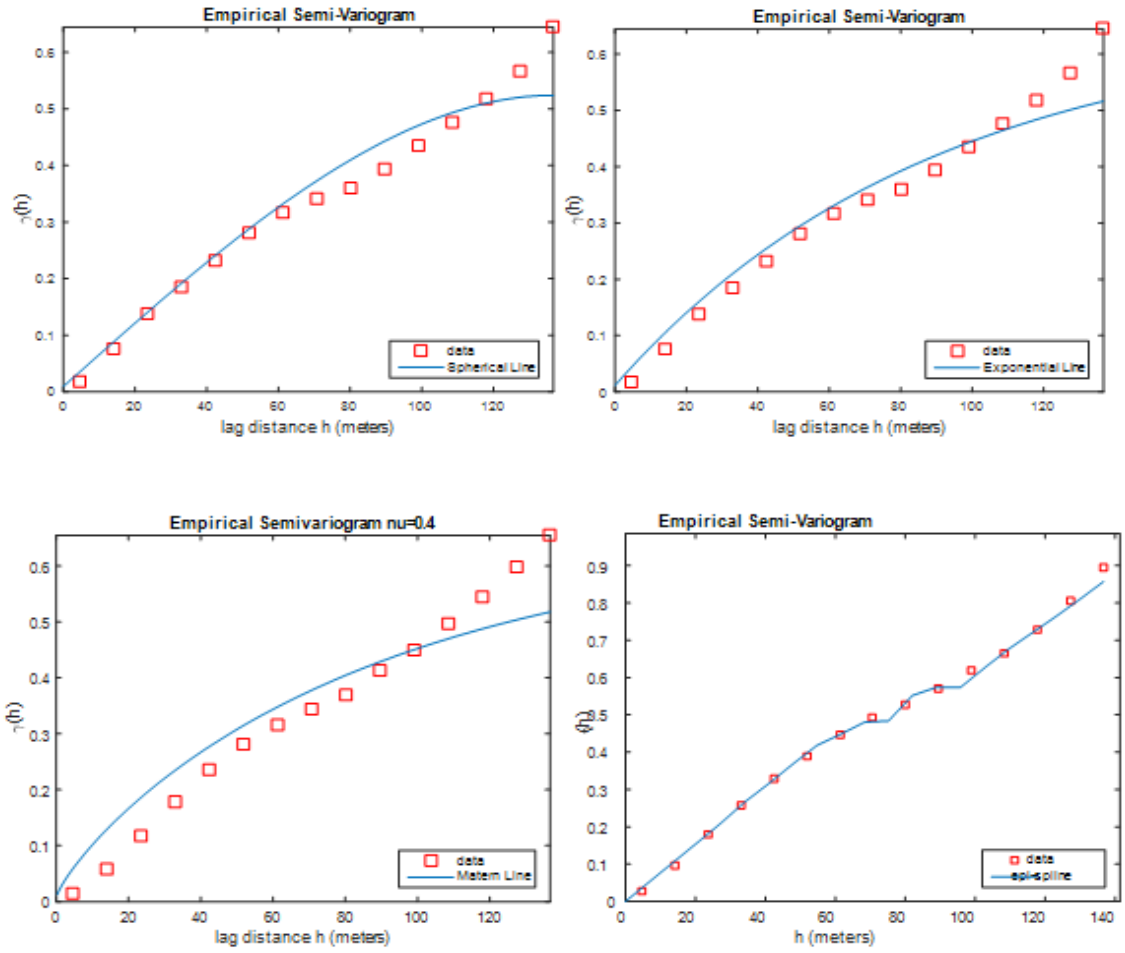

Figure 14. Semivariogram with Spherical, Exponential, Matérn, and E-fit Based Curve Applied in Region of Smooth Terrain.

As noted in the semivariograms for the region of rough terrain, the semivariogram based on the spherical formula clearly shows the range and sill for the data but does not fit the data as well as the E-fit based semivariogram. In this region, the terrain elevation was near constant so the data points, represented by the red squares, behaved in a linear fashion. Due to the linear behavior of the data in this region, the exponential and Matérn models are able to perform much better compared to the semivariograms they produced in the region rough terrain. Despite the improvement, the E-fit model again demonstrates an improved ability to better model the data. Again, the L2 norm test was conducted to determine what improvement was realized by the E-fit compared to the Matérn model. The average over 100 simulations was 59.73. This improvement achieved by the E-fit

model in this region is much smaller because the Matérn model is able to more closely fit the linear behavior of the data. Nonetheless, it still indicates an improved ability to model the data by the E-fit model.

The information presented in this section demonstrates the improvements E-fits offer. As noted in [21], before attempting to fit a periodic function to a set of data points, the user should ask what evidence there is to support the underlying phenomenon being investigated. While this makes sense, it highlights the advantage of E-fit because of its ability to incorporate such evidence into the model via soft constraints.

C. KRIGING RESULTS

Upon creation of the semivariograms and the completion of the curve fitting process, 10% of the data is randomly chosen and set aside as validation data. The 200-meter grid is then broken into 20-meter square sub-regions, resulting in 100 20-meter squared sub-regions. An assumption associated with this step is that each 20-meter sub-region is quasi-stationary. Six random data points were then selected from each 20-meter sub-region and ordinary kriging was used to calculate the midpoint of each square.

For each 20-meter sub-region, comparison is made between the validation data and the mid-point estimates within the region. This resulted in 32000 comparisons for each model applied to the 200-meter region. A histogram of the differences between the estimated midpoint and the validation data was then created. The range, MAE, and MSE are calculated to provide additional analysis. This process is then repeated employing the remaining models for curve fitting. The entire process, using the Matérn, spherical, exponential, and E-fit based models is then applied to the region of smooth terrain.

1. Model Application to a Region of Rough Terrain

Upon creation of the semivariogram for the 200-meter sub-region and removal of the validation data, the remaining data was divided into 20-meter sub-regions. Kriging was then used to produce an estimate of the midpoint for each 20-meter sub-region. This procedure was replicated for the spherical, exponential, Matérn, and E-fit models. The total number of simulations for each model was 100. Figure 15 is a side-by-side

comparison of a 3D plot of the original 200-meter region, broken into 20-meter squares, using the known measurements and a 3D plot of the same 200-meter region and the associated predictions using the four models. The plots in Figure 15 represent just one of the 100 simulations performed.

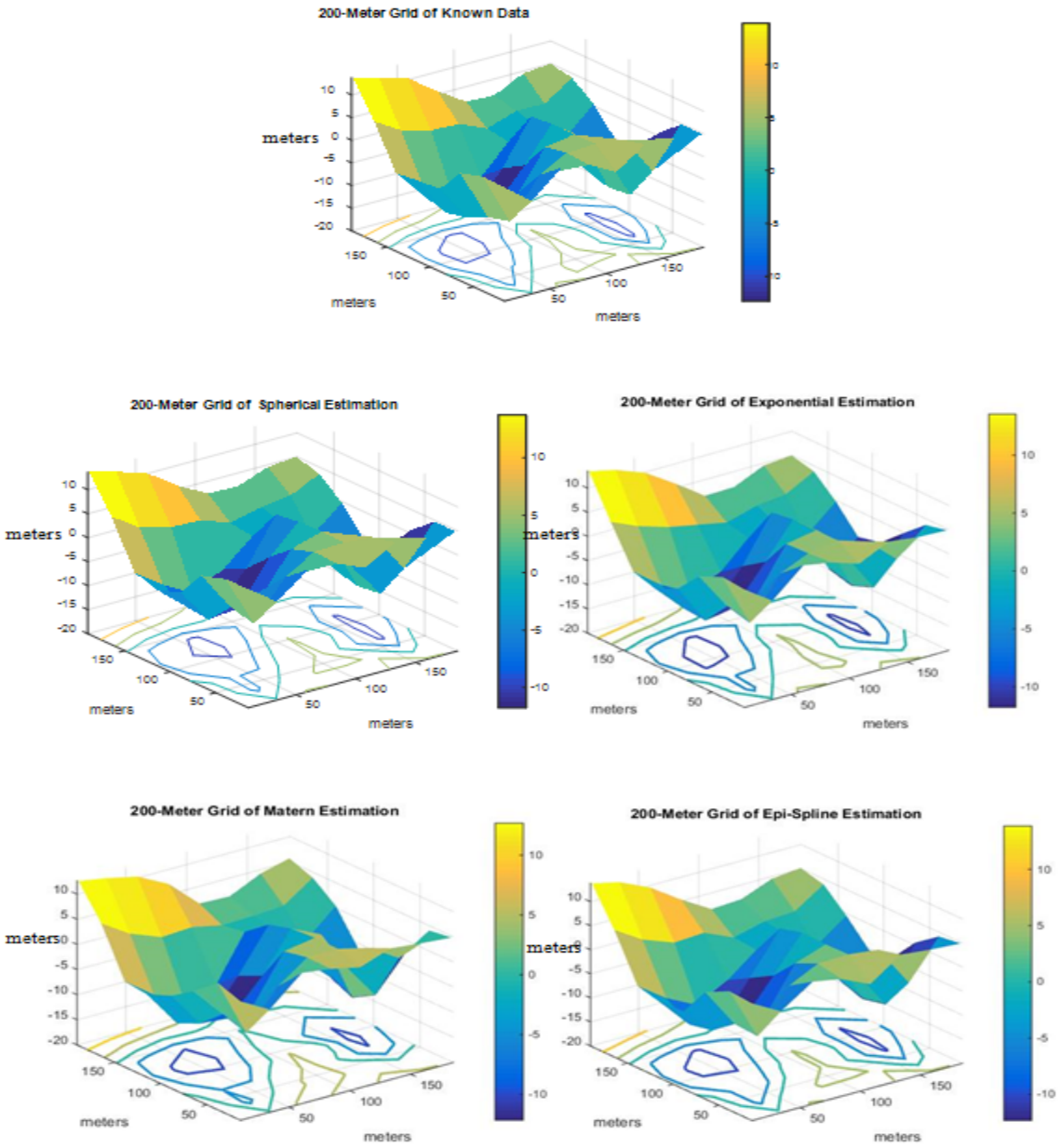

Figure 15. Comparison of Known Data Grid and the Resultant Estimation Produced by the Spherical, Exponential, Matérn, and E-fit Based Models in the Region of Rough Terrain.

These simple plots show that there is variation between the four models. The E-fit model appears to be more accurate over the entire region compared to the other models. Table 1 is a comparison of the MAE, MSE, and range between the models. The histograms in Figure 16 show a comparison of the models. The histograms represent the distribution of the error between the mid-point estimation and the data points set aside for validation within the same 20-meter sub-region. Each histogram represents just one of 100 simulations. A Gaussian distribution is superimposed onto the histograms for illustrative purposes only.

The data in Table 1 demonstrates the quantitative improvement by the E-fit based model across the three statistics. This information supports the idea that the accuracy of semivariograms directly improves the resultant kriging estimations. In terms of meters and in the context of the accuracy of a map used of underwater navigation, these numbers are valuable improvements. The histograms demonstrate the distribution of the errors and how closely they fit a Gaussian distribution.

Table 1. Comparison of MSE, MAE, and Range for Spherical, Exponential, Matérn, and E-fit Based Models in Region of Rough Terrain for 100 Replications.

	Spherical	**Exponential**	**Matérn**	**E-fit**
MAE	1.741	1.742	1.836	1.738
MSE	2.378	2.385	2.523	2.366
Range	14.573	14.489	15.747	14.242

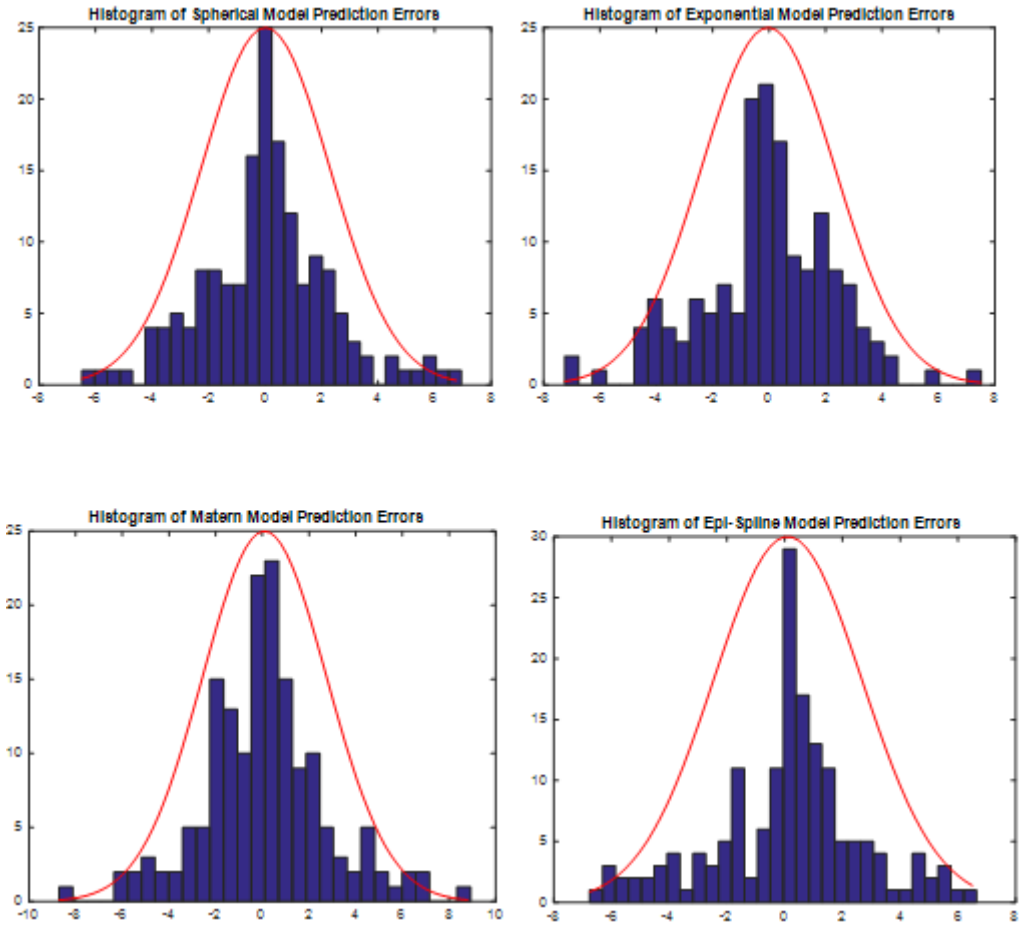

Figure 16. Histogram of Prediction Errors for Spherical, Exponential, Matérn, and E-fit Based Models.

2. Model Application to a Region of Smooth Terrain

The simulation was next applied to the region of smooth terrain and again employed the spherical, exponential, Matérn, and E-fit models. Figure 17 is a comparison of 3D plots of the 200 meter region and 3D plots of the estimation produced by the models.

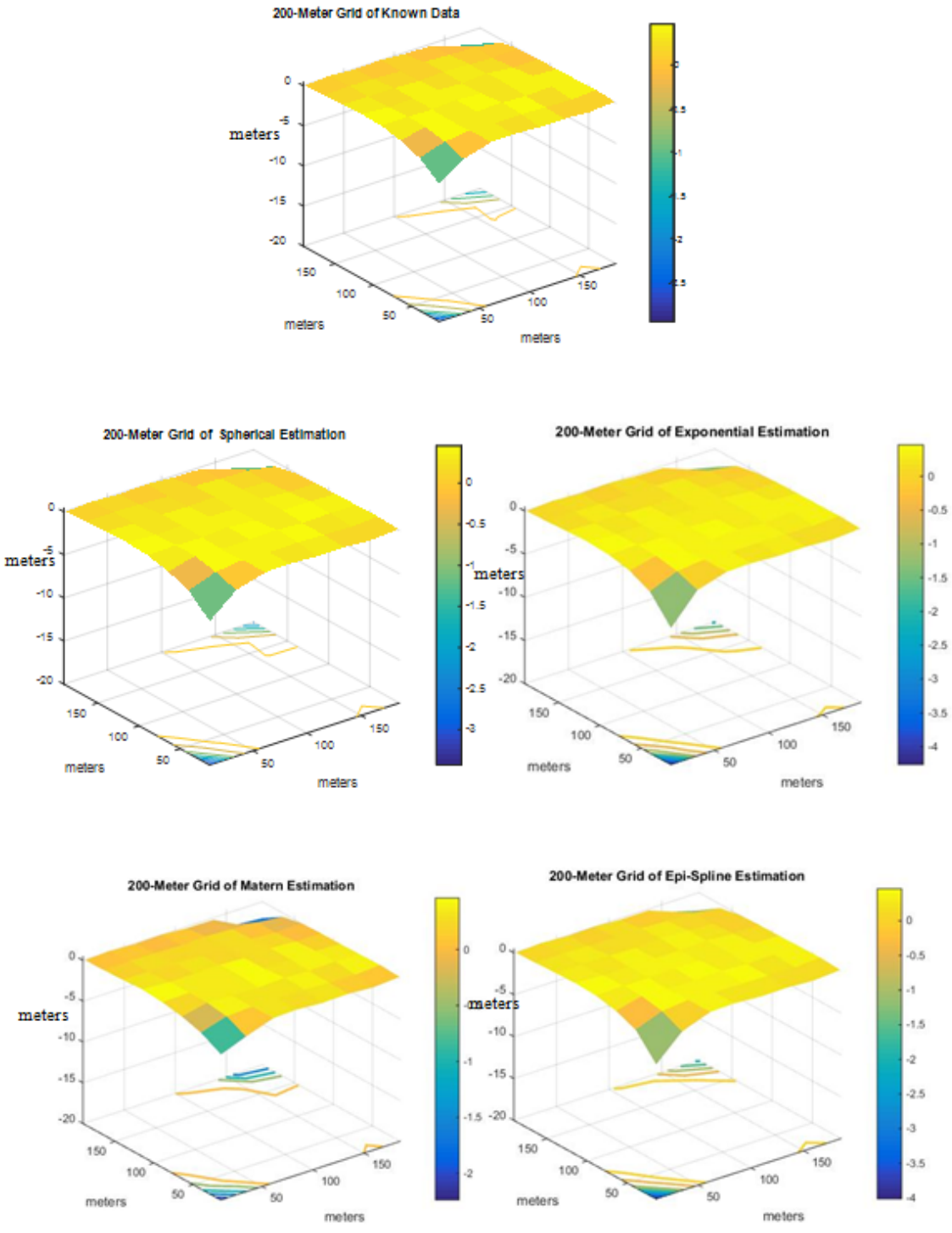

Figure 17. Comparison of Known Data and Spherical, Exponential, Matérn, and E-fit Based Estimation in Region of Smooth Terrain.

As was the case in the region of rough terrain, the images in Figure 17 indicate similar performance between the models. Visually, the E-fit model does a better job of estimation across the entire region compared to the traditional models. Table 2 represents the MAE, MSE, and range for the four models as it relates to the region of smooth terrain. Figures 18 and 19 present the histograms of the resultant errors between the predicted midpoint and the validation data for the four models. These histograms demonstrate the errors from one iteration of the simulation. Again, a Gaussian distribution is superimposed for illustrative purposes.

Table 2. Comparison of MSE, MAE, and Range Between Spherical, Exponential, Matérn, and E-fit Based Models in Region of Smooth Terrain for 100 Replications.

	Spherical	**Exponential**	**Matérn**	**E-fit**
MAE	0.172	0.169	0.185	0.168
MSE	0.412	0.407	0.451	0.396
Range	4.920	4.846	5.193	4.782

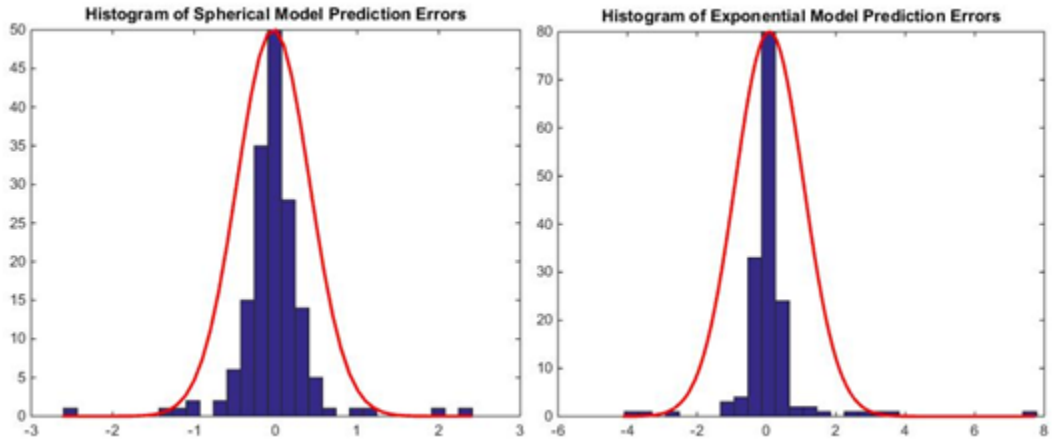

Figure 18. Histogram of Errors from Spherical and Exponential Based Models in Region of Smooth Terrain.

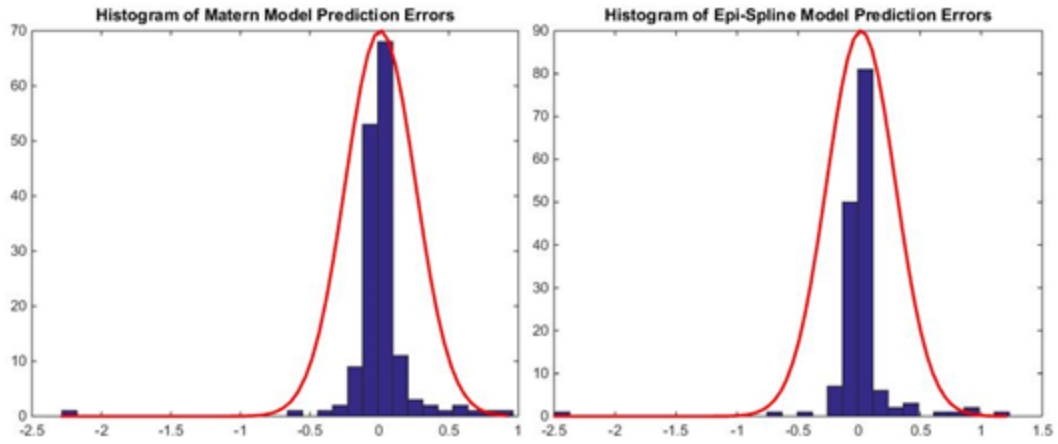

Figure 19. Histogram of Prediction Errors from Matérn, and E-fit Based Models in Region of Smooth Terrain.

The information in Table 2 demonstrates the improvement produced by the E-fit model across the three statistics. The histograms demonstrate that the errors associated with the predictions may be summarized as a Gaussian distribution and are centered at zero. Compared to the region of rough terrain, all of the models show improved performance in this region. It is also important to realize that the E-fit models are able to improve upon the resultant estimations and mapping process when the covariance is less desirable than traditional models.

This data supports the idea that E-fits used to model empirical semivariograms are an improvement to traditional methods. In some cases, the covariance of the E-fit based model is worse. However, the improvement is seen in the mean error and the resulting map. The improved map can be utilized to improve TAN.

VI. CONCLUSIONS AND RECOMMENDATIONS

As the interest in navigation without access to GPS continues to grow, the topic of OSE will continue to be investigated and improved upon. Operations in environments void of GPS signals necessitate the continued improvement upon technologies like TAN. As the data has shown, E-fit is one method of improving upon the OSE process which will result in an improvement in maps created for TAN. It is not suggested that the proposed model is the end all solution, rather, E-fit appears to be just one piece of the puzzle.

As demonstrated by the traditional models of curve fitting, goodness of fit does not necessarily guarantee improved results. However, the E-fit model made improvements over all metrics investigated within this thesis. This improvement can be seen both visually and quantitatively. It is evident that if the relationship of the data is as depicted by the semivariogram produced by E-fit, even in cases where the covariance is worse, it is still a better model. It all depends on the confidence of the semi-variance relationship, which is a function of the semivariogram model, which as demonstrated, is a function of terrain roughness

An important lesson learned while investigating this thesis relates to mesh size. Increasing the number of mesh-points provided an improved capability of E-fit to accurately model the data, which was expected based on the assertions in [12]. E-fit, as a non-parametric procedure of curve fitting are an improvement upon the parametric models used as comparison for this thesis.

It is recommended that follow on research investigate multiple questions that were unable to be explored by this thesis. First, an attempt to use the proposed model on an AUV/UUV for real time operations should be made. Such an attempt should make an effort to simultaneously create the map and perform TAN. Next, an investigation into the assumption of quasi-stationarity and its validity when doing real time operations is important. After all, there is no grid on the ocean floor. A recommendation would be to utilize some type of tree or quad-tree type of structure for this process. It is also recommend that a data set of greater resolution be employed. Such a data set would require more powerful computing hardware and time than were available for this thesis.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

[1] Unmanned systems integrated roadmap FY2013-2038, (n.d.). Dept. of Defense. [Online]. Available: http://www.defense.gov/Portals/1/Documents/pubs/DOD-USRM-2013.pdf. Accessed: Jun. 03, 2016.

[2] What is kriging. (n.d.). Kriging.com [Online]. Available: www.kriging.com/what iskriging.html. Accessed Jun. 01, 2016.

[3] M. Hammond, S. Houts, & S. Dektor. (2012, Oct. 30). Terrain relative navigation [Online]. Available: https://web.stanford.edu/group/arl/projects/ terrain-relative-navigation

[4] D. K. Meduna, S.M. Rock, & R.S. McEwen. (n.d.). Closed-terrain relative navigation for AUVs with non-inertial grade navigation sensors [Online]. Available: http://web.stanford.edu/group/arl/cgi-bin/drupal/sites/default/files/public/publications/MedunaRM_2010.pdf. Accessed Jul. 20, 2016.

[5] N.A.C. Cressie, *Statistics For Spatial Data*, Revised ed. Hoboken, NJ: John Wiley and Sons, 1993.

[6] J. P. Chilés & P. Delfiner, *Geostatistics: Modeling Spatial Uncertainty*, 2nd ed. Hoboken, NJ: John Wiley and Sons, 2012, pp. 1–192.

[7] G. Bohling. (n.d.). Kriging [Online]. Available: http://people.ku.edu/~gbohling/cpe940/Kriging.pdf. Accessed Jul. 18, 2016.

[8] G. Bohling. (2005, Oct. 17). Introduction to geostatistics and variogram Analysis [Online]. Available: http://people.ku.edu/~gbohling/cpe940/Variograms.pdf

[9] V. Gandhi, (n.d.). Semivariogram modeling—key readings [Online]. Available: http://www-users.cs.umn.edu/~gandhi/courses/CS8701/g4_e2_semi variogram.pdf. Accessed Jul. 06, 2016.

[10] M. L. Stein, *Interpolation of Spatial Data, Some Theory for Kriging.* New York: Springer, 1999.

[11] J. O. Royset, N. Sukumar, and R. J-B. Wets " Uncertainty quantification using exponential epi-splines," in *Proceedings of the International Conference on Structural Safety and Reliability*, New York, NY, 2013.

[12] J. O. Royset and R. J. B. Wets. (2014). From data to assessments and decisions: epi-spline technology. INFORMS tutorial, INFORMS [Online]. pp. 27–53. Available: http://pubsonline.informs.org/doi/abs/10.1287/educ.2014.0126

[13] G. G. Brown and R. F. Dell, (n.d.). "Formulating integer linear programs: A rogues' gallery," Faculty publication, Operations Research Department, Naval Postgraduate School, Monterey, CA, Apr. 2006.

[14] M. K. Chung. (2007, Feb. 1). Introduction to Random Fields. [Online]. Available: http://www.stat.wisc.edu/~mchung/teaching/MIA/theories/randomfield.feb.02.2007.pdf

[15] P. Abrahamsen. (1997, Apr.) . A review of Gaussian Random Fields and Correlation Functions. Available: https://www.nr.no/en/nrpublication?query=/file/917_Rapport.pdf

[16] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press, 2006, pp. 12–198.

[17] A. G. Journel and Ch. J. Huijbregts, *Mining Geostatistics.* Caldwell, NJ: Blackburn Press, 1991, pp. 32–35.

[18] Using interpolations and geostatistics for spatial modelling. (n.d.). University of the Western Cape [Online]. Available: http://planet.botany.uwc.ac.za/nisl/GIS/spatial/chap_1_30.htm. Accessed Jul. 18, 2016.

[19] C.V. Deutsch, "Correcting for negative weights in ordinary kriging," *Computers & Geosciences*, vol.22, no. 7, pp 765–773, 1996.

[20] A. Ron, (2010, Oct. 26). CS412: Introduction to numerical analysis, Lecture 14: Linear algebra, [Online]. Available: http://pages.cs.wisc.edu/~amos/412/lecture-notes/lecture14.pdf

[21] R. Webster and M. A. Oliver, *Geostatistics for Environmental Scientists*, 2nd ed. Hoboken, NJ: John Wiley & Sons, 2007, pp. 99.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California