A Model for Placement of Modular Pump Storage Hydroelectricity Systems

by

Joseph Warren Rosenbery II

A Thesis Presented to the
Faculty of the USC Graduate School
University of Southern California
In Partial Fulfillment of the
Requirements for the Degree
Master of Science
(Geographic Information Science and Technology)

December 2019

To Kelly and Olivia, I love you.

Table of Contents

List of Figures vii

List of Tables x

Acknowledgments xi

List of Abbreviations xii

Abstract xiii

Chapter 1 Introduction 1

  1.1. Project Goal 3

  1.2. Project Workflow 5

  1.3. Report Organization 6

Chapter 2 Background 8

  2.1. The Electrical Grid and Renewable Energy 8

  2.2. Hydroelectric Reservoirs and Pump Storage 10

  2.3. Overlay Analysis for Suitability Modeling 12

    2.3.1. Boolean Overlay Analysis 13

    2.3.2. Weighted Overlay Analysis 15

    2.3.3. Fuzzy Overlay 17

  2.4. Applications of Overlay Analysis in Renewable Energy Siting 19

Chapter 3 Modeling Framework 24

  3.1. Model Objective 24

  3.2. Engineering Requirements 24

  3.3. The Primary and Secondary Models 28

  3.4. Modeling Products 30

  3.5. Modeling Context 31

    3.5.1. Modeling Environment 31

3.5.2. Spatial and Temporal Scale 32

Chapter 4 Model Description 34

4.1. Data Requirements 34

4.1.1. Primary Model Data Requirements 34

4.1.2. Secondary Model Data Requirement 37

4.2. Primary Model Processes 37

4.2.1. Construction Area Analysis 38

4.2.2. Refining the Search Area 40

4.2.3. Extracting a subset of the DEM 41

4.2.4. Relief Calculations 42

4.2.5. Filtering Suitable Locations 44

4.2.6. Matching Points 44

4.2.7. Filtering Connections 45

4.2.8. Filtering Reservoir Locations 48

4.2.9. Model Results 49

4.3. Secondary Model Processes 50

Chapter 5 Case Study: Model Processing, Outputs, and Evaluation of Results 52

5.1. Preliminary Steps 52

5.1.1. Processing the DEM 53

5.1.2. Binary Screen Creation 54

5.1.3. Restricted Lines 55

5.2. Intermediate Results 56

5.2.1. Construction area identification 57

5.2.2. Creating the Subset DEM 61

5.2.3. Searching for Relief 62

        - 5.2.4. Making Connections 64
    - 5.3. Final Primary Model Results 65
    - 5.4. Secondary Model 67
    - 5.5. Model Performance 71
        - 5.5.1. Primary Model 72
        - 5.5.2. Secondary Model 82
        - 5.5.3. Alternative Study Areas 84
- Chapter 6 Conclusions 89
    - 6.1. Model Assessment 89
    - 6.2. Uses 94
        - 6.2.1. Viability in Market 94
    - 6.3. Future Work 95
- References 97
- Appendix A – Primary Model Process Table 101

List of Figures

Figure 1 – Basic Model Design 3

Figure 2 – A graphic flowchart showing the basic model design workflow 6

Figure 3 – The traditional topographically constrained configuration for pump storage hydroelectricity systems. Source: ClimateTechWiki.org 2006. 11

Figure 4 – Esri Fuzzy Membership Function Plots (Esri 2016a)... 18

Figure 5 – A simplified diagram showing the engineering parameters for relief and maximum lateral runout used in the demonstration of the Primary Model. 25

Figure 6 – The reservoir site leveling scale concept... 27

Figure 7 – The relationship between slope angle and cut/ fill volume needed to create a flat construction surface. 28

Figure 8 – A graphical depiction of the conceptual Primary Model.. 29

Figure 9 – A graphical depiction of the conceptual Secondary Model.. 30

Figure 10 – The portion of the model flow chart showing the steps to determine areas suitable for construction. 39

Figure 11 – A graphical depiction of the slope processes of construction area analysis............. 40

Figure 12 – Modeling process for construction area identification. 41

Figure 13 – The portion of the model flow chart showing the steps to create a new subset DEM dataset.. 42

Figure 14 – Model processes for upslope and downslope relief calculations.............................. 43

Figure 15 – The portion of the model flow chart showing the steps to pair upper reservoir location with lower reservoir locations.. 45

Figure 16 – The portion of the model flow chart showing the steps to remove connections between the upper reservoir and lower reservoir locations that cross named streams. 46

Figure 17 – A graphic depiction of conditions identified as false and real matched connections where relief must be at least 300 m. 47

Figure 18 – The portion of the model flow chart showing the steps to identify and eliminate false match connections. 48

Figure 19 – The portion of the model flow chart showing the steps to eliminate upper and lower reservoir locations that do not have a connection after the connection filtering steps 49

Figure 20 – The portion of the Secondary Model flow chart showing the steps that apply the values of the fuzzy ranking dataset to the upper and lower reservoir locations. 51

Figure 21 – The processed DEM for Los Angeles County used in the model example. 54

Figure 22 – Areas not for placement consideration (left) and *Binary Screening Layer* (right) 55

Figure 23 – *Restricted Lines* dataset for Los Angeles County 56

Figure 24 – Slope in degrees for Los Angeles County 57

Figure 25 – The raster dataset showing the maximum slope for each nine-cell neighborhood in the study area. 58

Figure 26 – The areas identified by the model as being suitable for the placement of a terminal reservoir. 59

Figure 27 – The areas identified as suitable for construction of a terminal reservoir after the application of the binary screening layer. 60

Figure 28 – The subset DEM created by using the areas deemed suitable for construction of a terminal reservoir to extract elevation values from the study area DEM 62

Figure 29 – Intermediate datasets showing the maximum downhill (left) and upslope (right) relief within 1,500m search radius. 63

Figure 30 – Preliminary upper and lower reservoir location. 64

Figure 31 – A summary of final Primary Model outputs for Upper and Lower Reservoir. Locations exaggerated for visual effect 65

Figure 32 – A selection of final Primary Model results, La Canada Flintridge. 66

Figure 33 – The *Fuzzy Layer* for Los Angeles County. 69

Figure 34 – A selection of final model results near La Crescenta, California, after the integration of the fuzzy layer. 70

Figure 35 – Histograms showing the distribution of fuzzy membership values for upper and lower reservoir locations. The red vertical line on the histograms indicates the mean membership value for each dataset. 71

Figure 36 – Distribution of connection lengths 73

Figure 37 – Distribution of point to point relief magnitude 74

Figure 38 – A sample of Primary Model Results for upper and Lower reservoir locations......... 75

Figure 39 – Model results near Palmdale, California. 76

Figure 40 – Example results from the Primary Model outputs near Universal City, Los Angeles County, California. Note there is only one upper reservoir site indicated.......... 77

Figure 41 – A graphical example of the one to many relationships between upper and lower reservoir locations 78

Figure 42 – Complex relationships in clustered reservoir locations.. 79

Figure 43 – Results filtered to originate from a single upper reservoir location 81

Figure 44 – Distribution of alternative study areas in California 85

List of Tables

Table 1 – Binary Criteria for Siting Wind Turbines (after Sparkes and Kidner 1996) 13

Table 2 – Esri Raster Overlay Tool Summary (Esri 2016c) 16

Table 3 – Fuzzy Overlay Functions (Esri 2016b) 18

Table 4 – Modified AHP results (Uyan 2013, 11-17) 20

Table 5 – Suitability criteria used to determine pump storage conversion potential for existing hydroelectric facilities. 22

Table 6 – An example portion of the attribute table for one of the Secondary Model outputs (*n=40,868)* 51

Table 7 – Variables used to generate the fuzzy dataset used in the Secondary Model for the Los Angeles County case study 67

Table 8 – Summary of Primary Model Results for all Study Areas 86

Acknowledgments

First, I would like to thank my advisor Dr. Karen Kemp for all of her support and flexibility while I worked through this project. I was glad that we were able not only to interact as professor and student but through the thesis writing process as well.

I would also like to thank my thesis committee, Dr. John Wilson and Dr. Robert Vos, for their input, encouragement, and support.

I would like to sincerely thank Dave Guis and Andy Wallace and the rest of the team at Wallace Kuhl and Associates for supporting me during my tenure at USC and affording me the opportunity to further my education. I would also like to thank Kurt Balasek for giving me the idea for this research project and running through ideas with me.

Finally, I would like to thank my family: my Dad for always believing in whatever I wanted to pursue; my wife Kelly, for putting up with all of the time I spent staring into the computer and still giving me her unwavering support, and my daughter, Olivia for keeping me smiling.

List of Abbreviations

3DEP	3D Elevation Program
CGS	California Geologic Survey
DEM	Digital Elevation Model
GIS	Geographic information system
GISci	Geographic information science
MPSHS	Modular pump storage hydroelectricity System
SSI	Spatial Sciences Institute
USC	University of Southern California
USGS	United States Geologic Survey

Abstract

As the global energy market pushes toward the further development and integration of renewable energy and reduced reliance on fossil fuels, the energy industry has looked to innovative solutions to solve the shortcomings of green energy production. Diurnal fluctuation in electrical production potential in solar and wind sources creates a need to develop ways to store surplus energy resources for later deployment. Pump storage hydroelectricity, in which surplus energy is used to pump water uphill to recharge a hydroelectric reservoir, holds a great deal of potential when used in conjunction with other types of renewable energy. This report documents the design and development of a two-phase analytical spatial model that identifies suitable locations for the placement of paired top and bottom terminal reservoirs of a modular pump storage hydroelectricity system (MPSHS). The first phase of the model applies user-defined search criteria to identify locations for the construction of terminal reservoirs that meet the relief and lateral run distance requirements. Further refinement of results from the first modeling phase using secondary information can be used to rank suitable locations based on user-supplied environmental, economic, and socio-demographic constraints and preferences. This thesis presents details of model function as well as case study results for Los Angeles County.

Chapter 1 Introduction

With global concern over the deleterious effects of climate change, there has been a concerted effort to generate a more significant percentage of electrical energy from renewable sources, such as hydroelectricity, wind and solar, reducing the production of harmful greenhouse gas emissions (Rosenberg 2008). To address this concern, the state of California has set agressive goals concerning electricity and emissions. In 2018, the California State Senate passed Senate Bill 100, which states that energy production in the State of California should achieve net-zero greenhouse emissions by 2045 by focusing efforts on the development of renewable energy sources (California Senate 2018).

Temporal fluctuations in the potential availability of conditions favorable to green energy production such as wind and solar mean that green energy supply is often out of phase with consumer demand. Without a way to store energy produced by renewable energy sources during their peak productivity, fossil fuel generation must be available to meet the demand when it exceeds the renewable energy production potential (Kaplan 2009).

Hydroelectric power is one suitable remedy. In 2017, it supplied nearly 17% of the electrical power to the world by storing water in such a way that it can be used later for the generation of electricity. However, financial, political, and practical constraints on the construction of new water reservoirs for hydroelectric energy production largely hinder the further development and exploitation of the resource (USGS 2018).

Traditional hydroelectric power generation facilities are a one-way system, meaning all the water passing through the system moves downstream. In these systems, the electricity also flows in only one direction, outward to the consumer (USGS 2018). Thus, the generating potential is finite as the water used to turn the turbine-generator has to come from upstream in the watershed and be stored in the reservoir (Madani, Guégan, and Uvo 2014, 153-163). The Hoover Dam on the Colorado River is an example of this type of system.

One method that has been developed to store renewable energy for use on-demand is pump storage hydroelectricity (Yang and Jackson 2011, 839-844). At its root, pump storage hydroelectricity is a concept built on the more traditional hydroelectric model, where water stored in a reservoir is released through a conduit (penstock) which feeds a turbine connected to a turbine generator. However, pump storage systems have a unique attribute; they can move water back uphill.

Pump storage hydroelectricity is a method by which the finite storage capacity is augmented by the ability to pump water back to a higher potential energy state, ready for reuse in the energy production cycle. While pumping requires energy, it can be supplied on demand, only when there is unused and otherwise wasted energy available. With the implementation of pump storage hydroelectricity production, the overall grid not only becomes less reliant on more traditional fossil fuel-based forms of energy conversion but also provides alternatives to those working to improve grid stability (Rehman, Al-Hadhrami, and Alam 2015, 586-598).

The research reported in this document supports efforts to identify potential solutions to the problems currently facing the renewable energy market. How can these systems provide the energy needed on demand without relying on traditional fossil fuel facilities for energy production when environmental conditions limit the capabilities of renewable sources? Modular Pump Storage Hydroelectricity Systems (MPSHS) can take advantage of the benefits afforded by the traditional pump storage hydroelectricity model while avoiding the environmental and political constraints of their larger counterpart. This type of technology has the potential to work in tandem with other types of renewable energy products to reduce harmful greenhouse gas emissions and move toward energy sustainability.

1.1. Project Goal

The goal of this research project was to develop a GIS model package within an Esri ArcGIS ModelBuilder application that can be used by project developers and engineers (end-user) in support of a decision-making process to determine the placement of MPSHS. This model provides a preliminary assessment tool for identifying locations ideal for terminal reservoirs of MPSHS. The model has two components. The Primary Model explores terrain within the designated study area to identify locations suitable for the construction of reservoir tanks. The Secondary Model builds on the Primary Model outputs by assigning aggregate suitability values to each potential reservoir location using fuzzy logic datasets provided by the end-user (Figure 1).

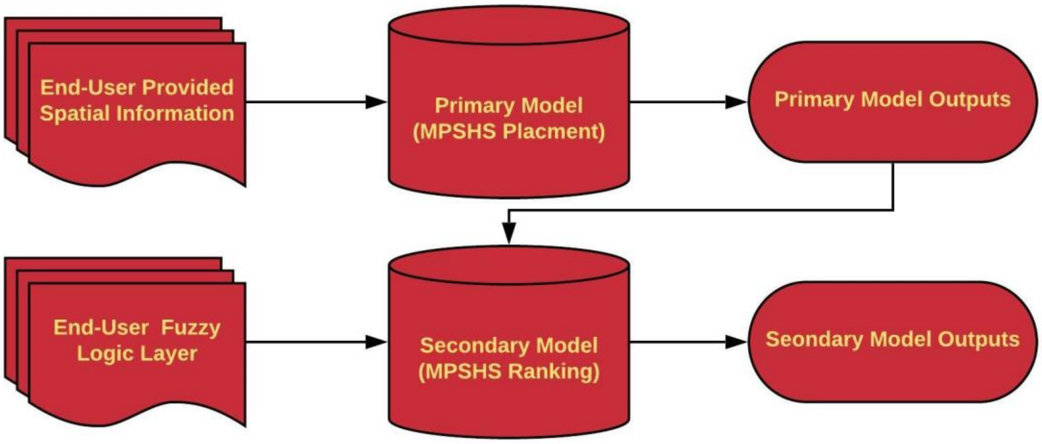


Figure 1 – Basic Model Design

Using physical parameters provided by design engineers in the form of the constraints of 1) minimum head requirements, 2) maximum lateral run distances and 3) minimum tank footprint size, topographically suitable terminal reservoir locations are identified by the Primary Model. Accommodations for additional environmental variables that effectively eliminate areas known to be unavailable for development have been incorporated into the model as two strategic sets of variables acting as areal or linear prohibitions to construction.

The first set of environmental variables produce a binary screening layer compiled from end-user provided datasets that identify areas known to be unsuitable for construction. For example, the areas within the boundary of a National Park are likely not suitable for construction. The binary screening data must be converted into a binary screening raster in which areas that are suitable are coded 1 and areas that are not suitable are coded 0. All such areal features are identified in advance by the model end-user and combined into one raster dataset for use as a single model variable. This process is designed to fine-tune suitable location identification and reduce the processing load.

The second set of environmental variables is used to assess the viability of the reservoir connections. One of the fundamental components of the MPSHS is the connection, called a penstock, that serves as the conduit from the upper reservoir to the lower reservoir. There are many natural and man-made linear features that serve as continuous barriers to the construction of penstocks, such as large streams and roadways. After the model has established the complete set of possible connections between the upper and lower reservoirs, the model searches for connections that cross linear features that cannot be crossed by a penstock and removes them from inclusion in the Primary Model results.

Finally, an optional Secondary Model, further enhancing the products of the Primary Model, provides the capability for the end-user to develop and apply their own additional suitability layer using fuzzy logic, enabling further refinement in site selection capabilities.

As a core design requirement, this model was developed to use freely and widely available spatial data. This approach allows the end-user of the model to gather all the required data necessary to run the model to completion with minimal investment in time and capital with respect to data procurement.

The datasets produced by the Primary Model include two 30m raster datasets, one for each of the upper and lower reservoirs of the MPSHS, and a vector dataset of lines representing viable connections between paired reservoirs. The Secondary Model applies the fuzzy logic raster dataset to the two reservoir location datasets and produces a set of points at the center of all raster cells that are suitable reservoir locations, each point attributed with its aggregate fuzzy membership value.

1.2. Project Workflow

To accomplish the goals set forth in the section above, the model was developed using an interactive approach common to projects of this type (Figure 2). After the development of the research question, background research was conducted to understand better the environment within which the problem was set. Following the research phase, a conceptual model was developed for the study that incorporated the primary design elements required by project engineers. With this basic understanding of the project goals, the models were developed and evaluated through a versioning process that allowed for assessment through incremental progress. The final product was then tested in multiple geographic locations to evaluate model processes for issues caused by spatial and topographic variability, thus allowing for further refinement of the final model.

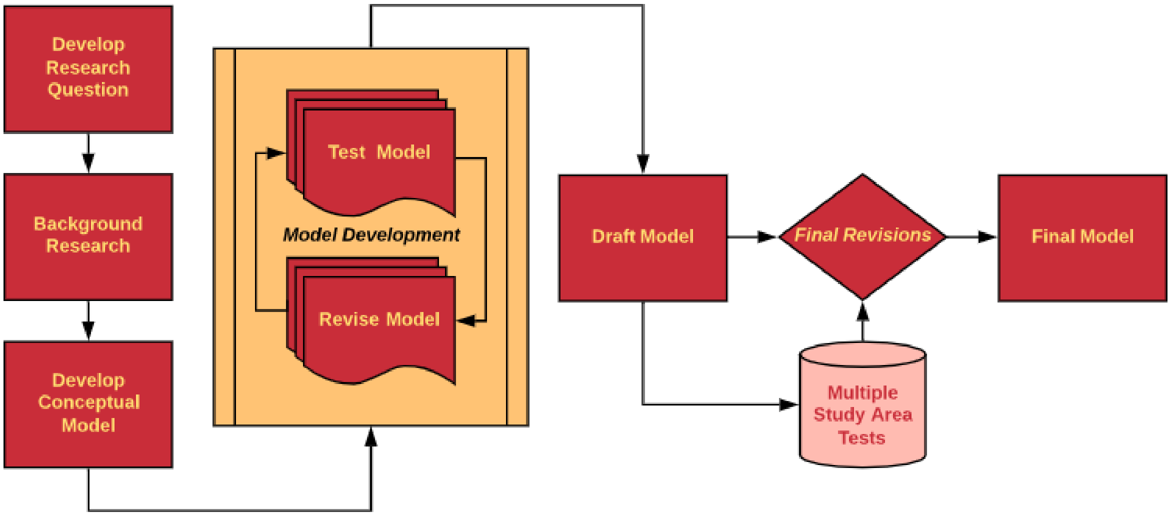


Figure 2 - A graphic flowchart showing the basic model design workflow

1.3. Thesis Organization

This document is laid out as follows. First, Chapter 2 explores the basic components and systems of the electrical grid in the United States including the incorporation of green energy, the role that different types of renewable energy play in the energy production network, and some of their significant disadvantages. Second, pump storage hydroelectricity is explored as a method for recovery and storage of renewable energy as potential energy. Finally, the role of GIS and suitability modeling in identifying locations for renewable resource projects is examined.

The modeling framework is discussed in Chapter 3. This includes a detailed examination of the conceptual model on which the computational model was developed. The conceptual model details the engineering requirements governing the MPSHS, the project constraints, a description of model design, and the primary model outputs. Additionally, Chapter 3 discusses the modeling context employed, including the modeling environment and relevant spatial and temporal scales. This chapter also discusses the project design footprint and the advantages of using county boundaries as the preferred study area limits.

Chapter 4 describes the structure of the model. Beginning with data gathering, this section outlines the required data, processing steps, and a walkthrough of the processes comprising the Primary Model developed for this research. The next section details the Secondary Model components and processes.

While this model was developed over multiple study areas, Chapter 5 walks through the model steps using Los Angeles County as an example modeling unit. First, the inputs to the model are explored prior to model execution. The subsequent sections include a walkthrough of the intermediate process output datasets and final model outputs for both the Primary and Secondary models. A detailed evaluation of the final model outputs for both the Primary and Secondary models completes the chapter.

Chapter 6 provides conclusions in the form of an overview of the model, its performance, limitations, and ultimate usefulness. This chapter also provides details of opportunities for continued research into this topic and examines the applicability of techniques developed to other applications such as recreation, transportation, and engineering.

Chapter 2 Background

This chapter provides background information, setting the context for the research presented. Beginning with an overview of the consumer electrical grid and the role of renewable energy, the context of this chapter focuses on a discussion of the applications of pump storage hydroelectricity. Additionally, this chapter provides a review of the application of GIS modeling to the development of renewable energy infrastructure.

2.1. The Electrical Grid and Renewable Energy

The electrical grid in the United States is comprised of multiple components whose role is to generate electricity from a variety of sources and distribute electricity to consumers. The collection of systems that comprise the electrical grid is the byproduct of a multitude of small systems that were built to meet the needs of local and regional customers. Over time, these small systems have grown together to create the modern grid, which is comprised of three major interconnected units. These cover the western states and western Canadian provinces (Western Interconnection), the eastern states and eastern Canadian provinces (Eastern Interconnection), and most of the state of Texas (ERCOT Interconnection) (Kaplan 2009).

With the rise of public and political awareness surrounding the need for increased efficiency concerning energy generation and consumption, the U.S. is developing a technology-driven grid management system currently being integrated into the existing system, called the Smart Grid. Goals of the Smart Grid program allow for increased efficiencies in the demand/ supply curve, provide tools for end-user management, improve quality and reliability, and enable the incorporation of renewable energy sources into the grid (Heirman 2012).

The U.S. electrical grid works on a demand-supply routine, which means energy production must match demand and be available when demand for energy is high. The fluctuation of demand on the grid is called load cycling. Thus, the supply of electrical energy fluctuates continuously to meet the current demands of the electrical distribution grid. For some types of renewable energy, diurnal fluctuations in generation potential may be temporally displaced from that of the load cycle (Denholm et al. 2010).

Solar energy is a prime example of this effect as the peak generating potential is in the middle of the day. The influence of solar energy on the grid produces a challenge for grid managers due to its inconsistent contribution to the system. This is particularly difficult to manage when solar energy is approaching the end of its diurnal cycle. At these times, demand is typically increasing while conditions favorable to solar energy production decrease. The implication is that alternative energy sources must compensate to stabilize the grid (California Independent System Operator 2016).

One solution to the problem of load balancing is the application of energy storage on the grid that can be tapped as needed (Suul, Uhlen, and Undeland 2008). The development of viable long-term energy storage solutions could affect the broader implementation of renewable energy by 20% (Benitez, Benitez, and van Kooten 2008). Current grid technology supports the momentary deficit in electrical supply and captures oversupply with large capacitors that provide temporary supply and storage, which allows for supply to match the demand curve (Chu and Majumdar 2012). When the demand is displaced temporally from the renewable source peak load cycle, the supply typically comes from traditional fossil fuel sources. To solve the electrical grid storage problem, the widespread implementation of pump storage hydroelectricity facilities has been a cost-effective and efficient tool to supply on-demand energy when needed (Denholm et al. 2010; Chu and Majumdar 2012).

2.2. Hydroelectric Reservoirs and Pump Storage

Traditional hydroelectric energy facilities use water stored in a reservoir to generate power by releasing water through a turbine which turns a generator. This technology is widely used around the world and in 2017 accounted for approximately 17% of energy production worldwide (USGS 2018). These systems not only produce clean, reliable energy, but they also serve as storage reservoirs, allowing the rationing of water for agricultural and domestic consumption, providing recreational space, and protecting large populations against devastating floods. For all the benefits to be gained from building reservoirs and equipping them with hydroelectric generation facilities, large topographically-constrained reservoirs can have a deleterious effect on the environment (The National Geographic Society 2011).

There are also many reasons that building a new reservoir can be problematic. Large-scale reservoirs have a high initial investment cost and can lead to extensive habitat loss and truncation of riparian fisheries (USGS 2018). Downstream of reservoirs, the rivers are starved of sediment crucial to natural habitats, and natural flow regimes are disrupted. Modification to a naturally-regulated system in equilibrium can cause channel incision and irreparable damage to the ecosystem (Pasternack, Wang, and Merz 2004).

The primary limitation for broader implementation of traditional pump storage systems is the project scale. These traditional systems typically have a large, topographically-constrained reservoir and a lower elevation impoundment for retaining water to be pumped up for reuse (Rehman, Al-Hadhrami, and Alam 2015). Furthermore, many large reservoir projects are confronted with fierce public opposition (Napier, Carter, and Bryant 1986). Environmentalist organizations such as Friends of the River and the Sierra Club regularly form protests, work with state and federal lobbyists, and file lawsuits in opposition of new reservoir construction making the challenge even greater (Friends of the River 2018). While the proposed infrastructure required for MPSHS would likely go through a rigorous public notification process prior to construction, similar tank constructions exist in many urban areas without significant public protest.

The concept of pump-storage hydroelectricity is based on the concepts of traditional hydroelectricity but includes a design for reversing flows to recharge the system (Figure 3) (Rehman, Al-Hadhrami, and Alam 2015, 586-598). Like traditional hydroelectric systems, when demand for energy on the grid becomes high, stored water in the upper reservoir flows down through a turbine generator and into a second, lower reservoir (USGS 2018). However, when energy production from other sources of green energy is higher than off-peak demands, the load can be balanced by activating the pump storage system, moving water back uphill, “recharging” the system for the next deployment (Rehman, Al-Hadhrami, and Alam 2015, 586-598).

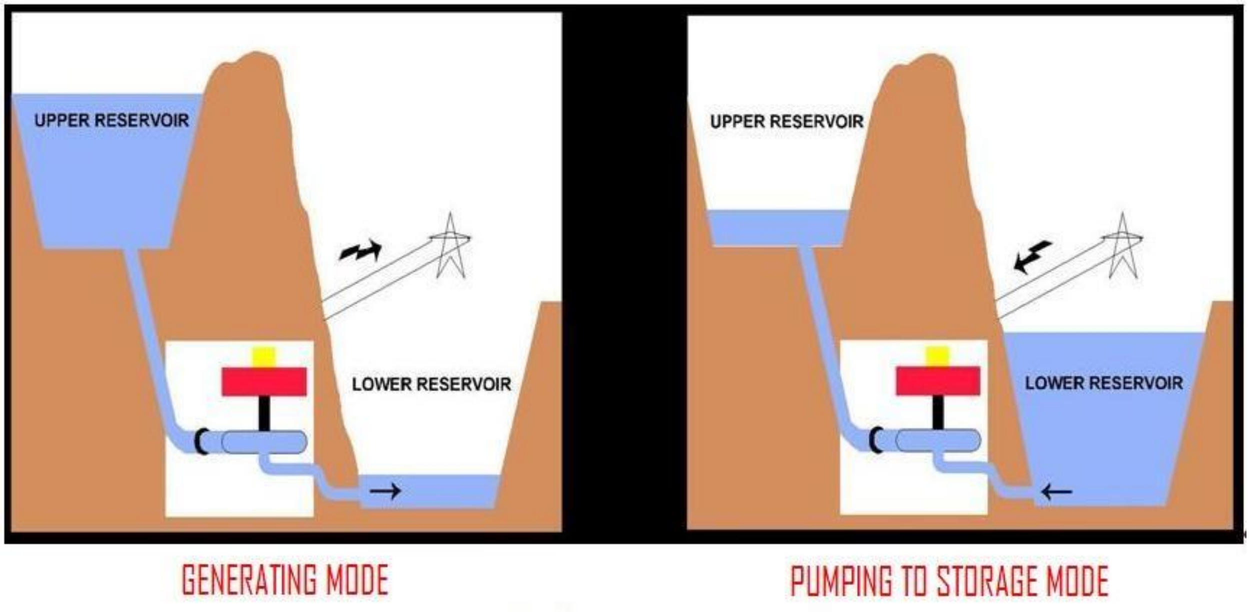


Figure 3 - The traditional topographically constrained configuration for pump storage hydroelectricity systems. Source: ClimateTechWiki.org 2006.

To support the further development of green energy as a viable solution to the use of fossil fuels, pump storage hydroelectricity works in tandem with renewable energy sources such as wind and solar (Yang and Jackson 2011). The use of pump storage hydroelectricity generation facilities complements other types of renewable energy sources by storing electrical energy as potential energy. The ability to store energy on the grid produced by a renewable source would dramatically increase the viability and eventually penetration of renewable energy technology (Bueno and Carta 2006).

2.3. Overlay Analysis for Suitability Modeling

The application of Geographic Information Systems (GIS) in land-use suitability analysis has been revolutionary to land use planning (Malczewski 2004). Furthermore, site suitability analysis using GIS has been a powerful tool for the discovery of suitable sites on which to place renewable energy infrastructure. At its most basic, site suitability models use environmental variables, assign them values for suitability, and combine those values to determine suitable locations (Mitchell 2012).

With a strong public, industrial, and political momentum behind the transition away from fossil fuels, there is a significant push toward understanding what role spatial sciences can play in renewable resource development and deployment. These proven renewable energy technologies have precise and well-understood criteria regarding suitable locations with which to design the most effective suitability models. Thus, site suitability is especially useful at siting wind and solar generating potential (Henning Sten Hansen 2005).

Fundamental in the suitability modeling framework, overlay analysis is the process of stacking spatial data and combining the layers to achieve a meaningful output value for all locations using data found to be relevant to the objective (Bolstad 2005). In GIS, an overlay can be accomplished with raster or vector datasets. In both cases, multiple datasets are combined through arithmetic processes to produce a single output dataset containing a conclusion from a combination of the inputs (O'Sullivan and Unwin 2010).

2.3.1. Boolean Overlay Analysis

As an example of the use of GIS-based suitability analysis for siting renewable energy facilities, Sparks and Kinder combined 19 spatial datasets using simple Boolean overlay processes to identify suitable locations for the constructions of wind turbine generators in England. Their approach used simple proximity buffers surrounding spatial phenomena that constrained the suitability of locations and assigned a simple binary qualifier for suitability, meaning any location in the study area was either suitable, or it was not (Sparkes and Kidner 1996). Table 1 below shows the criteria used by Sparkes and Kidner to identify suitable locations.

Table 1- Binary Criteria for Siting Wind Turbines (after Sparkes and Kidner 1996)

Feature	**Distance to Site must be greater than**
Airports	3 km
National Parks	1 km
National trust property	1 km
Military danger zone	3 km
Scenic area	1 km
Forest park	1 km
Built-up area	2 km
City centroid	5 km
An urban centroid	2.5 km
Town centroid	1.5 km
Small town or village centroid	1 km
Small village, hamlet or isolated settlement	750 meters
Lake, marsh or reservoir	250 meters
Motorway, A-road or B-road	300 meters
Railway	250 meters
River or canal	200 meters
Radio or TV mast	250 meters

Feature	**Distance to Site must be greater than**
'Picturesque' or scenic feature	1 km
Elevation	Site must be above 100 meters elevation

Creating the binary raster datasets shown above requires a multiple-step reclassification process. In the case of suitability, where proximity to a known feature is the primary determining factor in suitability, the distance must be determined for each criterion separately (Mitchell 2012). Whether data are provided in vector or raster format, a geoprocessing tool such as Esri's Euclidean Distance tool can be applied (Esri 2016e). This produces a raster dataset where each cell is assigned a value equal to the distance of its center point from the closest input feature (vector input) or non-null cell (raster input).

With this raster dataset indicating the distance values from the target phenomena, the distance raster can then be reclassified based on the suitability criteria. Esri's reclassification tool interrogates the dataset and reassigns each cell with a value based on the reclassification parameters defined in the tool (Esri 2016f). For example, if only the areas that are not within 200 meters of a river or canal are to be considered suitable, then the reclassification tool would assign all those values within 200 meters of a river or canal with a zero, a negative suitability response. Conversely, those areas outside of that buffer would be deemed suitable and assigned a value of 1, a suitable response (Mitchell 2012).

Combination of multiple binary datasets such as those outlined above by a multiplication method similarly returns a binary response. This means that the resulting overlay dataset is comprised of only those areas that returned a positive response for all of the considered criteria (O'Sullivan and Unwin 2010).

Beyond binary overlay, the other two most common types of overlay are the weighted overlay and the fuzzy overlay (Mitchell 2012). These methods move away from the Boolean determination and provide results on a gradational basis in terms of suitability. By using these more advanced methodologies, suitability can be graded, and suitable sites can be determined relative suitability criteria (O'Sullivan and Unwin 2010).

2.3.2. Weighted Overlay Analysis

Weighted overlay has been used successfully to support decision making in suitability modeling for the placement of renewable energy resources. The primary purposed of weighting is to leverage the relative influence of individual criteria against each other. Applications for these techniques have been applied in circumstances where many criteria are to be considered. This becomes especially critical when considering a hierarchy of importance with respect to criteria.

Aydin et al. (2013) used a weighted overlay to examine placement for hybrid renewable energy systems. Their research produced a complex configuration of environmental variables pertaining to both wind and solar. The overlay process accounted for the variables' influence on the generation potential for each energy type before combining the criteria to establish suitability for the hybrid system.

Weighted overlay uses the same techniques as the Boolean method; however, this data-driven approach assigns relative importance to each component considered in the overlay process. This process is referred to as indexing (O'Sullivan and Unwin 2010). This provides a better approach when looking at criteria that include conflicting attributes or objectives (Carver 1991).

The process by which multiple types of geospatial data are brought together in support of decision making is called the multi-criteria evaluation method. While MCE is useful for combining multiple datasets in suitability analysis, the process is complicated by not only the selection of the component feature set but also by how the criteria are weighted in the overlay process. Thus, the most critical component in weighted overlay analysis, also called weighted linear combination, is the determination of layer weights prior to the overlay process.

Weights represent the relative perceived importance of the components of the overlay (Carver 1991). The decisions regarding component weight in the overlay process are critical to the proper use of the method and are often misapplied; this is called the multicriteria decision-making problem. The relative importance of criteria should be founded in sound data-driven reasoning, not *ad hoc* estimation (Malczewski 2000). The methods for determining component weights should be determined through a process of data interrogation and evaluation. This process typically employs a process to determine a hierarchy of importance within the criteria considered.

Weighted overlay in GIS can be accomplished in multiple ways. Overlay tools provided by Esri include Zonal statistics, combine, weighted overlay, and weighted sum. Table 2 provides a brief description of each tool.

Table 2 – Esri Raster Overlay Tool Summary (Esri 2016c)

Tool	**Purpose**
Tool	Summarizes values in a raster layer by zones (categories) in another layer—for example, calculate the mean elevation for each vegetation category.
Zonal Statistics	Assigns a value to each cell in the output layer based on unique combinations of values from several input layers.
Combine	Automates the raster overlay process and lets you assign weights to each layer before adding (you can also specify equal influence to create an unweighted overlay).
Weighted Overlay	Overlays several rasters, multiplying each by their given weight and summing them together.

2.3.3. Fuzzy Overlay

Many objects have hard physical boundaries. For example, a building site has definite boundaries, and an electrical transmission line has a precise linear pathway. This defined boundaries approach is apparent in the Boolean overlay example presented above by Sparkes and Kidner (1996) where they could not consider areas within the buffer distance of the criteria selected. On one side of the buffer, the area is suitable, while on the other, it is not. The boundary is, therefore, defined as sharp.

However, other phenomena have attributes and extents that vary with respect to location and definition. The term for these conditions is Fuzzy. Fuzziness is a way of representing a gradational property of spatial phenomena, such as soil, which varies across its boundaries, transitioning from one type or category to another over a distance. In other words, the boundaries between values are not defined by a definite border (O'Sullivan and Unwin 2010).

In overlay analysis, fuzzy datasets provide criteria that have suitability values that change gradually with a change in location. For example, the suitability of a soil class for the construction of a building may vary as its attribute values move towards the center of the class' defined attribute range. Thus, as a location of interest moves away from the boundary between soil classes, fuzziness can be stated as a value related to the strength of membership in a suitability set related to distance from the boundary.

The fuzzy membership values can be determined by a membership function that examines the range of values in the fuzzy dataset and assigns a fuzzy value ranging from zero to one, where one has the highest membership. Common types of membership functions can assign high membership to low values, high values, or values centered around an ideal value. Esri's fuzzy membership tool contains a variety of membership functions for assigning membership

(Esri 2016a). Resulting raster datasets containing the fuzzy membership data are referred to as fuzzy set datasets. Figure 4 shows membership plots for these functions.

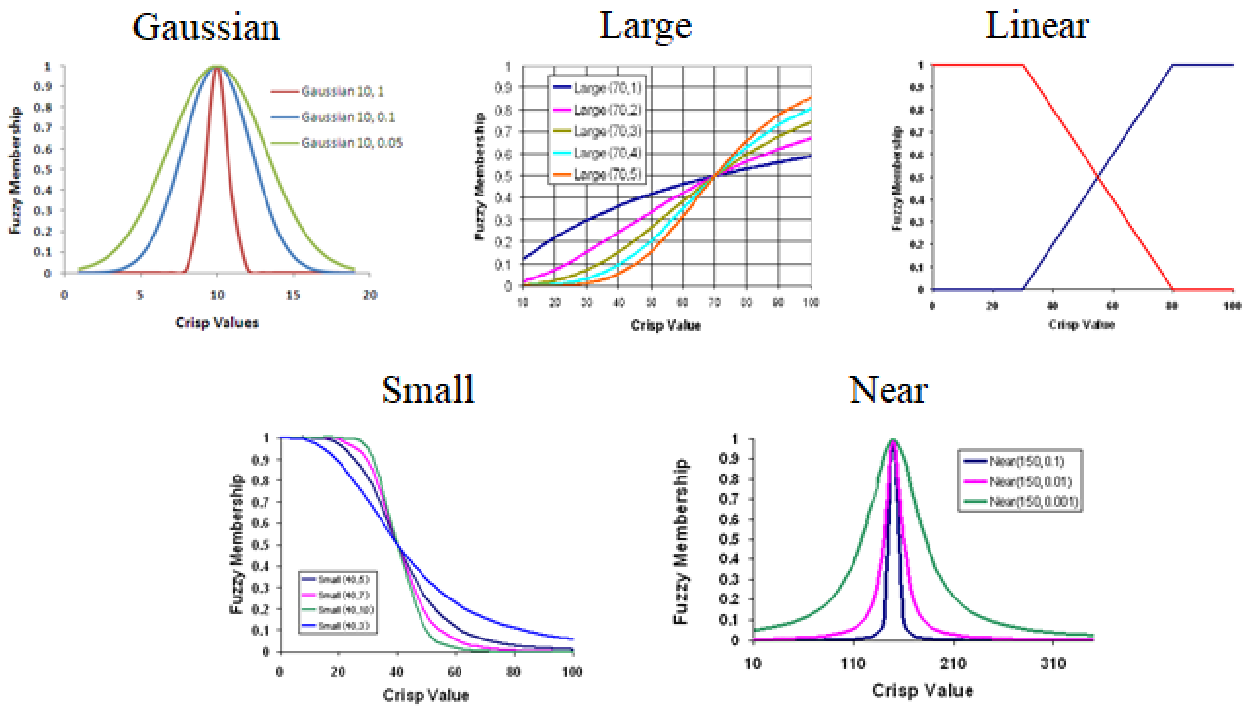


Figure 4 - Esri Fuzzy Membership Function Plots (Esri 2016a)

The fuzzy overlay is the process of combining the individual fuzzy sets based on a predefined method. As with the creation of the fuzzy sets, there are multiple ways to combine data using the fuzzy overlay tool. Table 3 identifies the methods and their applicability in the Esri fuzzy overlay tool (Esri 2016b).

Table 3 – Fuzzy Overlay Functions (Esri 2016b)

Type	**Function**	**Uses**
Fuzzy And	Returns the minimum value of the sets considered	Useful for determining the least common denominator for suitability criteria
Fuzzy Or	Returns the maximum value of the sets considered	Useful for identifying the highest membership value of the input criteria

Type	Function	Uses
Fuzzy Product	Returns the product of the values considered	Not often used, this value will be lower than all of the individual input criteria, as each variable will be a fraction of the full membership (1)
Fuzzy Sum	Returns the sum of the values considered	Not often used, this value will be a linear combination of all of the input criteria. A sum of the inputs of a fuzzy overlay will not necessarily produce a dataset where the highest membership values are the most universally suitable.
Fuzzy Gamma	Returns an algebraic product of the fuzzy sum and the fuzzy product, both raised to the power of gamma	This can be used to produce a value typically intermediate to that of the product and sum overlays. However, this is a compromise function between those two approaches.

2.4. Applications of Overlay Analysis in Renewable Energy Siting

According to the United States Energy Information Administration (EIA), renewable energy production accounted for approximately 11.5 % of the total energy consumed in the U.S. Of that 11.5 %, solar (0.95 %), wind (2.5 %), and hydroelectricity (2.7%) make of more than half (6.5%) (United States Energy Information Administration 2019). The remainder of the renewable market is divided between biomass energy (5.1%) and geothermal (0.2%).

Of the primary renewable energy sources currently utilized for production, solar, wind, and hydroelectricity have the most direct application scenarios for Site suitability analyses. Each of these types of energy infrastructure has their own unique circumstances to which suitability modeling can be approached and, in every case, multiple criteria need to be evaluated in order to develop a functional operation successfully.

Solar projects on a commercial scale have enormous structures that often cover huge tracts of land. MCE has been shown to be an increasingly crucial component of site suitability analyses for siting solar projects. However, as previously mentioned, the relative importance of each criterion considered in the analysis is crucial to successful implementation. Uyan applied the analytical hierarchy process (AHP) to help determines the relative importance of each component used in the study (Uyan 2013).

The study used five criteria in the suitability analysis. These criteria included the distance from residential areas, land use, distance from roads, slope, and distance to transmission lines. Additionally, binary constrains as buffers were placed around residential areas, roads, hydrologic features, and environmentally protected areas. An essential consideration in this study is that because of the uniform distribution of solar energy potential in the study area; it was not considered as an important criterion in the suitability analysis.

The criteria were processed using AHP, a mathematical method used in multicriteria decision making. In the AHP analysis, each criterion is compared pairwise with all other criteria, eventually determining the influence on the outcome, to which each criterion contributes by assigning an unbiased weight to each criterion. Table 4 presents a modified list of criteria and weights based on the AHP analysis from this study.

Table 4 – Modified AHP results (Uyan 2013, 11-17)

Criteria	**Weight**
Distance to Residential	0.14
Land Use	0.41
Distance to Roads	0.03
Slope	0.08
Distance to Transmission Lines	0.34

These results indicate that of the five criteria used in the study, weights were distributed asymmetrically. Weights show that the most influential component of the analysis was land use, then the distance to transmission lines. The remainder of the criteria accounted for roughly 25% of the remaining influence, with the distance to roads being weighted at just 3% (Uyan 2013).

Aydin applies a similar approach to suitability modeling for wind generating facilities but employs the use of fuzzy logic criteria. First and foremost, wind production must be an evaluation of wind energy potential, which was developed as a fuzzy set as wind energy varies across space. The second set of criteria were developed using clearly defined environmental objectives. These objectives included distance limitation s from nature reserves, residential areas, and airports, in addition to habitat concerns (Aydin, Kentel, and Duzgun 2010).

Wind energy was quantified into a fuzzy dataset which needed only modification into a fuzzy set for the purposes of integration into the model, where the higher potential for wind resulted in higher membership values. The remainder of the objective environmental datasets had to be transformed into fuzzy sets based on membership functions. In each case, a linear membership function was used to assign membership, such that criteria showed zero membership until the value was at least half of the distance limit for that particular criteria. Thereafter, a linear membership was assigned until reaching the distance limit, where all values exceeding that limit were granted full membership (Aydin, Kentel, and Duzgun 2010).

Through a process of multicriteria decision making for analysis of the environmental variables, two different outcomes resulted. One operation produced a worst-case scenario, and the other produced the best-case. This was accomplished by using different fuzzy overlay operators. Fuzzy “and” resulted in a scenario where locations were ranked by the lowest valued objective. Fuzzy “or” results in locations ranked by their highest value objective.

Each of the two environmental impact scenarios was then combined with the wind energy potential membership set to produce two, scenario-based final evaluations of suitability. The final step in the study was to validate the model determined suitability values by assessing the values at the locations of existing wind farms in the region. Aydin et al. concluded their model successfully uses the presented criteria to estimate the locations with high potential for use as locations for generating wind energy (Aydin, Kentel, and Duzgun 2010).

Suitability modeling can be used for identification of areas suitable for placement of new energy infrastructure, but it can also be used to evaluate existing infrastructure. In the research presented by Fitzgerald, GIS is used to explore the potential for the conversion of traditional unidirectional hydroelectric facilities into pump storage hydroelectric generating facilities (Fitzgerald et al. 2012).

The criteria considered for this study focused on topographic analysis surrounding existing dams and focused on the potential for placement of a lower reservoir within 5 kilometers of the dam in an area meeting the design criteria. Lower reservoir location was selected based on average slope and elevation over the area. Additionally, there needed to be at least 150 m of head between the dam and the selected lower reservoir location. The parameters and constraints used in this model are described in Table 5.

Table 5 – Suitability criteria used to determine pump storage conversion potential for existing hydroelectric facilities.

Transformation, Topography & Physical Characteristics	
Minimum volume of existing reservoir	1,000,000 m³
Maximum distance between existing reservoir and potential lower reservoir site	5 km
Minimum head	150 m
Maximum slope of second reservoir area	5°
Assumed minimum of new, second reservoir surface area	70,000 m²
Minimum distance from new reservoir to inhabited sites	500 m
Minimum distance from new reservoir to existing transportation infrastructure	200m
Minimum distance from new reservoir to an UNESCO site 5 km	5 km

The model outputs identify all of the reservoirs where an area exists that meets the criteria set forth in Table 4 and identifies an areal footprint for the lower reservoir. The research shows potential for the use of GIS in suitability analysis surrounding preexisting hydroelectric generation facilities and the conversion to pump storage capabilities. However, the authors acknowledge that the limitations to the construction of large reservoirs in the environment remain a challenge.

Chapter 3 Modeling Framework

This chapter describes the workflow of the model developed in this study. This chapter also explores the conceptual model used to develop the model and discusses the modeling context. Both of these elements are integral to model development and understanding.

3.1. Model Objective

The model developed for this study was designed to assist in identifying suitable locations for MPSHS. At its core, the model was expected to find a potential location for the placement of paired upper and lower reservoir components of the MPSHS given parameters provided by project development engineers in conjunction with other spatial components and optional pathways for further refinement of location selection. This model is intended to be used as a core component of a broader site selection process to be expanded upon and field verified by project engineers. It is designed to use freely and easily obtainable spatial data with coverage spanning the continental United States.

3.2. Engineering Requirements

Based on the fundamental principles of pump storage hydroelectricity systems previously described, the MPSHS has basic physical requirements for the placement of its primary components, the upper and lower reservoirs. The foundational spatial relationship between these components controls the potential energy stored in the system and is defined by the change in head over distance. While conceptual designs are available, for the purposes of this study, the system parameters used in the model are hypothetical and do not reflect specific construction requirements, acknowledging only that these parameters exist as variables. However, the values used throughout, are based loosely on the conceptual design parameters. Thus, the two core requirements that the primary model uses for the demonstration implementation are (Figure 5):

1. Relief or the change in elevation between the upper and lower reservoirs must be greater than 300 meters; and,
2. The lateral distance between the upper and lower reservoirs cannot exceed 1,500 meters.

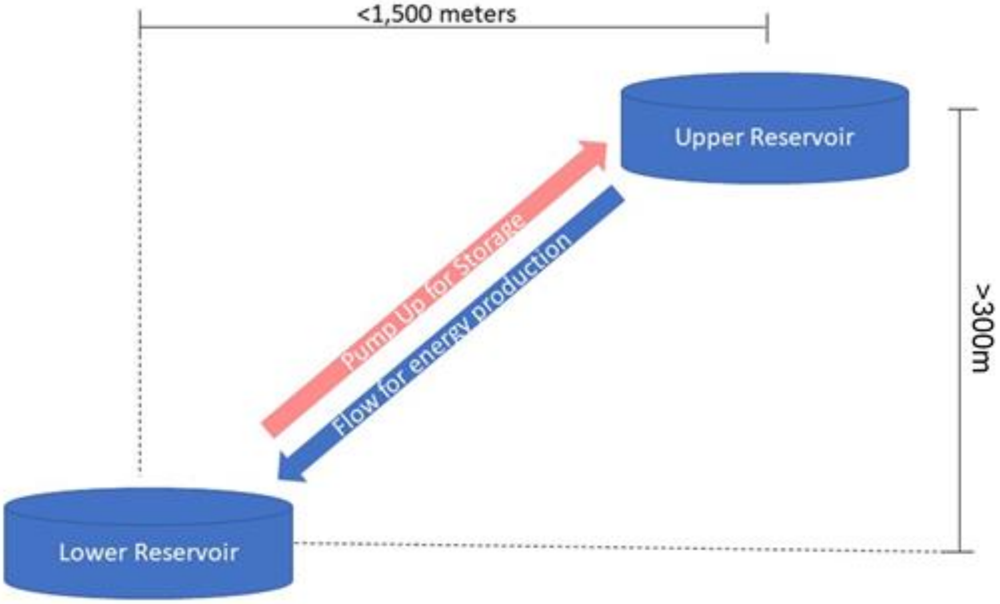


Figure 5 - A simplified diagram showing the engineering parameters for relief and maximum lateral runout used in the demonstration of the Primary Model.

For the system to function, the reservoirs must be large enough to contain sufficient water to support the sustained operation of the hydroelectric generator. This implies that a specific areal footprint is necessary to construct tanks of the required size. For the purposes of the demonstration of this model, the area available for construction of each tank pad must be at least 8,100 square meters (90 meters by 90 meters). Additionally, for construction to be viable based on slope stability and earthwork considerations, the area to be selected for construction of the tank pad cannot exceed a slope angle of 15°.

The restriction on slope is governed by a feasibility problem with respect to both geotechnical limitations and cost of construction. If it is assumed that the reservoir tank must be placed on level ground, then any candidate site must be leveled prior to construction. This process requires soil to be moved from one part of the site to the other. Leveling a site could be accomplished by cutting a portion of the upslope soil material and placing it on the downslope side (Figure 6). Thus, the slope angle has a direct relationship to cut and fill volumes and downslope stability concerns (Connolly, MacLaughlin, and Leahy 2009).

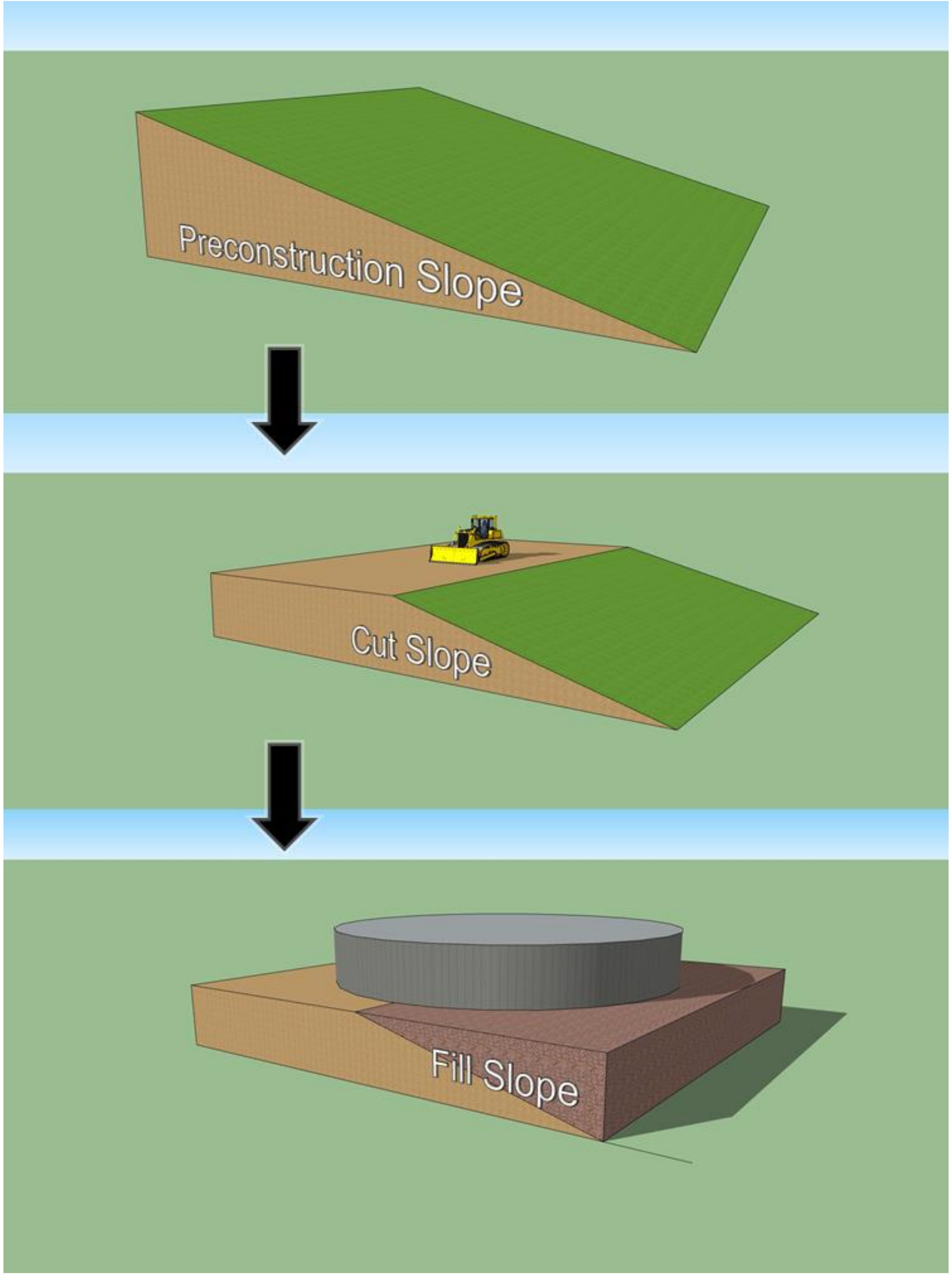


Figure 6 – The reservoir site leveling scale concept

Figure 7 demonstrates the relationship between slope angle and the volume of soil needed to be moved to level a 900m² area (30m by 30m), which is approximately 1/9th of the modeled construction area, or one cell in the suggested resolution for model analysis. When these volumes are applied to an entire 8,100 area, volumes of soil that may require excavation, transportation, and re-compaction on steeper slopes will become prohibitively expensive. Thus, the demonstration model's slope limitation of 15° is deemed an appropriate starting point.

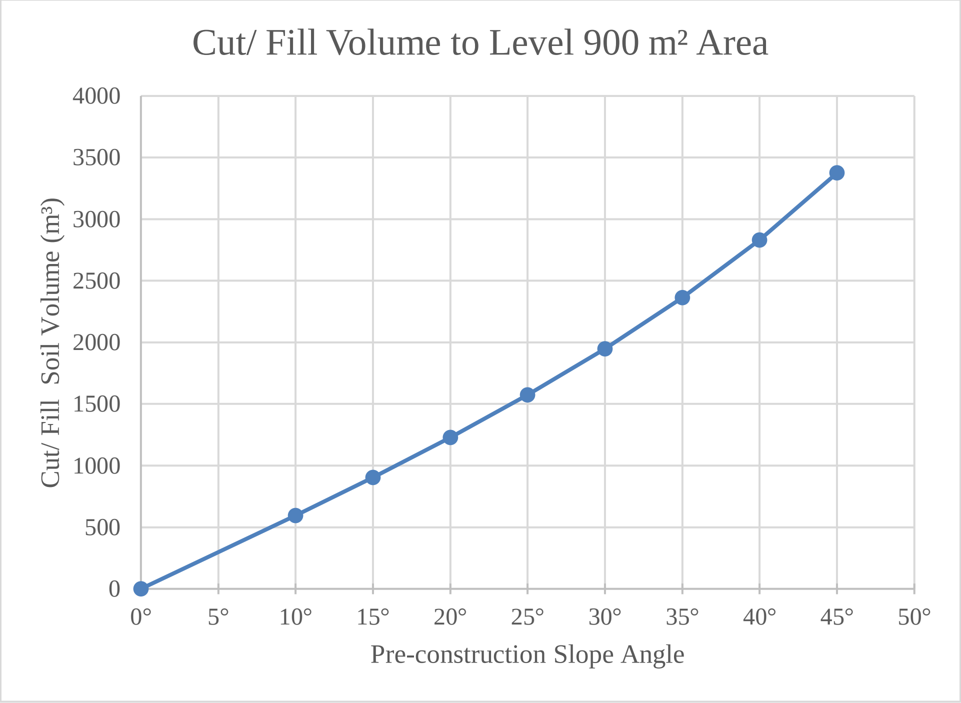


Figure 7 - The relationship between slope angle and cut/ fill volume needed to create a flat construction surface.

3.3. The Primary and Secondary Models

This model was designed to determine potential areas suitable for the construction of terminal reservoirs in the MPSHS. In the Primary Model, suitable locations for upper reservoirs that, within the model search area, have an associated area suitable for the construction of a lower reservoir are identified. Figure 8 depicts the major functional components of the Primary Model.

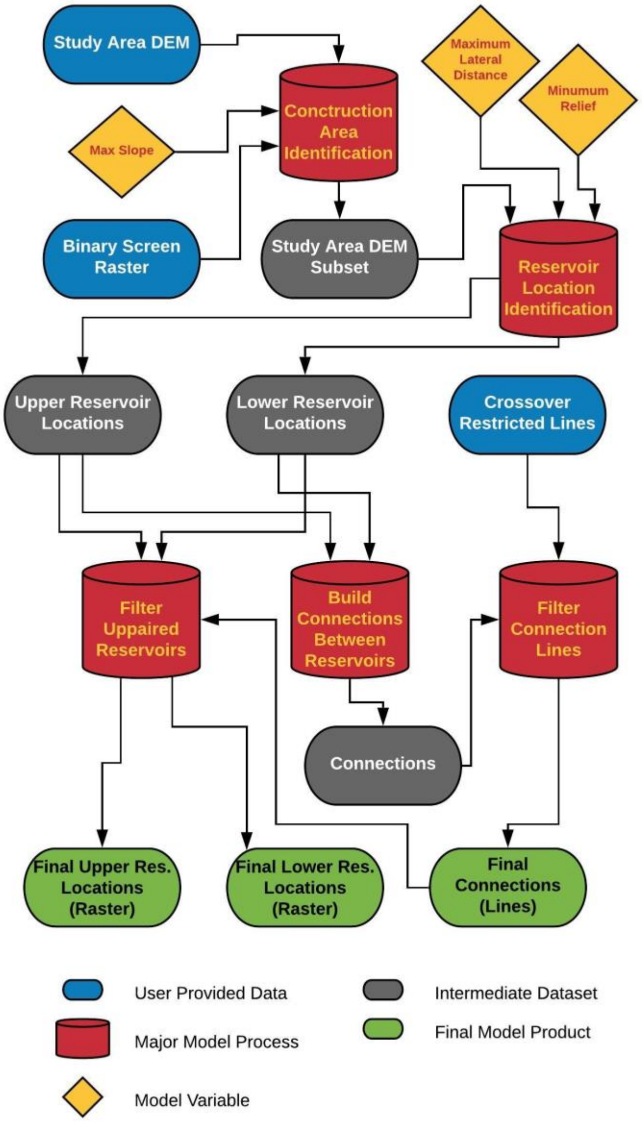


Figure 8 – A graphical depiction of the conceptual Primary Model

The core dataset required for the analysis is a Digital Elevation Model (DEM) dataset for the study area. The Primary Model also accommodates two optional screening components to be provided by the model end-user. These include 1) a binary screening raster layer identifying parts of the landscape that are unsuitable for construction of reservoirs (e.g., built-up areas or national parks) and 2) a vector line dataset representing crossover restricted lines (e.g., major highways or rivers) that connections cannot cross (called "Restricted Lines" in Figure 8).

The Secondary Model uses a dataset derived using fuzzy logic that is applied to the Primary Model outputs to determine the best candidate sites for deployment of MPSHS. The *Fuzzy Layer* is composed of end-user-provided data, which is combined by application of a fuzzy overlay then converted into points with attributes noting their suitability relative to their fuzzy membership (Figure 9).

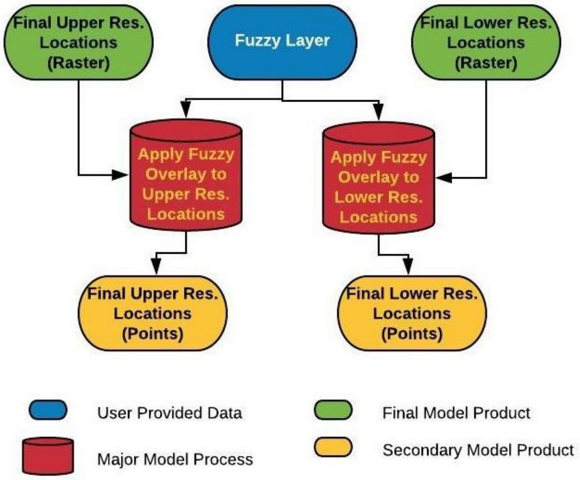


Figure 9 - A graphical depiction of the conceptual Secondary Model

3.4. Modeling Products

The purpose of the Primary Model is to find locations that meet the siting requirements identified by project engineers for the placement of the upper and lower reservoirs of the system. Each cell in the Primary Model output raster datasets represents the center of an area suitable for either upper or lower reservoir construction. Additionally, each cell in the output raster identified as a suitable location for the construction of either an upper or lower reservoir is within the required proximal distance of its companion reservoir location. The model also identifies connections such that each upper reservoir location is paired with all of the possible lower reservoir locations within the designated search area, a one-to-many scenario.

Finally, application of the Secondary Model assesses further the suitability of both the upper and lower reservoir locations based on end-user-defined criteria. This process is accomplished by application of a fuzzy overlay and returns vector points for both upper and lower reservoir locations that are assigned attribute data regarding their aggregate suitability relative to the individual components of the fuzzy dataset. The conversion to vector points was found to facilitate better data interrogation when examining final results at the single system level.

3.5. Modeling Context

This model is intended to be used as a core component of a broader site selection process. This section discusses the context of the model.

3.5.1. Modeling Environment

This model was built using Esri GIS software. Working with the Esri ArcGIS Pro software suite, analysis tools offered within the spatial analysis and spatial statistics extensions were utilized for the terrain analysis and data processing portions of this research. While these methods are not unique to Esri GIS software, the model developed for this thesis was created using Esri's GIS environment and their ModelBuilder application. ModelBuilder is an application developed to create and manage models as workflow routines. These link a series of geoprocessing tools together using a visual programing language (Esri 2016g). The choice to work exclusively within the ArcGIS Pro environment was due mainly to its ubiquity in industry and ease of use with respect to both model construction and distribution.

3.5.2. Spatial and Temporal Scale

The spatial resolution for this analysis was determined based on commonly available DEM data, which has a spatial resolution of 30 meters. The Primary Model uses DEM data easily obtained from a source such as the USGS 3DEP which has a vertical accuracy of 3.04 meters at a 95% confidence interval (Gesch et al. 2014). At this scale, regional terrain characteristics relevant to this model are accurately captured. Small scale geomorphic processes that may operate over short time-scales are largely masked at this coarser scale, but the temporal currency of the DEM is not critical.

Additional data selected by the end-user for inclusion in the model for either the binary screening layer or the restricted lines dataset could potentially have a significant effect on the end result. Therefore, careful consideration of their temporal currency and spatial accuracy should be made when selecting these datasets. For example, if the lands contained within the boundary of a National Park were to be considered unsuitable for deployment, it is vital to confirm that the input datasets have spatial accuracy appropriate for use with a 30m DEM and that the data are current. Using Death Valley National Park as an example, in 2019, the park expanded by 90,000 acres. Working today with an older boundary layer would have a significant impact on the size of the area to be evaluated for suitable MPSHS sites.

This model was developed to operate at the scale of an average county in California, approximately . During the development process, it was found that at this scale, the model performed nominally with respect both to the quality of the model outputs and the processing resources required to run the model. Designing the model for use at a county extent also determines the extent of the model outputs. By adopting this extent as the unit of analysis, the spatial datasets and spatial extent can be standardized, allowing the model to be broadly applicable. Furthermore, the county is often the extent of existing political and administrative constraints on the implementation of MPSHS technology. These political and administrative constraints could range from land use permitting to construction regulations. Keeping the MPSHS within a single county jurisdiction may also cut down on costs associated with long term operation of the facility.

However, while his model is designed to be implemented at the county scale, there is no built-in limitation for the size or shape of the modeled area. Average county sizes vary nationally. In California, for instance, counties range in area from 121 square kilometers to 51,948 square kilometers. In areas where counties are smaller and denser, such as those in the eastern United States, multiple counties could be considered for analysis by merely merging county-level datasets.

Chapter 4 Model Description

This section describes the model developed in this study. Principal model components are separated into groups representing the main functions of the model. While the model was constructed within the Esri environment, the terms used below describe the GIS processing steps underlying the packaged tools provided by Esri. These sections describe the major steps in the model to generate the final product.

This model uses Esri ArcGIS Pro software and the Esri ModelBuilder application to create two packaged models. The Primary Model outputs the set of locations suitable for reservoir construction and the set of lines between paired upper and lower reservoirs. The Secondary Model combines the Primary Model outputs with a fuzzy joint membership function layer to further refine the output by identifying the most suitable candidate sites. Thirty-four individual processes are combined to form the Primary Model, and four are used in the Secondary Model.

4.1. Data Requirements

This model was developed to utilize data that are widely and freely available throughout the continental United States in order to facilitate its widespread deployment.

4.1.1. Primary Model Data Requirements

The Primary Model examines the terrain for the placement of the system. The following sections describe the necessary data and the related variables used as model inputs required to generate model parameters.

4.1.1.1. Study Boundary

The study area boundary defines the lateral limits of the area to be examined. As discussed in the previous chapter, the model was developed to work at a county scale. Study boundary data used in the model will typically be vector files of county boundaries obtained from a government source. For use in the model, this data should be projected with as little areal distortion as possible. This requires the use of an equal area type projection such as an Albers projection. Once a suitable projected Coordinate System is selected, all spatial data should be re-projected into that system and registered against the root DEM dataset.

4.1.1.2. DEM

The base dataset of the Primary Model is a DEM. To ensure that the model has applicability over as wide an area as possible, it was designed to utilize the United States Geological Survey's 3D Elevation Program (3DEP) 1/3 arc-second (10 meters) to 1 arc second (30 meters) elevation products. The 3DEP data provided at this scale is a seamless DEM dataset with full coverage of the continental US. The 3DEP products are distributed by the USGS in 1-degree panels and are seamlessly combinable to cover large areas comprised of multiple panels (USGS 2019).

For integration into the model, the DEM panels must be combined to create a single DEM dataset with coverage over the entire study area. This is accomplished by generating a mosaic dataset from multiple panels. Once combined, the dataset must be projected into the designated projected coordinate system for the model and converted into the designated project resolution of 30 meters by 30 meters using a bilinear interpolation for resampling. The elevations in the new raster DEM dataset should be presented in meters relative to mean sea level. Finally, the DEM is cropped to the lateral limits of the study area to eliminate unnecessary processing.

It is important to note that the DEM layer selected for the *Study Area DEM* resolution and position sets the “Model spatial framework” that everything else must be registered to.

4.1.1.3. Binary Screen Components

The *Binary Screening Raster* is an optional component of the Primary Model. It is included to provide the end-user with an opportunity to eliminate areas predetermined to be unavailable for construction of the MPSHS. This data can include administrative and other areas that will not be considered in the analysis. By default, the model incorporates a blank screening layer.

The *Binary Screening Raster* must be a raster dataset in the same extent and spatial resolution as the input DEM. For areas that are not to be considered in the suitability analysis, cell values should equal zero; for areas considered, cell values should equal one. If the end-user does not want to incorporate a binary screening dataset into the workflow, a constant value raster (all values = 1) can be substituted.

4.1.1.4. Restricted Crossings

The Primary Model provides a method for recognizing linear features that exist in the study area that connections (i.e., penstocks) cannot cross. This type of data would be presented by a vector line dataset and could represent streams, utility corridors, political boundaries, or other barriers that cannot be crossed. In the Primary Model, this feature dataset is referred to as *Restricted Lines*.

The *Restricted Lines* dataset is to be provided by the end-user and is to be composed of a single vector line dataset. In this dataset, each line represents a feature that a connection cannot cross when connecting the two terminal reservoirs. The dataset can be comprised of any number of different linear components or phenomena.

4.1.2. Secondary Model Data Requirement

The *Fuzzy Layer* is the input dataset that is applied to the Primary Model outputs by the Secondary Model. While the data used in the binary screening layer in the Primary Model eliminate certain areas from consideration as reservoir sites, the fuzzy layer is comprised of spatial variables that have varying impact on the suitability of locations for the placement of the MPSHS components. For example, if the proximity to a roadway improves the suitability of a potential reservoir location, a fuzzy membership layer could be created such that cells closer to roadways have a higher fuzzy membership than those further away. This dataset could be one dataset or a combination of several spatial datasets in which phenomena have been given a fuzzy membership value and combined into a single fuzzy set for incorporation into the Secondary Model through a joint membership function such as a fuzzy overlay. The cell values of the fuzzy dataset will be closer to one for those areas that have high suitability for reservoir development, and less suitable areas will have cell values approaching zero.

As with the binary dataset, the fuzzy dataset must be in the same spatial extent and spatial resolution as the input DEM.

4.2. Primary Model Processes

The Primary Model performs the terrain analysis that examines the selected *Study Area DEM* for locations suitable for the construction of MPSHS. As noted in the chapter on model design, the components of the model answer four basic questions 1) where is the terrain suitable for construction of a reservoir? 2) where are suitable locations for the upper reservoirs? 3) where are suitable locations for lower reservoirs? and 4) which upper locations are paired with which lower locations?

Both the Primary and Secondary Models were constructed in Esri ArcGIS Pro using the ModelBuilder application. The Primary Model is comprised of a series of component processes and produces three final outputs. The only required input for the model is a 30-meter DEM in an appropriate coordinate system with elevations provided in meters. Optional inclusions into the model are a *Binary Screening Layer* and a *Restricted Lines* dataset.

The following sections provide a detailed look at the major steps in the Primary Model. A complete presentation of Primary Model processes along with related inputs, outputs, tools, and parameters is presented as Appendix A.

4.2.1. Construction Area Analysis

The construction area requirements are provided by the design engineer in the form of an areal footprint. For example, a reservoir of appropriate size may be a cylindrical reservoir 70 meters in diameter. Application of a 10m radial buffer for necessary equipment dictates that the minimum size requirement for the area of construction must contain a circle of at least 90m. At the spatial resolution of 30m, the area required to support the construction of the design reservoir is a nine-cell Moore neighborhood, with a square footprint measuring approximately 90 meters by 90 meters, or 3x3 30m cells.

Additional constraints on construction dictate the maximum slope angle that can be considered suitable for construction as provided by the design engineer. As a model default, 15° is used as the slope angle limit. This means that for a location to be considered suitable, the slope between the center focal cell and the center of the eight surrounding cells cannot exceed the slope angle limit. To allow for this model to be used in varying scenarios, the search area (size and shape of the neighborhood) and the maximum slope criteria are model parameters that can be adjusted to suit the user's needs.

This analysis is completed in a three-step process, shown in Figure 10, and is based on the slope in the study area. First, the slope is calculated from the Study Area DEM using the slope tool. This tool identifies the maximum difference in elevation value between each cell and each of its eight neighbors (Esri 2016d). The tool produces a raster dataset (*Study Area Slope Raster*) where each cell value represents the maximum slope in degrees at each cell in the dataset.

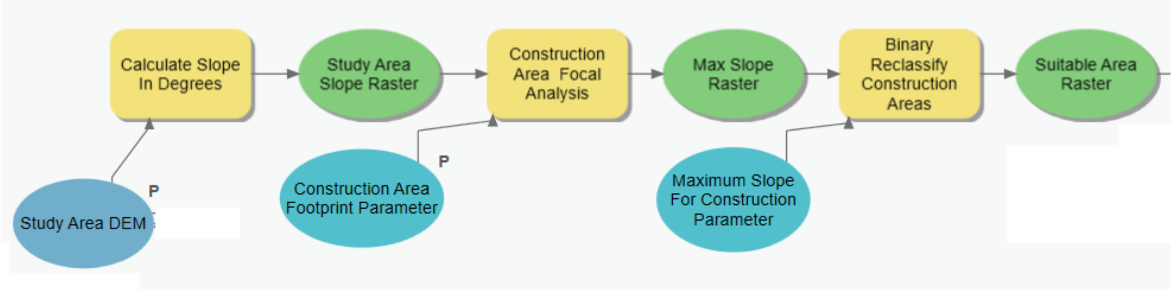
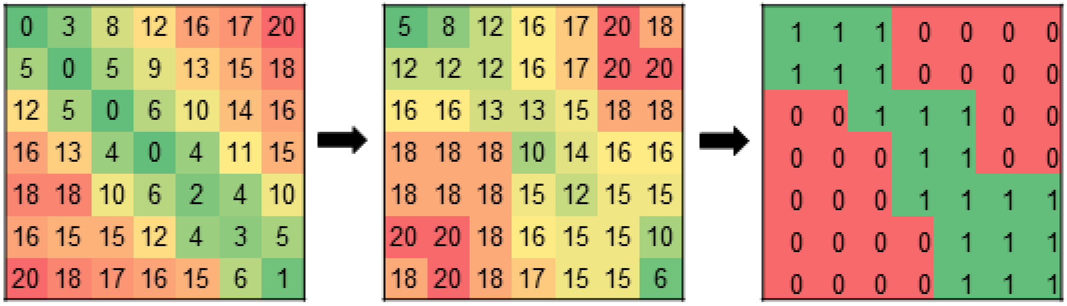


Figure 10 - The portion of the model flow chart showing the steps to determine areas suitable for construction.

Next, using the *Study Area Slope Raster* derived from the input *Study Area DEM*, a moving window analysis is performed over the entire study area using the focal statistics tool to identify the maximum slope in the neighborhood of each cell. This is accomplished by use of the focal statistics tool. Parameters for this stage in the model include the radius of the search area for the moving window analysis and the shape of the search area. The search area is defined as the radius in cells around the focal cell. For example, given the 8,100 square meter footprint demonstration requirement (3x3 30m cells), for each cell in the grid, by default, this tool looks at values of the focal cell and the eight neighbor cells (a two-cell radius) and assigns the maximum observed slope angle to the focal cell in the output raster dataset. Values in the newly created *Max Slope Raster* represent the maximum slope within the search neighborhood (equivalent to the construction area footprint) for each focal cell in the study area.

The maximum neighborhood slope raster is then reclassified into a binary raster dataset where focal cells with all neighbors not exceeding the *Slope Limit* are assigned a value of 1, and those with exceedances are assigned a value of 0 (Figure 11). The Slope Limit is a key parameter in the model, and the End-User is given the ability to modify this parameter. The resultant *Suitable Area Raster* identifies those cells within the study area that are the center of a neighborhood where the slope and areal extent is suitable for the construction of reservoir components. This estimation is based solely on the local topography expressed by the DEM.



0	3	8	12	16	17	20
5	0	5	9	13	15	18
12	5	0	6	10	14	16
16	13	4	0	4	11	15
18	18	10	6	2	4	10
16	15	15	12	4	3	5
20	18	17	16	15	6	1

Study Area
Slope

5	8	12	16	17	20	18
12	12	12	16	17	20	20
16	16	13	13	15	18	18
18	18	18	10	14	16	16
18	18	18	15	12	15	15
20	20	18	16	15	15	10
18	20	18	17	15	15	6

Max Slope

1	1	1	0	0	0	0
1	1	1	0	0	0	0
0	0	1	1	1	0	0
0	0	0	1	1	0	0
0	0	0	1	1	1	1
0	0	0	0	1	1	1
0	0	0	0	1	1	1

Reclassified
Slope

Figure 11 – A graphical depiction of the slope processes of construction area analysis

4.2.2. Refining the Search Area

Upon completion of the construction area analysis, the optional *Binary Screening Raster* is applied. The *Binary Screening Raster* is provided by the user and identifies areas that are excluded from the area to be considered for development. These areas are represented in the user-provided binary raster as zero values. Areas that are to be considered should have a value of one.

Application of the *Binary Screening Raster* to the *Suitable Area Raster* is accomplished by multiplying the construction area raster by the binary screen. The net effect eliminates from consideration areas that were previously identified as suitable by the Construction Area Focal Analysis, but which should not be included in the final set of suitable sites.

The product of this stage of the model is the *Construction Area Raster*, where areas comprised of grid cells fit for the construction of the technology are characterized by a value of 1 and those that do not have a value of 0. Using the raster to vector conversion tool, the model then converts the *Construction Area Raster* to a vector polygon layer, which is then reduced to only those areas deemed suitable through the slope analysis by using the select tool to create a new vector dataset, *Construction Area Polygons*, which are comprised of only areas deemed suitable for construction of the terminal reservoirs (Figure 12).

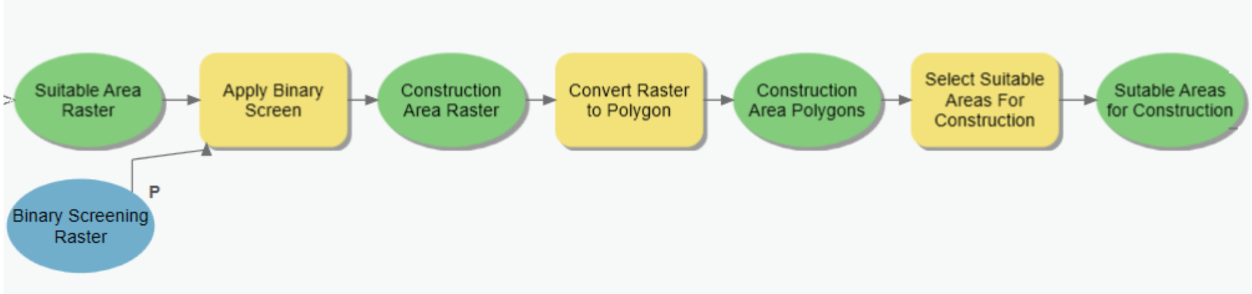


Figure 12 – Modeling process for construction area identification.

4.2.3. Extracting a subset of the DEM

To reduce processing time, the *Construction Area Polygons* dataset is used as a mask feature to extract from the *Study Area DEM* a raster layer whose data includes elevation for only those areas suitable for construction. The result is the *Study Area DEM Subset,* as shown in Figure 13. Additionally, the polygons representing areas suitable for construction are exported. While not a critical component of the model outputs, the *Final Suitable Areas Polygons* provide the end-user with a graphical representation of the areas considered that could be useful in presentation of the model outputs or independent validation.

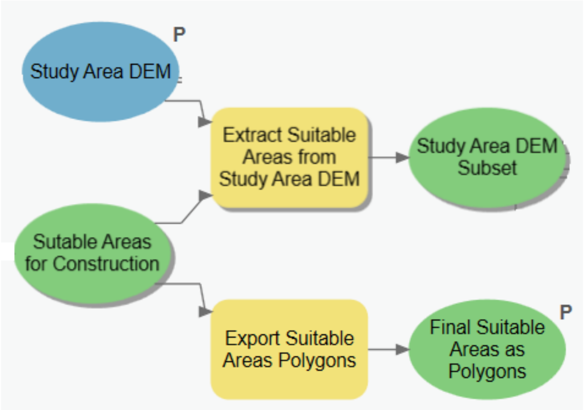


Figure 13 - The portion of the model flow chart showing the steps to create a new subset DEM dataset.

4.2.4. Relief Calculations

Relief calculations are the critical engineering component of the suitability model. As shown in Figure 14, parallel modeling processes identify locations where the downslope and upslope relief within the given search distance meet the system design requirement. The net result is two vector point datasets, *Upper Locations as Points* and *Lower Locations as Points*, where each point represents the center of a 90m by 90m square that is either a potential upper or lower reservoir location as determined by the maximum change in elevation within the maximum connection distance identified by the *Relief Search Radius Parameter*.

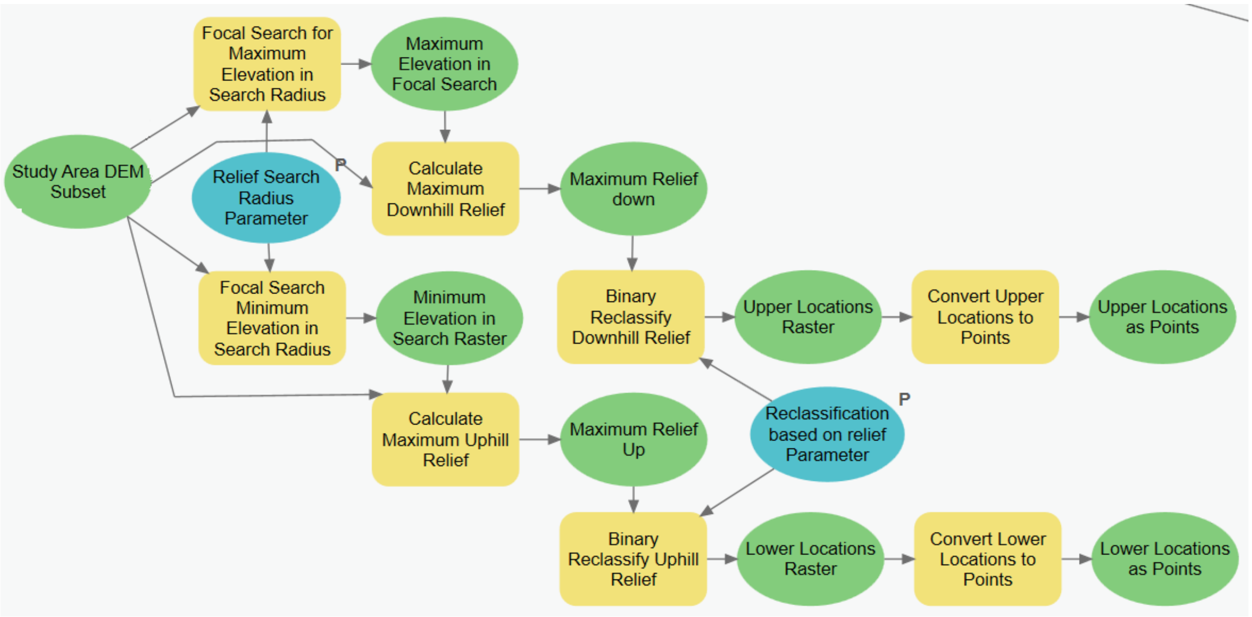


Figure 14 – Model processes for upslope and downslope relief calculations.

For upper reservoir citing, calculations are applied to only the *Study Area DEM Subset* extracted by the previous step. A moving window analysis is applied to each focal cell in the *Study Area DEM Subset*. Each cell within the 1,500m relief search radius around the focal cell is examined to determine the minimum elevation therein. Finally, the value of the elevation of the lowest cell within each cell's search area is assigned to the focal cell, creating the *Maximum Elevation in Focal Search Raster*. The model steps for identifying potential lower reservoir locations from amongst the suitable building sites is accomplished the same way by modifying the search criteria to identify the maximum elevation within the search area.

To identify the maximum potential relief for each cell in an area suitable area for construction of the system, maximum upslope relief is calculated by application of raster math, where, for each cell, the *Maximum Elevation In Focal Search Raster* is subtracted from the elevation of the same cell in the *Study Area DEM Subset*, producing a new raster dataset where each cell contains a value indicating the maximum downslope relief within the search area. This process is mirrored and repeated to generate the raster dataset for maximum Downhill relief.

4.2.5. Filtering Suitable Locations

Following the relief calculations, the model now has two raster datasets that demonstrate potential relief within the study area for locations suitable for construction. This dataset must then be filtered to eliminate those cells in which the maximum (upslope or downslope) potential relief does not meet the design requirements for the system. Filtering is accomplished by reclassifying both the upslope and downslope maximum relief raster such that cells in which the value meets or exceeds the system design specification for relief are given a value of 1 and cells that do not meet the requirement are reclassified as NODATA. Using the NODATA eliminates irrelevant results from each dataset as they will not be needed in later calculations.

Each raster is then converted into a vector points dataset such that each cell that meets the relief requirements and represents the centroid of an area suitable for construction is represented by a single point. Again, two datasets are created identifying potential reservoir locations as points, and these are referred to in the model as the *Upper Location Points* and *Lower Location Points* (Figure 14). The purpose of this step is to provide two vector point datasets representing potential reservoir locations to serve as inputs into the near analysis.

4.2.6. Matching Points

The next stage in the model works to identify and pair potential upper reservoir sites to their respective lower reservoir sites (Figure 15). The previous model components have produced two sets of points, the *Upper Location Points* and *Lower Location Points,* that have undefined spatial relationships to each other. The relief calculation and associated processes have created these datasets independently, and it is necessary that each potential upper reservoir point is paired with at least one potential lower reservoir point.

To identify these one-to-many relationships, a table is generated, using the Near tool, that identifies all of the lower reservoir points within the previously established *Relief Search Radius* for each upper reservoir point. This table has two fields for upper reservoir location coordinate pairs (X and Y) and two fields for lower reservoir location coordinate pairs. Each row in the table represents a unique connection between an upper reservoir point and a lower reservoir point that fall within the lateral search radius of each other.

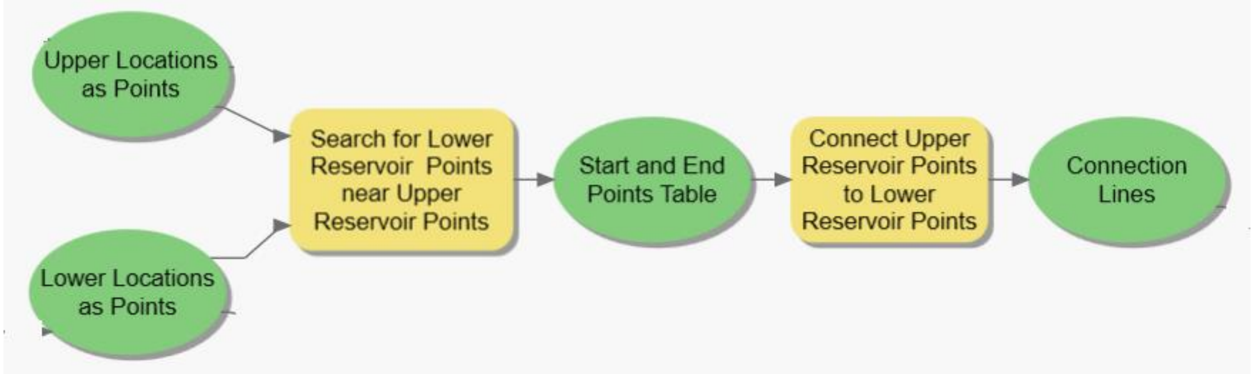


Figure 15 - The portion of the model flow chart showing the steps to pair upper reservoir location with lower reservoir locations

Using the relationship table generated by the near analysis, lines are generated using a tool that creates a straight line from each upper reservoir location to all of the paired lower reservoir locations within the search area. The resulting output is a vector line dataset (*Connection Lines*) representing all of these connections.

4.2.7. Filtering Connections

At this stage in the workflow, the one-to-many connections created by pairing all of the upper reservoir locations to their respective lower locations has two significant complications. The first problem is that there are likely linear features such as roads or streams that a system connection cannot cross. The second problem is that not all lower reservoir points located within designated search areas are from the upper reservoir locations that genuinely meet the requirement because the near analysis discussed above indiscriminately identifies all target points within the search radius as a match. In other words, just because a potential lower reservoir is within the search radius of a potential upper reservoir, that pair may not meet the relief requirement.

4.2.7.1. Filtering Restricted Line Crossings

Using the *Restricted Lines* dataset, the model identifies reservoir connections that intersect barriers to connection. The selected connections are then removed from the connection dataset by selecting the inverse of the connections that cross *Restricted Lines*, then copies those features to a new dataset. The result is a dataset is comprised of only *Connections That Do Not cross Restricted Lines* (Figure 16).

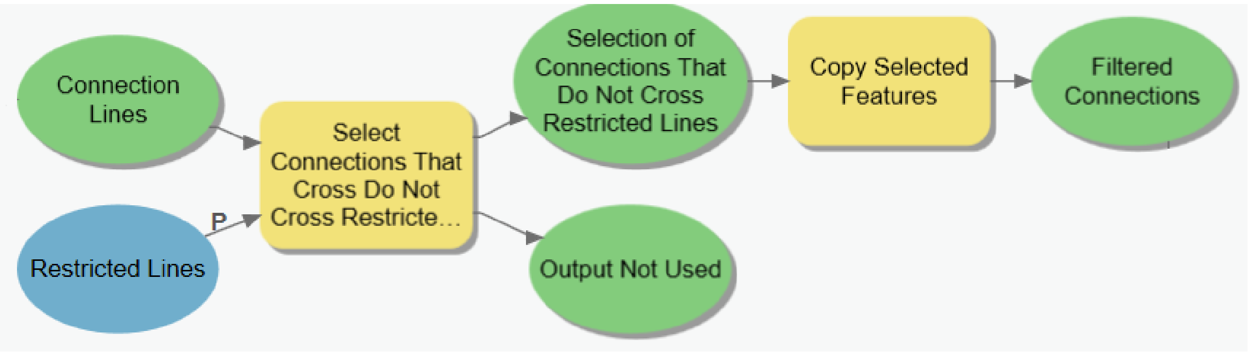


Figure 16 - The portion of the model flow chart showing the steps to remove connections between the upper reservoir and lower reservoir locations that cross named streams.

4.2.7.2. Filtering False Match Connections

The processes creating the connection and removing those that cross *Restricted Lines* still leaves the second problem to solve. Application of the Near analysis has identified lines that connect each of the upper reservoirs to all lower reservoir locations within the *Relief Search Area* distance; false match connections are made. Figure 17 demonstrates this concept.

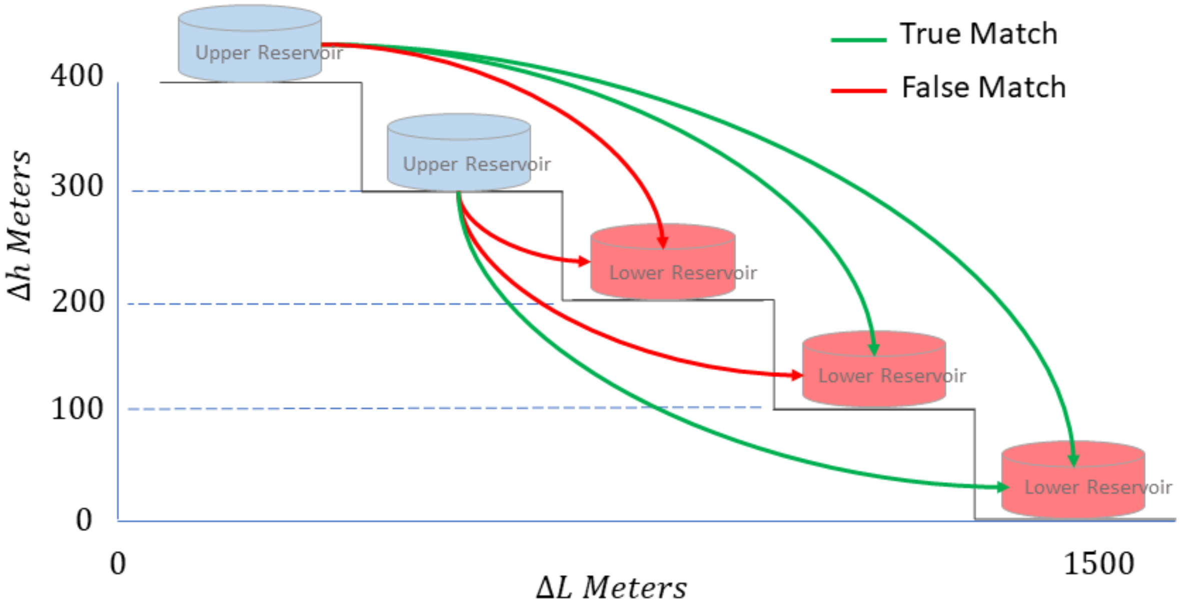


Figure 17 - A graphic depiction of conditions identified as false and real matched connections where relief must be at least 300 m.

False match connections occur when a potential location for a lower reservoir location is within the near search for an upper reservoir location, but that connection lacks the relief required by the model design parameters. To correct this problem, a model component was developed to identify and eliminate these false match connections (Figure 18).

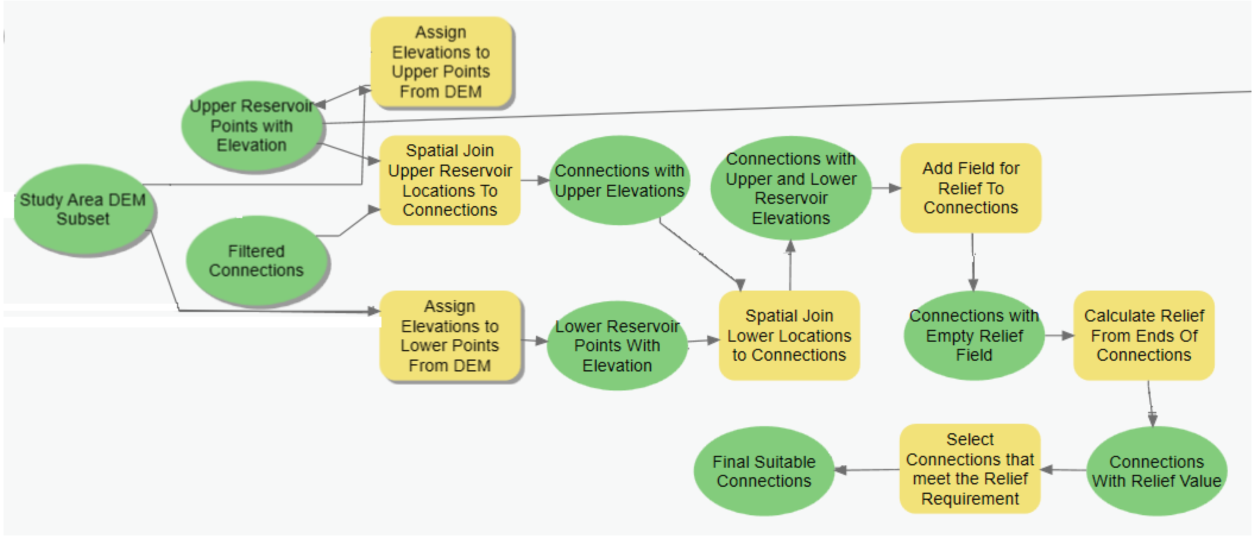


Figure 18 - The portion of the model flow chart showing the steps to identify and eliminate false match connections.

While not a complicated computation, there are many steps in the process. The elevation data from suitable locations must be joined with connections. After each connection has elevations for both terminal ends, the relief can be calculated by subtracting the two values. The final step in this process is selecting those connection lines that meet or exceed the design relief requirement to generate a *Final Suitable Connections* dataset.

4.2.8. Filtering Reservoir Locations

With the filtering of the connection lines completed, some locations previously identified as being suitable for the placement of the upper or lower reservoir locations may no longer have a connection to a reservoir pair. Using the reservoir locations represented as points (*Lower/ Upper Reservoir Points with Elevations*), the connections are used to select all of the points in each dataset that do not intersect a remaining connection endpoint. The selected points are then deleted from each of the point datasets. The newly created *Final Upper Reservoir Locations Points* and *Final Lower Reservoir Locations Points* datasets are then converted into new raster datasets, *Final Upper Reservoir Locations Raster*, and *Final Lower Reservoir Locations Raster*, respectively (Figure 19).

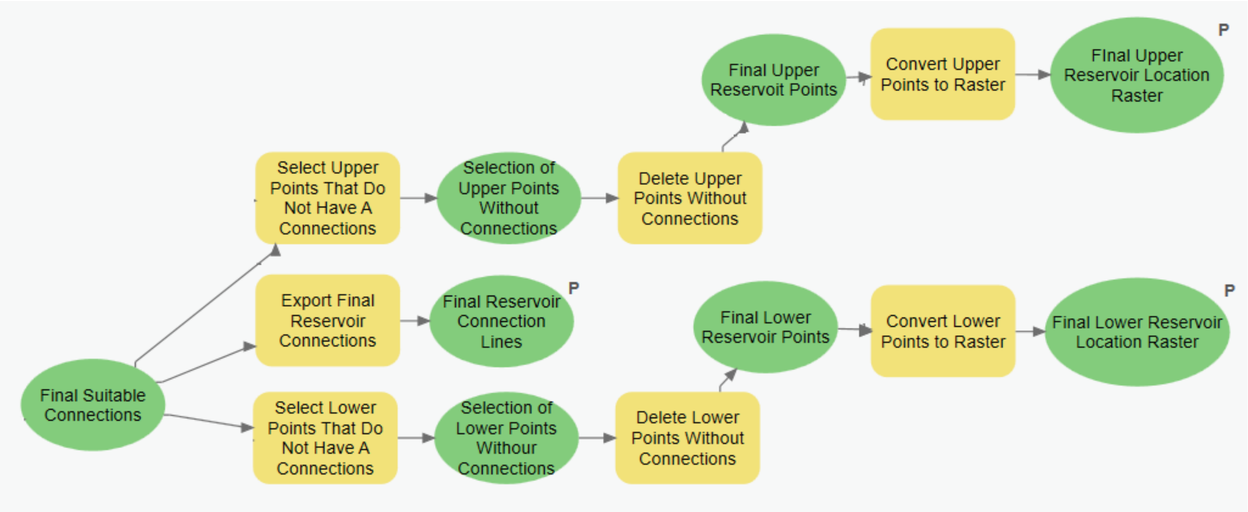


Figure 19 - The portion of the model flow chart showing the steps to eliminate upper and lower reservoir locations that do not have a connection after the connection filtering steps.

4.2.9. Model Results

The final Primary Model outputs are the *Final Lower Reservoir Location Raster, Final Upper Reservoir Location Raster*, *Construction Area Polygons*, and *Final Reservoir Connections*. Reservoir locations are provided as raster datasets at the same resolution as the input DEM dataset. The centroid of each cell represents the center of a 90m by 90m area with potential for use as a modular reservoir location.

The *Construction Area Polygons* are simply areas where the slope limit criteria have been met, and the construction of the modular reservoir is possible.

Final Reservoir Connections are exported separately in a parallel process as vector lines with attributes indicating relief and distance between the paired reservoirs. Since these connections are an integral component of the filtering of reservoir locations, they are considered intermediate data and not preserved in the ModelBuilder workflow.

4.3. Secondary Model Processes

The Secondary Model builds on the outputs from the Primary Model and employs a user-provided fuzzy dataset to analyze the suitability of the paired upper and lower reservoir dataset. The user-provided data selected for this analysis can be comprised of many types of spatial phenomena. Each fuzzy set represents a spatial factor that has been assigned a fuzzy membership value. All of the separate fuzzy sets are then combined using the fuzzy overlay tool to produce the *Fuzzy Layer*, used in the Secondary Model.

The Secondary Model uses the *Final Model Lower Reservoir Location Raster* and the *Final Model Lower Reservoir Location Raster* from the Primary Model as the two primary input raster datasets. It should be recalled that the Primary Model assigns the value of 1 to all suitable grid cells within both the upper and lower reservoir location raster datasets, while all other cells have a "NODATA" value. The *Fuzzy Raster* created by the end-user, with values ranging from 0 to 1, is the third input dataset. In parallel processes, the upper and lower reservoir raster datasets are multiplied by the fuzzy overlay raster. The resulting intermediate datasets are comprised of raster layers where the grid value in areas previously identified as being suitable for a reservoir location is equal to its corresponding fuzzy membership value (Figure 20).

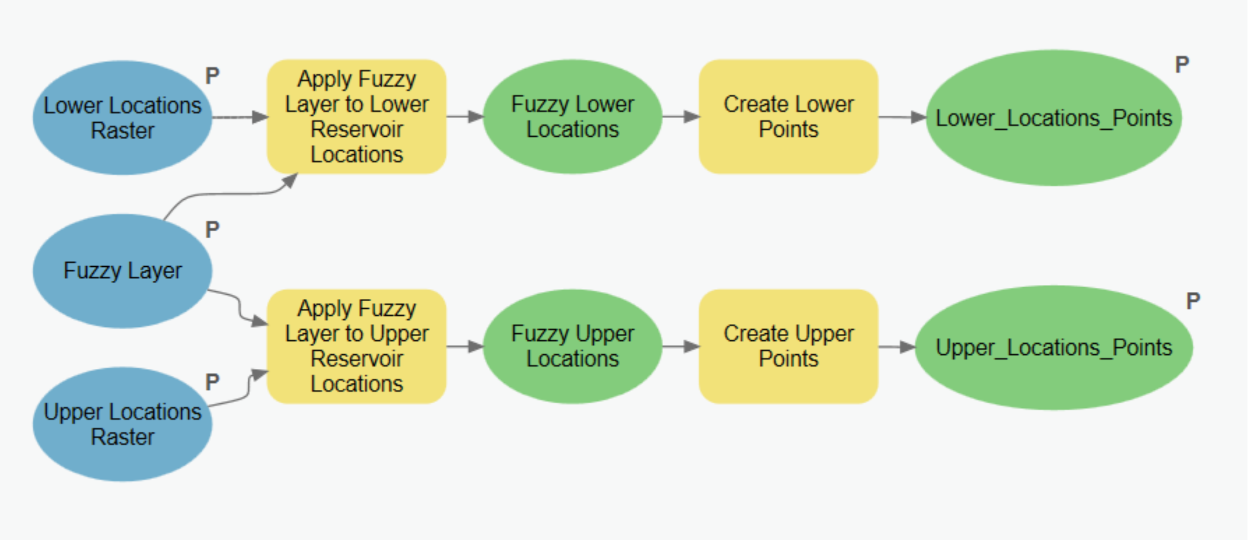


Figure 20 - The portion of the Secondary Model flow chart showing the steps that apply the values of the fuzzy ranking dataset to the upper and lower reservoir locations.

The final step in each of the parallel processes transforms the *Final Model Lower Reservoir Location Raster* and the *Final Model Lower Reservoir Location Raster* to points containing the membership value corresponding with the co-located cell in the *Fuzzy Layer*. The vector points now represent the centroid of an area suitable for construction of a reservoir that meets the relief requirement when matched with a paired reservoir and contains attribute values indicating that location's fitness for use when considering the variables present in the fuzzy overlay. Table 6 presents a typical attribute field layout for Secondary Model products.

Table 6 – An example portion of the attribute table for one of the Secondary Model outputs *(n=40,868)*

OBJECT_ID	**point_id**	**Fuzzy Membership Value (Grid_Value)**
1	1	.99554
2…	2	.75456
…40,868	40,868	.01213

Chapter 5 Case Study: Model Processing, Outputs, and Evaluation of Results

This chapter examines the intermediate and final outputs of the Primary and Secondary Models applied to the county of Los Angeles. This area was chosen because of its proximity to USC and the potential for terrain suitable for deployment of the technology.

During the model development process, several areas in California were used as test areas. These areas included Los Angeles County, Mono County, Santa Clara County, Butte County, and Yolo County. These areas provided a diverse cross-section of topography and landform geomorphology. Los Angeles County, California, served as the primary study area and the subject of the case study provided in this chapter. The remaining study areas are briefly discussed at the end of this chapter.

This chapter takes a step by step approach to examine model processes, intermediate data, and final products.

5.1. Preliminary Steps

Beyond data procurement, raw data must be converted into data types and formats suitable for use in the model. This includes processing the *Study Area DEM*, assembling the *Binary Screening Layer* and the *Restricted Lines* dataset.

Additionally, because the model was designed and tested using study areas located within the state of California, the projected coordinate system used herein is the North American Datum 1983 (2011) California Teale Albers Coordinate system in meters. This coordinate system was chosen because it is an equal-area projection with minimal areal distortion, covering the entire state.

5.1.1. Processing the DEM

To serve as the principal input into the Primary Model, the DEM must be in a single raster dataset at the correct spatial resolution and in the correct coordinate system. The DEM created for the Los Angeles County example is comprised of four 3DEP 1-degree panels. The 3DEP panels were combined to create a single mosaic DEM dataset and converted into the North American Datum 1983 (2011) California Teale Albers Coordinate system with horizontal units in meters. The elevation is provided by the USGS 3DEP program in meters by default; thus, no modification is required, as the engineering requirements and projected coordinate system use meters as the unit of measure. The final input DEM was then extracted such that the extent of coverage is coincident with the county boundary (Figure 21).

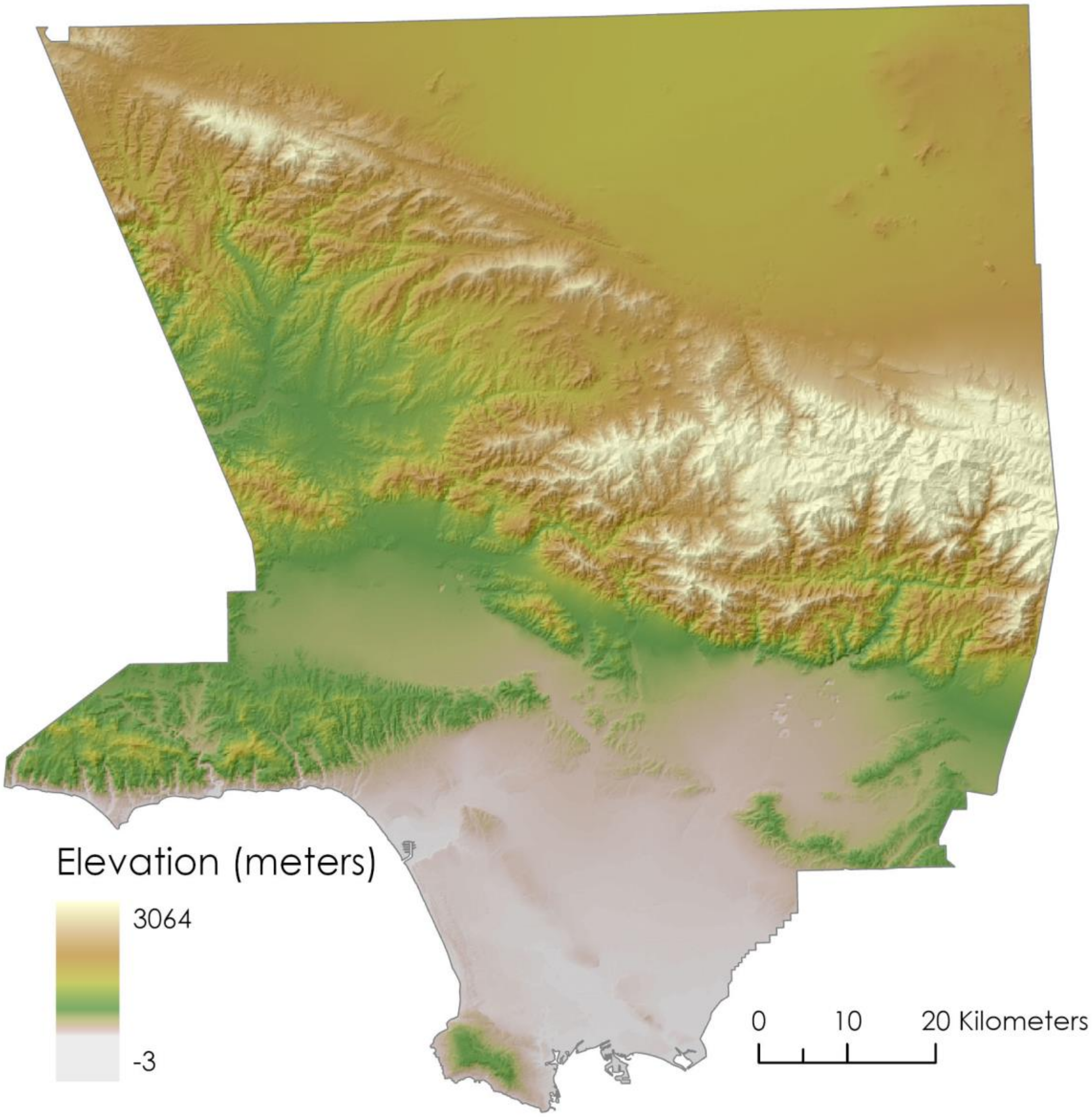


Figure 21 - The processed DEM for Los Angeles County used in the model example.

5.1.2. Binary Screen Creation

The binary screen is an optional element in the model that serves to help the end-user eliminate areas from consideration in the early stages of the modeling process. This served two major functions. First, eliminating areas from consideration reduces the area which needs to be processed, in turn, reducing model run times. Second, it reduced the potential for false match results. For example, without eliminating waterbodies from consideration, the model can identify lakes as large flat areas suitable for placement of the lower reservoir.

For the Los Angeles County example, NHD waterbodies, National Parks, and State Parks were used to create a single binary screening layer using a simple presence or absence test (Figure 22; left). From these vector areas, a raster dataset was created in the same coordinate system and spatial resolution (30m) as the DEM. Where an exclusion feature was present, the raster was given a value of 0, if no exclusion feature was present, that cell was assigned a value of 1 (Figure 22; right).

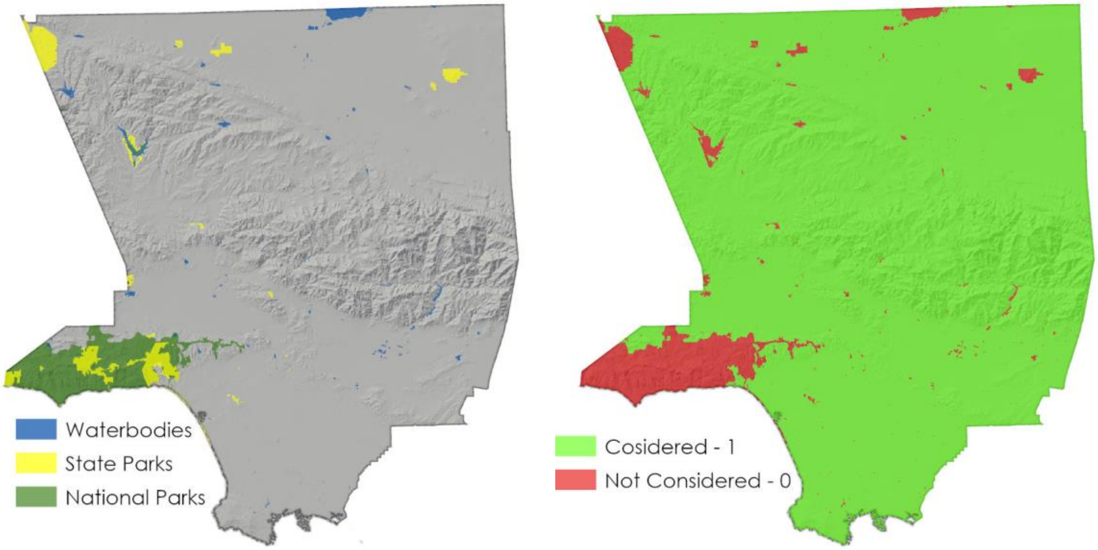


Figure 22 – Areas not for placement consideration (left) and *Binary Screening Layer* (right)

5.1.3. Restricted Lines

The final input into the Primary Model is the *Restricted Lines* dataset consisting of linear features such as roads, streams, or utility corridors that the reservoir connections cannot cross. For this example, a filtered version of the NHD streamline dataset was used.

The logic used to filter the stream line dataset for this case study assumed that all features where the name field is not null are protected by the Water Quality Act based on their prominence in the dataset. Thus, these streams cannot be crossed without submitting to the Environmental Impact Assessment process. Smaller tributaries or drainages, while still having the potential to cause problems during project development, are less likely to have a significant impact. Figure 23 shows the named stream features for Los Angeles County.

Figure 23 – *Restricted Lines* dataset for Los Angeles County

5.2. Intermediate Results

The Primary Model produces many intermediate datasets that are superfluous and not preserved in the final model outputs. However, understanding these data in the context of their application is essential to understanding how the model functions as a whole.

5.2.1. Construction area identification

The first stage of the model uses the DEM covering the entire study area (*Study Area DEM)* and identifies areas suitable for construction of the terminal reservoirs. The first step produces a slope raster from the *Study Area DEM* for Los Angeles County, where slope angles are represented in degrees (Figure 24).

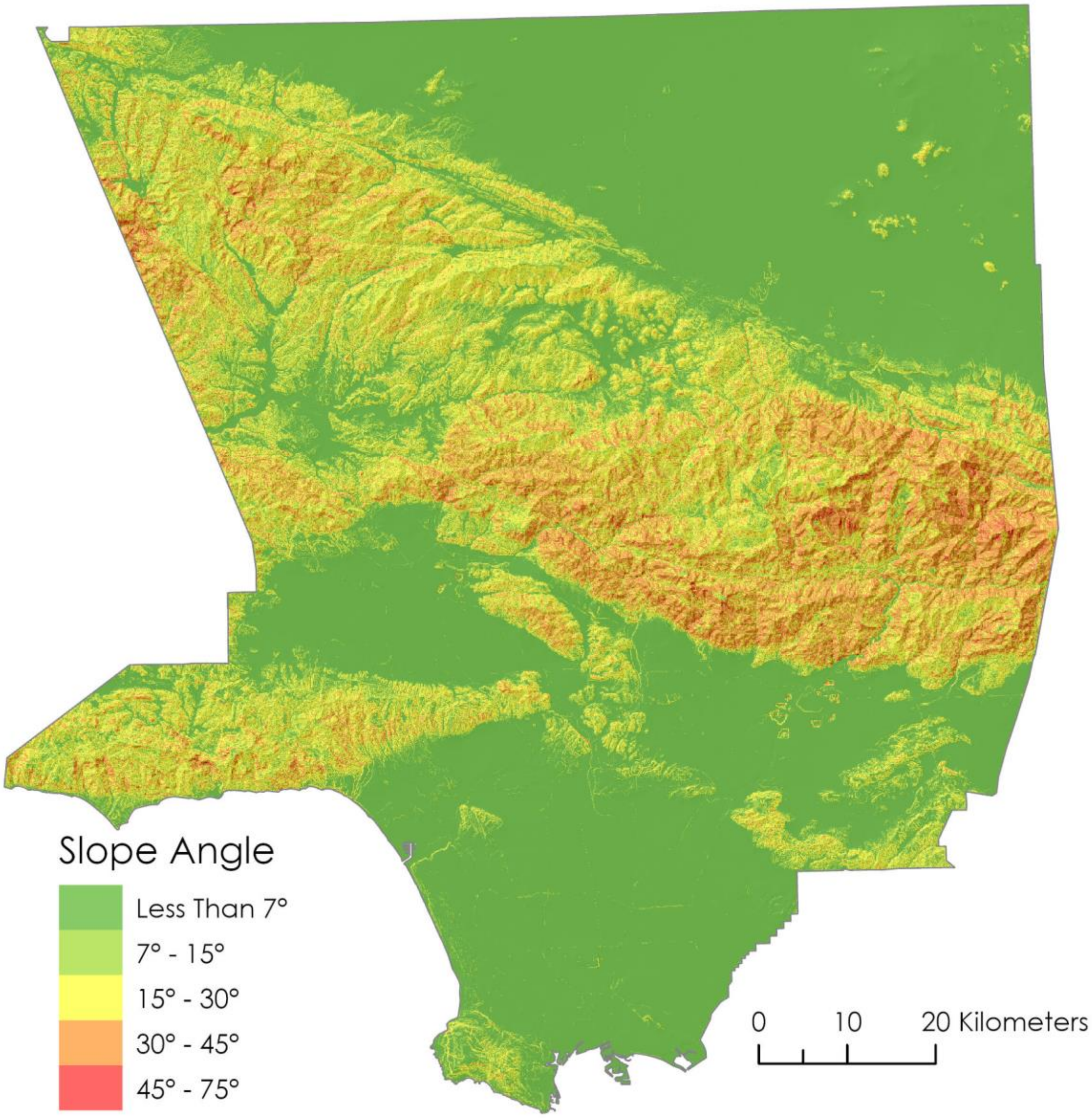


Figure 24 - Slope in degrees for Los Angeles County

The second step performs the moving window analysis producing a new raster dataset where each cell is given the value of the maximum slope angle in its 9 cell Moore neighborhood.

Each cell in the *Maximum Slope Raster* identifies the maximum slope in an area approximately 90m by 90m required for reservoir construction(Figure 25).

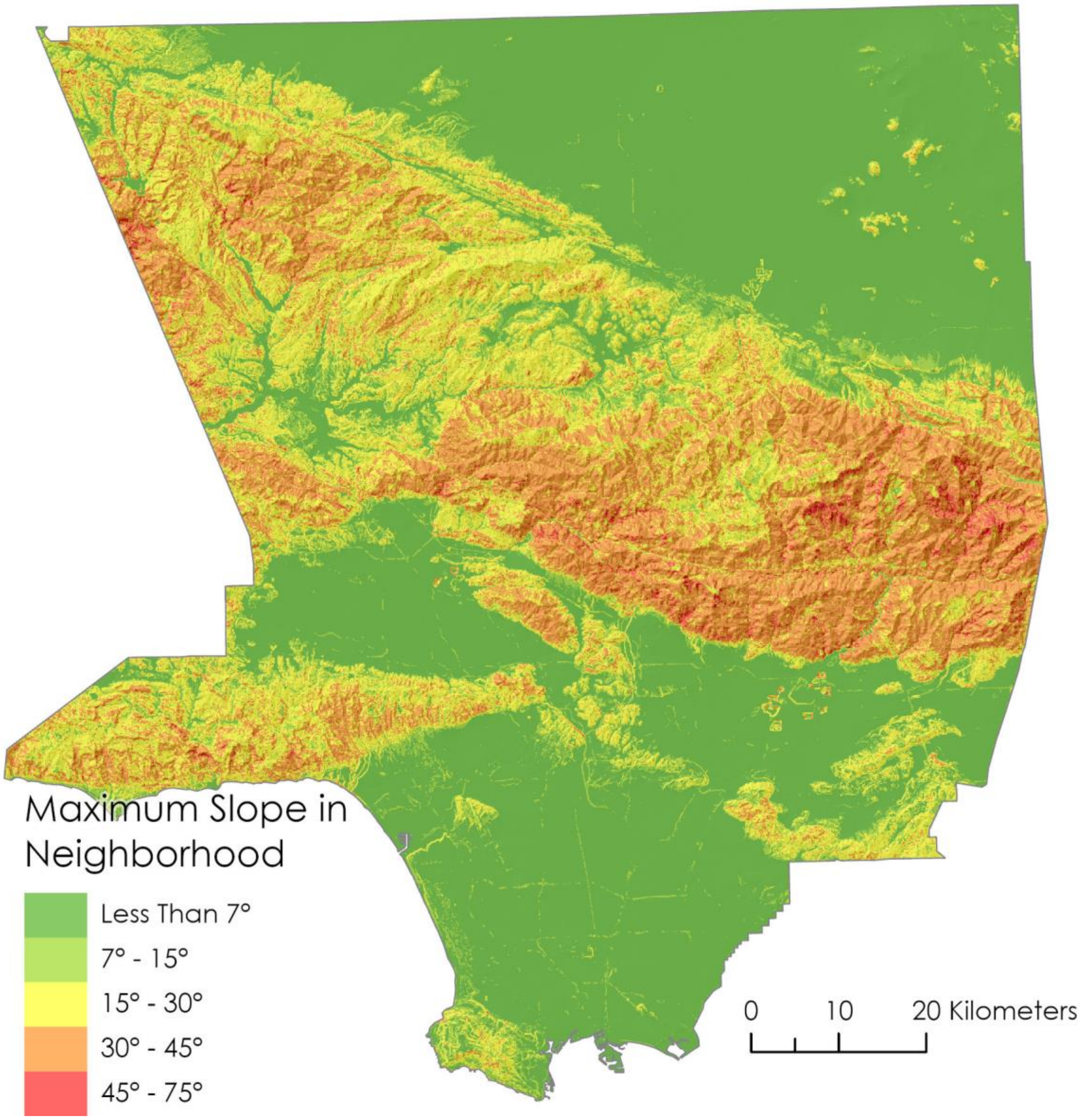


Figure 25 - The raster dataset showing the maximum slope for each nine-cell neighborhood in the study area.

Reclassifying the *Maximum Slope Raster* performs a binary pass/fail test on the dataset where cell values are modified to represent the suitability of an area for construction. In this example, if a cell value is found to exceed the design specification of 15° for maximum slope in the neighborhood, the cell is given a value of zero, and if it does not exceed that value, the cell is reassigned a value of one (Figure 26).

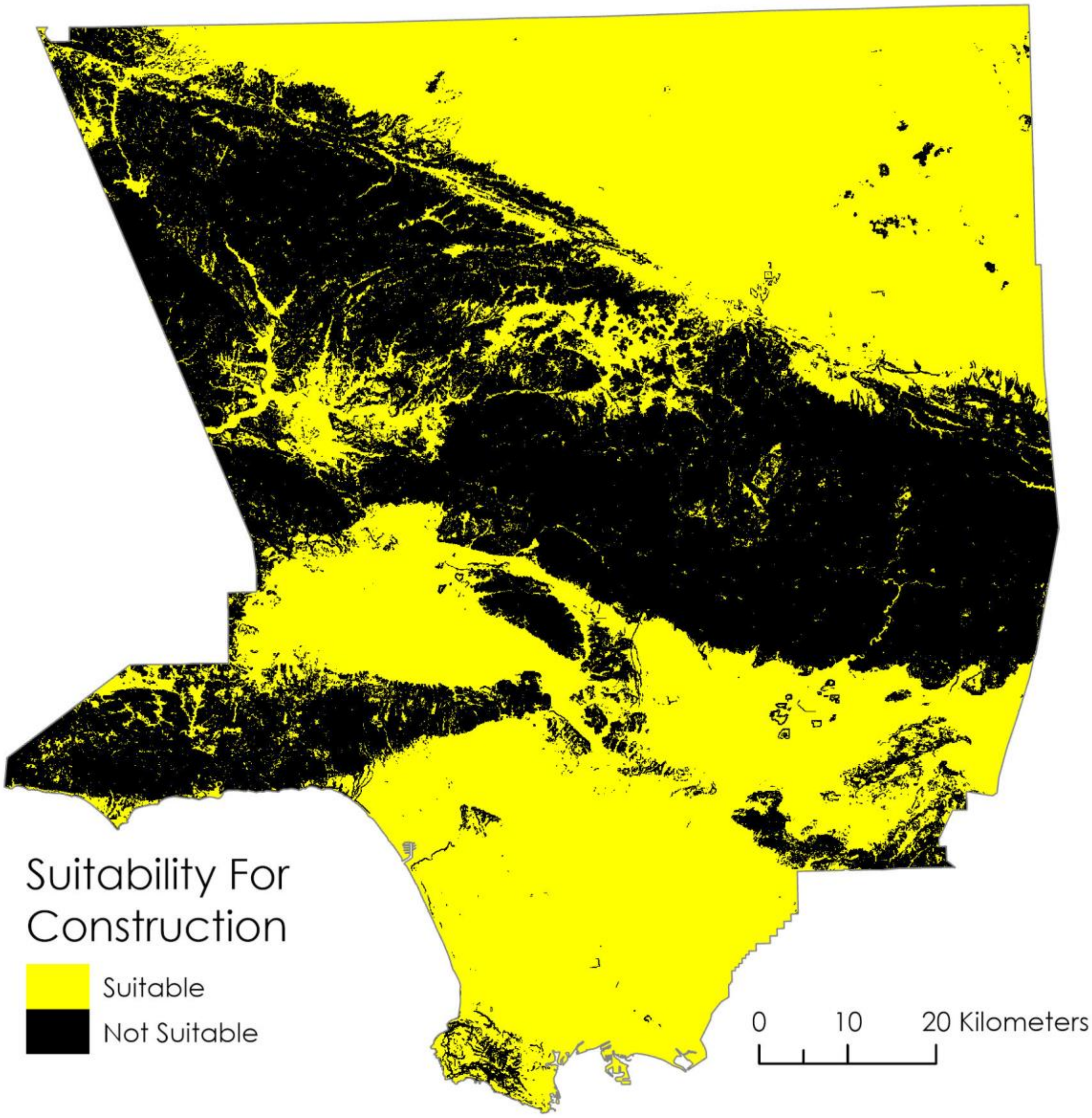


Figure 26 - The areas identified by the model as being suitable for the placement of a terminal reservoir.

The *Binary Screening Raster* is multiplied by the *Reclassified Maximum Slope Raster* (binary), where both datasets are comprised of values of either zero or one, thereby eliminating the areas that would be suitable for construction but fall within an area deemed not suitable by a component of the binary screening dataset. Areas identified as suitable in both datasets are given a value of 1. Figure 27 shows the effects of this overlay on the areas considered suitable for construction of a terminal reservoir.

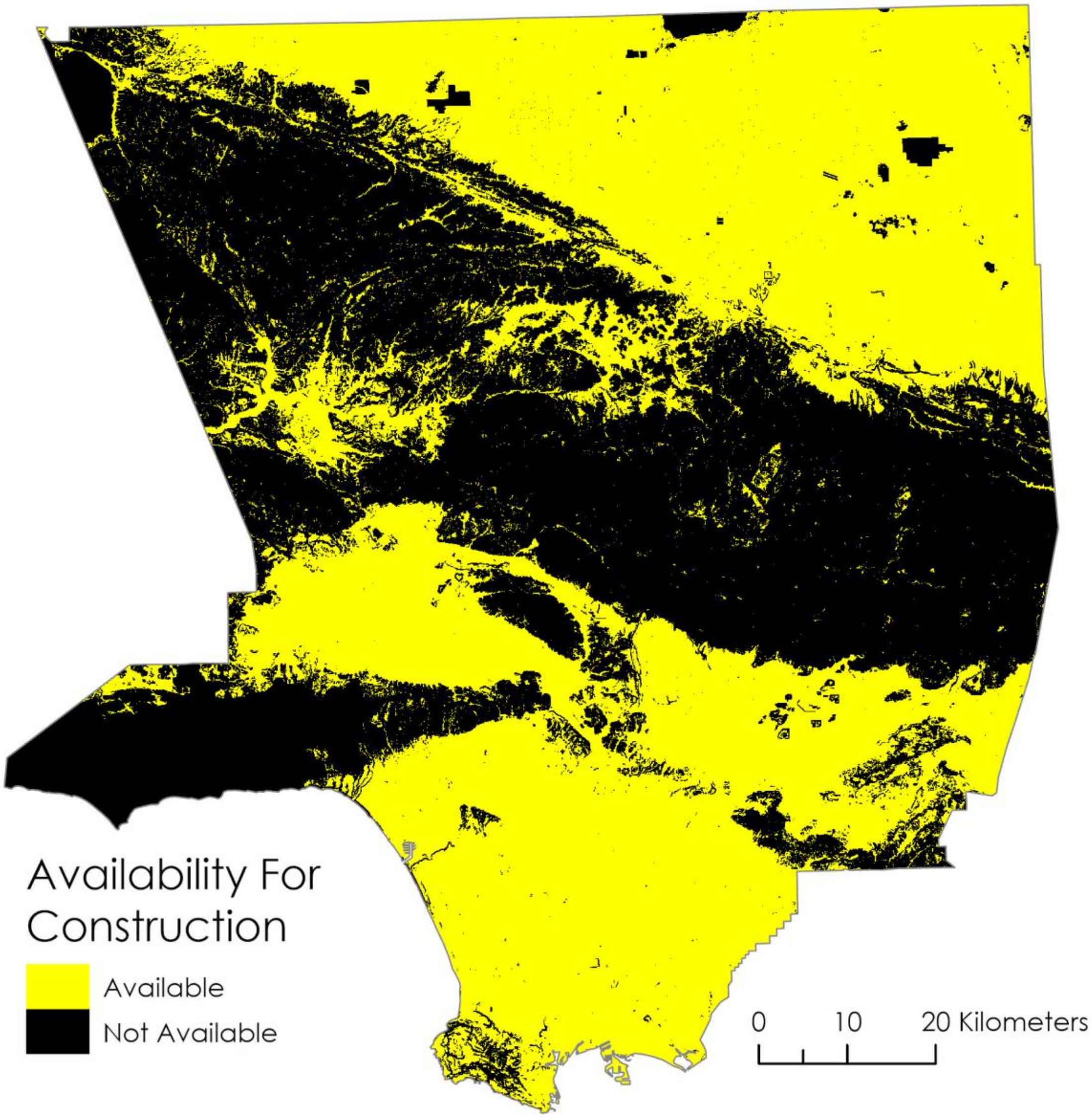


Figure 27 - The areas identified as suitable for construction of a terminal reservoir after the application of the binary screening layer.

The final step in this stage is comprised of two processes. The first process converts the areas deemed suitable for construction to vector polygons. The second process removes from the polygon dataset those polygons representing areas not suitable for construction, resulting in a set of polygons covering all of the areas suitable for constructing either the upper or lower reservoir components. This dataset is used to extracting the *Subset DEM (described below).*

5.2.2. Creating the Subset DEM

The *Subset DEM* for Los Angeles County was used to perform the remainder of the terrain analysis. Because only the areas identified as being suitable for construction area relevant to the remainder of the model processes, eliminating the extraneous information from the *Study Area DEM* reduced the processing load. Figure 28 shows the new *Subset DEM* extracted using the areas identified as suitable for construction. The subset DEM demonstrated the drastic reduction in areas to be considered, which eliminated most areas located in mountainous areas.

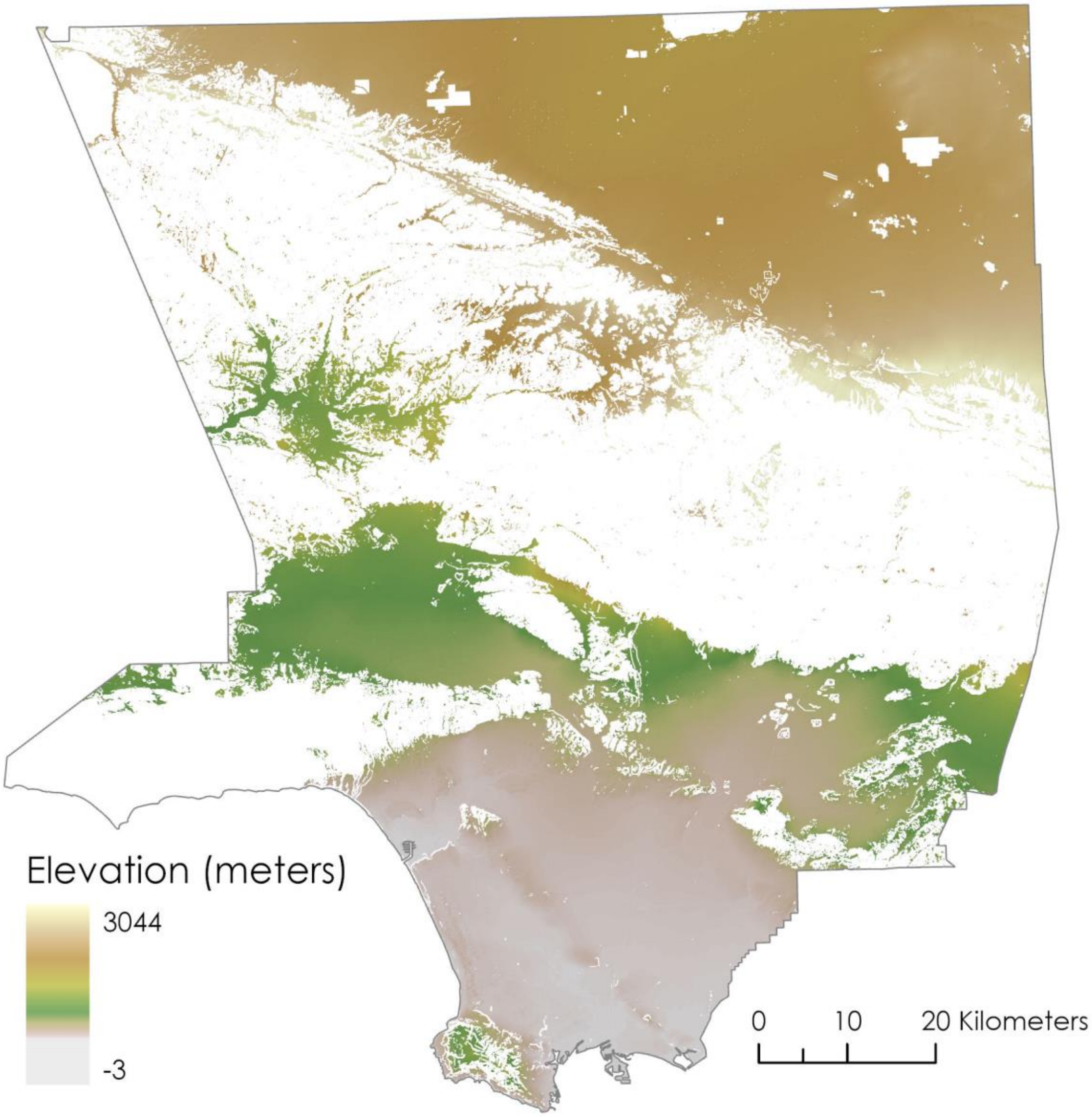


Figure 28 - The subset DEM created by using the areas deemed suitable for construction of a terminal reservoir to extract elevation values from the study area DEM

5.2.3. Searching for Relief

Working with the *subset DEM*, relief is calculated by a moving window analysis which looks in a 50 cell (1500m) radius from the focal cell to identify the lowest point within the search radius. A new raster dataset is then created where the focal cell assumes the value of the lowest point within the search radius. The same function is repeated in mirror, where the focal cell searches the same radius for the highest point, creating another raster dataset. Each raster is then compared to the subset DEM using a raster math function yielding two raster datasets that describe the maximum relief in both the upslope and downslope directions (Figure 29).

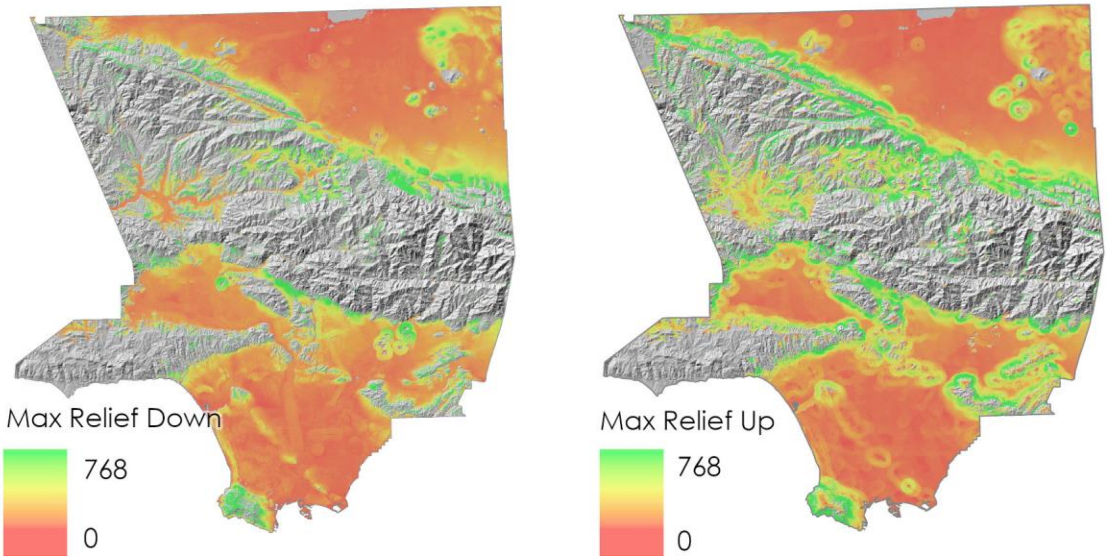


Figure 29 - Intermediate datasets showing the maximum downhill (left) and upslope (right) relief within 1,500m search radius.

The relief datasets were then reclassified using the design parameters for the system. In this example, the minimum relief from the upper reservoir to the lower reservoir is 300 meters. Thus, each of the datasets was processed to remove all cells where relief did not meet the minimum requirement. Figure 30 shows the reclassified maximum relief layers for both downslope (left) and upslope (right). These areas identify the preliminary reservoir locations for both the upper and lower reservoir sites, independently.

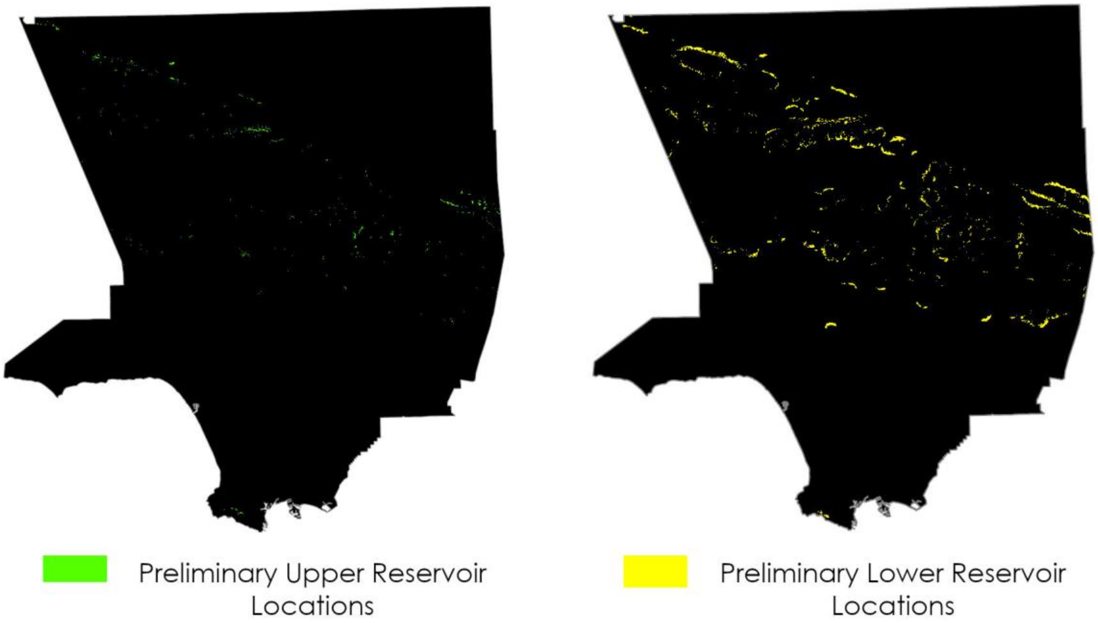


Figure 30 – Preliminary upper and lower reservoir location.

5.2.4. Making Connections

The process of making connections utilizes a near analysis which applies a many to one search, eventually producing a table matching each potential upper reservoir location to all of the potential lower reservoir locations within its 1,500-meter search radius. For the Los Angeles County example, the table included 3,156,665 lower reservoir location matches for 18,338 upper reservoir locations.

First, this step eliminates those connections that cross restricted line features. In this case study, as described above, a subset of the NHD streamlines dataset was used as the only component of the *Restricted Lines* dataset. Second, the model assigns the elevations at each of the connection’s terminal ends as attributes to each connection, then calculates relief for each connection. Connections that do not meet the relief requirement are identified as false match connections and eliminated from the dataset. This filtering process reduced the potential connections in Los Angeles County from 3,156,665 to 2,207,844.

5.3. Final Primary Model Results

The final steps in the model use an intersect function to examine the potential tank locations to ensure they are coincident with a previously identified as suitable is found not to be coincident with a connection, it is eliminated from the population. The filtering of the upper and lower reservoir locations (Figure 30) further reduces the locations identified as suitable for reservoir placement by 5% and 8% for the upper and lower reservoir locations, respectively (Figure 31).

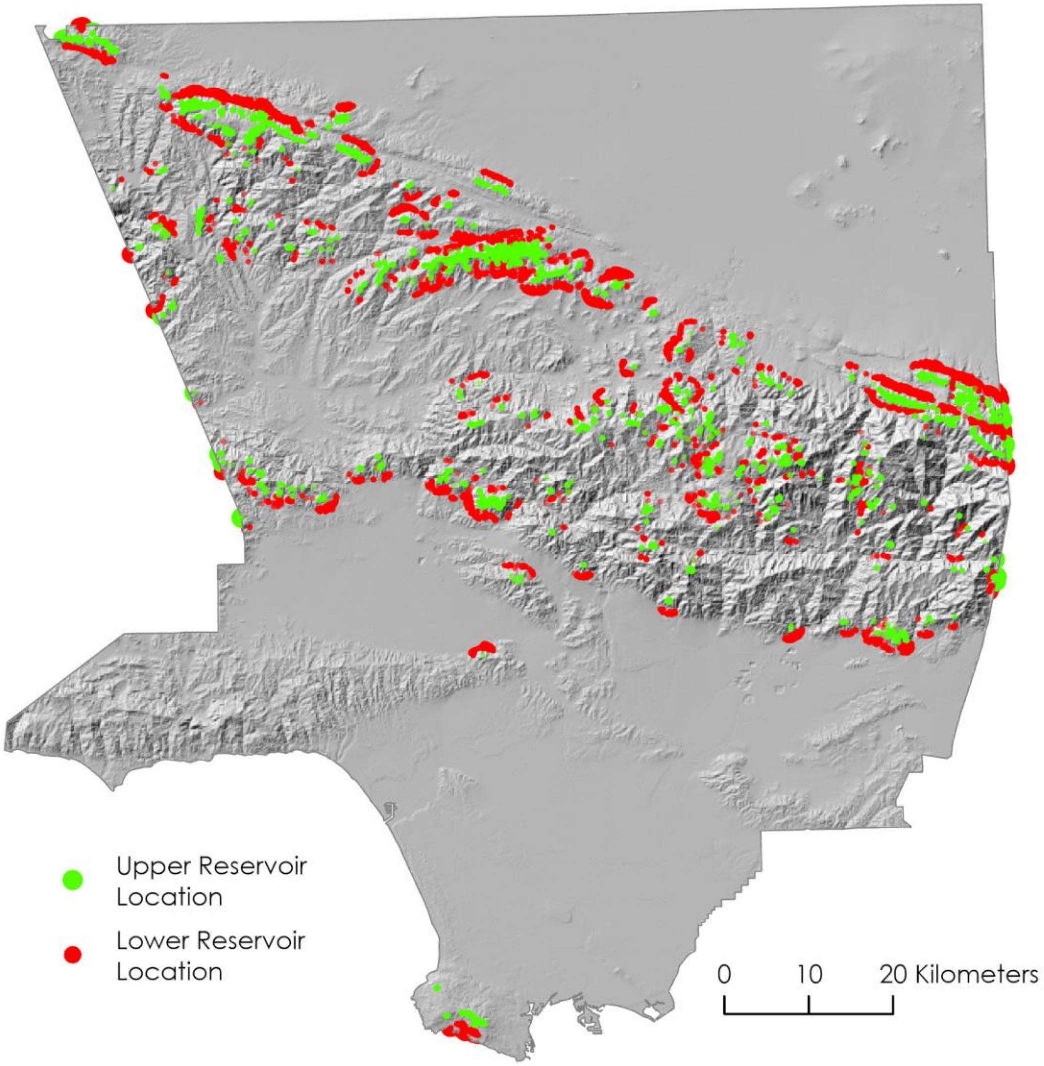


Figure 31 – A summary of final Primary Model outputs for Upper and Lower Reservoir. Locations exaggerated for visual effect

The Primary Model produces three final datasets covering the entire County of Los Angeles. Figure 32 shows a close up of these results for a small part of LA County. This perspective shows an array from a cluster of upper reservoir locations to a broader area identified for lower reservoirs. Additionally, Figure 32 demonstrates the nature of reservoir placement relative to the terrain.

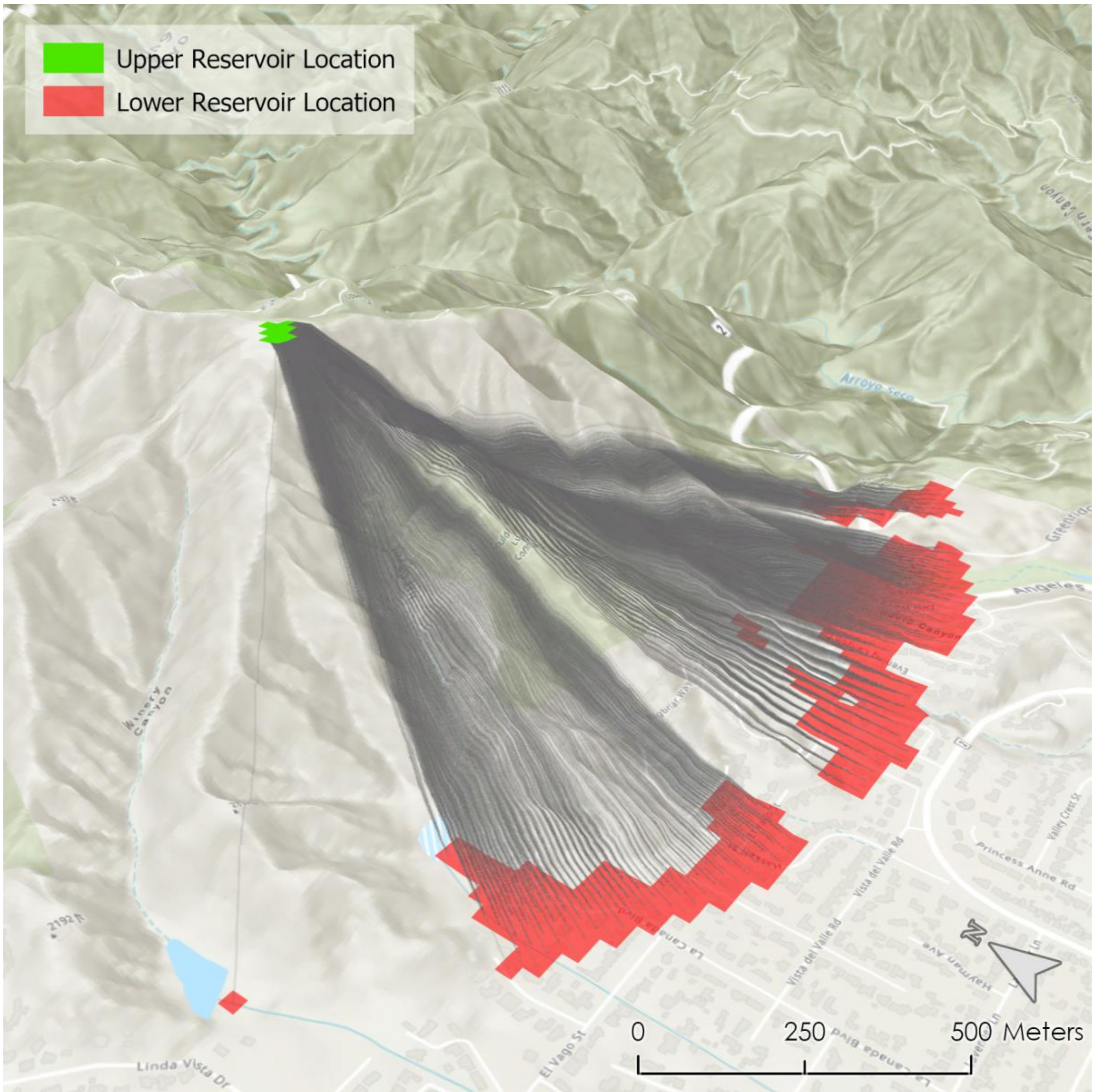


Figure 32 – A selection of final Primary Model results, La Canada Flintridge.

5.4. Secondary Model

For the example used herein, a fuzzy dataset was constructed using four spatial components. These include epicenter locations for historical earthquakes, green energy-producing facilities, roadways, and landslide susceptibility.

Each of these datasets was converted into fuzzy data using the conditions outlined in Table 6. Landslide susceptibility was obtained as a raster ranked from 1 to 10 where 1 is low susceptibility, and 10 is high. It was converted into a fuzzy set by reclassifying the value of areas less susceptible to landslides having a higher membership. Each of the vector datasets (Historical Earthquakes, Green Energy Facilities, and Roads) was converted into a distance raster where cell values denote the Euclidian distance from each feature.

Table 7 - Variables used to generate the fuzzy dataset used in the Secondary Model for the Los Angeles County case study

Input Dataset	Primary Data Type	Spatial Scale	Fuzzy Membership	Membership Type	Midpoint
Distance to Historic Earthquakes	Points	California	Higher membership with increased distance from Earthquakes	Large	30 km
Distance to Green Energy Production Facilities	Points	California	Lower membership with increased distance from Facilities	Small	20 km
Distance to Roads	Lines	California	Lower membership with increased distance from Roads	Small	1,000 m
Landslide Susceptibility	Raster	California	Dataset ranked 1 to 10; fuzzy membership ranked .1 to 1, where higher values have a lower membership	Small	5

The resulting *Fuzzy Raster* incorporates the four fuzzy sets into a single raster dataset using a with values ranging from 0 to 1, infinitely.

Review of both the fuzzy sum and fuzzy product outputs indicated that these functions provide the extreme ends of the membership spectrum. The fuzzy product output provides a dataset where low membership dominates, while the sum function provides the opposite where most of the study area has higher membership.

Ultimately, the overlay used in the case study was produced using a fuzzy gamma function, where gamma was equal to 0.9. Documentation for the Fuzzy Overlay Tool provided by Esri reports that the gamma function performs an algebraic combination of both the product and sum fuzzy overlay functions serves as a compromise function, and therefore, may provide adequate results for the demonstration of this model feature (Esri 2016b).

Figure 33 shows the *Fuzzy Layer* for Los Angeles County before integrating it with the Primary Model results. Using the fuzzy sets considered, the overlay shows that large portions of low-lying areas have high membership, while areas in the mountainous areas show lower membership. Membership values for the county range from near zero to 0.99. Isolated points of low membership in the Los Angeles basin correspond to epicenters of historical earthquakes.

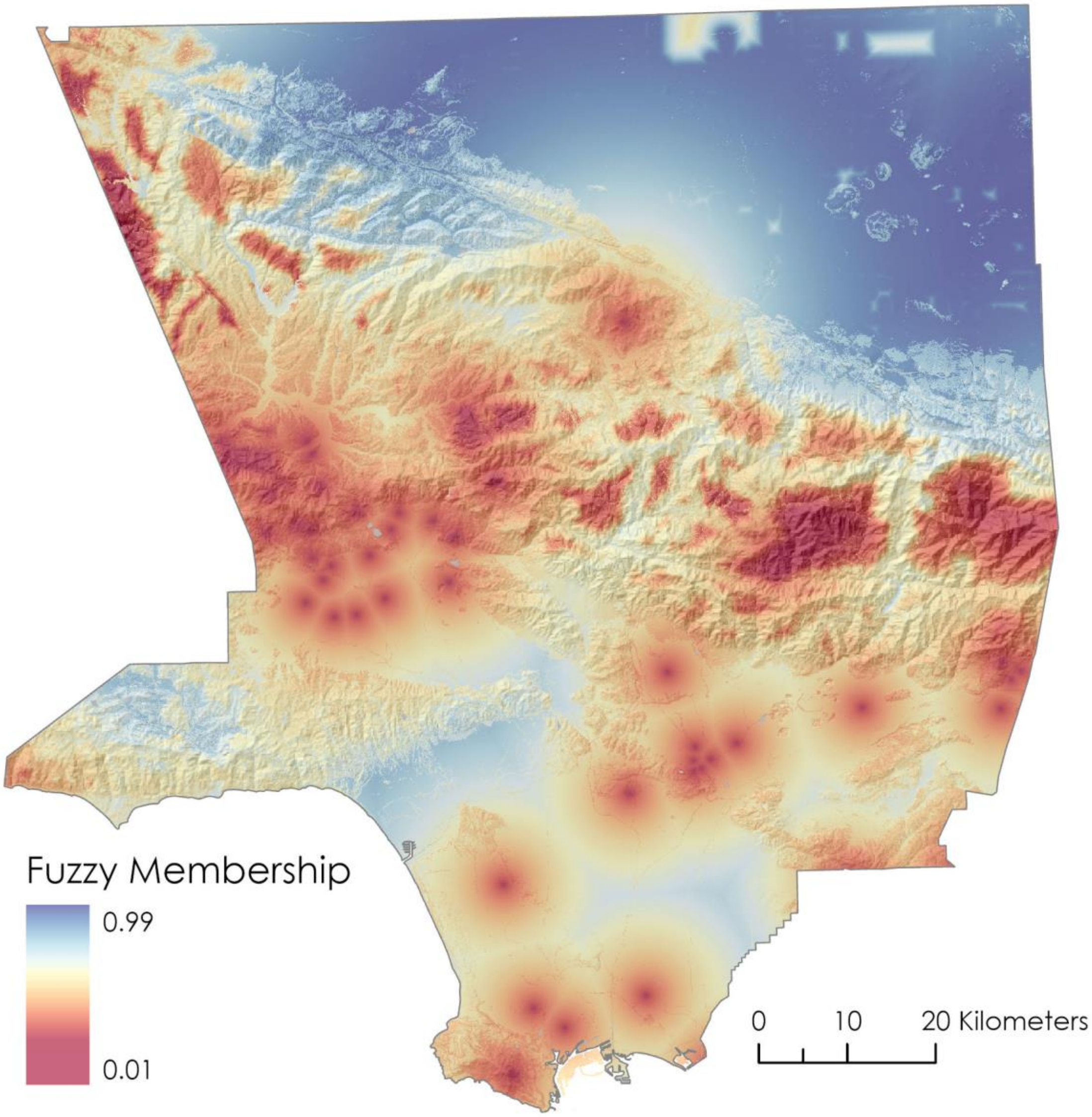


Figure 33 - The *Fuzzy Layer* for Los Angeles County.

Figure 34 shows results of the Secondary Model outputs for both the upper and lower reservoir locations in a selected area of Los Angeles County. Membership value for model outputs indicates the fitness of that location to support the placement of the designated terminal reservoir. In the area shown, all areas identified as suitable for upper reservoir locations have high membership values for suitability. For lower reservoir location, most areas have been identified as having low membership.

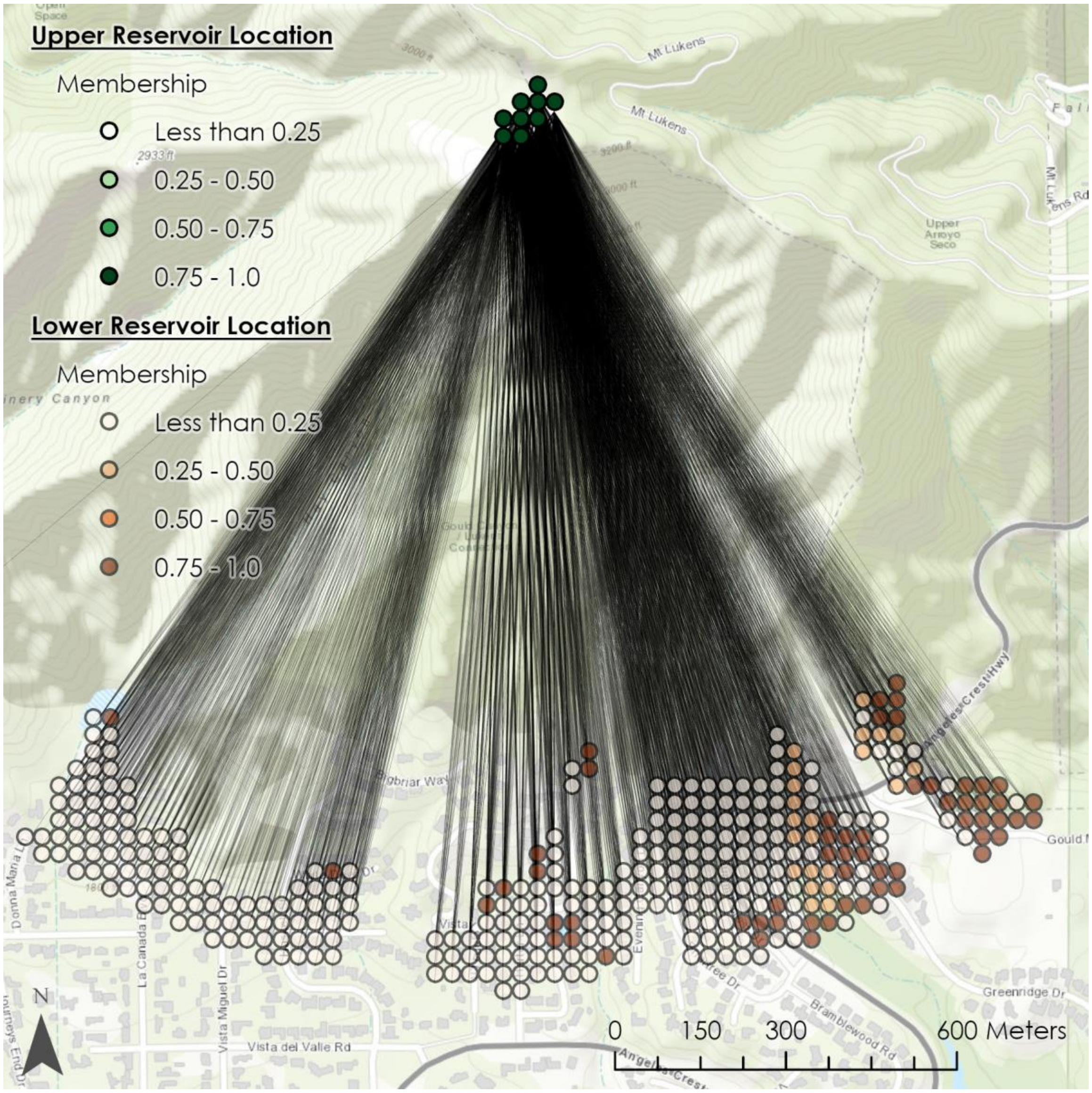


Figure 34 – A selection of final model results near La Crescenta, California, after the integration of the fuzzy layer.

In this case study the distribution of membership for the reservoirs, both upper and lower, occurred in a trimodal pattern, where membership for reservoir locations was clustered near zero, near one, or near the center. The mean membership for upper reservoirs was slightly higher than center while the mean for lower reservoir membership was slightly less than center (Figure 35).

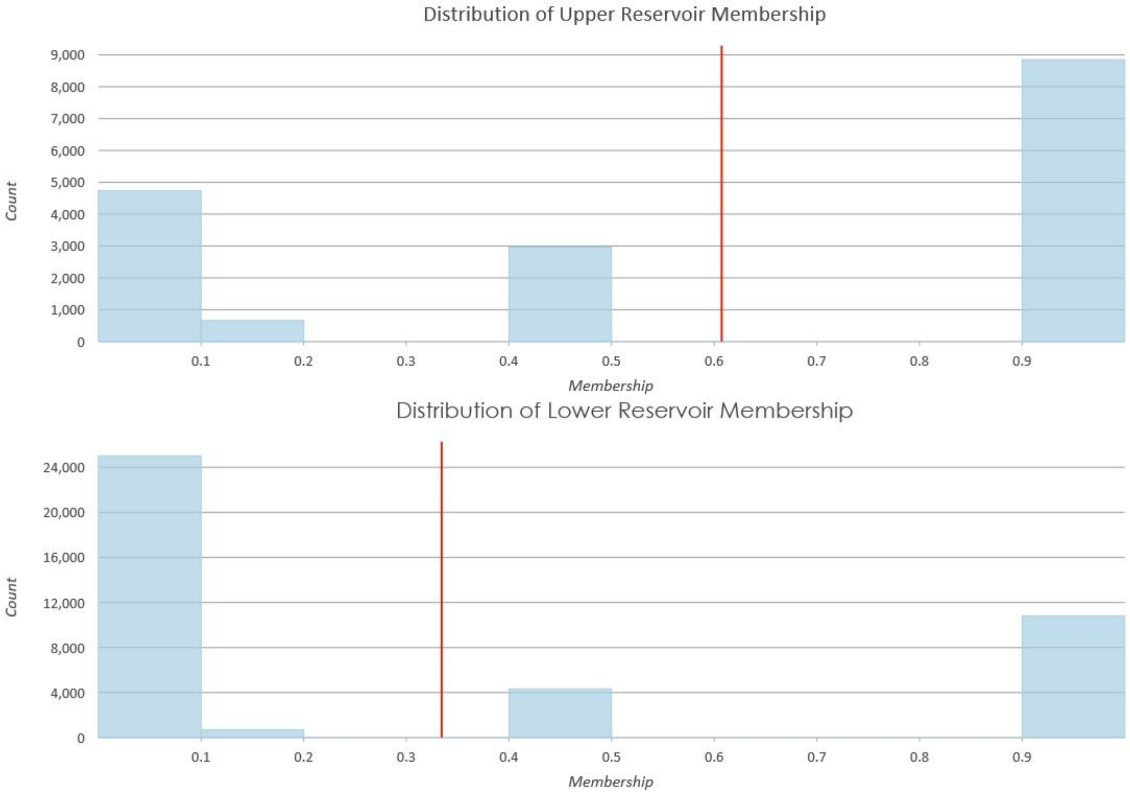


Figure 35 – Histograms showing the distribution of fuzzy membership values for upper and lower reservoir locations. The red vertical line on the histograms indicates the mean membership value for each dataset.

5.5. Model Performance

Review of the final Los Angeles County data outputs for both the Primary and Secondary Models shows that there are many possibilities for the deployment of MPSHS. The model identified 12,716 locations suitable for the placement of the upper reservoir and 40,868 potential lower reservoir locations. The disparity of these populations due primarily to the fact that in this geomorphic setting the toe of most slopes and valley floors are geomorphically more stable and therefore have a higher likelihood of supporting land suitable for construction of the required infrastructure.

Geomorphic controls on terrain vary by study area location and have a significant impact on the expression of terrain features. In California, there are many different geomorphic provinces that each contain unique topographic features that this model can exploit. Using examples from Los Angeles County, it is apparent that the model placement of lower reservoir location tends to prefer the wide valley floors, where sediment has accumulated as alluvium. Due to its arid nature, low precipitation and corresponding low amounts of runoff create the extensive alluvial plains abutting steep, rugged slopes with minimal soil development (Norris and Webb 1990). This corresponds to an abundance of lower reservoir location relative to the population of corresponding upper reservoirs.

5.5.1. Primary Model

In Los Angeles County, the Primary Model identified 2,207,884 possible connections linking upper reservoir locations to lower reservoir locations. Relief calculated for these connections ranged from the minimum 300m to a maximum of 768m. The mean relief for the population of connections was 343m, indicating that the bulk of the connections identified by the model occurred near the design limit.

Another useful indicator of viability is the relief gradient, or the lateral distance needed to achieve the desired change in head. For the Los Angeles County example, the mean lateral distance required was found to be 1,352m with a standard deviation of 127m. These statistics show that placement of upper and lower reservoirs in Los Angeles County approached the engineering limitations set forth by the system design, with 414,589 of the 2,207,884 connections existing at the lateral limit of 1,500m. In other words, the number of viable connections increased as the lateral connection distance increased. Figure 36 demonstrates the relationship between the number of connections meeting the relief requirement and the lateral connection distance.

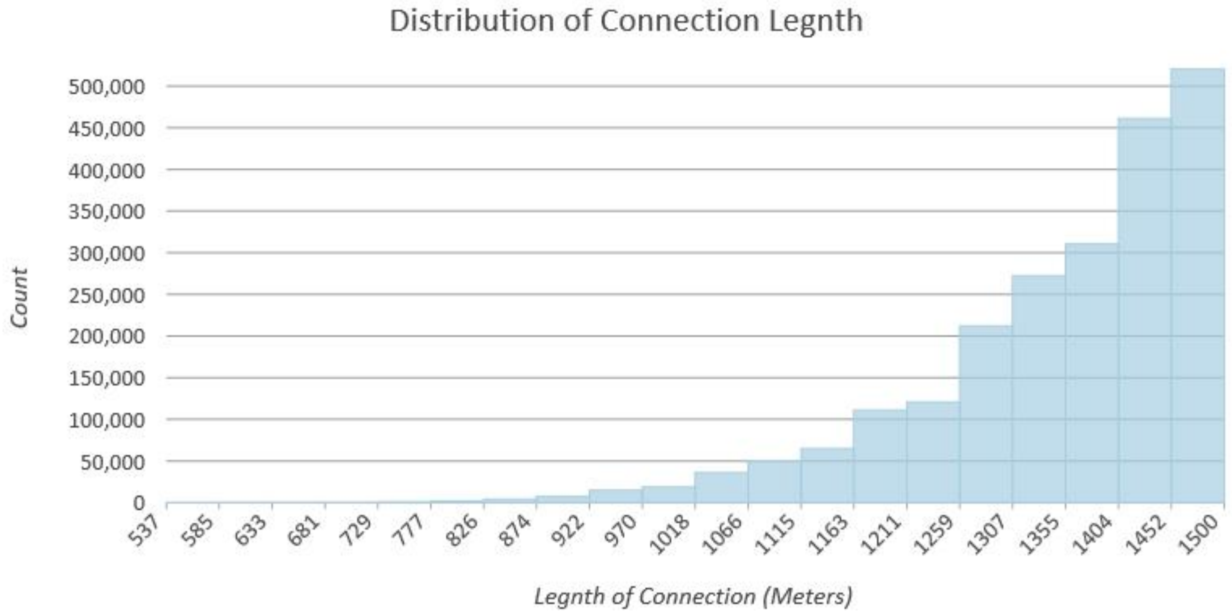


Figure 36 – Distribution of connection lengths

Intuitively, this relationship is logical. As the design parameters allow for increased distances between the upper and lower reservoirs, the number of potential reservoir locations and in turn connections will increase. In other words, as allowable lateral distances increase, so does the number of potential reservoir locations; thus, more placement opportunities. Conversely, a similar trend can be found for the minimum relief parameter. As the minimum required head (relief) decreases, the number of potential reservoir locations increases (Figure 37).

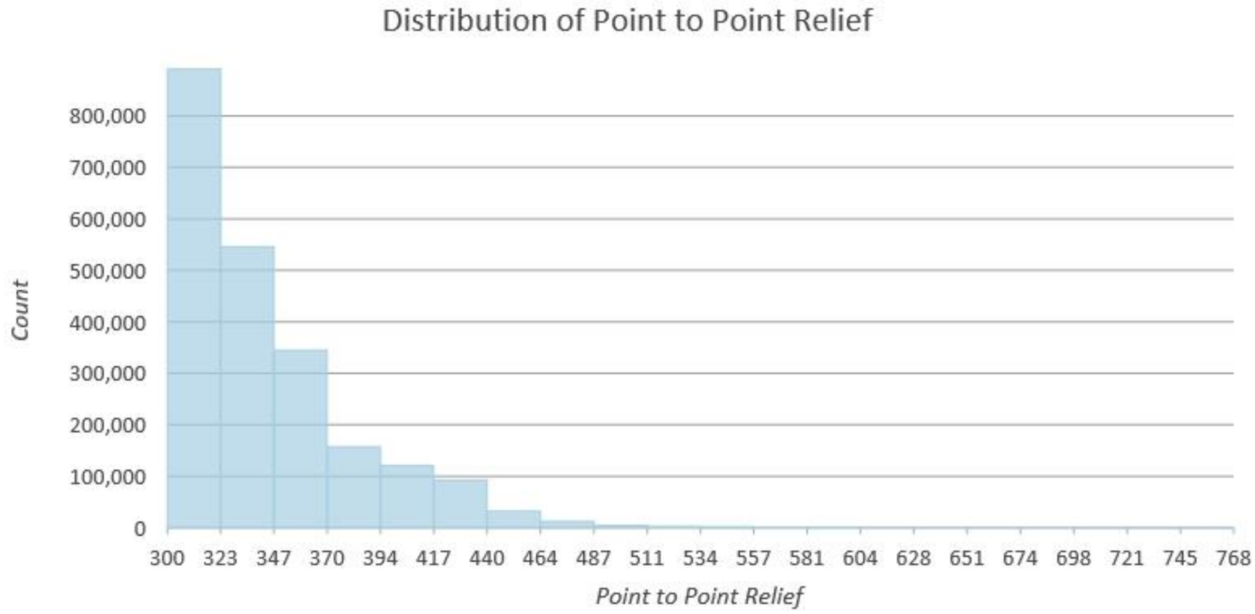


Figure 37 – Distribution of point to point relief magnitude

This observation shows that the lower the minimum relief requirement demanded by the MPSHS design, the higher the potential for suitable locations. The relationship between the minimum relief required for the MPSHS to function and the number of potential locations for placement is an essential component when considering the spatial extent for the analysis. For example, in an area where relief gradients are less, connections may need to be longer than 1,500 meters. Thus, engineering requirements would need to be modified, and system components reevaluated.

The model results for the upper and lower reservoirs were compared to topographic data and features to determine the validity of selections. Figure 38 demonstrates a collection of these observations in an area near Palmdale, California. It should be noted that each node, represents the center of a 90m by 90m area, suitable for construction of a terminal reservoir.

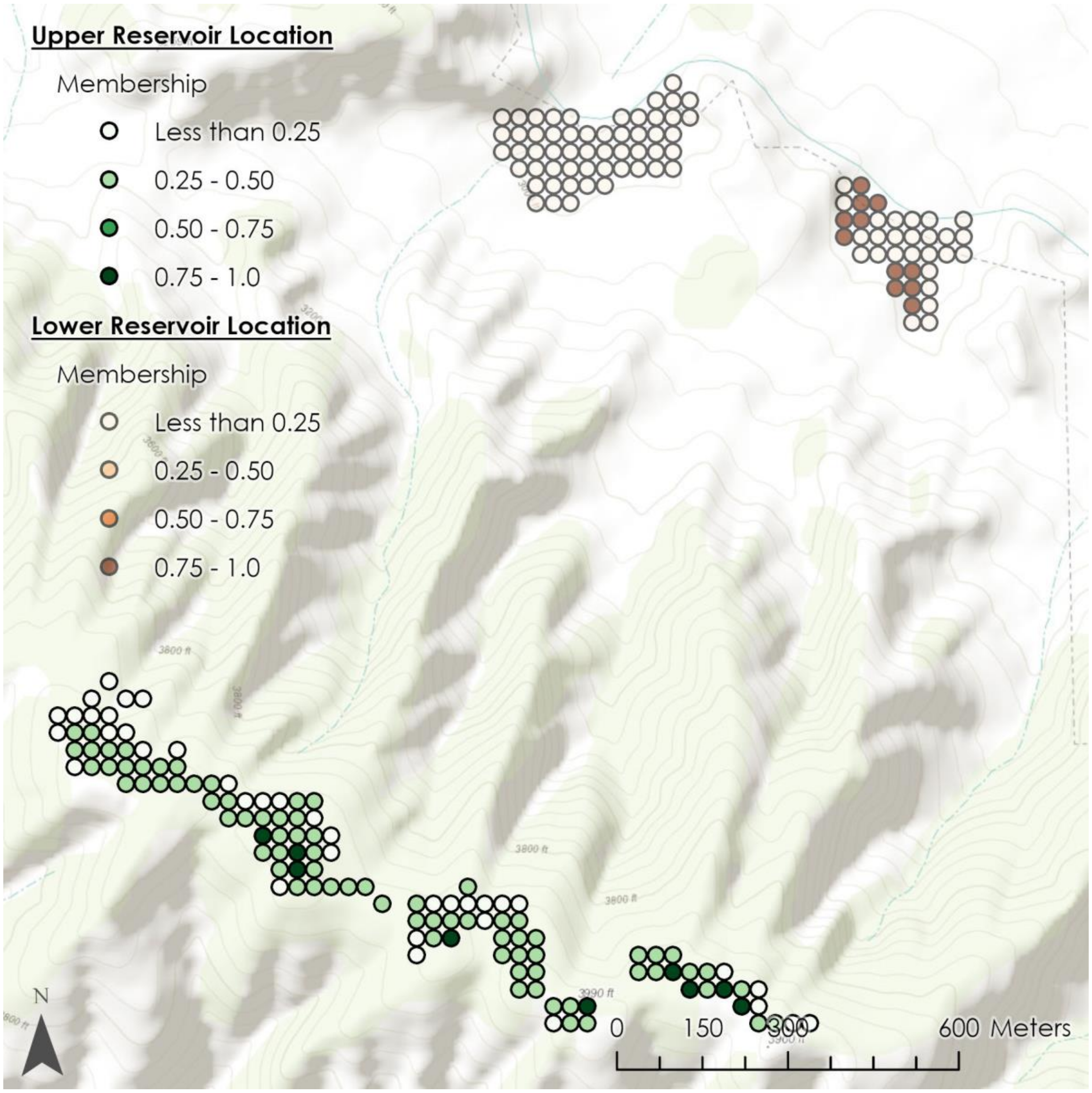


Figure 38 – A sample of Primary Model Results for upper and Lower reservoir locations

The placement of upper reservoir locations occurs along a ridgeline where the topography is relatively flat. Some upper reservoir locations occur on the top of the slope while others occur on both flanks. While all of these meet the relief requirement of 300m to the toe of the slope where the cluster of lower reservoir locations, it would likely be impractical to utilize suitable locations on the backslope of the ridge (the side opposite the paired lower reservoir). Figure 39 demonstrates the application of the restricted line constraints as there are no lower locations present on the side opposite the stream feature; in this case, the California Aqueduct. Additionally, proximity to renewable energy is demonstrated as the photovoltaic solar panels are shown adjacent to a cluster of suitable sites in the lower right of Figure 39.

Figure 39 – Model results near Palmdale, California.

Lower reservoir locations are determined first based on the identification of paired upper reservoir locations through the use of the near analysis, as most study areas are dominated by flat terrain suitable for construction of a lower reservoir. Conversely, upper reservoir locations are often limited to geographically isolated areas such as ridge tops (Figure 39) or, more confining features such as hilltops (Figure 40). Based on this relationship, clusters of lower reservoir locations often occur 1) at the toes of the slopes on which their paired upper reservoir is located and 2) form a fan shape with the leading edge representing the maximum lateral distance for a viable connection.

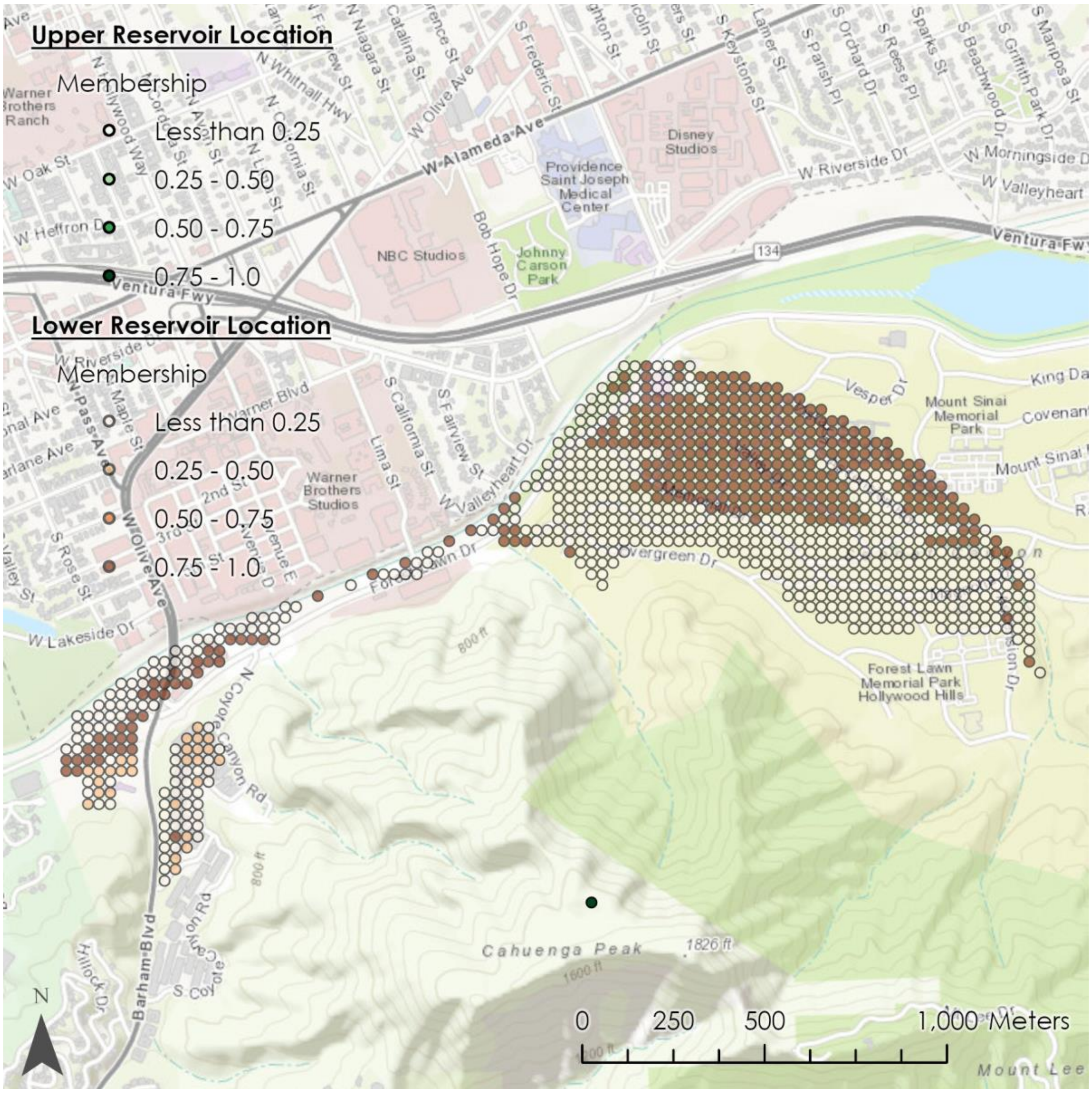


Figure 40 – Example results from the Primary Model outputs near Universal City, Los Angeles County, California. Note there is only one upper reservoir site indicated.

The modeled locations for lower reservoir connections shown in Figure 40 also demonstrate the effectiveness of the Construction Area Focal Analysis, whereby any area with topography not suitable for construction is eliminated from placement consideration.

The number of connections in the Los Angeles outputs is enormous, with over 2 million connections. This large quantity of connections is due to the relationships between the upper and lower reservoir locations, as it is only in rare circumstances where a single reservoir node exists in isolation. Typically, reservoir locations occur in clusters and have one-to-many relationships with other clustered reservoir pairs that meet the criteria. In this case, one upper reservoir may serve as a match for many lower reservoir locations. Such is also true for other upper locations in the cluster. The result is an array of connections from each upper location to each of the paired lower reservoir locations (Figure 41).

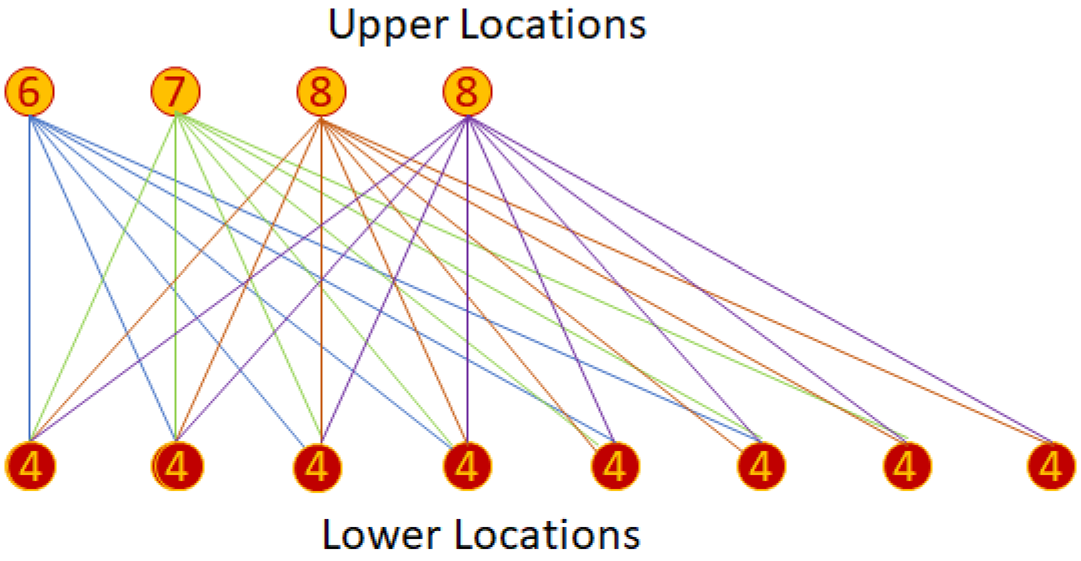


Figure 41 – A graphical example of the one to many relationships between upper and lower reservoir locations

The usability of the model results is complicated by the enormous dataset produced for the connections. As described above, there are over 2 million connections for the Los Angeles

County study area. A closer look at the results shows how clustered results overlay one another such that any one of the connections of the upper reservoir and paired lower connections are difficult to discern making the practicality of this dataset questionable without an interpretation workflow focused on refinement and filtering of the results (Figure 42). The large areas of black shown in the figure are the overlapping connections linking all of the upper reservoir locations to all the possible matched lower reservoir locations.

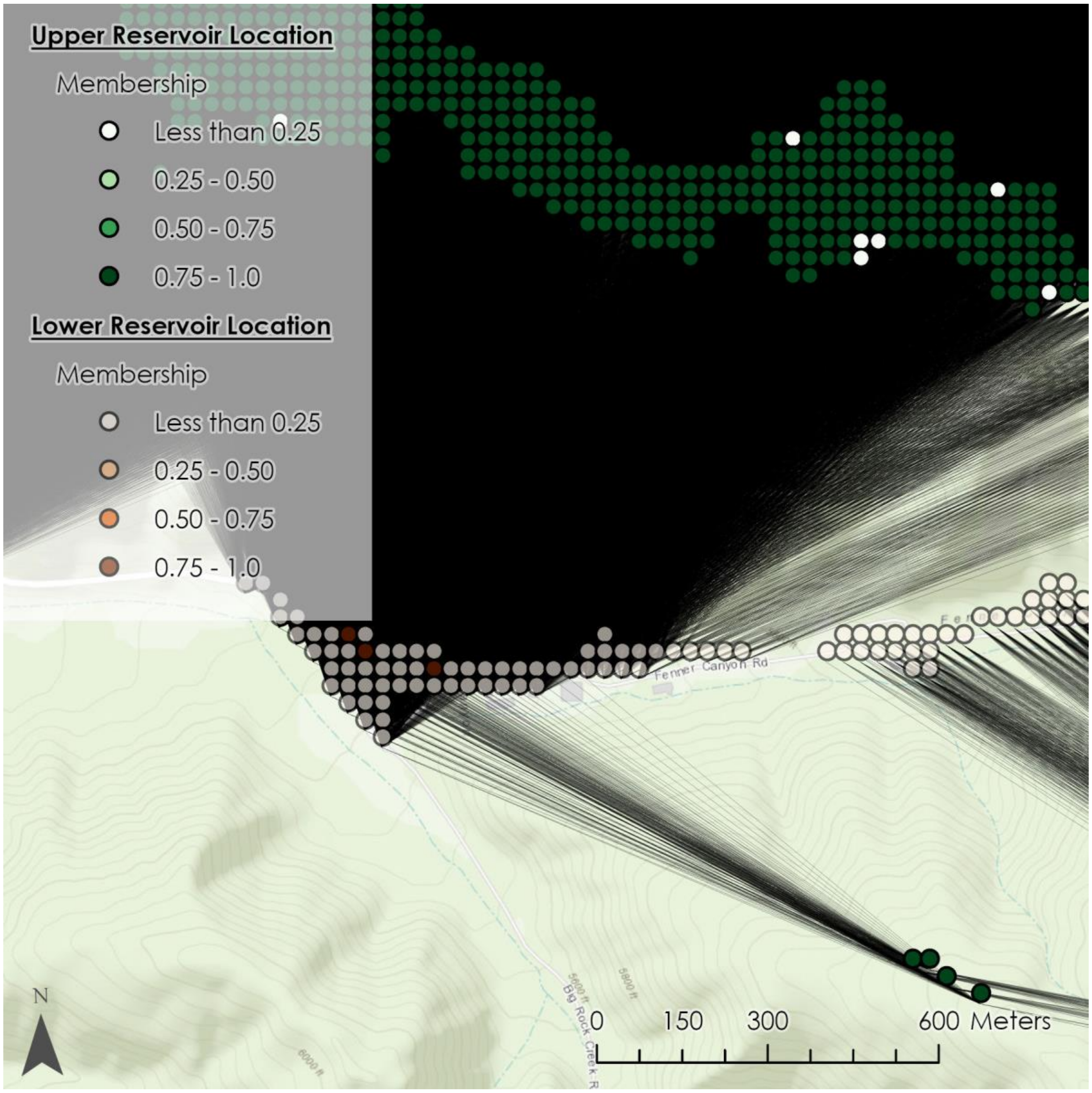


Figure 42 – Complex relationships in clustered reservoir locations

Selecting a viable location for the MPSHS could use a top-down process similar to the model itself, which may be ideal in study areas where the limiting factor in suitability is upper reservoir placement. In other words, first, find a suitable upper location from the final dataset and then filter out all of the connections that do not originate from that point. The final selection can then proceed from the paired lower reservoirs. As shown in Figure 43, this filtering process would eliminate a significant portion of the possible lower reservoir locations. Of course, the process could also be done in reverse where all of the possible upper pairs are found using a lower reservoir location.

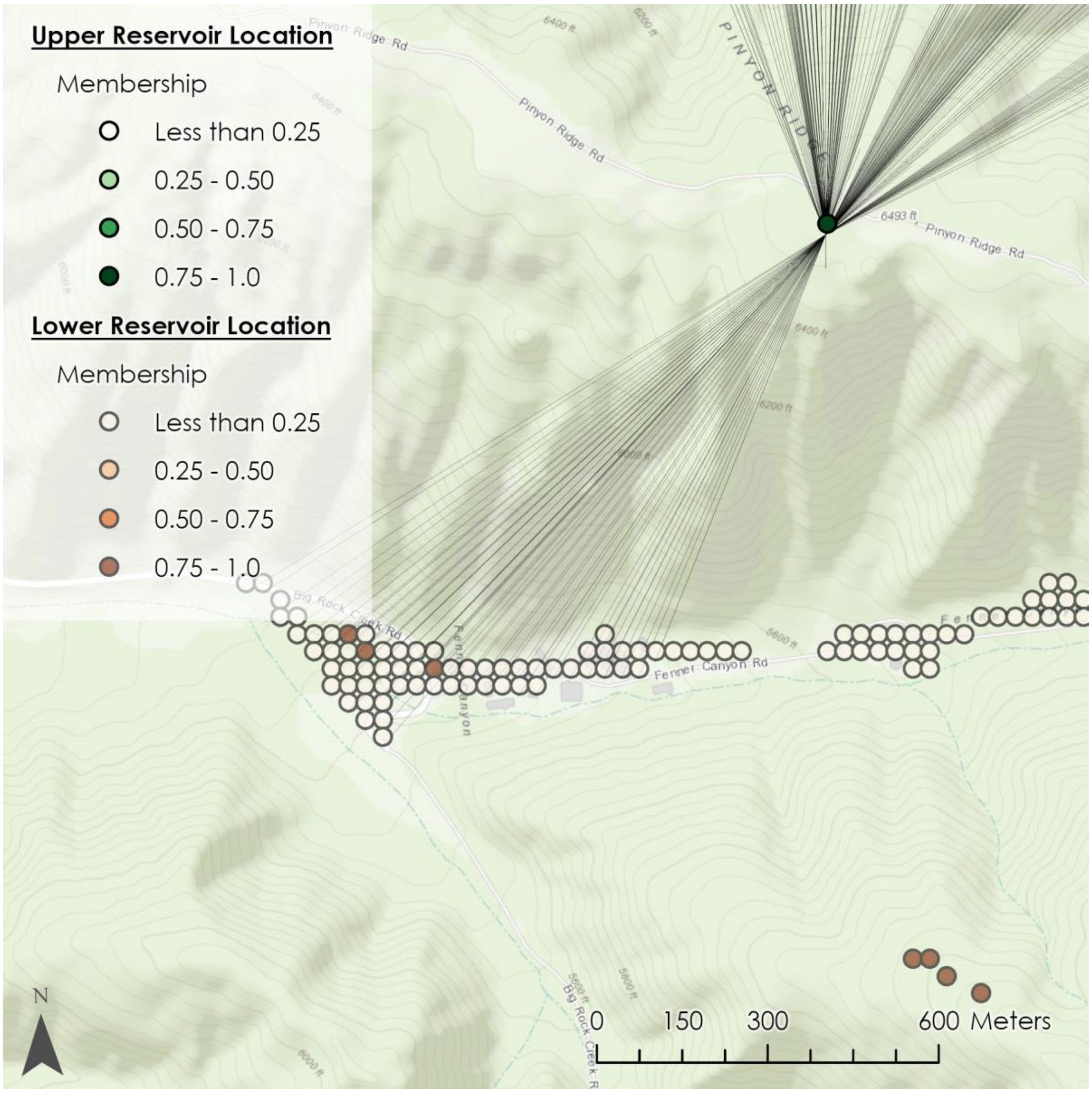


Figure 43 – Results filtered to originate from a single upper reservoir location

The filtering process can be easily implemented using the attributes of the connection lines themselves. Each line has the feature ID for both its origin (upper locations) and its endpoint (Lower Locations). By running a query to select all connections that are associated with the target feature, either an upper or lower reservoir location, a simplified set of connections can be identified.

Examining the results for connections further indicate that there is an unknown component with regard to the character of the terrain between the reservoir locations. While this is obviously a variable in the construction feasibility for construction, it is outside of the scope of this study.

5.5.2. Secondary Model

Secondary Model results for the Los Angeles study area show strong clustering of membership values in the fuzzy dataset when applied to the locations identified for reservoir placement. Using the *Fuzzy Layer* dataset described above, the Secondary Model for Los Angeles County resulted in reasonable results with respect to suitability. As expected, in the greater Los Angeles area, distance to roads was not a major limiting factor within the fuzzy dataset. Historical seismic activity expressed as isolated incidences of low membership. Additionally, green energy production criteria graded over a long distance in its fuzzy set, so this variable would be unlikely to affect small scale suitability (i.e., reservoir location within a cluster). Therefore, the dominant influencer on differences in the membership of clustered reservoir locations in the Fuzzy Layer is most likely landslide potential dataset.

Additionally, the scale of the fuzzy sets applied the membership function over an expansive area with highly variable distributions in phenomena. For example, historic earthquakes were used as a base dataset. Conversion into a fuzzy set was accomplished by applying a higher membership based on distance from the epicenter.

While sound logically, the resulting fuzzy set was biased toward those parts of the state, not in seismically active areas. Furthermore, the midpoint function was based on a static 30km radius, decreasing from that point. This ensured that the closer to an epicenter (zero to 30km) membership would be below 0.5. Beyond 30km the membership would be higher at a rate normalized to the furthest distance in the Euclidian input distance to epicenter raster.

If the *Fuzzy Layer* and its parent fuzzy sets were created at the study area scale, the outcome might have been not only different but more applicable to the area of interest. In these cases, more study area-specific choices in variables could be considered.

Another critical element of the Secondary Model is the selection of the Fuzzy Sets themselves. The four variables selected for use in this case study were selected for their simplicity of presentation and processing. The parameters by which they were transformed into fuzzy sets were equally as simple. Primarily, the functions looked at a membership based on simple distance parameters (i.e., higher membership values for proximity to roads). This likely contributed to the uniformity in outcome with respect to the statistical distribution of membership values, especially in the chosen study area, where the density of the selected variables is significantly different from the remainder of the state.

Another consideration for potential future use of the Secondary Model is the joint membership function used in the overlay process. For this example, a gamma function (gamma = 0.9) was used to provide a result that acted as a compromise between the product and sum functions. This used an algebraic combination of both the sum and product functions raised to the power of gamma. As with other aspects of overlay analysis, careful consideration should be applied to all components when selecting the parameters for the Joint Membership Function.

Although the development of the Fuzzy Overlay used in the model case study was intended for demonstration only, it does expose some of the difficulties with respect to scale, fuzzy criteria, and fuzzy membership function selections. The *Fuzzy Layer* created from the overlay process was constructed to apply a uniform fuzzy overlay to the Los Angeles County study area and each of the alternate study areas to which the model was applied during development. This means that the fuzzy membership functions were applied uniformly to the entire statewide datasets, rather than on a study area-specific scale. Parameters of the fuzzy membership functions were chosen based on limited data and many assumptions. In a practical use scenario, there should much more consideration placed on the creation of fuzzy sets and the decisions regarding the fuzzy membership functions for the overlay analysis.

Thus, the *Fuzzy Layer* used in the Secondary Model case study for Los Angeles County is a simple version of an end-user provided *Fuzzy Layer* and was developed to show feasibility, rather than present defensible real-world application.

5.5.3. Alternative Study Areas

During the development process, the model was tested in multiple locations within the State of California to ensure that it would work under various geographic conditions. These locations were Mono County, Butte County, Santa Clara County, and Yolo County. These alternative study areas are shown in Figure 44.

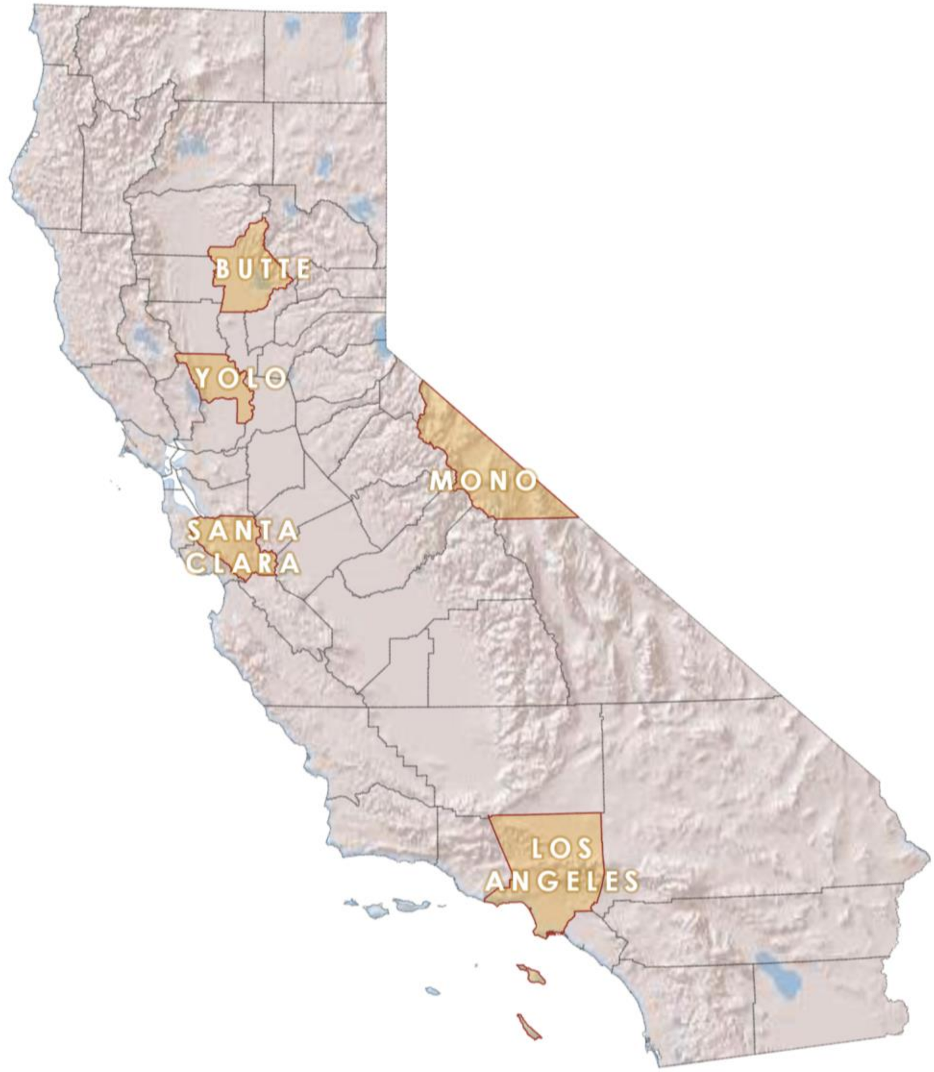


Figure 44 – Distribution of alternative study areas in California

The same process described above for Los Angeles County was applied to the other study areas. To maintain consistency in results for each study area, the *Binary Screening Layer, Restricted Lines* dataset, and the *Fuzzy Layer* were created to cover the entire state. For each study area, the county boundary was used to extract the data to the study area limits.

Collectively, the test areas presented a diverse range of terrain types to test the Primary Model. They span the geomorphic provinces of the Central Valley, Coast Range, Basin and Range, Sierra Nevada, and Transverse ranges, which provide a diverse collection of landforms and topography. Additionally, they cover regions from densely populated, highly developed areas to mostly rural undeveloped counties that also presented variation in the fuzzy sets included in the *Fuzzy Layer* used in the Secondary Model. For each study area, the county boundary was used to extract the data and to define the study area limits. Table 7 presents a summary of the results in different case study areas.

Table 8 – Summary of Primary Model Results for all Study Areas

Study Area	**Elevation Range (Meters)**	**County Area (Square Kilometers)**	**Potential Upper Reservoir Locations**	**Potential Lower Reservoir Locations**	**Number of Connections**	**Upper/ Lower Reservoir Ratio**
Los Angeles County	0 – 3,064	12,310	17,216	40,868	2,207,844	0.42
Santa Clara County	0 - 1,336	3,377	0	0	0	-
Yolo County	0 - 955	2,652	1,388	1,060	75,196	1.31
Mono County	1,273 – 4,342	8,110	County datasets too large for computational resources; over 7 million connections			
Butte County	14 – 2,175	4,340	187,537	52,564	4,809,203	3.56

The results demonstrate that the model produces a large number of locations for both upper and lower reservoir locations in each study area. A closer examination of the results relative to the geomorphological terrain of each study area yields an insight into the suitability of specific settings to MPSHS deployment. Additionally, these results show that the total change in elevation within a study area is not alone, an indicator of MPSHS suitability potential.

In Santa Clara County, there were no locations that were found to be suitable for the construction of the MPSHS. This was likely due to the low countywide relief and the components of the *Binary Screening Layer*, which occluded a significant portion of the higher elevations from considerations.

Results for Yolo County found areas suitable for MPSHS on the eastern margin of the Coast Ranges. Yolo County occupies an area with moderate relief in the hills and canyons on its western flank and flat alluvial valleys and delta lands of the Sacramento Valley. Topography in the western portion is characterized by moderately steep slopes with peaks and ridgelines lining leaner fault-controlled valleys. Model results indicate that the terrain is suitable for MPSHS placement and there were around 75,000 suitable connections.

Mono County is situated on the east side of California at the margin between the Sierra Nevada and the Basin and Range Province. The Sierra Nevada mountains are a westward dipping fault block characterizes by high relief on its eastern boundary. Due to the massive potential for relief meeting the criteria of the model, the intermediate datasets proved too large for the computational resources at the scale chosen, with the near analysis producing over 7 ½ million matches. For an area of this type, it would be beneficial to refine further the *Binary Screening Layer* to minimize the areas considered prior to using this model. Due to the limitation in time, further efforts to complete the analysis in this study area were abandoned.

Butte County offers unique terrain when compared to the other test areas selected. Located at the northern limits of the Sierra Nevada, Butte County is often referred to as the tablelands. This is due to its large flat westward dipping buttes comprised of volcanic mudflows. Rivers and streams have incised through the mudflow to create a topography of plateaus and river canyons, which flow westward toward the central valley. This unique topographic profile provided a difference in outcome for the model results. The results showed a high number of upper reservoir locations could be placed on the tops of the laterally extensive plateaus while limiting the number of lower reservoir locations due to the steeply sloped canyons cross-cutting the terrain.

Overall the Primary Model provides results that are satisfactory to the goals set forth. The model outputs identify areas likely suitable for placement of both upper and lower reservoir locations that meet the engineering design specifications. Binary raster screening and restricted line screening processes effectively eliminate from contention MPSHS configurations that intersect features identified as not being suitable for MPSHS placement. The Secondary Model applies the attributes of the user-provided *Fuzzy Layer* to the areas identified as suitable for placement of a terminal reservoir, adding further value to the model results.

Chapter 6 Conclusions

This chapter provides an assessment of the success of the model developed during this study, including the overall model outcomes, potential uses, implications of the final product, and viability in the marketplace. This section also discusses areas for future development and other applications where the concepts developed herein could be applied.

Overall the Primary Model was found to succeed in identifying areas likely suitable for the construction of the MPSHS. Areas selected as suitable were found to meet the design specifications in both having an areal footprint which meets the slope limitations supporting construction feasibility of terminal reservoirs and meets or exceeds the relief requirements from the upper reservoir to the lower reservoir.

6.1. Model Assessment

Examination of intermediate datasets provides an explicit acknowledgment that within each study area, there are many locations which are suitable for placement of a reservoir, without meeting the additional engineering criteria. Using the design footprint method of screening proved to be the most significant topographic indicator for placement with respect to construction feasibility. While other topographic variables such as topographic position index (TPI) and terrain roughness were initially considered, they were found to be too limiting and difficult to apply over a diverse range of geomorphic conditions without calibration.

The moving window analysis is used for the construction footprint exploration (see Section 4.2.1) and both relief analyses (see Section 4.2.4). For the construction footprint analysis, this technique is able to identify all locations within the study area that are “flat” enough to be considered suitable over the entire design footprint. Review of intermediate datasets that identify areas meeting the construction footprint criteria indicates, as expected that most low-lying areas within the study area are identified as suitable. This is contrasted by results from upland areas, where suitability is typically far less common. Additionally, by making the search radius for the moving window analysis a model parameter, the end-user can modify the required footprint for construction as well as the slope angle limits, adding situational flexibility to the application. Applying basic concepts of geomorphology to modeled results, this method was shown to be capable of producing acceptable results under limited review. With little effort, the model could be modified to support the search for upper and lower reservoir locations where the area required for construction was not identical.

Application of the moving window analysis for the relief calculations worked to identify all grid cells in the *subset DEM* that met or exceeded the minimum design relief. By using the method to identify both upslope relief and downslope relief, the model was able to identify both the upper reservoir locations and the lower reservoir locations, independently. After identification of the upper and lower reservoir sites, connections could be made linking upper reservoir locations to lower reservoir locations.

Early in the model development process, it was apparent that the connections had to be filtered to provide more viable results. The connection culling process began by identifying the apparent flaws of connection identification by examining model results with respect to mapped features such as connections that crossed streams due to meandering channels and topographic features caused by fluvial morphology. Another glaring problem, as discussed in Section 4.2.7.2, was false match connections which were found to be a byproduct of the connection process linking upper reservoir locations to lower reservoir locations, without regard for point to point relief, to which a solution was created and integrated into the Primary Model.

A fundamental component of the hydroelectric generating potential is the motive force produced by the difference in head pressures between the upper and lower reservoirs. As this is an essential consideration for the suitability of one reservoir pair over another, connections linking reservoir pairs contain attribute data indicating the relief by connection allowing the user to filter and rank the connections by relief. In conjunction with connection length, this information can be further summarized by gradient or the ratio of the change in head over the connection distance. While these are important considerations for the selection of MPSHS placement, they are not incorporated into the model developed in this study as there are currently no engineering design criteria for these variables.

While auditing model outputs for observable breaches of the conceptual model, it became apparent water bodies such as lakes and oceans were identified as suitable for reservoir placement. This selection was due to the nature of water bodies to appear as large flat areas. It should be noted that in the case study discussed in Chapter 5, water bodies were included in the binary screening dataset, not as a standard model parameter. While this condition can be seen as a deviation from the conceptual model, case studies in the literature have shown that some viable projects can consider natural water bodies as lower reservoir options. Recognizing this reality, the inclusion of water bodies in the binary screening layer remains a recommendation rather than a standard component (Bueno and Carta 2006, 312-340).

Incorporation of the binary screening option into the model workflow was added as a method of reducing the area to be considered for placement of the system before the more resource-heavy computations were performed. This was found to be an essential step when the model moved from smaller study areas where terrain capable of supporting the technology was limited to a small fraction of the overall study area like Los Angeles County and Yolo County to larger, more mountainous areas like those in Mono County, located on the western boundary of the Basin and Range geomorphic province.

Furthermore, when considering data to be included in the binary screening dataset, potential areas to be excluded from the overall analysis were found to be highly variable and regionally specific. For example, in Los Angeles County, the user may wish to exclude developed areas due to the high population density while in Inyo County, lands belonging to Death Valley National Park may be essential to include in the binary layer. The same is true for the restricted line crossings, which will change drastically based on location and experience.

In Santa Clara County, one of the study areas considered as an example; there were no potentially viable solutions to the analysis. This was likely due to a large amount of terrain likely suitable being screened by the Binary Screening component and the low countywide relief.

Results for other study areas varied with respect to the ratio of upper reservoir location to lower reservoir locations. Some underlying correlations can be made between geomorphic expression, erosional patterns in geologically unique terrains, and total study area relief.

A clear example of terrain effects on the Primary Model outcome is in the comparison between the results from Los Angeles County and those of Butte County. Los Angeles County is an arid climate where the topography is dominated by the San Gabriel Mountains, a faulted and steeply sloped terrain with little soil development creating jagged peaks and ridgelines abutting nearly flat alluvial valleys. In contrast, the Sierra Nevada foothills of Butte County are capped by volcanic mudflows called the tablelands, expressed as flat-topped, shallowly westward dipping features cut by deeply incised east-west trending canyons. The topography of Butte County provides ample room for upper reservoirs while limiting potential lower reservoir site. Los Angeles County’s Upper Reservoir to Lower Reservoir ratio is 0.42, while Butte County's is

3.56. This implies that the terrain in Butte County is more suitable for upper reservoirs, while the opposite is true for Los Angeles County.

Due to the nature of the MPSHS, which relies on gravity and hydraulic head pressures to function, it is intuitive that the topographic character of a study area would have a dramatic influence on the model outcomes. While a surface-level examination of the data implies that an empirically quantifiable correlation likely exists between landform development and suitability for MPSHS, a thorough examination was not included in this study.

The Secondary Model provided integration of a fuzzy dataset to determine further suitability for the final Primary Model results. This provides the end-user an opportunity to further refine the suitability analysis by applying additional variables to the locations identified by the Primary Model. Similar to the binary screening dataset, the fuzzy dataset would likely incorporate a diverse and regionally-specific collection of variables. In the example provided above, four datasets with spatially variable degrees of influence are considered and are shown to have a diverse impact on the suitability of output features. It is important to note that the data used to develop the fuzzy dataset used in the case study was not weighted with respect to the relative importance of each input dataset. As a result, although accounted for in the creation of each fussy set, each fuzzy input had an equal degree of influence over the final ranking value.

The membership functions and the joint membership function of the fuzzy analysis were found to have an enormous degree of influence on the suitability analysis when incorporated into the project workflow. However, the application of defensible criteria selection in the fuzzy datasets was outside of the scope of this study.

6.2. Uses

The model developed for this study identifies potential locations for the development of MPSHS. Based on model parameters, reservoirs must be centered on an area 90 meters by 90 meters where the maximum slope cannot exceed the design specification. The upper reservoir must be located in such a position that the relief to a location suitable for the placement of a lower reservoir cannot be less than the design parameter of 300 meters. Additionally, the lateral distance between the upper and lower reservoirs cannot exceed 1,500 meters. In order to make the model as useful as possible with respect to the project-specific engineering constraints, the search area parameters for the construction footprint, the lateral relief search radius and the minimum relief requirements are modifiable by the model user.

This model produces output products to assist in solving the design problem of MPSHS placement. Components of this model may have application in other areas of engineering and suitability analysis. The moving window analyses used herein demonstrate the ability to characterize large datasets into easily digestible outputs designed to answer a specific problem based on terrain features. For example, relief, used herein as an average gradient between upslope and downslope points, may have applications in other material transport problems such as large conveyor design or fluid transport problems.

6.2.1. Viability in Market

The concept of MPSHS is an emerging market, and the scenario explored herein has yet to be put into action. While the current market for this model may be limited, the potential for this suitability modeling to be utilized as a component of a broader push toward energy sustainability is promising. Coupled with broader political and environmental motivation, this model could be used to support further consideration MPSHS as an alternative or complement to other resource types.

6.3. Future Work

This model focused on the initial steps of suitability analysis for MPSHS. The outputs are limited in scope to only locations where there is a potential for deployment of the technology. Future development in this area can work toward further refinement of the output products and refinement in suitability analysis.

While the model develops the connections between upper and lower reservoir locations, with the exception of the consideration of restricted lines, the character of the terrain between those connections is not considered. Project viability may be significantly impacted by the terrain roughness or topographic barriers between paired reservoirs. For example, if there is a ridgeline between terminal reservoirs, it may be impractical to construct the system at that location due to construction costs required to overcome that type of topographic features. The incorporation of further terrain analysis techniques could be applied to overcome these problems and further refine the model products to account for topographic barriers to placement not accounted for in the model developed to date.

Additional work in terrain analysis could expand on the suitability criteria, specifically, with respect to connections and suitability for placement. One significant variable in the viability for construction of MPSHS is the ability to install all of the system components as cost-effectively as possible. An addition to the model could work to characterize the terrain underlying the connection for suitability.

Additionally, the model could be further enhanced by a process that evaluates the suitability of the connections with respect to their path. One option may be to perform a buffer analysis for connections and examine the effects of external variables spanning the distance between terminal reservoirs. Careful consideration will be needed, as the factors affecting the connections may be drastically different from those considered for reservoir placement.

Reservoir locations presented as outputs of this model are provided as raster values in the Primary Model and then as points once the fuzzy dataset is applied. While these both give an idea of the location identified as suitable, further work could be undertaken to aggregate these outputs into areas that can be summarized and graded based on a larger spatial footprint to provide more tangible results to model end-users.

The model developed in this study has presented a reasonable solution to the design problem presented. At the county scale, freely available elevation data products with coverage across the contiguous United States allow this model to identify suitable locations for the placement MPSHS adequately and with flexibility for variations in engineering and design criteria. Parameters built into the model allow adjustments for system design modifications and incorporation of third-party screening and suitability layers for further refinement of the model results.

References

Aydin, Nazli Yonca, Elcin Kentel, and Sebnem Duzgun. 2010. "GIS-based environmental assessment of wind energy systems for spatial planning: A case study from Western Turkey."*Renewable and Sustainable Energy Reviews* 14(1): 364-373.

Aydin, Nazli Yonca, Elcin Kentel, and H. Sebnem Duzgun. 2013. "GIS-based site selection methodology for hybrid renewable energy systems: A case study from western Turkey." *Energy Conversion and Management* 70: 90-106.

Benitez, Pablo C., Liliana E. Benitez, and G. Cornelis van Kooten. 2008. "The Economics of Wind Power with Energy Storage." *Energy Economics* 30(4): 1973-1989.

Bolstad, Paul. 2005. *GIS Fundamentals: A First Text on Geographic Information Systems*. 2nd ed. White Bear Lake, Minnesota: Eider Press.

Bueno, C., and Jose A. Carta. 2006. "Wind powered pumped hydro storage systems, a means of increasing the penetration of renewable energy in the Canary Islands." *Renewable and Sustainable Energy Reviews* 10(4): 312-340.

California Independent System Operator. 2018. "What the Duck Curve Tells Us About Managing a Green Grid." Accessed 11/22/18, https://www.caiso.com/Documents/FlexibleResourcesHelpRenewables_FastFacts.pdf.

Carver, Stephen J. 1991. "Integrating multi-criteria evaluation with geographical information systems." *International Journal of Geographical Information System* 5(3): 321-339.

Chu, Steven, and Arun Majumdar. 2012. "Opportunities and challenges for a sustainable energy future." *Nature* 488(7411): 294.

ClimateTechWiki.org. 2019. "Energy Storage: Pumped Storage." Accessed 4/4/19, https://www.climatetechwiki.org/technology/jiqweb-ph.

Connolly, David, Shane MacLaughlin, and Martin Leahy. 2009. "Locating Potential Sites for Pumped Hydroelectric Energy Storage." In *Proc. of the Electrical Energy Storage Applications and Technologies (EESAT) Conference*, Seattle, USA, pp. 4-7.

California Senate. 2018. *SB-100 California Renewables Portfolio Standard Program: emissions of greenhouse gases* (September 10, 2018). Sacramento, CA.

Denholm, Paul, Erik Ela, Brendan Kirby, and Michael Milligan. 2010. *Role of Energy Storage with Renewable Electricity Generation (Report Summary)(Presentation)*. Report NREL/PR-6A2-49396. National Renewable Energy Lab, Golden, CO.

Esri. 2019a. "How Fuzzy Membership Works." Accessed 7/14/2019, http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-fuzzy-membership-works.htm.

Esri. 2019b. "How Fuzzy Overlay Works." Accessed 6/16/19, http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-fuzzy-overlay-works.htm.

Esri. 2019c. "Overlay Analysis." Accessed 7/14/2019, http://desktop.arcgis.com/en/arcmap/10.3/analyze/commonly-used-tools/overlay-analysis.htm.

Esri. 2019d. "Tool Help: Slope." Accessed 6/9/2019, http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/slope.htm.

Esri. 2019e. "Understanding Euclidean Distance Analysis." Accessed 7/14/2019, http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/understanding-euclidean-distance-analysis.htm.

Esri. 2019f. "Understanding Reclassification." Accessed 7/14/2019, https://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/understanding-reclassification.htm.

Esri. 2019g. "What is ModelBuilder?" Accessed 6/15/19, http://desktop.arcgis.com/en/arcmap/10.3/analyze/modelbuilder/what-is-modelbuilder.htm.

Fitzgerald, Niall, Roberto Lacal Arántegui, Eamon McKeogh, and Paul Leahy. 2012. "A GIS-based model to calculate the potential for transforming conventional hydropower schemes and non-hydro reservoirs to pumped hydropower schemes." *Energy* 41(1): 483-490.

Friends of the River. 2018. "Rivers Under Threat." Accessed 4/1/2018, https://www.friendsoftheriver.org/our-work/rivers-under-threat/.

Gesch, D.B., Oimoen, M.J., and Evans, G.A., 2014, *Accuracy assessment of the U.S. Geological Survey National Elevation Dataset, and comparison with other large-area elevation datasets—SRTM and ASTER*. U.S. Geological Survey Open-File Report 2014–1008.

Heirman, D., 2012. "What makes Smart Grid—Smart—And who is in the "game"?" *IEEE Electromagnetic Compatibility Magazine* 1(2): 95-99.

Hansen, Henning Sten. 2005. "GIS-based multi-criteria analysis of wind farm development." In *Proceedings of the 10th Scandinavian Research Conference on Geographical Information Sciences*, *13th-15th June 2005*, pp. 75-87.

Kaplan, Stan M. 2009. "Electric power transmission: background and policy issues." Library of Congress, Congressional Research Service.

Madani, Kaveh, Marion Guégan, and Cintia B. Uvo. 2014. "Climate change impacts on high-elevation hydroelectricity in California." *Journal of Hydrology* 510: 153-163.

Malczewski, Jacek. 2000. "On the use of weighted linear combination method in GIS: common and best practice approaches." *Transactions in GIS* 4(1): 5-22.

Malczewski, Jacek. 2004. "GIS-based land-use suitability analysis: a critical overview." *Progress in planning* 62(1) : 3-65.

Mitchell, Andy. 2012. *The ESRI Guide to GIS analysis: modeling suitability, movement, and interaction*. Vol. 3. Redlands, CA: Esri Press.

Montello, Daniel, and Paul Sutton. 2006. *An introduction to scientific research methods in geography*. SAGE Publications, Inc.

Napier, Ted L., Michael V. Carter, and Elizabeth G. Bryant. 1986. "Local perceptions of reservoir impacts: A test of vested interests." *American Journal of Community Psychology* 14(1): 17-37.

Norris, Robert Matheson, and Robert W. Webb. 1990. *Geology of California*. 2nd ed. Hoboken, NY: John Wiley & Sons.

O'Sullivan, David, and George LW Perry. 2013. *Spatial simulation: exploring pattern and process*. Hoboken, NY: John Wiley & Sons.

O’Sullivan, David, and David J. Unwin. 2010. *Geographic Information Analysis,* 2nd ed. Hoboken, NY: John Wiley & Sons.

Pasternack, Gregory B., C. Lau Wang, and Joseph E. Merz. 2004. "Application of a 2D hydrodynamic model to design of reach-scale spawning gravel replenishment on the Mokelumne River, California." *River Research and Applications* 20(2): 205-225.

Rehman, Shafiqur, Luai M. Al-Hadhrami, and Md Mahbub Alam. 2015. "Pumped hydro energy storage system: A technological review." *Renewable and Sustainable Energy Reviews* 44: 586-598.

Rosenberg, Ronald H. 2008. "Diversifying America's energy future: The future of renewable wind power." *Virginia Environmental Law Journal* 26: 505.

Sparkes, Andrew, and David Kidner. 1996. "A GIS for the environmental impact assessment of wind farms." In *ESRI European User Conference, London, 1996,* pp. 2-4.

Suul, Jon Are, Kjetil Uhlen, and Tore Undeland. 2008. "Variable speed pumped storage hydropower for integration of wind energy in isolated grids: case description and control strategies." In *Nordic Workshop on Power and Industrial Electronics (NORPIE/2008),* June 9-11, 2008, Espoo, Finland. Helsinki University of Technology.

The National Geographic Society. 2011. "Reservoir." Accessed 3/24/2019, https://www.nationalgeographic.org/encyclopedia/reservoir/.

United States Energy Information Administration. 2019. "Total Energy Consumption of Renewable and Nonrenewable Energy Sources." Accessed 7/14/2019, https://www.eia.gov/renewable/data.php.

USGS. 2018. "Hydroelectric Power Water Use." Accessed 3/24/2019, https://water.usgs.gov/edu/wuhy.html.

USGS. 2019. "About 3DEP Products and Services." Accessed 3/5/2019, https://www.usgs.gov/core-science-systems/ngp/3dep/about-3dep-products-services.

Uyan, Mevlut. 2013. "GIS-based solar farms site selection using analytic hierarchy process (AHP) in Karapinar region, Konya/Turkey." *Renewable and Sustainable Energy Reviews* 28: 11-17.

Yang, Chi-Jen, and Robert B. Jackson. 2011. "Opportunities and barriers to pumped-hydro energy storage in the United States." *Renewable and Sustainable Energy Reviews* 15(1): 839-844.

Appendix A – Primary Model Process Table

Process ID	Process Name	Input Data Name	Input Data Type	Tool	Major Parameter(s)	Output Data Name	Output Data Type	Purpose
1	Slope	*Study Area DEM*	30m Elevation Raster	Slope	Slope in degrees	*Study Area Slope Raster*	30m Slope Raster	Creates a Slope raster for use in identifying suitable locations for construction of terminal reservoirs
2	Construction Area Focal Analysis	*Study Area Slope Raster*	30m Slope Raster	Focal Statistics	Search Distance: 1 Cell; Search Pattern: Rectangle	*Max. Neighborhood Slope Raster*	30m Raster	Examines each of the eight neighbor cells for each cell and assigns the parent cell value equal to the highest cell in the neighborhood
3	Binary Reclassify Construction Areas	*Max. Neighborhood Slope Raster & Reclass Expression*	30m Raster	Reclassify	Reclass Expression: If the slope is greater than 15 then 0; else 1	*Max Slope Binary Raster*	30m Raster	Reclassifies each cell into binary code to eliminate areas that are not suitable for construction based on slope.
4	Apply Binary Screen	*Max Slope Binary Raster & Binary Screening Raster*	30m Rasters	Times	Multiply cells	*Construction Area Raster*	30m Raster	Creates a raster dataset that additionally eliminates the areas identified by the binary screening raster from consideration in terminal reservoir placement
5	Convert Raster to Polygon	*Construction Area Raster*	30m Raster	Raster to Polygon	Based on the value field	*Construction Area Polygons*	Vector Polygons	creates a vector polygon dataset that identifies all of the areas deemed suitable for the construction of a terminal reservoir independent of relief (1) and areas that are not suitable (0).

Process ID	Process Name	Input Data Name	Input Data Type	Tool	Major Parameter(s)	Output Data Name	Output Data Type	Purpose
6	Select Suitable Areas for Construction	*Construction Area Polygons*	Vector Polygons	Select	Based on the value field	*Suitable Areas for Construction*	Vector Polygons	Creates a vector dataset of polygons representing ONLY those areas suitable for placement of a terminal reservoir
7	Extract Suitable Areas from Study Area DEM	*Suitable Areas for Construction (Mask) & Study Area DEM (Target Layer)*	30m Raster & Vector Polygons	Extract by Mask	Extract from Study Area DEM by Suitable Areas for Construction	*Study Areas DEM Subset*	30m Raster	Extracts the portions of the DEM that are coincident with
8	Export Suitable Areas Polygons	*Suitable Areas for Construction*	Vector Polygons	Export	None	*Final Suitable Areas*	Vector Polygons	Exports intermediate dataset
9	Search for Maximum Elevation	*Study Areas DEM Subset*	30m Raster	Focal Statistics	Search Distance: 50 Cells (1500m); Search Pattern: circular	*Maximum Elevation in Focal Search Raster*	30m Raster	Searches a specified radius for the maximum elevation and assigns the target cell that value
10	Search for Minimum Elevation	*Study Areas DEM Subset*	30m Raster	Focal Statistics	Search Distance: 50 Cells (1500m); Search Pattern: circular	*Minimum Elevation in Focal Search Raster*	30m Raster	Searches a specified radius for the minimum elevation and assigns the target cell that value

Process ID	**Process Name**	**Input Data Name**	**Input Data Type**	**Tool**	**Major Parameter(s)**	**Output Data Name**	**Output Data Type**	**Purpose**
11	Calculate Maximum Uphill Relief	*Maximum Elevation in Focal Search Raster & Study Areas DEM Subset*	30m Raster	Minus	(Maximum Elevation in Focal Search Raster) - (Study Areas DEM)	*Maximum Uphill Relief*	30m Raster	Subtracts the maximum elevation within the search area by the true elevation at each location.
12	Calculate Maximum Downhill Relief	*Minimum Elevation in Focal Search Raster & Study Areas DEM Subset*	30m Raster	Minus	(Study Areas DEM Subset) - (Minimum Elevation in Focal Search Raster)	*Maximum Downhill Relief*	30m Raster	subtracts the true elevation at each location by the minimum elevation in the focal search area.
13	Reclassify Uphill Relief	*Maximum Uphill Relief*	30m Raster	Reclassify	Reclass Expression: If relief is greater than 300 then 1; else NODATA	*Lower Locations Raster*	30m Raster	Creates a rater dataset comprised only of cells that meet the design relief requirement in the uphill direction.
14	Reclassify Downhill Relief	*Maximum Downhill Relief*	30m Raster	Reclassify	Reclass Expression: If relief is greater than 300 then 1; else NODATA	*Upper Locations Raster*	30m Raster	Creates a rater dataset comprised only of cells that meet the design relief requirement in the downhill direction.
15	Convert Lower Locations to Points	*Lower Locations Raster*	30m Raster	Raster to point	None	*Lower Location Points*	Vector Points	Creates a point vector dataset from the raster dataset

Process ID	Process Name	Input Data Name	Input Data Type	Tool	Major Parameter(s)	Output Data Name	Output Data Type	Purpose
16	Convert Upper Locations to Points	*Upper Locations Raster*	30m Raster	Raster to point	None	*Upper Location Points*	Vector Points	Creates a point vector dataset from the raster dataset
17	Search for Lower Reservoir points near upper reservoir points	*Upper Location Points & Lower Location Points*	Vector Points	Near	Search distance: 1500m	*Start and End Points Table*	Table	Generates a one to many tables where each upper reservoir location is paired with every lower reservoir point within the prescribed search radius
18	Connect Upper Reservoir Points to Lower Reservoir Points	*Start and End Points Table*	Table	XY to Line	Start Point (X,Y) and end Point (X,Y)	*Connection Lines*	Vector Lines	Creates a vector line dataset of lines connecting the upper reservoir to lower reservoir locations
19	Select Connections that Do Not Cross Named Streams	*Connection Lines & Restricted Lines*	Vector Lines	Select	Select the invert of lines that intersect Restricted Lines	*Selection of Connection Lines that Do not Cross*	Vector Lines	Selects the inverse of features that cross the restricted lines
20	Copy Selected Features	*Selection of Connection Lines that Do not Cross*	Vector Lines	Copy Features	Copy Selected	*Connections that Do not cross Restricted Lines*	Vector Lines	Creates a new vector dataset comprised of features that DO NOT cross the Restricted Lines Dataset

Process ID	**Process Name**	**Input Data Name**	**Input Data Type**	**Tool**	**Major Parameter(s)**	**Output Data Name**	**Output Data Type**	**Purpose**
21	Assign Elevations to Lower Points	*Study Areas DEM Subset & Lower Location Points*	Vector Points	Extract Values to Points	None	*Lower Reservoir Points with elevations*	Vector Points	adds the attribute of elevation from the DEM subset to each point selected as suitable for construction of a lower reservoir
22	Assign Elevations to Upper Points	*Study Areas DEM Subset & Upper Location Points*	Vector Points	Extract Values to Points	None	*Upper Reservoir Points with elevations*	Vector Points	adds the attribute of elevation from the DEM subset to each point selected as suitable for construction of an upper reservoir
23	Spatial Join Upper Locations to Connections	*Connections that Do not Cross Streams & Upper Location Points*	Vector Lines & points	Spatial Join	Join one to many	*Connections with upper elevations*	Vector Lines	Adds the attribute of upper reservoir elevation to each connection line
24	Spatial Join Lower Locations to Connections	*Connections with upper elevations*	Vector Lines & points	Spatial Join	Join one to many	*Connections with upper and lower elevations*	Vector Lines	Adds the attribute of lower reservoir elevation to each connection line
25	Add a field for Relief	*Connections with upper and lower elevations*	Vector Lines	Add Field	Field Type : Float	*Connections with Relief Field*	Vector Lines	Adds a field for relied on the connection dataset
26	Calculate Relief	*Connections with Relief Field*	Vector Lines	Calculate Field	(Upper Location Elevation) - (Lower Location Elevation)	*Connections with Relief*	Vector Lines	Calculates the relief value for each connection

Process ID	**Process Name**	**Input Data Name**	**Input Data Type**	**Tool**	**Major Parameter(s)**	**Output Data Name**	**Output Data Type**	**Purpose**
27	Select Connections that meet the Relief Requirement	*Connections with Relief*	Vector Lines	Select by Attribute	Select Connections where Relief exceeds 300m	*Final Suitable Connections*	Vector Lines	creates a new dataset with only the connections that meet the relief requirement
28	Export Connections	*Final Suitable Connections*	Vector Lines	Export	None	*Final Reservoir Connections*	Vector Lines	Exports Final Modeled Connections
29	Select upper points that DO NOT have a connection	*Final Suitable Connections & Lower Reservoir Points with elevations*	Vector Points	Select	Invert Select points that intersect connections	*Selected lower points without connections*	Vector Points	Selects all points in the lower points dataset that do not intersect with a connection
30	Select lower points that DO NOT have a connection	*Final Suitable Connections & Upper Reservoir Points with elevations*	Vector Points	Select	Invert Select points that intersect connections	*Selected Upper Points Without Connections*	Vector Points	Selects all points in the upper points dataset that do not intersect with a connection
31	Delete lower points without connections	*Selected lower points without connections*	Vector Points	Delete	Delete Selected	*Final Lower reservoir points*	Vector Points	Deletes points from the dataset that do not interest a connection line
32	Delete Upper points without connections	*Selected Upper Points Without Connections*	Vector Points	Delete	Delete Selected	*Final Upper reservoir points*	Vector Points	Deletes points from the dataset that do not interest a connection line

Process ID	Process Name	Input Data Name	Input Data Type	Tool	Major Parameter(s)	Output Data Name	Output Data Type	Purpose
33	Convert Lower Reservoir Points to Raster	*Final Lower reservoir points*	Vector Points	Point to Raster	None	*Final Lower Reservoir Location Raster*	30m Raster	Converts the final modeled lower reservoir locations to a raster dataset at the Study Area DEM Resolution
34	Convert Upper Reservoir Points to Raster	*Final Upper reservoir points*	Vector Points	Point to Raster	None	*Final Upper Reservoir Location Raster*	30m Raster	Converts the final modeled Upper reservoir locations to a raster dataset at the Study Area DEM Resolution