INVESTIGATION OF DIGITAL ELEVATION MODEL UNCERTAINTY IN GIS-BASED SOLAR RADIATION MODELS USING MARKOV CHAIN MONTE CARLO SIMULATION

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ABDUR-RAHMAN BELEL ISMAILA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
GEODETIC AND GEOGRAPHIC INFORMATION TECHNOLOGIES

SEPTEMBER 2013

Approval of the thesis:

INVESTIGATION OF DIGITAL ELEVATION MODEL UNCERTAINTY IN GIS-BASED SOLAR RADIATION MODELS USING MARKOV CHAIN MONTE CARLO SIMULATION

submitted by **ABDUR-RAHMAN BELEL ISMAILA** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Geodetic and Geographic Information Technologies Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen _____________________
Dean, Graduate School of **Natural and Applied Sciences**

Assoc. Prof. Dr. Ahmet Coşar _____________________
Head of Dept., **Geodetic & Geographic Infor. Technologies**

Prof. Dr. Şebnem H. Düzgün _____________________
Supervisor, **Mining Engineering Dept., METU**

Prof. Dr. Volkan Ş. Ediger _____________________
Co-supervisor, **Energy Systems Eng. Dept., Kadir Has Univ.**

Examining Committee Members:

Assoc. Prof. Dr. Ayşegül Aksoy _____________________
Environmental Engineering Dept., METU

Prof. Dr. Şebnem H. Düzgün _____________________
Mining Engineering Dept., METU

Prof. Dr. Oğuz Işik _____________________
City and Regional Planning Dept., METU

Prof. Dr. Gerhard-Wilhelm Weber _____________________
Institute of Applied Mathematics, METU

Asst. Prof. Dr. Serkan Kemeç _____________________
City and Regional Planning Dept., Yüzüncü Yıl Univ.

Date: 17.09.2013

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Abdur-Rahman Belel, Ismaila

Signature :

ABSTRACT

INVESTIGATION OF DIGITAL ELEVATION MODEL UNCERTAINTY IN GIS-BASED SOLAR RADIATION MODELS USING MARKOV CHAIN MONTE CARLO SIMULATION

Ismaila, Abdur-Rahman Belel
Ph.D., Department of Geodetic and Geographic Information Technologies
Supervisor: Prof. Dr. Şebnem H. Düzgün
Co-supervisor: Prof. Dr. Volkan Ş. Ediger

September 2013, 198 pages

In this study, a *Markov chain Monte Carlo (MCMC)* simulation approach that incorporates digital elevation model's (DEM) spatial autocorrelation is developed with the aim of assessing the impact of DEM uncertainty in *GIS-based solar radiation models*.

The method utilized the error probability distribution function (pdf) of Shuttle Radar Topography Mission (SRTM) DEM, and variogram model parameters of the study area's SRTM DEM as *a priori* information for the formulation of *Metropolis-Hasting algorithm*. Statistical analysis of the data extracted from the literature revealed that SRTM DEM error exhibit lognormal distribution with a mean of 4.209 m and standard deviation of 0.054 m. An exponential variogram model with sill of 125.74, range of 1,556.85) and nugget of 0 represent the spatial autocorrelation of the study area DEM. A total of 1,080 simulations is executed using $2m$ chains and the initial *burn-in* period of 80, representing 7.41% of the simulation is discarded. Multivariate potential scale reduction factor (MPSRF) of 0.99 is obtained after executing 1,080 MCMC simulations which indicates that the MCMC sampler has converged to a stationary distribution, being less than 1.1.

Thus, the results are assumed to be drawn from lognormal pdf. Whereas, the check for variogram reproduction based on 95% confidence level indicate that the variogram simulation remains valid since T^2=1.751 is less than the corresponding *F*-statistic of 23.19. The proposed methodology is coded and executed in MaTLAB Environment. Based on the simulation results, it is observed that the proposed framework allows better representation of the DEM data.

The realized DEMs together with other inputs were used to run the *Solar Analyst* and *r.sun* models. The results of both models showed a better performance using the realized DEMs than the original SRTM DEM. For *Solar Analyst*, DEM uncertainty has greater effect on diffuse radiation and direct duration, while, direct and global radiations are less affected. For *r.sun*, the DEM uncertainty has less influence on solar radiation outputs. Comparison of the two models shows that *Solar Analyst* is more sensitive to uncertainty than *r. sun*. Interestingly, the study reveals that relatively flat terrains where DEM uncertainty seems to

be low also exhibit high uncertainty in solar radiation estimates. This indicates that DEM may not be the only input associated with uncertainty.

Keywords: DEM, Uncertainty analysis, GIS-based solar radiation models, MCMC, spatial autocorrelation.

ÖZ

MARKOV ZINCIRI MONTE CARLO BENZETIŞIMI KULLANARAK CBS'YE DAYALI GÜNEŞIŞINIMI MODELLERINDEKI SAYISAL YÜKSEKLIK MODELI BELIRSIZLIKLERININ İNCELENMESI

Ismaila, Abdur-Rahman Belel
Doktora, Jeodezi ve Coğrafi Bilgi Teknolojileri Bölümü
Tez Yöneticisi: Prof. Dr. Şebnem H. Düzgün
Yardımcı Tez Yöneticisi: Prof. Dr. Volkan Ş. Ediger

Eylül 2013, 198 sayfa

Bu çalışmada, *CBS'ye dayalı güneş ışınımı modellerindeki* SYM belirsizliğinin etkisinin saptanması için sayısal yükseklik modelindeki (SYM) mekansal otokorelasyonu göz önüne alan bir Markov zinciri Monte Carlo benzetişimi yaklaşımı geliştirilmiştir.

Bu yöntem, *Metropolis-Hasting algoritmasının* oluşturulmasında, önsel (a priori) bilgi olarak Shuttle Radar Topography Mission (SRTM) SYM haritası hata olasılığıdağılımı fonksiyonunu ve çalışma alanının SRTM SYM'nin variogram model parametrelerini kullanmaktadır. Literatürden elde edilen verilerin istatistiksel analizi, SRTM SYM verisinin ortalaması 4.209 m ve standart sapması 0.054 m olan lognormal dağılıma sahip olduğunu göstermiştir. Eşik değeri "125.74", menzili "1,556.85" ve nugget değeri "0" olan bir eksponansiyel variogram modeli çalışma alanının SYM haritasının mekansal otokorelasyonun betimlemektedir. Toplamda 1.080 adet benzetişim *2m'lik* zincirler ve 80'lik bellek süresi ile çalıştırılmış benzetişimlerin %7,41'i gözardı edilmiştir. Çok değişkenli potansiyel ölçek azaltma faktörü olarak 0.99 değeri, 1.080 adet benzetişim sonrasında elde edilmiştir ki bu Markov zincirinin 1.1 değeri ile duraylı bir dağılıma yaklaştığını göstermektedir.

Bu nedenle elde edilen sonuçların lognormal pdf dağılımdan elde edildiği anlaşılmıştır. Variogram üretiminin %95 güven aralığı kontrol edildiğinde, T^2=1.751 değerinin 23.19 *F*-istatistik değerinden küçük olması nedeni ile variogram benzetişimlerinin geçerli olduğu görülmüştür. Önerilen bu yöntem MaTLAB ortamında kodlanmış ve çalıştırılmıştır. Benzetişim sonuçlarına göre, önerilen çerçeve, daha iyi SYM haritalarını temsiline olanak sağlamaktadır.

Benzetişim ile elde edilen SYM ve diğer girdiler *Solar Analyst* ve *r.sun* modellerinin çalıştırılmasında kullanılmıştır. Her iki modelden de alınan sonuçlar gösterdiki özgün SYM haritası yerine benzetişim ile elde edilen SYM haritasının kullanarak daha iyi performanslar elde edilmiştir. *Solar Analyst* için, doğrudan ve küreselışınım değerlerinin SYM'deki belirsizlikten daha az etkilenediği görülürken, yaygın ışınım ve doğrudan güneşlenme süresinin daha çok etkilendiği belirlenmiştir. *"r.sun"* için, SYM'deki belirsizliklerin güneş ışınımı çıktılarına daha az etkisi olduğu belirlenmiştir. İki modelin karşılaştırılmasında *Solar*

Analyst modelinin SYM'deki belirsizliklere daha duyarlı olduğu görülmüştür. İlginç birşekilde, çalışmada SYM belirsizliklerinin daha düşük olduğu nispeten düz arazilerde güneş ışınımı tahminlerinde yüksek belirsizlikolduğu bulunmuştur. Bu durum SYM belirsizliklerinin tek ilgili parametre olmadığına isaret etmektedir.

Anahtar kelimeler: SYN, Belirsizlik analizi, CBS'ye dayalı güneş ışınımı modelleri, Markov zinciri Monte Carlo, mekansal otokorelasyon

To My Parents

Late Mallam Ismaila Yerima Bello and Hajiya Aishatu Sa'adu Pullo

In love, gratitude, and respect.

ACKNOWLEDGEMENTS

I first thank Allah (SWT) for his infinite mercies, and strength for enabling me to complete my doctoral degree successfully. I pray to Allah (SWT) to bless the education I acquired and make it beneficial not only to myself and my family but to the Muslim Umma and humanity in general, amen.

I would like to express my deep heartfelt thanks to my supervisor, Prof. Dr. Şebnem H. Düzgün for her keen and intelligent guidance throughout the research. Her great personality, patience and thoughtfulness made this study possible. Special thanks go to my co-supervisor Prof. Dr. Volkan Ş. Ediger for his outstanding contributions. I also extend my appreciation to my thesis committee and jury members, Assoc. Prof. Dr. Ayşegül Aksoy, Prof. Dr. Oğuz Işik Prof. Dr. Gerhard-Wilhelm Weber and Asst. Prof. Dr. Serkan Kemeç. Their thoughtful comments and suggestions made a significant improvement of my dissertation.

I wish to express my sincere gratitude to my beloved mother Hajiya Aishatu Sa'adu Pullo for her unconditional love, morale support and prayers. I am also deeply indebted to all my brothers and sisters for their support and prayers. I owe a great deal of appreciation to Mallam Mohammed Ismaila, Barrister Ahmed Ismaila and Mallam Dauda Ismaila for their constant and endless support towards my education. My thanks also go to all my uncles and aunties. Especially, I extend my deep appreciation to Hajiya Rahimatu Yerima Bello and Alhaji Muhammad Sa'adu Pullo for their special care and support.

My deepest gratitude goes to my beloved wife Zainab. Words alone cannot express what I owe you for your encouragement, prayers and patience throughout this thesis. I am indeed very grateful to Almighty Allah (SWT) for blessing me with a humble, obedient and respectful wife of your kind.

Special thanks to Barrister Ahmadu Toungo Zubairu and his wife Hajiya Hafsat for their support and prayers. Special thanks also go to Alhaji Abubakar Sharu Abbas for his constant encouragement, advice and support. I am also indebted to Ozgür Erinmez for his friendship, guidance on MaTLAB and assistance during the coding of my work.

I would like to sincerely acknowledge The Scientific and Technological Research Council of Turkey (TÜBİTAK) for providing a doctoral research fellowship. I am also grateful to the Federal Government of Nigeria and the Republic of Turkey for the opportunity and support given to me to pursue this great achievement.

Thanks are also due to all the academic staff of Geodetic and Geographic Information Technologies (GGIT) department. Special thanks go to Assoc. Prof. Dr. Nurünnisa Usul and Prof. Dr. Mahmut Onur Karslıoğlu for their support. I cordially thank my friends, colleagues, and Research Assistants at GGIT; Fatih Kara, Ertuğrul Boza, Semih Kuter, Dr. Gülcan Sarp, Dr. Reşat Geçen, Dr. Fatih Keskin, Kenan Bolat, Serdar Sürer, and Erkan Kalayci for their valuable support and encouragement during my study at METU. Finally, my deepest thanks are due to all those who contributed towards the successful completion of this research and my studies in general.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

ABREVIATIONS

T_{LK}	Linke Turbidity Factor
AERONET	AErosol RObotic NETwork
ANN	Artificial Neural Network
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
CPV	Concentrated Photovoltaic
CSD	Census Subdivisions
CSP	Concentrated Solar Power
DEM	Digital Elevation Model
DGPS	Differential Global Positioning System
DNI	Direct Normal Irradiance
DTM	Digital Terrain Model
EM	Expectation-Maximization Algorithm
GCP	Ground Control Point
GHI	Global Horizontal Irradiation
GIS	Geographic Information System
GIScience	Geographic Information Science
GMM	Generalized Method of Moments
GPS	Global Positioning System
GRASS	Geographic Resources Analysis Support System
IDRISI	An integrated feature-rich GIS and Image Processing software system for the Analysis and display of spatial data
IPCC	Intergovernmental Panel on Climate Change
$K\downarrow$	Solar Radiation
Kc	Clear Sky Index
KDMD	Klein's definition of mean day
kWh/m^2	Kilowatt-Hours Per Square Meter
m/s	Meters Per Second
MAE	Mean Absolute Error
MC	Monte Carlo
MCMC	Markov chain Monte Carlo
ME	Mean Error
MH	Metropolis-Hastings
MPSRF	multivariate potential scale reduction factor
MW	Megawatt
NASA	National Aeronautics and Space Administration
PDF	Probability Distribution Function
PV	Photovoltaic
PVGSIS	Photovoltaic Geographic Information System
rMBE	Relative Mean Bias Error
RMSE	Root Mean Squared Error
rRMSE	Relative Root Mean Squared Error
RS	Remote Sensing
SAR	Synthetic Aperture Radar

SDTS	Spatial Data Transfer Standard
SRTM	Shuttle Radar Topography Mission
TIN	Triangulated Irregular Network
UA	Uncertainty Analysis
UCGIS	University Consortium for Geographic Information Science
USGS	United States Geological Survey
$W.m^{-2}$	Watt Per Meter Square
Wh/m^2	Watt Hours Per Square Meter

CHAPTER 1

INTRODUCTION

1.1 Background and Motivation

Geographic Information System (GIS)-based solar radiation models (Alsamamra et al. 2009) are developed in order to predict both short- and long-wave solar energy fluxes at and near the Earth’s surface (Sheng et al. 2009) and gives the outputs in the form of solar resource or potential maps. Generally, solar radiation is a major source of energy that supports most of the physical and biochemical processes on the Earth's surface (Li, 2008). It is also a key resource in supporting renewable energy policies on concentrated solar power (CSP) and concentrated photovoltaic (CPV), and above all, an indispensable input for several models used in the field of earth and environmental sciences namely; crop models (Stockle et al. 2003), ecological models (Haxeltine and Prentice, 1996), global circulation models (IPCC, 1996), biological process models (Huang and Fu, 2009), soil-vegetation-atmosphere transfer models (Hansen, 1999; Mariscal et al. 2000), soil temperature models (Lei et al. 2010), and crop growth models (Hodges et al. 1987; Jones et al. 2003). In essence, solar radiation maps derived using GIS-based models and other methods are vital to many fields of human endeavors ranging from renewable energy applications, earth and environmental sciences, city planning, architecture, agriculture, forestry, horticulture, natural resource management, engineering among others.

Despite the significance of solar radiation resource data to many facets of human life, this vital resource is not readily available (Huang and Fu, 2009; Ruiz-Arias et al. 2010); and where it exists, there are uncertainties associated with it. Those derived using GIS models are not an exception, therefore, seriously hinders decision-making process in terms of policy formulation, solar energy applications and investments, and scientific studies that may require solar radiation as a key input.

Traditionally, long term field observations of some meteorological parameters such as sunshine, temperature, and relative humidity, are used in assessing the solar radiation that falls in a specific location on the Earth's surface. However, Her (2011) noted that field monitoring is expensive to implement and requires a long time-series data to provide statistically useful information. As such, *GIS-based solar radiation models* are a necessary complement to field monitoring and in some cases a better option because it is inexpensive, fast, and allow analysis of what-if scenarios, which is not practical in field monitoring. Moreover, a model is reusable in other relevant cases and various functions of a model can be designed with regard to the modeler’s needs. In recent years, the significant progress in the field of GIS has led to the development of various models for estimation of solar radiation. These include, (1) Solar radiation flux (*SolarFlux)* model by Hetrick et al. (1993), which was further modified by Dubayah and Rich (1995, 1996), (2) *Solei-32* model by

Miklánek and Mészároš (1993) and then improved by Mészároš et al. (2002), (3) *SRAD* model developed by Wilson and Gallant (2000), (4) *Solar Analyst* model by Fu and Rich (1999, 2000), and (5) *r.sun* model developed by Hofierka (1997) and further modified by Hofierka and Šúri (2002).

For instance, *Solar Analyst* is a model evolved from the earlier models, specifically the *SolarFlux* model which was primarily based on the Canopy model (Rich et al. 1995). It is presently available in ArcGIS Spatial Analyst extension and enables the analyses of the sun effects over any geographic location for a particular period of time (Charabi and Gastli, 2010). The major input for the model include digital elevation model (DEM), slope and aspect. In addition, some parameters can be set as default or changed to suit the condition of geographic location under study. The tool, take into consideration the effects of atmosphere, site latitude, elevation, slope, aspect, sun angle shift and terrain shadows (Huang and Fu, 2009). For details of the model see Chapter 2.

On the other hand, the *r.sun* model was developed for Geographic Resources Analysis Support System (GRASS) Open Source GIS software by Hofierka (1997) and further improved by Hofierka and Šúri (2002) and Šúri and Hofierka (2004). Conceptually, the *r.sun* model according to Kryza et al. (2010) is based on the equations developed for the European Solar Radiation Atlas by Scharmer and Greif (2000) and on previous research of Hofierka (1997). The model operates in 2 modes; Mode 1 predicts the instantaneous direct, diffuse, and reflected radiation values in ($W.m^{-2}$), and the solar incidence angle (degrees) for any specified day and time. Whereas, the total amount of daily radiation (Wh/m^2) and direct radiation duration in (minutes) are calculated in Mode 2. The model requires a few main inputs comprises of DEM, slope, aspect, and local solar time (for Mode 1) and specific day number (for both Modes). Other inputs including linke turbidity factor (T_{LK}) and ground albedo are either set as default or changed to suit the area under study. Spatially varied inputs can be defined as grid maps. Furthermore, the model takes into account sky obstructions caused by terrain features from the DEM. The astronomical parameters, for instance, solar declination, are automatically calculated by the model (Šúri and Hofierka, 2004). More details are provided in Chapter 2.

These models are integrated with GIS and indeed serve as powerful tools for understanding landscape processes, rapid and cost-effective method for describing the solar radiation distribution at all scales (Fu and Rich, 1999) while taking into account the slope inclination, aspect and shadow effects by producing solar radiation maps (Dubayah and Rich, 1995; Fu and Rich, 2002; Tovar-Pescador et al. 2006). The use of these models has increased and their input datasets also improved dramatically. Yet, there has been a notable gap with respect to reliability information of the models outputs.

Uncertainty is defined as lack of knowledge regarding measurement reliability (Fu and Rich, 1999; Wechsler, 1999). The maps generated by these models often have various uncertainties since the current models did not provide uncertainty information or an inbuilt analysis tool. As the knowledge of solar radiation potential is vital for planning, operation and maintenance of solar energy technology systems (Hoyer-Klick et al. 2009) and several other applications, the level of uncertainty can be critical in the decision making process and guide for scientific research efforts (Harrison et al. 2011).

Generally, good modeling practice suggests the provision of model's confidence level in terms of uncertainties related to the model output (Crosetto et al. 2000). Moreover, since models are only an approximation of reality, their inputs are rarely, if ever, exactly known. Thus, their outputs are also likely to deviate from reality or errors that are contained in models and their inputs will propagate to their output (Heuvelink, 1998a). Model inputs as reported by Crosetto et al. (2000) are subject to several sources of uncertainty, for instance, errors of measurement, inadequate sampling. In addition, there is model uncertainty, which relates to structures, assumptions and simplifications of a model. Thus, an error in any model produces uncertainty.

DEM related uncertainties are due to several sources, structures and methods of DEM generations. These can be grouped mainly into; (1) systematic, (2) blunders, and (3) random (USGS, 1997; Wechsler, 2007). Systematic uncertainties are due to the processes involved in generating the DEM and generally follow specific patterns that can lead to bias or artifacts in the final DEM product (Fu and Rich, 1999; Wechsler, 1999). Systematic uncertainties can be minimized or eliminated especially when the source is known. Wechsler (1999) noted that blunders refer to the vertical errors related to data collection method which are mostly identified and eliminated before the data is released. On the other hand, random uncertainties remain in the data after the removal of known blunders and systematic uncertainties. In the case of *GIS-based solar radiation models*, topography is a major factor determining the solar energy incident on any location found on the Earth (Dubayah and Rich, 1995). In other words, Fu and Rich (2002) noted that variation in elevation, slope, aspect and obstructions of terrain features generates great local gradients of solar radiation. However, the main input data for *GIS-based solar radiation models* are the DEM and its derivatives (slope, aspect). DEMs are models that represent the Earth's surface elevations. This form of spatial data provides a model of reality that contains deviations from the truth (Wechsler, 2007), or inherent errors that constitute uncertainty (Holmes et al. 2000; Wechsler and Kroll, 2006; Hunter and Goodchid, 1997).

Approaches to DEM uncertainty modeling have been developed over the years. However, a number of shortcomings exist, for example, only global DEM accuracy figures are available (Hebeler, 2008). In addition, the same DEM with different spatial resolutions usually generates different estimates of elevation, slope, and aspect, particularly in complex terrains (Ruiz-Arias et al. 2009). In general, Fisher (1998) and Holmes et al. (2000) have come to the conclusion that DEMs contain errors to a certain degree that have great consequences on slope and aspect. Similarly, DEM errors generate varied solar radiation estimates. Therefore, if the solar radiation maps are used for solar energy applications, policy development, or scientific research, these uncertainties will propagate. Additionally, error propagation arising due to the processing of different sources of data layers within a GIS environment can also generate significant noise that impact the results interpretations (Heuvelink et al. 1989). These impose some limitations on the confidence of *GIS-based solar radiation models* outputs (Crosetto et al. 2000) and a challenge to their widespread application. Consequently, Jahanpeyma et al. (2007) noted that the presence of uncertainty in both spatial data and spatial analyses will potentially expose users to undesirable consequences in their decision-making process. In this regard, Zhang and Goodchild (2002) reported that the characterization (modeling and portrayal) of spatial uncertainty and its propagation to geographical modeling and its impact on spatial decision-making, has been identified as a critical research priority in GIScience, e.g., Heuvelink (1998), Heuvelink and Burrough

(2002), Sheng et al. (2009). Though, uncertainty in GIS is well known, and extensively investigated, the problem has not been solved (Goodchild, 1992; Devillers et al. 2002) rigorously.

One way of managing uncertainty is through error propagation models or uncertainty analysis (UA), which allows assessment of uncertainty in the model output as a result of error propagation through the model input data and uncertainties in the model itself (Heuvelink, 1998; Crosetto and Tarantola, 2001). In any field, UA is a prerequisite to model building (Crosetto et al. 2000). Similarly, in GIS, UA is necessary because different input data from different sources are associated with a wide range of errors (Thapa and Bossler, 1992; Karssenberg and De Jong, 2005). Heuvelink (1998) and Karssenberg and De Jong (2005) reported that modeling of error propagation via complex dynamic spatial models is assessed using powerful computational techniques where most of which model errors as stochastic variables. "Analytical solutions for error propagation as a result of spatial functions with stochastic variables also exist. However, these are only available for a limited number of relatively simple functions and do not support most dynamic models. As a result, error propagation in dynamic spatial environmental models is mostly computed by *Monte Carlo (MC)* simulation" (Karssenberg and De Jong, 2005). The MC approach to UA is based on adding a random noise factor (which can be positive or negative) to the data that is simulated using random number generators.

On the other hand, previous studies that examine the veracity of the custom-built *GIS-based solar radiation models* compared them to measured values from ground-based measurement sites (see for example Rich et al. 1995; Šúri et al. 2007; Ruiz-Arias et al. 2009; Kryza et al. 2010; Kumar, 2011). This approach of model validation based on ground-based measurements is not a viable option due to the dearth of ground-based measurement sites in many countries around the world. Others, for example, Thompson (2003); Ruiz-Arias et al. (2009) concentrated on the influence of DEM resolution on model output. However, it is important to note that *MC* simulation does not consider spatial autocorrelation, which is essential for spatial UA. Oliver et al. (1997) reported that among the techniques for modeling uncertainty, *Markov chain Monte Carlo (MCMC)* techniques hold most promise. Despite the advantages of this technique, it is observed from the literature that none of the previous studies have utilized it, and as well incorporated spatial autocorrelations despite the anecdotal and empirical evidence that shows DEM error is spatially variable and autocorrelated (Theobald, 1989; Weibel and Brändli, 1995; Ehlschlaeger and Shortridge, 1997; Hunter and Goodchild, 1997; Fisher, 1998; Kyriakidis et al. 1999; Carlisle, 2005). However, very little research has attempted to model this spatial autocorrelation. Wechsler (1999) and Raaflauband Collins (2006) stressed that uncertainty model that did not take into account the spatial autocorrelation can be seen as the ‘worst-case scenario’.

Therefore, new methodologies are required to fill these gaps, more especially, an *MCMC* simulation based method. The *MCMC* technique is a general Bayesian inference method for simulation of stochastic processes. It generates equally probable realizations from the target distribution using the transition probability of a Markov process with the property that its limiting invariant distribution is the target distribution. The Markov chain is then iterated in a computer-generated *MC* simulation, and the output, after a transient phase and under various sets of conditions, is a sample from the target distribution (Chib and Greenberg, 1995). The approach differs fundamentally from other methods, such as, *Taylor series*, or *MC* simulation, because it captures both the uncertainty in the correlation and the dependencies between the posterior correlations, variances, and means that are induced by their joint estimation from data and then correctly propagates them through the decision model (Ades and Lu, 2003). Additionally, it copes with modeling non-Gaussian nature of the parameter uncertainty (Minasny et al. 2011). Thus, Bayesian inference is particularly suited to the estimation of parameter uncertainty in GIS or environmental models because *a priori* information on the values of parameters can be conveniently incorporated into the parameter estimation process, in the form of informative *prior* distributions.

1.2 Purpose and Scope of the Research

The main purpose of this study is to investigate the effect of spatially autocorrelated DEM (elevation) errors on the *GIS-based solar radiation models* outputs using *MCMC* simulation based on *Metropolis-Hastings (MH) algorithm*. Spatial autocorrelation of elevation errors as noted by Guth (1992), Ehlschlaeger et al. (1997), Hunter and Goodchild is critical for any model of spatial uncertainty because errors at a certain location are found to influence errors at neighboring locations positively or negatively (Campbell, 1981). There are several *MCMC* algorithms which include *Metropolis-Hastings* (Metropolis et al. 1953) and *Gibbs sampler* (Geman and Geman, 1984); further details of these algorithms are given in Chapter 3. However, in this study, the *MH algorithm* is adopted because of its generality, simplicity and powerfulness (Robert and Casella, 2010).

Technology and in particular, GIS simplified the process of data capture, analysis, and presentation of outputs, yet there still remain many pitfalls, and users need to be able to think critically about what they are doing and the reliability of results obtained from this novel technology (Brimicombe, 2010). Though, efforts have obviously been taken to reduce uncertainty for any model, it is necessary to be mindful of the fact that uncertainty cannot be eliminated completely, and therefore it needs to be considered when deriving, distributing or displaying the results. Therefore, the scope of this study is focused on accounting for DEM uncertainties, their propagation and assessing their impact on the outputs of *GIS-based solar radiation models* specifically the *Solar Analyst* and *r.sun* models. In addition, the comparisons of the outputs from these two models are made.

1.3 The Main Contributions

For decades, uncertainties in both spatial data and spatial models are considered important for GIScience, for instance, see Goodchild and Gopal (1989); and the references contained therein (Atkinson, 1999). Uncertainty is extremely important in the process of model parameterization and assessment (Refsgaard et al. 2007; Wang et al. 2009; Wang and Chen, 2012). Generally, uncertainties arise from (1) measurement variations of spatial and temporal data; (2) non-uniqueness of model parameters and interaction among different parameters and various processes; and (3) imperfect model structure like processes description (Refsgaard et al. 2007; Wang and Chen, 2012). The impact of uncertainty is even more significant as the modeling results are often used in the policy making and decision.

Most studies pertaining to uncertainty in GIScience are currently concentrated on uncertainty due to individual sources, e.g., inputs. Moreover, the methods are confined to simple sensitivity analysis and MC simulations. More effective techniques should be developed to quantify the uncertainties in input data, model parameters and model structure. UA does not only give the uncertainty from different sources, but also gives an evaluation of model performance and limitations.

Hence, the basic contributions of this research include;

i. Prior to this study, an application of the MCMC method to investigate the effects of DEM uncertainties in *GIS-based solar radiation model's* outputs has never been addressed, thus, this research is the first of its kind to use MCMC method. The method captures both the uncertainty in the correlation and the dependencies between the posterior correlations, variances, and means that are caused by their joint estimation from the data and uses decision-model to propagate them (Ades and Lu, 2003). In addition, it copes with modeling non-Gaussian nature of the parameter uncertainty (Minasny et al. 2011),

ii. Unlike previous studies which assume that spatial autocorrelation errors in DEM are normally distributed, the current study incorporates spatial autocorrelation by generating a chain of correlated realizations through the rejection of perturbations or transitions that do not meet the acceptance criteria,

iii. Quantification of uncertainties allows better decision making practices. Consider for example, the solar radiation potential which is the main factor to be considered when locating solar power plants such as CSP and CPV. Thus, renewable energy policy-makers and investors can use such information in finding the most suitable locations for the site selection of solar power plants.

1.4 Research Outline

This thesis is structured in Seven Chapters. Chapter 2 gives a review on the theoretical basis of *GIS-based solar radiation models* and their applications, sources and structures of DEM data, DEM uncertainty and sources, and DEM uncertainty metrics. Chapter 3 starts with an exploration of the concept of uncertainty, uncertainty classification and types of error, and currently employed techniques for analysis of error propagation in GIScience. The Chapter further introduces an *MCMC* simulation technique, including relevant theory for Markov chain, general properties of Markov chains, convergence diagnosis and then concludes with spatial autocorrelation. Chapter 4 covers the detailed description of the proposed *MCMC* simulation methodological framework. Chapter 5 presents an implementation of the proposed methodology using a case study area of Abuja, Nigeria. Chapter 6 discusses results of the implementation of the proposed methodology based on the case study. Chapter 7 summarizes the main findings and conclusions of this thesis and future outlook.

CHAPTER 2

GEOGRAPHIC INFORMATION SYSTEM (GIS)-BASED SOLAR RADIATION MODELS

This chapter provides main conceptual background for the research by reviewing the relevant literature on solar energy models, theory, design, and implementation of the two *GIS-based solar radiation models*, namely *Solar Analyst* (Rich et al. 1995; Fu and Rich, 1999) and *r.sun* (Hofierka, 1997; Hofierka and Šúri, 2002). In addition, sources and structures of digital elevation model (DEM) uncertainty, and DEM uncertainty metrics are overviewed.

2.1 Solar Energy Models

Solar radiation is the fundamental renewable energy source that sustains the biosphere and drives its self-organization. Thus, reliable knowledge of solar radiation data sets is essential to several domains, for instance, energy planners, engineers, urban planners, architects, renewable energy, climatology, ecology and hydrology (Ertekin and Yalziz, 2000). The solar radiation data is useful, e.g., for optimum site-selection of solar energy generation facilities (Charabi and Gastli, 2010), assessment of future energy, economical benefits, and formulation of effective policies (Šúri et al. 2007).

In the literature, there are a wide variety of methodologies/models used for estimating solar radiation and these can be categorized into; *empirical methods* (for example, Angstrom, 1924; Winslow et al. 2001; Isikwue et al. 2012), *artificial neural network (ANN) models* (for instance, Alawi and Hinai, 1998; Ahmed and Adam, 2013); *satellite based models* (for example, Cano et al. 1986; Dubayah, 1992; Perez et al. 2002; Sorapipatana, 2010). Lastly, the *GIS-based solar radiation models* (for instance, Hetrick et al. 1993; Miklánek and Mészároš, 1993; Dubayah and Rich, 1995; Fu and Rich, 1999, 2000; Wilson and Gallant, 2000; Mészároš et al. 2002; Hofierka and Šúri, 2002). The *empirical models* usually consist of a few measurable meteorological parameters and require the development of a set of equations that relates it and other meteorological parameters (Donatelli et al. 2003). *ANN models* provide prediction with a reliable degree depending on the availability of input parameters (Azadeh et al. 2009). *Satellite models* according to Perez et al. (2002) derive a cloud index from the satellite (e.g., Meteorological Satellite, Geostationary Operational Environmental Satellite) visible channel and use this index to construct a clear sky global irradiance model that may be adjusted for ground elevation and linke turbidity factor (T_{LK}). Hofierka and Zlocha (2012) and Tovar-Pescador et al. (2006) noted that *GIS-based solar radiation models* predict the amount of incoming radiation using topographic information contained in DEM, for instance, elevation, slope, aspect, shadow effect, and latitude, as well as other different physical parameterization. However, this study considered only the *GIS-based solar radiation models*.

2.2 GIS-Based Solar Radiation Models

The *GIS-based solar radiation models* have been developed since 1990s. The models utilize terrain features available in DEM in order to calculate elevation, slope, aspect, and latitude. Based on this information and different physical parameters, these models are able to predict the spatial distribution of the incoming radiation (Tovar-Pescador et al. 2006). The integration of solar radiation models with GIS provides an efficient means of computing radiation over large geographic locations while taking into account the effects of local terrain, incorporation of environmental and socio-economic variables, and generation of scenarios to policy-makers (Dubayah and Rich, 1995; Hofierka and Šúri, 2002; Nguyen and Pearce, 2010). Šúri et al. (2007) further stressed that a tool of this nature could effectively help in assessing solar energy potential at various levels, for example, local, national or supra-national. Tovar-Pescador et al. (2006) and Fu and Rich (2002) noted that GIS along with DEM have favored recently the development of solar radiation models, especially in complex terrains. Additionally, the mutual interaction of GIS with remote sensing (RS), as well as the inclusion of the methods of statistics and geostatistics as tools for spatial analysis has contributed greatly towards this success. The first generation of the models was based on simple empirical algorithms which describes the solar radiation atmospheric attenuation, for instance, *SolarFlux* model developed for ArcInfo GIS by Hetrick et al. (1993) and modified by Dubayah and Rich (1995), *Solei-32* for IDRISI by Miklanek (1993), and *Genasys* for (Kumar et al. 1997). The second generations of GIS-based solar radiation models are more advanced, for example, *SRAD* by McKenney et al. (1999), *Solar Analyst* by Fu and Rich (2000), and *r.sun* model of Hofierka and Šúri (2002). The *Solar Analyst* and *SRAD* models are designed for estimating the interactions of both short- and long- waves that falls on the surface of the Earth. However, despite being suitable for fine-scale analysis, these models have some limitations with respect to diffuse radiation computation and application for large areas (Šúri and Hofierka, 2002; 2004). On the other hand, Geographic Resources Analysis Support System (GRASS) Development Team (2006) reported that the *r.sun* model, available in the open source GIS GRASS environment is free from the aforesaid limitations. Thus, it operates quickly and effectively for both larger scale areas and higher spatial resolution data. However, it is important to note that none of these models provide a level of uncertainty in the outputs. Among these *GIS-based models*, only *Solar Analyst* and *r.sun* models are taken into account for review here, as they are considered for the thesis work.

2.2.1 Solar Radiation Modeling with ArcGIS

The *Solar Analyst* is a model evolved from earlier models, specifically, the *SolarFlux model* which was primarily based on the *Canopy model* (Rich et al. 1995). It is presently available in ArcGIS Spatial Analyst extension and enables the analyses of the effects of the sun over a geographic area for specific time periods. It computes direct radiation, diffuse radiation, and global radiation; and reports the outputs in watt hours per square meter (Wh/m^2). However, reflective radiation, while technically part of global radiation, is not included in *Solar Analyst* because of its complexity and the relatively low influence on total radiation (Fu and Rich, 1999; Huang and Fu, 2009). The model also generates direct radiation duration (i.e., duration of direct incoming solar radiation) in hours. The major input for the model is DEM and its derivatives (slope and aspect), and atmospheric transmittivity (Table 2.1).

Table 2.1. *Solar Analyst* inputs (ESRI, 2008)

Solar Analyst inputs	Unit
DEM	Meters
Latitude	Single value
Slope and Aspect	decimal degrees
Z factor	Dimensionless
Day interval	Days
Hour interval	Hours
Sky size/resolution	Cells
Time configuration	Numbers
Zenith divisions	Dimensionless
Azimuth divisions	Dimensionless
Calculation of directions	Dimensionless
Diffuse model type	Dimensionless
Transmitivity	Dimensionless

The model takes into account the effects of atmospheric conditions, site latitude, elevation, slope, aspect, sun angle shift, and topography shadows. The model analyses a landscape or specific location using two methods, namely, area and point. The area solar radiation analysis is used to calculate the insolation across an entire landscape. For each specific location, the computation is repeated in the input terrain surface and produces maps of radiation estimates for the whole area. While, the point solar radiation analysis is used to calculate the amount of radiant energy for a given location. Locations can be stored as point features or as x, y coordinates in a location table. Solar radiation calculations can be performed for specified locations only. These calculations, performed for point specific or area based, are executed based on the following steps below;

(1) terrain is used to compute an upward-looking hemispherical viewshed,
(2) direct radiation is computed by overlaying the viewshed on a direct sunmap,
(3) diffuse radiation is computed by overlaying the viewshed on a diffuse skymap,
(4) the process is repeated for all the interested locations to generate solar radiation potential map. Computation for hemispherical viewshed (upward-looking) of all specific locations is carried out from the DEM. A hemispherical viewshed is like a fish-eye photograph and generates a view of the entire sky from ground level (Huang and Fu, 2009).

The methods from the hemispherical viewshed algorithm upon which the *Solar Analyst* is based on, was developed by Rich et al. (1990; 1994) and extended by Fu and Rich (2000, 2002) is hereby summarized below.

2.2.1.1 Global Radiation Computation

Global radiation ($Global_{tot}$) is computed as the summation of direct (Dir_{tot}) and diffuse (Dif_{tot}) radiation of all sunmap and skymap sectors, respectively. The Global radiation is computed using Equation (2.1):

$$Global_{tot} = Dir_{tot} + Dif_{tot}. \tag{2.1}$$

The sunmap is a raster layer that defines the suntracks, i.e., sun's position as it moves over time, for instance, hourly, daily, monthly, and seasonally (Fu and Rich, 1999). On the other hand, skymaps are grid maps produced by splitting the entire sky into several sectors based on zenith and azimuth (Fu and Rich, 1999; Ruiz-Arias et al. 2009; Martínez-Durbán et al. 2009).

2.2.1.2 Direct Radiation Computation

Total direct radiation (Dir_{tot}) is the summation of all the direct radiation ($Dir_{\theta,\alpha}$) from all the sunmap sectors:

$$Dir_{tot} = \sum Dir_{\theta,\alpha}. \tag{2.2}$$

The direct radiation from the sunmap sector ($Dir_{\theta,\alpha}$) with a centroid at zenith angle (θ) and azimuth angle (α) is computed using Equation (2.3):

$$Dir_{\theta,\alpha} = S_{const}.\beta^{m(\theta)}.SunDur_{\theta,\alpha}.SunGap_{\theta,\alpha}.\cos\left(AngIn_{\theta,\alpha}\right), \tag{2.3}$$

where

S_{const} represents the solar flux outside the atmosphere at the mean earth-sun distance, usually refers to solar constant. β denotes atmospheric transmitivity (i.e., mean of all wavelengths) for the shortest path (in the direction of the zenith); $m(\theta)$ denotes the relative optical path length, measured as a proportion relative to the zenith path length, refer to Equation (2.4). $SunDur_{\theta,\alpha}$ represents the time frame of the sky sector, usually, equal to the day interval (for instance, a month) multiplied by the hour interval (for example, a half hour). For partial sectors (near the horizon), the duration is computed using spherical geometry; $SunGap_{\theta,\alpha}$ represents sunmap sector gap fraction; $AngIn_{\theta,\alpha}$ denotes incidence angle between the sky sector centroid and the axis normal to the surface, refer to Equation (2.5). Zenith angle and elevation/height are used to calculate the relative optical length($m(\theta)$). Equation (2.4) is used to compute zenith angles less than 80°:

$$m(\theta) = EXP\left(-0.000118.Elev - 1.638.10^{-9}.Elev^2\right) / \cos(\theta), \tag{2.4}$$

where

θ denotes the solar zenith angle, Elev represent the elevation/height (m).

Angle of incidence ($AngInSky_{\theta,\alpha}$) is the angle between the intercepting surface and a given sky sector with a centroid at zenith angle and azimuth angle. The Angle of incidence is computed using Equation (2.5):

$$AngInSky_{\theta,\alpha} = a\cos\left[Cos(\theta).Cos(G_z) + Sin(\theta).Sin(G_z).Cos(\alpha - G_a)\right], \tag{2.5}$$

where

G_z denotes zenith angle of the surface,
G_a represents azimuth angle of the surface,
Note that refraction is important especially for zenith angles greater than 80°.

2.2.1.3 Diffuse Radiation Computation

The diffuse radiation at its centroid (Dif) is computed and incorporated based on the time interval for each specific sky sector, and then corrected using the gap fraction and the angle of incidence as shown in Equation (2.6):

$$Difn_{\theta,\alpha} = R_{\mathrm{glb}}.P_{dif}.Dur.SkyGap_{\theta,\alpha}.Weight_{\theta,\alpha}.\cos\left(AngIn_{\theta,\alpha}\right), \tag{2.6}$$

where

R_{glb} is the global normal radiation, see Equation (2.7); P_{dif} represents the amount of diffused global normal radiation flux. Usually, 0.2 and 0.7 are used for very clear sky and very cloudy sky conditions, respectively; Dur denotes the time interval; $SkyGap_{\theta,\alpha}$ represents the gap fraction (proportion of visible sky) for the sky sector; $Weight_{\theta,\alpha}$ denotes the amount of diffuse radiation originating from a particular sky sector relative to all sectors, refer to Equations (2.8)-(2.9); $AngIn_{\theta,\alpha}$ represents the angle of incidence between the centroid of the sky sector and the intercepting surface. R_{glb} is computed using Equation (2.7). It is the summation of the direct radiation for each and every sector (obstructed sectors included) without correction for angle of incidence, then correcting for proportion of direct radiation, which equals to $1 - Pdif$.

$$R_{\mathrm{glb}} = \left(S_{const}\sum\left(\beta^{m(\theta)}\right)\right)\Big/\left(1 - Pdif\right). \tag{2.7}$$

For the uniform sky diffuse model, $Weight_{\theta,\alpha}$ is computed using Equation (2.8):

$$Weight_{\theta,\alpha} = \left(\cos\theta_2 - \cos\theta_1 / Div_{azi}\right), \tag{2.8}$$

where

θ_1 and θ_2 are the bounding zenith angles of the sky sector,
Div_{azi} is the number of azimuthal divisions in the skymap.

For the standard overcast sky model, $Weight_{\theta,\alpha}$ is calculated as follows:

$$Weight_{\theta,\alpha} = \left(2\cos\theta_2 + \cos 2\theta_2 - 2\cos\theta_1 - \cos 2\theta_1\right)\Big/4.Div_{azi}. \tag{2.9}$$

The computation of total diffuse solar radiation for any specific location (Dif_{tot}) is done by summing all the diffuse solar radiation (Dif) derived from each and every skymap sectors as shown in Equation (2.10):

$$Dif_{tot} = \sum Dif_{\theta,\alpha}. \tag{2.10}$$

Having understood the theoretical basis of the Solar Analyst, the review now refocused on the application of the model. Charabi and Gastli (2010) investigated the solar electricity prospects in Oman using the *Solar Analyst* model. They emphasized that in order to minimize the model's computational time without affecting the accuracy of the results, appropriate DEM resolutions must be applied. A 100 m x 100 m DEM which shows a relatively short execution time of the model is adopted. The study concludes that the in-situ measurements realized with pyrometers will not be able to capture the spatial variability in radiation caused by topography as the GIS model does. Similarly, for the same study area (Daqum, Oman), Gastli and Charabi (2010) utilized DEM with different resolutions ranging between (1,000 m x 1,000 m and 5,000 m x 5,000 m) and obtained results that showed no significant difference.

Clifton and Boruff (2010) assessed the potential for concentrated solar power (CSP) development in Western Australia's Wheatbelt using ArcGIS. They first generated the solar radiation potential using the *Solar Analyst* and then overlayed several datasets which comprises of environmental variables and electricity infrastructures. However, the authors were not contented with a default of 0.5 for general clear sky conditions and a value of 1 signifying complete transmission advised by ESRI (2008) to represent solar radiation transmissivity of the study area. To determine the most appropriate transmissivity value for the region, they performed an analysis of variance based on the average monthly direct normal irradiance (DNI) measurements collected for the two bureaus of meteorology stations in the closest proximity to the Wheatbelt and the National Aeronautics and Space Administration's (NASA) satellite data (NASA, 2009) for Australia. Following the determination of appropriate transmissivity value of 0.85, *Solar Analyst* was used to calculate DNI at the Earth's surface using the Shuttle Radar Topography Mission (SRTM) 90 m x 90 m DEM. The study concludes that CSP facilities can be suitably located on a large portion of the Wheatbelt which can be tailored to local patterns of supply and demand.

Tovar-Pescador et al. (2006) evaluated the reliability of *Solar Analyst* results over a complex topography and then compared the results against experimental data observed in Sierra Nevada National Park, Spain. The study employed a DEM of 20 m resolution and radiation data from 14 radiometric stations. The results indicate that the model performed well but with minor overestimation. The study, therefore, concluded that terrain parameters, such as elevation, slope, and aspect greatly affect the solar radiation, particularly in areas of relatively short distances. Moreover, the topographic parameters obtained from the DEM were compared with global positioning system (GPS) measurements and the results show that there are relatively high differences in slope and aspect values than that of the elevation. They also noted that the lack of incorporation of ground albedo by the model is another factor that may lead to incorrect estimates. Similarly, Ruiz-Arias et al. (2009) evaluated the reliability of daily solar radiation estimates of *Solar Analyst*, *r.sun*, *SRAD* and *Solei-32* models in a complex topography of Sierra Nevada National Park, Spain. The estimates of these models were tested against the field observation data from 14 meteorological stations. In addition, the effect of DEM resolution on solar radiation outputs of the models was also examined using a 20 m and 100 m DEM resolution. The result shows that *Solar Analyst* provides the least estimate when compared with *r.sun* model but yields a better performance with the 20 m DEM resolution, whereas *r.sun* shows better results using the 100 m resolution DEM. Unlike the *Solar Analyst*, the *r.sun* outputs are more similar to the ground measurements despite using different DEM resolutions. Atmospheric transmittivity; one of the basic input of *Solar Analyst* is sensitive to the presence of clouds thereby affecting the model output. Table 2.2 shows the summary of previous studies using *Solar Analyst*.

Table 2.2. Summary of previous studies that utilized the *Solar Analyst* model

Author(s)	Study Area	Input inputs	General remark
Charabi and Gastli (2010)	Daqum, Oman	DEM (100 m x 100 m)	Selection of appropriate DEM is important. The in-situ measurement realized with pyrometers will not be able to capture the spatial variability caused by topography as GIS models do.
Gastli and Charabi (2010)	Daqum, Oman	DEM resolutions ranging between (1,000 m x 1,000 m and 5,000 m x 5,000 m).	No significant difference in solar radiation outputs by using different DEM resolutions.
Clifton and Boruff (2010)	Western Australia's Wheatbelt	DEM (90 m x 90 m) and transmissivity value of 0.85.	Expressed doubt over the transmissivity value of (0.5 for general clear sky conditions and a value of 1 for complete transmission) advised by ESRI to represent solar radiation transmissivity of the study area.
Ruiz-Arias et al. (2009)	Sierra Nevada National Park, Spain	DEM (20 and 100 m) and radiation data from 14 radiometric stations.	*(i)* DEM resolution tends to affect the quality of model results; *(ii)* Atmospheric transmittivity is sensitive to the presence of clouds.
Tovar-Pescador et al. (2006)	Sierra Nevada National Park, Spain	DEM (20 m) and radiation data from 14 radiometric stations.	*(i)* Topographic parameters (elevation, slope and aspect) strongly modify the solar radiation; *(ii)* DEM resolution plays a vital role in obtaining accurate solar radiation estimates; and *(iii)* the model does not take into account the albedo.

While *Solar Analyst* is a sophisticated model, it is not without some limitations. The most notable limitation is that the model generalizes overcast conditions. Cloud cover is an important factor when determining incoming solar radiation of an area and is addressed through radiation parameters in *Solar Analyst* by estimating the proportion of radiation that passes through overcast skies, and the proportion of diffuse radiation. These parameters, however, do not directly account for the local overcast.

2.2.2 The Solar Radiation Model for GRASS GIS

The *r.sun* clear-sky solar radiation model was developed for GRASS GIS software (open source) by Hofierka and Šúri (2002) based on the previous work of Hofierka (1997). The model calculates all the three components of solar irradiance: beam, diffuse and reflected for both clear-skies and overcast conditions. It is robust and flexible over various scales and considers all relevant inputs as spatially distributed entities to enable computations for large areas with complex terrain (Šúri and Hofierka, 2004). The model has the following key features:

(1) it is a grid-based GIS tool with varied spatial inputs and outputs. Table 2.3 presents a list of all the inputs,
(2) it is an open source model with available source code for further improvement,
(3) all the clear-sky and real-sky solar radiation components for irradiation and irradiance values are provided,

(4) computation can be done assuming solar or civil time,
(5) it has a large scalability for various spatial resolutions and region sizes. Memory management and code optimization allow use of high resolution data, and
(6) integrating the model with GRASS GIS provide an opportunity to process both the input and output data within one computer environment (Hofierka and Cebecauer, 2008).

Table 2.3. *r.sun* inputs (Hofierka and Šúri, 2002)

Input name	Type of input	Description	Mode	Units	Interval of values
Eleven	Raster	Elevation	1, 2	Meters	0 – 8900
Aspin	Raster	Aspect (panel azimuth)	1,2	Decimal degrees	0 – 360
Slopein	Raster	Slope (panel inclination)	1,2	Decimal degrees	0 – 90
Linkein	Raster	T_{LK}	1,2	Dimensionless	0 – ≈7
Lin	Single value	T_{LK}	1,2	Dimensionless	0 –≈7
Albedo	Raster	Ground albedo	1,2	Dimensionless	0 – 1
Alb	Single value	Ground albedo	1,2	Dimensionless	0 – 1
Latin	Raster	Latitude	1,2	Decimal degrees	-90 – 90
Coefbh	Raster	Clear-sky index of beam component	1,2	Dimensionless	0 – 1
Coefdh	Raster	Clear-sky index of diffuse component	1,2	Dimensionless	0 – 1
Day	Single value	Day number	1,2	Dimensionless	0 – 366
Declin	Single value	Solar declination	1,2	Radians	-0.40928 – 0.40928
Time	Single value	Local (solar) time	1	Decimal hours	0 – 24
Step	Single value	Time step	2	Decimal hours	0.01 – 1.0
Dist	Single value	Sampling distance coefficient	1,2	Dimensionless	0.1 – 2.0

The *r.sun* output parameter's settings are automatically recognized between the Modes (1 & 2). For instance, performing computation in Mode 1 produces raster maps of the incident angle (incidout) and solar radiation components including beam (beamrad), diffuse (diffrad), and reflected (reflrad) radiations. On the other hand, computation in Mode 2 generates the sums of solar irradiation for the chosen global radiation components (beam_rad, diff_rad and refl_rad) with respect to the defined day period. In addition, a beam irradiation duration (insol_time) is calculated.

Apart from the clear-sky irradiances/irradiations, the *r.sun* model also computes the overcast radiation when supplied with beam and diffuse components of clear-sky index. By default, the incidence angle and irradiance/irradiation maps are generated without taking into account the shadow effects. However, settings must be changed to account for the shadow effects in complex terrains this may result to different estimates, particularly at low sun altitudes. Zero is recorded in the output maps for shadowed locations. The model's inputs are presented in Table 2.4. In addition to output raster maps, the model stores parameters utilized in the computations and provide r.sun_out.txt local text files. An example of the stored parameters includes day number, solar constant, interval of latitude, time step T_{LK} and ground albedo values.

Table 2.4. *r.sun* out raster maps (Hofierka and Šúri, 2002)

Parameter name	Description	Mode	Units
Incidout	Solar incidence angle	1	Decimal degrees
beam _rad	Beam irradiance	1	$W.m^{-2}$
diff_rad	Diffuse irradiance	1	$W.m^{-2}$
refl_rad	Ground reflected radiance	1	$W.m^{-2}$
insol_time	Duration of the beam irradiation	2	min.
beam_rad	Beam irradiation	2	$Wh.m^{-2}.day^{-1}$
diff_rad	Diffuse irradiation	2	$Wh.m^{-2}.day^{-1}$
refl_rad	Ground reflected irradiation	2	$Wh.m^{-2}.day^{-1}$

The main equations utilized for computation of direct and diffuse radiations that fall on the Earth's surface is given by Šúri and Hofierka (2004). The computation basically starts with the solar constant (I_o), equivalent to 1367 $W.m^{-2}$. The model considers the Earth's orbit eccentricity with the correction factor of e incorporated in the computation of extraterrestrial irradiance G_o normal to the solar beam:

$$G_o = I_o e. \tag{2.11}$$

where

$$e = 1 + 0.03344 \cos\left(j' - 0.048869\right), \tag{2.12}$$

the day angle j' is in radians:

$$j' = 2\pi/365.25, \tag{2.13}$$

and j represent the day number varying from 1st January to 31st December (365/366).

The beam radiation normal to the solar beam B_{oc} is computed using Equation (2.14):

$$B_{oc} = G_o \exp\left(-0.8662.L_{LK}.m.\delta_R\left(m\right)\right), \tag{2.14}$$

where

L_{LK} denotes linke turbidity factor,
m represents relative optical air mass,
δ_R is the Rayleigh optical thickness at air mass (m).

The beam radiation on a horizontal surface B_{hc} is computed as:

$$B_{hc} = B_{os} \sin\left(h_o\right), \tag{2.15}$$

where

h_o denotes the solar height.

Finally, the beam radiation on an inclined surface B_{ic} is given as:

$$B_{ic} = B_{hc} \sin \delta_{\exp} / \sin(h_o), \tag{2.16}$$

where

$\delta_{\exp}$ represents the measured solar incidence angle between the Sun and inclined surface.

The modeling of the diffuse component on a horizontal surface D_{hc} (W.m^{-2}) for a clear sky condition B_{hc} is computed as a product of the G_o and a diffuse transmission function T_n which dependent only on the L_{LK}. $F_d(h_o)$ is the solar altitude function, which also depends on L_{LK}, and a diffuse solar altitude function F_d depend only on the solar altitude h_o:

$$D_{hc} = G_o T_n(L_{LK}) F_d(h_o). \tag{2.17}$$

The diffuse irradiance on inclined surface D_{ic} (W.m^{-2}) is computed with Equation (2.18) and Equation (2.19), depending on whether the raster layer is a sunlit or shadowed surface:

sunlit surface:

$$D_{ic} = D_{hc}\left(F(\gamma_N)(1-K_b) + K_b \sin \delta_{\exp} / \sin \mathrm{h}_o\right), \tag{2.18}$$

shadowed surface:

$$D_{ic} = D_{hc}\left(F(\gamma_N)\right), \tag{2.19}$$

where

$F(\gamma_N)$ accounts for the diffuse sky irradiance, and K_b is a proportion between direct radiation and extraterrestrial solar radiation on horizontal surface. L_{LK} represents linke turbidity factor which increases the diffuse radiation and decreases the direct radiation.

The clear-sky diffuse ground reflected irradiance for inclined surfaces R_i is calculated using Equation (2.20):

$$R_i = \rho_g G_{hc} r_g(\gamma_N), \tag{2.20}$$

where

$$r_g(\gamma_N) = (1-\cos\gamma_N)/2, \tag{2.21}$$

and:

$$G_{hc} = B_{hc} + D_{hc}. \tag{2.22}$$

The global irradiance/irradiation on a horizontal surface for overcast conditions G_h is computed using Equation (2.23):

$$G_h = G_{hs} k_c. \tag{2.23}$$

where

k_c is the ratio between horizontal global radiation under overcast G_{hs} and clear-sky G_{hc}:

$$k_c = G_{hs}/G_{hc}. \tag{2.24}$$

For a more comprehensive explanation of *r.sun* algorithms refer to Hofierka and Šuri (2002) and Šuri and Hofierka (2004). However, to demonstrate the model's capabilities, Hofierka and Šúri (2002) utilized the model to assess the potential photovoltaic electricity in 10 countries from Eastern Europe. Spatial data are represented by raster maps with parameters representing terrain, latitude, L_{LK}, radiation, and clear-sky index. Šúri et al. (2005) developed a database of solar radiation for European countries based on the *r.sun* model, and interactive internet tools for accessing, displaying and estimating the productivity of photovoltaic (PV) electricity for any specific location. The methodology utilized the following inputs;

(1) average monthly daily global and diffuse radiations obtained from 566 meteorological stations between 1981 to 1990,
(2) L_{LK} of 611 locations obtained from the SoDa database and then interpolated,
(3) average monthly values of clear-sky index obtained from meteorological station and then interpolated, and
(4) a 1x1 km spatial resolution DEM generated from the United States Geological Survey (USGS) SRTM-30 data.

They concluded that such a comprehensive GIS database provides reliable information on solar radiation and other parameters for both the European and neighboring countries as compared to other sources of data available.

Hofierka and Cebecauer (2008) assessed the solar resources in Slovakia using *r.sun* model with the aim of producing a solar radiation database. The study utilized 100 m DEM resolution resampled from SRTM 90 m, measured global and diffuse irradiation from ground stations, L_{LK} map, an interpolated clear-sky index map, and constant albedo of 0.15. The quality of the model outputs reflects the source data and solar radiation methodology. A high resolution DEM improved the accuracy of topographical factors (relative air mass and shadowing). The real-sky irradiation modeling is more influenced by the lack of representative ground-based measurements reflecting local variations and dynamics of cloudiness. However, the overall accuracy of the database represented by monthly and annual values of irradiation was assessed using basic statistical measures, the relative mean bias error (rMBE) and relative root mean square error (rRMSE). The rMBE in six evaluated meteorological stations in Slovakia reached 3.1% and 2.7% for monthly and annual real-sky irradiation on a horizontal plane, respectively. The rRMSE for monthly and annual irradiation was a bit higher approaching 5.9% and 4.2%, respectively.

Kryza et al. (2010) estimated the global radiation based on clear sky condition for Wadel Jarsberg, Poland using *r.sun* and then evaluated the results by comparing to field observation data from the Polish Polar Station. The main data for the study area includes; 10 meter resolution DEM, slope, aspect, clear-sky index, and T_{LK}. T_{LK} is a key parameter influencing the solar radiation estimate with *r.sun* model and is defined as the ratio of the broad band extinction coefficient at unit air mass (δ) to Rayleigh's optical thickness (δ_R). Thus, the interpretation of T_{LK} is the number of clean dry atmospheres necessary to produce exact extraterrestrial radiation attenuation generated by the real atmosphere. The T_{LK} utilized by Kryza et al. (2010) is computed using an empirical formula proposed by Dogniaux (1984), and assumed to be constant over the study area. The study concludes that *r.sun* underestimates the global radiation due to mis-specification of the T_{LK}, aerosol optical thickness used for the research period, and the clear-sky condition, which is based on low and medium level cloudiness for model evaluation.

Nguyen and Pearce (2010) used *r.sun* model to compute insolation, and then identified candidate sites for establishing new solar parks in the South-eastern part of Ontario, Canada. To run the model, the following data sets were utilized; DEM (1 km GTOPO30, 90 m SRTM30), slope, aspect, latitude, albedo (constant 0.2), mean days corresponding to the sun's angular position, T_{LK}, clear sky index (Kc), field measurements of global horizontal irradiation (GHI). Whereas, census subdivisions (CSD), soil classifications for forestry, soil classifications for agriculture, and land use classifications (excluding agriculture) were utilized in the multi-criteria evaluation to identify large-scale PV sites. The authors used two different DEMs with different resolutions since Cebaucauer et al. (2007) suggested that decreasing the spatial resolution of DEM may lead to overestimation of annual electricity output. Table 2.5 shows summary of studies on *r.sun* application.

Table 2.5. Summary of the previous studies using *r.sun* model

Author(s)	Study Area	Inputs	General remarks
Hofierka and Šúri (2002)	Ten East European countries	- USGS GTOPO30 DTM (1km), - Slope, - Aspect, - Local solar time, - Latitude raster map, - L_{TK} map, - Ground albedo (0.15), - Clear-sky irradiance values.	The *r.sun* is a complex and flexible solar radiation model. It is particularly appropriate for modeling of large areas with complex terrain because all spatially variable solar parameters can be defined as raster maps.
Hofierka and Šúri (2002)	Slovakia	- DTM (100 m), - Latitude raster map, -T_{LK}, - Ground albedo (0.15), and - Clear-sky irradiance values.	The results show considerable differences in spatial pattern due to terrain shadows in mountainous regions.
Šúri et al. (2005)	European	- Average monthly global and diffuse irradiation for 566 ground meteorological stations covering a period 1981-1990, - L_{TK}, - Clear-sky index, and - DEM of 1x1km resolution derived from the USGS SRTM-30 data.	*(i)* A photovoltaic geographic information system (PVGSIS) was developed based on the *r.sun* model output. *(ii)* High resolution DEM can dramatically improve the spatial accuracy of the shadowing.
Hofierka and Cebecauer (2008)	Solar database for Slovakia	- DEM (100 m), - Global and diffuse solar radiation measured from ground stations, - T_{LK} and clear-sky index maps, - Constant albedo of 0.15.	*(i)* High resolution DEM improved the accuracy of terrain factors. *(ii)* The real-sky irradiation modeling is more affected by the lack of representative field data reflecting local variations and dynamics of cloudiness.

Table 2.5. Summary of the previous studies using *r.sun* model (continued)

Author(s)	Study Area	Inputs	General remarks
Kryza et al. (2010)	Wadel Jarlsberg, Poland	- DEM (10 m), - Slope, - Aspect, -T_{LK}, - Clear-sky index, and - Radiation data from meteorological stations	*(i)* T_{LK} is a key parameter influencing the solar radiation estimate with *r.sun* model.
Nguyen and Pearce (2010)	Southern-eastern Ontario, Canada	- DEM (1km GTOPO30, 90m, SRTM30 m), - Slope, - Aspect, - Latitude, - Albedo (constant 0.2), - Mean days and corresponding angular position of the sun, -T_{LK}, - Ground-measured values of GHI, - Kc, - CSD, - Soil types for forestry, - Soil types for agriculture, - Land use classification (exclusive of agriculture).	(*i)* Decrease of DEM resolution from 100 to 3000 m may potentially lead to overestimation of the regional means of the yearly electricity yield by 8.2% for the former system and 15.2% for the latter one.

2.3 Sources and Structures of DEM Data

DEM generally refers to the digital cartographic representation of the elevation or terrain of the Earth's surface as regularly spaced intervals in x, y directions; using z-values (elevation values) referenced to a common vertical datum and represented as a raster or a triangular irregular network (TIN) (Greve, 1996; Mikhail et al. 2001).

A number of techniques and technologies with various accuracies are used to capture DEMs. These include; *digitizing, total station, differential global positioning system (DGPS), photogrammetry* and *RS technologies: Light Detection and Ranging and Interferometric Synthetic Aperture Radar sensors* (Blomgren, 1999; Farah et al. 2008). *Digitization* is done by tracing elevation data from topographic map sources (see for example Gessler et al. 1993; Austin et al. 1996). This was often done in the United States before production of DEMs became widely available and remains a viable option. Some scientists have preferred to manually produce elevation data in areas where existing DEMs were subject to excessive artifact error (de Swart et al. 1994; Garbrecht and Starks, 1995). The increasing affordability and sophistication of *DGPS* technology has enabled the construction of project-specific DEMs, particularly for small study areas or for spot elevations at specific sites. Advantages of using the *DGPS* approach include that the researcher can specify the spatial resolution and extent of the survey, design the sampling campaign, and oversee the production of the final digital elevation model. The accuracy of both the final DEM and the *DGPS* data can be tested if the sampling campaign is carefully designed; leading to sophisticated uncertainty models (Oliver et al. 1989). However, Shortridge (2001) reported that calibrating the *DGPS*, particularly in rural/wilderness areas far from elevation benchmarks, can be difficult, time consuming, and expensive. Collecting densely sampled elevation data in a careful and systematic manner is a lengthy process, particularly if it is combined with a plan to model data uncertainty. Finally, interpolation schemes to generate an elevation surface from the collected point data introduce uncertainty into the final product, and the choice of method can seem arbitrary, see Lam (1983) for a description of spatial interpolation methods. On the other hand, the rapid advances in RS have made high resolution (i.e., less than 100 m) DEMs available for both small area and global DEMs available to the scientific community and general public users in the last decade, such as SRTM, Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), and altimetry DEMs (Chang et al. 2010).

Generally, DEMs are categorized into three structures; *grid based (raster), triangulated irregular network (TIN),* and *vector or contour based* (Moore et al. 1991; Wilson and Gallant, 2000), see Figure 2.1.

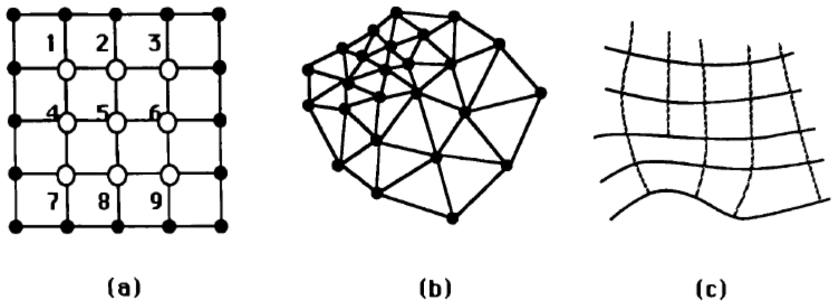


Figure 2.1. DEM structures: (a) square-grid network based on a moving 3 x 3 sub matrix centered on 5; (b) triangulated irregular network - TIN; (c) vector or contour-based network (Moore et al. 1991)

"*Grid-based methods* may use a regularly-spaced triangular, square, or rectangular mesh or a regular angular grid, such as the 3 arc-second spacing used by the US Defense Mapping Agency. The data can be stored in a variety of ways, but the most efficient is as *z*-coordinates corresponding to sequential points along a profile with the specified starting point and grid spacing. The elemental area is the cell bounded by three or four adjacent grid-points for regular triangular and rectangular grid-networks, respectively" (Moore et al. 1991). Square-grid DEMs are easier to implement in a computer due to its simplistic approach of representing elevation values in the form of matrices that implicitly record the topological relationships between the data points (Moore et al. 1991; Wilson and Gallant, 2000; and Zhao and Tang, 2008). Thus, these advantages have made the square-grid DEMs emerged as the most commonly used data structure in several GIS applications during the past decades for instance, solar radiation modeling (Tovar-Pescador et al. 2006; Charabi and Gastli, 2010; Clifton and Boruff, 2010), hydrology (Chang and Tsai, 1991; Chaubey et al. 2005), multi-criteria decision analysis (van Zyl and Labadie, 2011; Lawal et al. 2012), statistical and quantitative analysis of terrain (Iwahashi et al. 2001; Zhou and Liu, 2004; Miliaresis, 2008). The grid-based DEMs have been generated in two ways, direct photogrammetric measurements, or interpolation from digitized contours. Recently, methods based on RS such as interferrometry from SAR satellite data (Genz and van Genderen, 1996) and laser scanners (Ackermann, 1996) have been developed for generating DEMs. A *Triangulated irregular network (TIN)* is a triangulated mesh constructed on the (x, y, z) locations of a set of data points (Tate and Maidment, 1999). Thus, it is a data structure based on "a piece-wise linear interpolation of a set of points in x, y, z coordinates, that results in nonoverlapping triangular elements of varying size. Although several methods exist, the Delaunay triangulation is a preferred technique since it provides a nearly unique and optimal triangulation, e.g., Watson and Philip, 1984; Tsai, 1993. For a set of points, the Delaunay criterion ensures that a circle that passes through three points on any triangle contains no additional points" (Vivoni et al. 2004), see Figure 2.2.

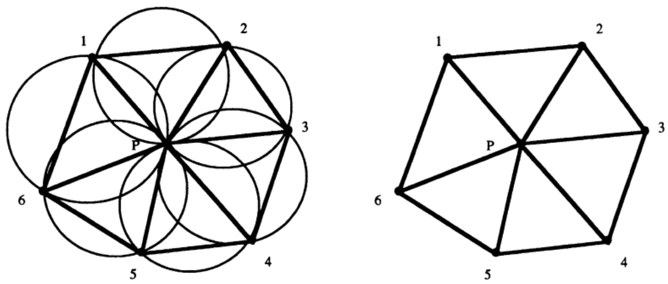


Figure 2.2. Local Delaunay triangulation about a point P (Jones et al. 1994)

TINs have also found widespread usage (for instance, Tajchman, 1981; Yu et al. 1997). It is reported that "the best TINs sample surface-specific points, such as peaks, ridges, and breaks in slope, and form an irregular network of points stored as a set of x, y, and z values together with pointers to their neighbors in the net. TINs can easily incorporate discontinuities and may constitute efficient data structures because the density of the triangles can be varied to match the roughness of the terrain" (Moore et al. 1991). Kunler (1994) noted that the method has the advantage of getting rid of the excess storage space incurred with respect to the topological relations. *Vector or contour-based method* is based on the stream path analogy first proposed by Onstad and Brakensiek (1968). It "consist of digitized contour lines and is stored as digital line graphs in the form of x, y coordinate pairs along each contour line of specified elevation. These can be used to subdivide an area into irregular polygons bounded by adjacent contour lines and adjacent streamlines" (Moore et al. 1991). The main disadvantages of the method include its requirement for an order of magnitude, large space for data storage, and computationally inefficient. In addition, the model has been criticized for over specifying the elevation surface at contour intervals and under specifying areas falling between the intervals. Digital contours are essentially verbatim copies of the paper world of pen-based cartography; for terrain modeling purposes, other digital models should be more effective at characterizing terrain (Moore et al. 1991). However, the vector or contour-based method is still in use in many countries because contour line data has a wider coverage and in different scales, thus representing an inexpensive data source (Ardiansyah and Yokoyama, 2002).

2.4 DEM Uncertainty and Sources

The process of DEM production is a type of abstraction, and contains uncertainty. As such, DEMs does not perfectly match the real-world terrain. The precise degree of this mismatch at every point is unknown, giving rise to uncertainty about the relationship between data and actual terrain (Shortridge, 2001). Therefore, caution must be taken when using DEM in spatial analysis for decision-making.

In any of the DEMs described in Section 2.3, Shortridge (2001) reported that two general sources of uncertainty may be specified: data-based uncertainty, due to the inaccuracy of measured elevations, and data model-based uncertainty, due to differences between the structural characteristics of the model and the landscape.

Data-based uncertainty is the difference between the elevation of a location specified in the data set and the true elevation at that location. In theory, this difference can be measured by ground survey. DEM files produced by the USGS and other mapping agencies typically are assigned a root mean square error (RMSE) from a (usually small) set of locations for which the true elevation is known (USGS, 1995). The true elevations at these locations are compared with the DEM estimate; RMSE is derived from the sum of the squared differences. Agencies engaged in DEM production use this kind of measure for reporting their data quality, and the DEM accuracy literature describes this approach in detail (Shearer, 1990; Bolstad and Stowe, 1994). Global measures like RMSE are inadequate for analysis of uncertainty since they provide no information about the local spatial structure.

Uncertainty arising from surface characterization depends very much upon the data model being used (Goodchild, 1992). Two of the terrain data models described earlier does not explicitly represent a continuous terrain surface. In an array of points, no assumption is made about the elevation of intermediate locations. Digitized contours exhaustively capture all elevations at the contour intervals, but do not specify elevations falling between contour intervals. Strictly speaking, several uncertainties exist regarding the elevation of any location not specified in either of these models. In practice, assumptions about intermediate values are frequently made, since it is usual and often critical to model terrain as a continuous surface, with an elevation specified at every point. Typically, contours and arrays of points are converted to data models that are used for surfaces, with defined values at every location, like rasters and TINs.

A raster DEM assigns a single elevation to every location within a cell. There is often uncertainty about what this elevation represents. Is it the elevation of the center of the cell? Is it the height of the lower left corner of the cell? Is it the mean elevation value for the area within the cell? Surface elevation is discontinuous at the cell boundaries. In contrast, the elevation surface is continuous across a TIN, though the slope surface is not continuous at facet edges if TIN facets are planar. Although elevations are specified for every location in these models, discrepancies can certainly arise between the real-world terrain and the structural characteristics of the model. The elevations at all vertices of a TIN could conceivably be without error, yet the facets fail to capture actual terrain characteristics. Similarly, elevations for all cells in a raster DEM could correctly characterize the mean cell elevation, but the fidelity of the model surface (flat-topped squares, like stacks of blocks) to the real world surface is very poor.

For topography-based models, that is, models characterizing and detecting topographic form, or models simulating processes that act upon this topography, DEMs are a potential source of uncertainty. DEMs consist of measured or digitized elevation values, and as such are subject to any error in the data capturing process. Widespread DEMs such as GLOBE or SRTM are distributed with accuracy figures that only give global measures such as RMSE lacking any information on the spatial distribution of error. Where uncertainty from DEM accuracy has to be modeled to assess its impact on the results of associated topographic models, assumptions have to be made about the spatial distribution of uncertainty. Several studies have shown that these assumptions influence the impact of uncertainty in spatial data models. Besides DEM accuracy, a number of factors in handling DEM data introduce additional uncertainty. These factors include the choice of data model, processing such as projecting and resampling of a DEM data, and algorithms used to extract and process elevation based information.

Generally, the interaction of several factors and processes influence the amount of solar radiation that falls on any location on the Earth. For instance, topography, ground albedo, and forest canopy must be taken into account when studying the spatial distribution of solar radiation at local level. Furthermore, terrain affects the amount of direct and diffuse radiations as a result of shadowing effects. Thus, as terrain complexity increases the diffuse and reflected components of solar radiation becomes vital (Ruiz-Arias et al. 2009).

2.5 DEM Uncertainty Metrics

DEM altitudes are given as interval data, and that statistical methods, for example, mean error (ME), mean absolute error (MAE), and root mean square error (RMSE) which is the square root of the variance of a distribution are used as an uncertainty metrics (Caruso, 1987; Kumler, 1994; Fisher, 1998; Weng, 2002; Carlisle; 2005; Gonga-Saholiariliva et al. 2011). The DEM uncertainty metrics are given by Gonga-Saholiariliva et al. (2011):

(a) Root mean square error and mean error

$$RMSE(z) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(z_{dem} - z_{ref}\right)^2}, \tag{2.25}$$

$$ME = \frac{1}{n}\sum_{i=1}^{n}\left(z_{dem} - z_{ref}\right), \tag{2.26}$$

where

z_{dem} represents DEM elevation,

z_{ref} denotes reference elevation,

n is the number of test points.

(b) Residual maps and standard deviation algorithms - pixel by pixel algebraic subtraction of elevation values provides a series of residual values *(Re)* that can be mapped:

$$\mathrm{Re}_{(i,j)} = z_{(i,j)dem} - z_{(i,j)ref}, \tag{2.27}$$

where

i and j are lines and columns of the grid, and are the elevations of the DEM and the reference grid, respectively.

The standard deviation of Re is calculated as follows:

$$Std = \sqrt{\frac{\sum_{i=1}^{n}\left[z_{dem} - z_{ref} - ME\right]^2}{n-1}}, \tag{2.28}$$

where

ME is the mean error of the DEM.

"Standard deviations were calculated each time for the entire image, cell by cell, yielding a map of *Re* values in which deviations in elevation values from those of the reference model could be located" (Gonga-Saholiariliva et al. 2011).

The advantage of these measures is that they summarize elevation errors in a DEM as a single value which is relatively quick to calculate and easy to report and makes comparisons of DEMs (Carlisle, 2005). However, the major limitation of single global value per DEM grid is that the finer-scale spatial variability of the error is ignored (Wechsler, 2003), which contains a great deal of useful information for terrain-based applications, and as such it does not allow identification of areas where error is the largest (Kyriakidis et al. 1999). Global indices are thus insufficient (Li and Wong, 2010) and any useful study of DEM accuracy should also investigate the spatial variation of error values (Carlisle, 2005). Burrough and McDonnell (1998) state that a single RMSE value implies that the error is uniform across the DEM. Whereas, many empirical studies show that DEM errors are spatially correlated and heteroscedastic (Hunter and Goodchild, 1997; Fisher, 1998; Kyriakidis et al. 1999; Carlisle, 2005).

Fisher (1999) calculated the probability of any DEMs grid cell being visible from the viewpoint as given in Equation 2.29:

$$p(X_{ij}) = \frac{\sum_{k=1}^{n} x_{ijk}}{n}, \tag{2.29}$$

where

$p(X_{ij})$ is the probability of a cell at row *i* and column *j* in the grid raster DEM is visible;

x_{ijk} is the value at the grid cell of the binary - coded viewshed in realization *k* (1 for "true" and 0 for "false"), such that *k* takes values 1 to *n* (the number of simulations).

The image produced in such a way consists of the pixels with $p(X_{ij})$ values, and it is known as probability map. It is used as a measure of "attribute" accuracy. Namely, for all thematic displays of some phenomena where quantitative parameters are used, probability map is the best indicator of reliability of such analysis.

CHAPTER 3

UNCERTAINTY ANALYSIS IN GEOGRAPHIC INFORMATION SYSTEM (GIS)

This chapter begins by discussing the concept of uncertainty, classification, types, and techniques used in modeling uncertainty in the field of GIS from Section 3.1 to 3.4. Then the Bayesian inference and *Markov chain Monte Carlo (MCMC)* simulation are given in Section 3.5. The remainder of the Sections deals with Markov chain theory and general properties of Markov chains, *Monte Carlo (MC)* simulation, *Metropolis-Hastings (MH) algorithm*, *Gibbs sampler*, Monitoring convergence of *MCMC* simulation, and spatial autocorrelation.

3.1 The Uncertainty Analysis in GIS

The term *uncertainty* is used to refer to the differences between the information provided by a spatial database and the corresponding information that would be available to someone able to observe and measure the real world directly. It includes the effects of errors created during the creation of database, and information loss that occurs during generalization (Hunter et al. 1995). Uncertainty in spatial data denotes "the lack of knowledge of the true value or the value that would be discovered if one were to visit the field and make an observation using a perfectly accurate instrument" (Hunter and Goodchild, 1997). In general, the term uncertainty refers to the unavoidable inaccuracies, inexactness, or inadequacies which are commonly found in spatial data sets and their resultant propagation through analyses (Brimicombe, 2010).

According to Heuvelink's (1998) formulation, uncertainty can be mathematically expressed as follows:

$$U(x) = g\left(A_1\left(b_1(x)+V(x)\right),\ldots,A_m\left(b_m(x)+V_m(x)\right)\right), \tag{3.1}$$

where

x = location of the grid cell,
$x \in D$ (domain of interest),
$U(\cdot)$ = output map containing all $x \in D, g(\cdot) =$ a GIS operation,
$A(\cdot)$ = an input map containing all $x \in D, b(\cdot) =$ value of x in the input map,
$V(\cdot)$ = a random field representing error or uncertainty.

The operation $g(\cdot)$ can be for instance, calculation of solar radiation, a standard filter operation to compute slope, aspect from a DEM, etc. Error propagation is aimed at estimating the amount of error in the output $U(\cdot)$, given the operation $g(\cdot)$ and errors in the random input attributes $A_i(\cdot)$. The output map $A_i(\cdot)$ is also a random field, with mean $\mu(\cdot)$ and variance $\tau^2(\cdot)$. From the view point of error propagation, the main interest is in uncertainty of $U(\cdot)$, as contained in its variance $\tau^2(\cdot)$.

However, it is important to note that uncertainty propagation problem is relatively easy if $g(.)$ is a linear function meaning that the mean and variance of $U(.)$ can be directly and analytically derived. The theory on functions of random variables also provides several analytical approaches to the problem for nonlinear $g(.)$, but few of these can be resolved by simple calculation (Abbaspour et al. 2003). Therefore, analytical methods are not very suitable because they are complex and not generally applicable and practically feasible (Heuvelink et al. 1989).

3.2 Types of Uncertainty in GIS

In the past decades, several studies attempted classification of uncertainty in spatial data (e.g., Burrough, 1986; Goodchild et al. 1992; Hunter, 1993; Burrough and McDonnell, 1998; Heuvelink, 1998; Veregin, 1999). Others used classification to investigate errors in spatial databases. For example, Burrough (1986) has categorized GIS errors into three (3) different classes, (1) error sources due to age and aerial coverage of the data, (2) errors due to original measurement or natural variations, and (3) errors due to complex computer processing, for instance, approximation (Abbaspour et al. 2003). On the other hand, Alai (1993) noted that GIS functionality is also used as basis to classify errors. However, according to Abbaspour et al. (2003) the most commonly used error taxonomy is embedded in the US spatial data transfer standard (SDTS) (NIST, 1992). Therefore, it can be understood that there is wide variation regarding the classification of uncertainties. Therefore, this study adopts the Alai's (1993) three level taxonomy (Figure 3.1).

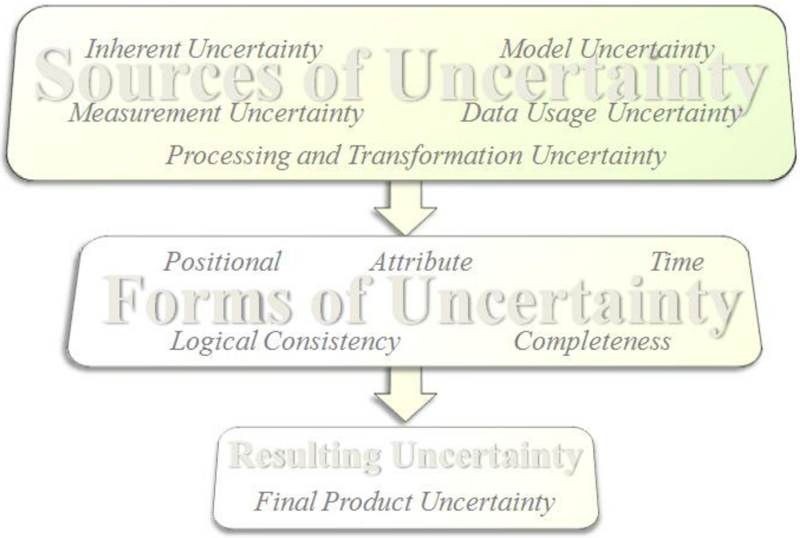


Figure 3.1. Taxonomy of uncertainty in GIS (Alai, 1993)

As indicated in Figure 3.1, the first level of the taxonomy deals with sources of uncertainty which have been classified into five categories: (1) the inherent uncertainty of the phenomena being mapped, (2) measurement uncertainty of spatial phenomena due to accuracy limitations of all observations, (3) model uncertainty which arises due to the models that are used to communicate the measurements, (4) processing and transformation uncertainty which refers to the secondary uncertainty caused during computer manipulation of data, following the data measurement, and (5) data usage uncertainty which has only recently received attention among researchers which is concerned with the manner in which spatial data are used. The second level of the taxonomy also classifies the forms of uncertainty into five aspects which are positional, attribute, time, logical consistency and completeness. This classification is mainly based on the SDTS. The third level refers to resulting uncertainty. The separation of the final product uncertainty from the forms of uncertainty is done because of the manner in which they may occur in the product (Abbaspour et al. 2003).

3.3 Uncertainties in the Models

According to Heuvelink et al. (1989), generally, there are three main areas where uncertainty affects a model. These include model input data, model parameters, and interaction within the model. Uncertainty is inherent in the model input data. It can arise from measurement error, local variation in the data or outdated observations. Each of these types of uncertainties has one common issue; they question the possibility of obtaining the same information suggested by the data for any given location. Uncertainties due to model parameters involve two important aspects. First, some uncertainties have to be random as they do not present a complete understanding of the phenomena or does not reflect the phenomenon as it is complex. Second, uncertainty in model parameters comes from the values assigned to the model parameters. Uncertainty due to interaction within the model arises due to the inherent error imposed by the GIS processes. When two or more datasets are combined in an operation, the error associated with the input datasets is compounded.

3.4 Modeling Uncertainty in GIS

Uncertainty information in GIS analysis is vital for effective decision-making that relies on geospatial data (UCGIS, 1998). However, most of the current GIS's do not provide this information (Hwang et al. 1998). Uncertainty propagation in GISs has been an active area of research for decades (see Heuvelink, 1998; Goodchild and Gopal, 1989; Hunter, 1999; Shortridge, 2001; Heuvelink and Burrough, 2002; Sheng et al. 2009; Fisher and Tate, 2006) and this has led to the development of several techniques which can be categorized into *analytical error models* and *stochastic simulation.* These techniques are therefore reviewed in this section.

3.4.1 Analytical Methods

The analytical error propagation methods use an explicit mathematical model to describe the mechanisms of error propagation for a particular multi-criteria decision rule (Eastman et al. 1993). The basis of the analytical method is well established in the general law of propagation of variances (Mikhail and Ackermann, 1976). However, analytical methods are much less attractive because they are cumbersome and require use of simplifying approximations (Heuvelink et al. 1989). It has been concluded that analytical approaches are suitable for situations where the sources of uncertainty are not significantly correlated and when the terrain analysis is sufficiently simple to allow arithmetic operations (Zhang and Goodchild, 2002). On the other hand, the analytical method has been considered unsuitable for decision-based local derivatives (Raaflaub and Collins, 2006), such as the D8 algorithm for determination of flow directions (O'Callaghan and Mark, 1984). However, the most frequently cited analytical model of error was developed by Hunter and Goodchild (1995). This basic model was founded on simple probability theory and the root mean square error (RMSE) of the DEM. The model assumes that any pixel follows a Gaussian distribution around the measured elevation value, and that DEM's RMSE is considered as the variance of the local error. Using this formulation, probabilities per-pixels are computed and mapped across a DEM with respect to any specific elevation value or contour line. The algorithm is incorporated in IDRISI GIS under Pclass operation and utilized by Eastman et al. (1993) to evaluate flooding due to sea-level rise. In addition, Huss and Pumar (1997) used the method to compute probability of visible areas. Another major contribution to the analytical error propagation analysis of the DEM was published by Florinsky (1998), in which he derived a number of variance equations for different calculation methods for slope, aspect and curvature. The only weakness of the study was that the DEM error was assumed to be uncorrelated. Recently, Zhou and Liu (2004) examined the accuracy of slope and aspect algorithms by focusing on their accuracy on artificial surfaces.

3.4.2 Stochastic Simulation Methods

Simulation is generally referred as the process of replicating reality using a model (Borisov et al. 2009), and appropriately applied to problems that are too difficult to solve analytically (Dagpunar, 2007). On the other hand, stochastic simulation is a method that "attempt to mimic or replicate the behavior of a system by exploiting randomness to obtain statistical sample of possible outcomes" (Heath, 2002). Such methods according to Heath (2002) and Liu (2009) are useful for studying, (1) non-deterministic (stochastic) processes (2) deterministic systems that are too complex to model analytically, and (3) deterministic problems whose high dimensionality make standard discretizations infeasible.

Previous studies have shown that stochastic models are very useful and advantageous tools for assessing uncertainty. Due to the availability of fast and inexpensive computational power, the best approach is to model a real phenomenon as faithfully as possible, and then rely on a simulation study to analyze it (Prodan and Prodan, 2001). A stochastic simulation method of representing uncertainty generates likely answers, from which a "good" answer is selected based on certain criteria. In other words, stochastic simulation is based on generation of equally probable realizations from which some specific statistics such as mean, standard deviation, coefficient of variation (CoV) are computed and then evaluated. Though, generation of the "real" map is not certain from the simulation process, it provides a distribution of results from which one can safely state that the "true" map lies (Wechsler, 1999; 2006). In the past decades, the stochastic approaches used in analyzing uncertainty in GIS-based studies include: *First and second order Taylor Series methods, Rosenblueth's method, MC simulation, and geostatistical simulation.*

Heuvelink et al. (1989) applied a *Taylor series expansion* to derive a model of errors in spatial data sets that are propagated by map overlay operations. Specific values and standard deviations associated with input values were used to obtain errors and to generate a map that displayed how model results and errors are distributed over space. Others utilized the *MC* simulation approach in analyzing spatial data uncertainty. For example, Hunter and Goodchild (1995); Heuvelink (1998); Holmes et al. (2000); Oksanen and Sarjakoski (2005b);

Wechsler and Kroll (2006); and Raaflaub and Collins (2006) studied uncertainty propagation from DEMs in the derived topographic parameters like slope and aspect. Liu and Bian (2008) utilized *MC* simulation and analyzed the impact of spatial autocorrelation on the accuracy of four different slope algorithms. Lee (1992) and Lee et al. (1996) studied the impact of DEM errors on hydrology features through simulation approach and concluded that hydrology features are tremendously affected by even minor amount of the DEM errors. Liu (1994) applied the *MC* technique to simulate errors in DEM and then evaluated uncertainty in a forest harvesting model. Heuvelink (1998) analyzed accuracy of two (2) different slope algorithms with *MC* simulation. Hunter and Goodchild (1997) evaluated the impact of spatially simulated elevation errors according to various degrees of spatial autocorrelation on the computations of slope and aspect. Wang et al. (2006) investigated the impacts of DEM uncertainty in the simulated outputs of TOPMODEL, a semi-distributed hydrological model using *MC* technique. Hebeler and Purves (2009) utilized *MC* simulation to generate uncertainty maps to indicate the impact of DEM uncertainty on an ice sheet model (ISM). Fisher (1991) examines the effects of DEM uncertainty propagation on viewshed using *MC* simulation. Biesemans et al. (2000) used an *MC* simulation technique to determine the uncertainty propagation of the calculated on-site soil losses and off-side sediment accumulation in Belgium and concluded that the technique explained the difference between the field observation and model output mainly due to the uncertainty of the model input parameters. Holmes et al. (2000) used high accuracy GPS data in sequential Gaussian simulation to provide an *MC* framework for quantifying the effect of measured USGS 30 m DEM error on digital terrain modeling. Abbaspour et al. (2003) and Jahanpeyma et al. (2007) adopted *MC* techniques and assessed uncertainty propagation in overlay analysis.

On the other hand, the geostatistical simulation is a spatial extension of the concept of *MC* simulation. In addition to reproducing the data histogram, geostatistical simulations also honor the spatial variability of data, usually characterized by a variogram model. If the simulations honor the data themselves, they are said to be 'conditional simulations' (Vann et al. 2002). Conditional simulation seeks to draw from the local conditional cumulative density function values which honor the survey data and a predefined variogram. Atkinson (1999) reported that among the different approaches of conditional simulation, the indicator simulation (IS) is the most preferred approach and that *MCMC* techniques is the most promising one (Oliver et al. 1997) because; (1) it offers an improved flexibility in simulating from complex probability density functions (pdf), even when the pdf is explicitly defined (Shekhar and Xiong, 2008), (2) the *MCMC* method can cope with model nonlinearity and non-Gaussianity of the parameter uncertainty (Minasny et al. 2011), and (3) it generates random samples based on predefined posterior distributions through the construction of Markov chain according to the desired properties.

3.5 Bayesian Inference and MCMC Simulation

"Bayesian inference is a probabilistic inferential method. In the last two decades, it has become more popular than ever due to affordable computing power and recent advances in MCMC methods for approximating high dimensional integrals. Bayesian inference can be traced back to Thomas Bayes (1764), who derived the inverse probability of the success probability θ in a sequence of independent Bernoulli trials, where θ was taken from the uniform distribution on the unit interval (0, 1) but treated as unobserved" (Liang et al. 2010).

Bayesian inference for spatial data often requires evaluation of a highly multivariate posterior density $\pi(x|y)$. Typically, the dimension of the parameter space is between 50 and 10^6, and often $\pi(x|y)$ is known only up to scale. Integrations necessary to compute functionals of $\pi(x)$ such as posterior means or credible intervals are intractable through standard analytical or numerical integration techniques. To evaluate these otherwise intractable posterior quantities and *MCMC* (Higdon, 1994) is employed.

MCMC is a simulation technique that enables one to draw a large number of samples from a posterior distribution of interest. Inference about the posterior distribution is then based on quantities calculated from the collection of samples, for instance, the median or percentiles. The feature that makes this distinct from the *MC* simulation is that under *MCMC*, successive samples are selected by moving through the parameter space according to a Markov chain rather than randomly selecting them. This is practical, much more efficient than random sampling and works better because it is relatively straightforward to construct a Markov chain such that the samples are drawn approximately from any posterior distribution of interest (Henderson and Bui, 2005).

3.6 Markov chain Theory

According to Wendy et al. (2007) Markov chain $\{X^t\}$, named after Andrey Markov, is seen as a stochastic process whereby any given state in the series $\{X^t\}$, dependent on the earlier state of the chain, $X^{(t-1)}$, thus, it is a sequence of dependent (correlated) random variables.

$$X^{(0)}, X^{(1)}, X^{(2)}, \ldots, X^{(t)}, \ldots,$$

such that the probability distribution of $X^{(t)}$ given the past variables depends only on $X^{(t-1)}$. This conditional probability distribution is called transition kernel or Markov Kernel *K*; that is,

$$X^{(t+1)} \Big| X^{(0)}, X^{(1)}, X^{(2)}, \ldots, X^{(t)} \sim K\left(X^{(t)}, X^{(t+1)}\right).$$

For example, a simple random walk Markov chain satisfies

$$X^{(t+1)} = X^{(t)} + \in_t,$$

where

$\epsilon_t \sim \mathcal{N}(0,1)$, independently of$X^{(t)}$; therefore, the Markov kernel $K(X^{(t)}, X^{(t+1)})$ corresponds to $\mathcal{N}(X^{(t)}, 1)$ density.

Robert and Casella (2010) noted that, "the Markov chains encountered in MCMC settings enjoy a very strong stability property. Indeed, a stationary probability distribution exists by construction of those chains; that is, there exists a probability distribution f such that if $X^{(t)} \sim f$, then $X^{(t+1)} \sim f$. Therefore, formally, the kernel and stationary distribution satisfy the Equation (3.2)":

$$\int_x K(x,y)\, f(x)\,dx = f(y). \tag{3.2}$$

"The existence of a stationary distribution imposes a preliminary constraint on K called irreducibility. In other words, the kernel K allows for free moves all over the state-space no matter the starting value of $X^{(0)}$, the sequence $\{X^t\}$ has a positive probability of eventually reaching any region of the state-space. (A sufficient condition is that $K(x,y) > 0$ everywhere). The existence of a stationary distribution has major effects on the behavior of the chain $\{X^t\}$. For instance, most of the chains involved in *MCMC* algorithms are recurrent, that is, they will return to any arbitrary non-negligible set in infinite number of times" (Robert and Casella, 2010).

"In the case of recurrent chains, the stationary distribution is also a limiting distribution in the sense that the limiting distribution of $X^{(t)}$ is f for almost any initial value $X^{(0)}$. This property is also called ergodicity, and it obviously has major consequences from a simulation point of view, if a given kernel K produces an ergodic Markov chain with stationary distribution f, generating a chain from this kernel, K will eventually produce simulations from f. In particular, for integrable function h, the standard average" (Robert and Casella, 2010), as given in Equation (3.3):

$$\frac{1}{T}\sum_{t=1}^{T} h\left(X^t\right) \to \mathbb{E}_f\left[h\left(X\right)\right], \tag{3.3}$$

which means that the Law of Large Numbers that lies at the basis of MC methods can also be applied in MCMC settings. It is then sometimes called the Ergodic Theorem.

3.6.1 General Properties of Markov chains

There are several properties of Markov chains that are especially important when describing convergence. These include homogeneous, recurrent, absorbing, irreducible, stationary distribution, periodicity and ergodic. Generally, the properties have intimidating names that are inherited from mathematical Markov chain theory, but in reality are fairly straightforward ideas. Generally, if one can describe the mathematical status of a particular chain, then one can often determine if it is producing useful sample from the target distribution of interest. The properties are only summarized briefly here but for more technical details refer to Nummelin (1984), Gamerman (1997); Gilks et al. (1996); Robert and Casella (1999, 2004).

The **homogeneous** property of a Markov chain is observed when the transition probabilities of a chain say at step *n* depend on the value of *n*. For instance, the samplers that we use/pick are known to obtain the property (Monogan, 2010). One reason that the *MH algorithm* and the Gibbs sampler (both given in detailed in subsections 3.6.4 and 3.6.5) dominate *MCMC* implementations is that the chains they define eventually obtain this property.

A Markov chain is said to be **recurrent** with respect to a particular state, A. The state can either be a single point or set of points (necessary especially for continuous case), if the probability that the chain remains in A indefinitely often over time is one (Monogan, 2010; Robert and Casella, 1999). This can also be restated as saying that if the chain is currently in

A, it will eventually return to A with probability one. If the mean time of a chain returning to A is bounded, then it is a positive recurrent; otherwise referred to as null recurrent (Monogan, 2010). Recurrence is a desirable property in Markov chains and there are also stricter forms such as Harris recurrence which stipulates the same condition for every possible starting value. See Robert and Casella (1999) Chapter 4, for details.

There are also properties directly associated with states. A state (say A) is said to be **absorbing** if a chain cannot leave the moment it enters that state (Monogan, 2010).zzz This is represented as: $p(A, A^c) = 0$. A state is transient if given that a chain currently occupies a state A, the probability of not returning to A is non-zero. A state, $B: p(A, B) = 0$.

A Markov chain is said to be **irreducible** if each specific point or set of points (a subspace, required for the continuous case), A, can be reached from any other point or set of points (Monogan, 2010). That is, $p(\theta_i, \theta_j) \neq 0, \ \forall i, j \in A$. Notice that irreducibility is a characteristic of both the chain and the subspace. This relationship is expressed as follows: if a subspace is closed, finite and irreducible, then all states within this subspace are recurrent.

If we take a set of recurrent states (they must be non-empty, and bounded or countable), then their union creates a new state which is closed and irreducible (Meyn and Tweedie, 1993). This means that Markov chain assures the exploration of all the subspace. This presents an important outcome especially whenever a chain wanders into a closed, irreducible set of recurrent states then remain there and visits each and every single state (eventually) with probability one (Monogan, 2010).

The Markov chain **stationary distribution** is defined as $\pi(\theta)$ for θ on the state space A. We denote $p(\theta_i, \theta_j)$ to indicate the probability that the chain will move from θ_i to θ_j at some arbitrary step t (Monogan, 2010) from the transition kernel, and $\pi^t(\theta)$ as the marginal distribution. This stationary distribution is then defined as satisfying:

$$\sum\nolimits_{\theta_i} \pi^t(\theta_i)\, p(\theta_i, \theta_j) = \pi^{t+1}(\theta_j) \qquad \text{(Discrete Case)},$$

$$\int \pi^t(\theta_i)\, p(\theta_i, \theta_j)\, d\theta_i = \pi^{t+1}(\theta_j) \qquad \text{(Continuous Case)}. \tag{3.4}$$

Therefore, multiplication by transition kernel and evaluating for the current point (the summation step for discrete sample spaces and the integration step for continuous sample spaces) produces the same marginal distribution: $\pi = \pi p$ in shorthand. This demonstrates that "the marginal distribution remains fixed when the chain reaches the stationary distribution and we might as well drop the subscript designation for iteration number and just use $\pi(\theta)$" (Monogan, 2010).

Once the chain reaches its **stationary distribution** (also called its invariant distribution, equilibrium distribution, or limiting distribution if discussed in asymptotic sense), it remains in this distribution and moves around (or “mixes”) throughout the subspace forever based on marginal distribution, $\pi(\theta)$ (Monogan, 2010). This is precisely what we want and expect from MCMC. If we can set up the Markov chain such that it reaches a stationary distribution that is the desired posterior distribution from our Bayesian model, then all we need to do is let it wander about this subspace for a while producing empirical samples to be summarized.

The good news is that the two forms of *MCMC* kernels that we will use have the property that they are guaranteed to eventually reach a stationary distribution which is the desired posterior distribution.

It is also possible to define the **periodicity** of a Markov chain. The periodicity refers to the length of time required to repeat chain values that are identical. It is desirable to have an aperiodic chain, that is, where the only length of time for which the chain repeats some cycle of values is the trivial case with cycle length equal to one. Why? It seems as though we would not necessarily care if there was some period to the chain values, particularly if the period was quite long or perhaps in the discrete state if it included every value in the state space. The result is that the recurrence property alone is not enough to ensure that the chain reaches a state where the marginal distribution remains fixed and identical to the posterior of interest.

If a chain is irreducible, positive recurrent, and aperiodic, then it is called **ergodic**. The ergodic Markov chains have property:

$$p^n_{n\to\infty}\lim\left(\theta_i,\theta_j\right)=\pi\left(\theta\right),$$

for all θ_i, and θ_j in the subspace (Nummelin, 1984). Therefore, in the limit the marginal distribution at one step is identical to the marginal distribution at all other steps. However, because of the recurrence requirement, this limiting distribution is now closed and irreducible meaning that the chain will never leave it and is guaranteed to visit every point in the subspace. "Once a specified chain is determined to have reached its ergodic state, sample values behave as if they were produced by the posterior of interest from the model" (Monogan, 2010).

Monogan (2010) noted that though ergodic theorem is related to Markov chains but similar to the law of large numbers. Any defined posterior distribution function can be estimated from Markov chain samples that is in its ergodic state since mean sample values strongly generate consistent parameter estimates. More formerly, suppose $\theta_{i+1},\ldots,\theta_{i+n}$ are n (not necessarily consecutive) values from a Markov chain that has reached its ergodic distribution, a statistic of interest, $h(\theta)$, is computed empirically:

$$\hat{h}\left(\theta_i\right)=\frac{1}{n}\sum_{i=i+1}^{i+n}h\left(\theta_i\right)\approx h\left(\theta\right), \tag{3.5}$$

and for finite quantities this converges almost surely: $p\left[\hat{h}(\theta_i)\to h(\theta), as\ n\to\infty\right]=1$ (Roberts and Smith, 1994).

"The remarkable result from ergodicity is that even though Markov chain values, by their very definition, have serial dependence, the mean of the chain values provides a strong consistent estimate of the true parameter" (Monogan, 2010). Furthermore, provided that the limiting variance of the empirical estimator $\hat{h}(\theta_i)$ is bounded, then subject to very general regularity conditions the central limit theorem also applies:

$$\sqrt{n}\,\frac{\hat{h}(\theta_i)-h(\theta)}{\sqrt{VAR\left(\hat{h}(\theta_i)\right)}}\xrightarrow[n\to\infty]{}N(0,1). \tag{3.6}$$

There is also the notion of geometric ergodicity: the geometric rate of reduction in time for the total variation distance between any arbitrarily time point and the convergence of the limiting distribution, see Mengerson and Tweedie (1996) for details.

3.7 Monte Carlo (MC) Simulation

MC simulations are numerical methods that use random numbers to compute quantities of interest through the creation of a random variable whose expected value is the desired quantity. The simulated random variables are then tabulated and used with sample mean and variance to construct probabilistic estimates (SFITZ, 2006). The *MC* techniques are used to assess complex stochastic systems that are too complex to be understood or controlled using analytical or numerical methods. Examples of such systems include weather and climate systems, telecommunications networks and financial markets. It is also widely used in GIS to analyze uncertainty (Wechsler and Kroll, 2006; Heuvelink, 1998; Hunter and Goodchild, 1995; Holmes et al. 2000; Oksanen and Sarjakoski, 2005b; Raaflaub and Collins, 2006).

The technique is an alternative uncertainty assessment method that involves re-running an analysis several times. Each time the analysis is re-run, the variables which are subject to uncertainty (stochastic variables), are perturbed, or altered somewhat, according to some underlying assumption (usually, a probability distribution function). Each alternative, equally probable result is termed as realization of the *MC* process. Repeatedly perturbing the values, then running the analysis for each set of perturbed values, produces a large set of equally likely alternative values for the outcome of the analysis (Mowrer, 1997). Using the results from all the realizations, it is possible to calculate a mean and standard deviation that correspond to the algorithm under analysis. The mean is considered as the best estimate that the algorithm would produce, while the standard deviation is considered as error (Heuvelink, 1998). *MC* simulation generally, involves two steps (Hammersley and Handscomb, 1979; Heuvelink, 1998).

(1) For each MC run s, $s = 1 \ldots \ldots k$ (below, lower-case letters represent realizations of random variables given as upper-case letters in Equation (3.1):
   a. generate realizations $a_1(.), \ldots, a_m(.)$ of each stochastic model parameter, and realizations $i_{1\ldots m}$ for each stochastic input variable.
   b. with these realizations, run $g(\cdot)$ (Equation 3.1), and store the realizations of the model output variables $u_{1\ldots\ldots n}$ in which the interest lies, in the case of a dynamic model for all time steps.

(2) Compute sample statistics (for example, mean, variance, skewness) from the k model outcomes, for each model variable $1 \ldots n$. This approach needs a stochastic description of model input variables with their associated errors, and a methodology to draw realizations from these stochastic model inputs, needed in (*step 1a*) of the MC procedure. "A representative sample from the joint distribution of uncertain inputs and model parameters, structure, and solution can be obtained using an appropriate pseudorandom number generator and a sufficiently large sample size. The accuracy of the *MC* method is inversely proportional to the square root of the

number of runs N and, therefore, increases gradually with N" (Robert and Casella, 2010).

However, it is important to note that the standard *MC* simulation technique produces a set of independent simulated values according to some desired probability distribution, whereas, an *MCMC* method generate chains in which each of the simulated values is mildly dependent on the preceding value.

3.8 Metropolis Hastings (MH) Algorithm

The *MH algorithm* which is the most popular example of an *MCMC* method was developed by Metropolis et al. (1953) and later generalized by Hastings (1970). The algorithm has been used extensively in physics and later statistics. For instance, spatial statisticians, inspired by Hastings (1970) and Hammersley and Clifford (1971), began experimenting with *MCMC* in the study of lattice systems and spatial point processes, both of which could be simulated via discrete or continuous time Markov chains. In the early 1980's, Donald and Stuart Geman forged a link between *MCMC* and digital image analysis, following earlier work of Ulf Grenander on general pattern theory and his maxim “Pattern analysis = Pattern synthesis" (Grenander, 1983). In particular, their seminal paper (Geman and Geman, 1984) adopts a Bayesian approach, with Markov random fields (for instance, Besag, 1974) as prior distributions, and either the *Metropolis algorithm* or the *Gibbs sampler* to synthesize the posterior distribution of image attributes; and it uses the closely related method of simulated annealing (Kirkpatrick et al. 1983) to determine an approximation to the most probable image. See Gelfand and Smith (1990); Gelfand et al. (1990); Chib and Greenberg (1995) for recent examples. The *MH algorithm* provides a general approach for generating a correlated sequence of draws from the target density that may be difficult to sample by a classical independence method. The earliest and general approach of *MCMC* method is the *random walk Metropolis algorithm* (Metropolis et al. 1953).

Based on Metropolis et al. (1953), the *Metropolis algorithm* constructs a Markov chain through the division of transition probabilities into, (1) proposal distribution, and (2) acceptance function. The proposal distribution is used when drawing the proposed state, whereas, the acceptance function is used in determining the probability with which the proposed state is accepted as the next state. "If the proposal is not accepted, then the current state is taken as the next state. This procedure defines a Markov chain, with the probability of the next state depending only on the current state, and it can be iterated until it converges to the stationary distribution. Through, careful choice of the proposal distribution and acceptance function, we can ensure that this stationary distribution is the target distribution" (Sandborn et al. 2010).

An adequate condition for defining both the proposal distribution and acceptance probability which generates the right stationary distribution is detailed balance as indicated in Equation (3.7):

$$\pi\left(x^*\right)q\left(x_n\left|x^*\right.\right)A\left(x_n;x^*\right)=\pi\left(x_n\right)q\left(x^*\left|x_n\right.\right)A\left(x^*;x_n\right), \tag{3.7}$$

where

$\pi(x^*)$ denotes the Markov chain (i.e., target distribution) stationary distribution,
$q(x_n|x^*)$ represents the probability of proposing a new state x_n given the current state x^*,
$A(x^*; x_n)$ is the acceptance probability of proposal state.

Intuitively, detailed balance ensures, that the probability with which a move from x^* to x_n is observed, is equal to the probability with which a move from x_n to x^* is observed, once the chain has reached the stationary distribution. The *Metropolis technique* defines an acceptance function that satisfies the requirements of the detailed balance for any symmetric proposal distribution, with $q(x_n|x^*) = q(x^*|x_n)$ for all x^* and x_n. Equation (3.8) shows the Metropolis acceptance function:

$$A\left(x_n; x^*\right) = \min\left(\frac{\pi\left(x_n\right)}{\pi\left(x^*\right)}, 1\right), \tag{3.8}$$

This means that a higher probability of the proposal states over current state is always accepted.

A general form of the acceptance function which satisfies detailed balance was proposed by Hastings (1970). He also extended the *Metropolis algorithm* by integrating asymmetric proposal distributions (Sandborn et al. 2010). The main observation is that Equation (3.7) is satisfied by any acceptance rule as shown in Equation (3.9):

$$A\left(x_n; x^*\right) = \frac{s\left(x_n, x^*\right)}{1 + \dfrac{\pi\left(x^*\right) q\left(x_n \middle| x^*\right)}{\pi\left(x_n\right) q\left(x^* \middle| x_n\right)}}, \tag{3.9}$$

where

$s(x_n, x^*)$ is a symmetric in x^* and x_n and $0 \leq A(x^*; x_n) \leq 1$ for all x^* and x_n.

Taking:

$$s\left(x_n, x^*\right) = \begin{cases} 1 + \dfrac{\pi\left(x^*\right) q\left(x_n \middle| x^*\right)}{\pi\left(x_n\right) q\left(x^* \middle| x_n\right)}, if \ \dfrac{\pi\left(x_n\right) q\left(x^* \middle| x_n\right)}{\pi\left(x^*\right) q\left(x_n \middle| x^*\right)} \geq 1, \\ 1 + \dfrac{\pi\left(x_n\right) q\left(x^* \middle| x_n\right)}{\pi\left(x^*\right) q\left(x_n \middle| x^*\right)}, if \ \dfrac{\pi\left(x_n\right) q\left(x^* \middle| x_n\right)}{\pi\left(x^*\right) q\left(x_n \middle| x^*\right)} \leq 1, \end{cases} \tag{3.10}$$

Equation (3.10) produce the Metropolis acceptance function for proposal distribution that is symmetric in nature, and generalized as in Equation (3.11):

$$A\left(x_n, x^*\right) = \min\left(\frac{\pi\left(x_n\right) q\left(x^* \middle| x_n\right)}{\pi\left(x^*\right) q\left(x_n \middle| x^*\right)}, 1\right), \tag{3.11}$$

for symmetric proposal distributions. With a symmetric proposal distribution, taking $s(x_n, x^*) = 1$ gives the Baker acceptance function:

$$A\left(x_n, x^*\right) = \frac{\pi\left(x_n\right)}{\pi\left(x_n\right) + \pi\left(x^*\right)}, \tag{3.12}$$

where the acceptance probability is proportional to the probability of the proposed and the current state under the target distribution (Barker, 1965; Sandborn et al. 2010). "The Metropolis acceptance function has been shown to result in lower asymptotic variance and faster convergence to the stationary distribution (Billera and Diaconis, 2001) than the Barker acceptance function" (Sandborn et al. 2010).

MCMC algorithm is a simple technique of generating samples based on probability distributions in which other methods are not feasible. "The basic procedure is to start a Markov chain at some initial state, chosen arbitrarily, and then apply one of the methods for generating transitions outlined above, proposing a change to the state of the Markov chain and then deciding whether or not to accept this change based on the probabilities of the different states under the target distribution. After allowing enough iterations for the Markov chain to converge to its stationary distribution (known as the "*burn-in*"), the states of the Markov chain can be used to answer questions about the target distribution in the same way as a set of samples from that distribution" (Sandborn et al. 2010).

3.9 Gibbs Sampler

The *Gibbs sampler* known as the heat bath algorithm is a special case of *MH algorithm* but uses a somewhat different methodology from the *MH algorithm* and is particularly useful for generating *n*-dimensional random vectors. The distinguishing feature of the Gibbs sampler is that the underlying Markov chain is constructed, in a deterministic or random fashion, from a sequence of conditional distributions (Rubinstein and Kroese, 2007). It is the most popular computational method for Bayesian inference and was in use in the statistical physics before the same method was used by Geman and Geman (1984) in image analysis to analyze Gibbs distributions on lattices. The paper by Gelfand and Smith (1990) demonstrated the usefulness of the *Gibbs sampler* for solving a wide - range of issues in Bayesian analysis and made the Gibbs sampler a popular computational tool for Bayesian computation.

According to Liang et al. (2010), technically, the *Gibbs sampler* can be viewed as a special method for overcoming the problem of dimensionality through conditioning. The basic idea is similar to the idea behind iterative conditional optimization methods. Suppose that we want to generate random numbers from the target density $f(x), x \in \chi \subseteq \mathbb{R}^d$. Partition the d −vector x into K blocks and write $x = (x_1, \ldots, x_\mathrm{K})'$, where $\mathrm{K} \leq d$ and $\dim(x_1) + \cdots + \dim(x_\mathrm{K}) = d$ with $\dim(x_\mathrm{K})$ representing the dimension of x_k. Denote by:

$$f_k\left(x_k | x_1, \ldots, x_{k-1}, \ldots, x_\mathrm{K}\right) \qquad (k = 1, \ldots, \mathrm{K}), \tag{3.13}$$

the corresponding full set of conditional distributions. Under mild conditions, this full set of conditionals, in turn, determines the target distribution $f(x)$; according to Hammersley-Clifford theorem (Besag, 1974; Gelman and Speed, 1993):

Theorem ... (Hammersley-Clifford) - If $f(x) > 0$ *for every* $x \in \mathbb{X}$, then the joint distribution $f(x)$ is uniquely determined by the full conditionals (3.5.5). More precisely,

$$f(x) = f(y) \prod^{k=1\mathrm{K}} \frac{f_{jk}\left(x_{jk} | x_{j1}, \ldots, x_{jk-1}, y_{jk+1}, \ldots, y_{j\mathrm{K}}\right)}{f_{jk}\left(y_{jk} | x_{j1}, \ldots, x_{jk-1}, y_{jk+1}, \ldots, y_{j\mathrm{K}}\right)} \qquad (x \in \chi). \tag{3.14}$$

For every permutation j on $\{1, \ldots, n\}$ and every $y \in \mathbb{X}$.

Algorithmically, the Gibbs sampler is an iterative sampling scheme. Starting with an arbitrary point $x^{(0)}$ in χ with the restriction that $f(x^{(0)}) > 0$, each iteration of the *Gibbs sampler* cycles through the full set of conditionals (Equation 3.13) to generate a random number from each $f_k(x_k | x_1, \ldots, x_{k-1}, \ldots, x_\mathrm{K})$ by setting $x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_\mathrm{K}$ at their most recently generated values.

The Gibbs sampler algorithm take $x^{(0)} = (x_1^{(0)}, \ldots, x_K^{(0)})$ from $f^{(0)}(x)$ with $f\left(x^{(0)}\right) > 0$, and iterate for $t = 1,2, \ldots$

1. Generate $x_1^{(t)} \sim f_1\left(x_1 \middle| x_2^{(t-1)}, \ldots, x_K^{(t-1)}\right)$.

$$\vdots$$

k. Generate $x_k^{t} \sim f_k\left(x_k \middle| x_1^{(t)}, \ldots, x_{k-1}^{(t)}, x_{K+1}^{(t-1)}, \ldots, x_K^{(t-1)}\right)$.

$$\vdots$$

K. Generate $x_K^{(t)} \sim f_K\left(x_K \middle| x_1^{(t)}, \ldots, x_{K-1}^{(t)}\right)$.

Under mild regularity conditions, the distribution of $x^{(t)} = \left(x_1^{(t)}, \ldots, x_K^{(t)}\right)'$, denoted by $f^{(t)}(x)$, will converge to $f(x)$.

3.10 Monitoring Convergence of MCMC Simulation

The convergence analysis refers to the statistical analysis of sampler output to show, either *a posteriori* or during run time, whether or not the chain converges during a particular sample run (Brooks, 1998). There are several approaches on *MCMC* simulation convergence analysis from the literature. These approaches may be classified as informal, ad hoc, more elaborate approach, and those that include additional information beyond the simulation draws themselves (Brooks, 1998).

The informal approach includes, the *thick pen technique* developed by Gelfand et al. (1990), and *quantile and autocorrelation plots* proposed by Gelfand and Smith (1990). Whereas, the ad hoc methods comprised of research works of Kimbler and Knight (1987) and Gafarain et al. (1978). Both informal and ad hoc methods are easily applied. The more elaborate methods include, *eigenvalue estimation, Gelman and Rubin's technique, spectral density* *estimation technique,* and *cumulative sum path plots*. The *eigenvalue estimation approach* estimates the convergence rate of the sampler through appropriate eigenvalues (e.g., Raftery and Lewis, 1992; Garren and Smith, 1995). The *Gelman and Rubin's technique* is based on a classical analysis of variance to estimate the advantage of running the chain continuously (Brooks, 1998). The method requires multiple-chain sampling from dispersed starting values and then compares the within- and between-chain variances. Brooks and Gelman (1997) further extended the approach. The *spectral density estimation method* is applied by Geweke (1992) and Heidelberger and Welch (1983) to perform hypothesis tests for stationarity. The *cumulative sum path plots* were utilized by Brooks (1996) for convergence evaluation.

The methods for convergence analysis that include additional information beyond the simulation draws themselves are, weighting-based methods (Ritter and Tanner, 1992; and Zellner and Min, 1995), *kernel-based techniques* (Liu et al. 1993; and Roberts, 1994). *Weighting-based methods* require a single run, monitor the ratio of the target density (up to a normalizing constant) and the current estimate of the target density; stability of the ratio indicates that the chain has converged. *Kernel-based techniques* estimate the L^2-distance between the *t*-step transition kernel and the stationary distribution. In addition, Yu (1994) and Brooks et al.(1997) developed an approach for computing the L^1-distances between relevant densities. The advantage of the methods is that they assess convergence of the full joint density. However, generally, the models proved to be computationally time demanding to implement and difficult to interpret (Brooks, 1998).

However, the most popular among these methods are Gelman and Rubin (1992) and Raftery and Lewis (1992) and this is due to the availability of computer programs that implement them (Cowles and Carlin, 1996). For this reason, an in depth review of one among them, specifically, Gelman and Rubin method is considered here.

3.10.1 Gelman and Rubin Convergence Diagnostic Criteria

The Gelman and Rubin's (1992) method focus on applied inference for Bayesian posterior distributions in real problem, which often tend toward normality after transformations and marginalization (Cowles and Carin, 1996). The approach comprises of two steps.

Step 1, involves creating an over dispersed estimate of the target distribution which is used to start several independent sequences. This can be done as follows:

(1) Identify the high-density areas of the target distribution of $\boldsymbol{x}$ and find the K modes.
(2) Approximate the high-density areas using generalized method of moments (GMM):

$$\overline{P}(\mathrm{x}) = \sum_{k=1}^{k} \omega_k (2\pi)^{-d/2} |\Sigma_k|^{-1/2} \exp\left(-\frac{1}{2} (\mathrm{x} - \mu_k)^t \sum_k^{-1} (\mathrm{x} - \mu_k) \right). \tag{3.15}$$

(3) Form an over dispersed distribution by first drawing from the GMM and then dividing each sample by a positive number, which results in a mixture t distribution:

$$\overline{P}(\mathrm{x}) \propto \sum_{k=1}^{k} \omega_k |\Sigma_k|^{-1/2} \left(\eta + (\mathrm{x} - \mu_k)^t \sum_k^{-1} (\mathrm{x} - \mu_k) \right)^{-(d+\eta)/2}. \tag{3.16}$$

(4) Sharpen the over dispersed approximation by down weighting areas which have a relatively low density using importance resampling for instance.

Step 2 on the other hand involves re-estimating the target distributions:

(1) Independently simulate *m* sequences of length $2n$ from the over dispersed distribution and discard the first *n* iterations.
(2) For each scalar parameter of interest, estimate the following quantity from the last *n* iterations of *m* sequences:
   - B: the variance between the means from *m* sequences;

$$B = \frac{n}{m-1}\sum_{j=1}^{m}\left(\bar{\theta}_j - \bar{\bar{\theta}}\right)^2, \tag{3.17}$$

   where
   $\bar{\bar{\theta}} = \frac{1}{n}\sum_{j=1}^{m}\bar{\theta}_j$. This is the variance of the chain means multiplied by n because each chain is based on n draws.
   - W: the average of the *m* within-sequence variances;

$$W = \frac{1}{m}\sum_{j=1}^{m} s_j^2, \tag{3.18}$$

   where
   $s_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(\theta_{ij} - \bar{\theta}_j\right)^2$, s_j^2 is just the formula for the variance of jth chain.
   W is then just the mean of the variance of each chain.
   - $\hat{\mu}$ denotes estimate of the target mean: mean of *mn* samples.
   - $\hat{\sigma}^2$ represents an estimate of target variance (unbiased):

$$\sigma^2 = \frac{n-1}{n}W + \frac{1}{n}B.$$

(3) Estimate the posterior of target distribution as a *t* distribution (considering variability of the estimates $\hat{\mu}$ and $\hat{\sigma}^2$) with center of $\hat{\mu}$ and scale of $\sqrt{\hat{V}} = \sqrt{\hat{\sigma}^2 + B/mn}$.

(4) Monitor the convergence by shrink factor$\sqrt{\hat{R}} = \sqrt{(\hat{V}/W)\, df/(df-2)}$, as it gets closer to 1 for all scalars, collect burn-out samples,

$$\sqrt{R} = \sqrt{\left(\frac{n-1}{n} + \frac{m+1}{mn}\frac{B}{W}\right)\frac{df}{df-2}}. \tag{3.19}$$

Although, *Gelman and Rubin's methodology* of convergence analysis was originally created for the Gibbs sampler, it is applicable to any *MCMC* algorithm. The method gives more emphasis on the reduction of biasness in estimation. Gelman and Rubin stated "that the "shrink factor" approaches 1 when the pooled within-chain variance dominates the between-chain variance to mean that at that point, all chains have escaped the influence of their starting points and have traversed all of the target distribution" (Cowles et al. 1996). They further asserted that it is impossible to determine convergence using a single chain. Thus, there is a need for additional independent chain(s) that starts from different dispersed initial values, to be used for comparison (Cowles et al. 1996).

There are several studies in the literature that criticized the Gelman and Rubin's approach. For instance, according to Cowles and Carlin (1996) the method depends greatly on the ability of the user to find out a starting distribution which is over dispersed with regards to the target distribution (a condition that needs user's knowledge for verification). Secondly, the method may be questionable due to its dependence on the normal approximation to analyze convergence to the true posterior. In addition, the method is univariate in nature.

"However, they suggested applying their procedure to -2 times the log of the posterior density as a way of summarizing the convergence of the joint density. Advocates of running a single long chain consider it very inefficient to run multiple chains and discard a substantial number of early iterations from each" (Cowles et al. 1996).

3.11 Spatial Autocorrelation

Spatial autocorrelation, long considered important in geostatistics, spatial analysis, and econometrics, is becoming more widely recognized in other fields (Borcard et al. 1992). The first law of geography attributed to Tobler states that “everything is related to everything else, but near things are more related than distant things” (Tobler, 1970), this refers to spatial autocorrelation. In other words, observations in close spatial proximity tend to be more similar than are observations at greater separation. Such a condition is important when dealing with UA because error at a specific location is expected to have an influence on the neighboring locations either positively or negatively (Campbell, 1981). In spatial analyses, Hunter and Goodchild (1997) asserted that neglecting the error’s spatial autocorrelation leads to a “worst-case scenario”, that is, a scenario that assumes errors in spatial data are completely random and not spatially autocorrelated. Wechsler (1999) presented a methodology for simulating the DEM error correlation through the *MC* simulation and examined the impact of DEM error correlation on elevation and three derived parameters that are frequently used in hydrological analyses. However, Oksanen and Sarjakoshi (2005a) in their study of error propagation of DEM-based surface derivatives with *MC* and *analytical methods* challenged the ‘worst-case scenario’ issue because none of the DEM derivatives investigated in their study had a maximum variation with spatially uncorrelated random error. They also concluded that *MC* method is appropriate for analyzing both constrained and unconstrained derivatives, whereas, analytical approach seems to be more useful for constrained derivatives. Wechsler (1999) developed four filter methods to represent spatial autocorrelation of error through random error fields, namely: neighborhood autocorrelation, mean spatial dependence, weighted spatial dependence, and interpolated spatial dependence. In order to incorporate the spatial dependence distance assumed to be the *range* derived from the semivariogram (variogram) analysis of the study area data, the methods utilized neighborhood or search radius of a filter. However, the approach employed by each method is entirely different.

Analysis of the semivariogram/semivariance (variogram) gives important information regarding the nature and structure of spatial dependency or variability in a random field (Kitanidis, 1997). However, there exist several tools available which are used to perform spatial variability analysis. These include correlation functions, covariance functions, and variograms. Although, these different tools provide different statistical parameters, they primarily describe the spatial relationship between variables. Of all these available tools, the variogram is a common choice for many earth science applications (Isaaks and Srivastava, 1989) since classical statistical methods are not always adequate to analyze space-structured phenomenon (Legendre and Fortin, 1989). Generally, variogram is a model that characterizes the spatial continuity or roughness of a data set (Barnes, 2011).

The mathematical definition of the variogram is expressed in Equation (3.20) (Houlding, 2000):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left[z(x_i) - z(x_i + h) \right]^2 , \tag{3.20}$$

where

$\gamma(h)$ is the estimated semivariance for the lag (distance) h

$N(h)$ is the number of measured point pairs in the distance class h,

$z(x_i + h)$ is the data value at cells separated from x_i by lag distance h in the chosen direction.

The resulting variogram according to Brimicombe (2010) has a number of components that are evident once a best-fit curve is established (see Figure 3.2).

- $\gamma(h)$ increases with lag (varying inversely with autocorrelation) to reach a *sill* beyond which there is no increase in $\gamma(h)$.
- The lag h at which the sill is reached, known as the *range*, represents the limit of spatial dependence. Presumably, spatial autocorrelation is essentially zero beyond the range (Bohling, 2005).
- An intercept of $\gamma(h) > 0$, or *nugget*, represent spatially uncorrelated error.

Variogram model generation involves two steps: the experimental variogram computed from the data and the theoretical variogram model fitted to the data (Figure 3.2). The experimental variogram is calculated by averaging one-half difference squared of the z-values over all pairs of observations with the specified separation distance and direction. It is plotted as a two-dimensional graph. The theoretical variogram model is selected from a set of mathematical functions which describe the spatial relationships. The appropriate model is chosen by matching the shape of the curve of the experimental variogram to the shape of the curve of the function (Barnes, 2011). The main aim of fitting a model to the experimental variogram is to give an algebraic expression for the relationship between values at fixed distances. There are many possible models to fit an experimental variogram. Some commonly used models include: *linear, spherical, exponential,* and *Gaussian models* (Figure 3.3). Mathematical expressions of these models are presented in Equations (3.21) - (3.29). Detailed descriptions of variogram models are presented in various publications (e.g., Issaks and Srivastava, 1989; Clark and Harper, 2002; Houlding, 2000).

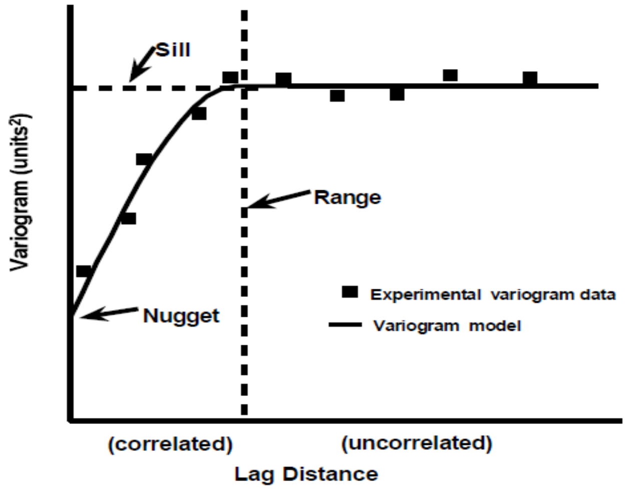


Figure 3.2. Experimental variogram and variogram model (Murray et al. 2004)

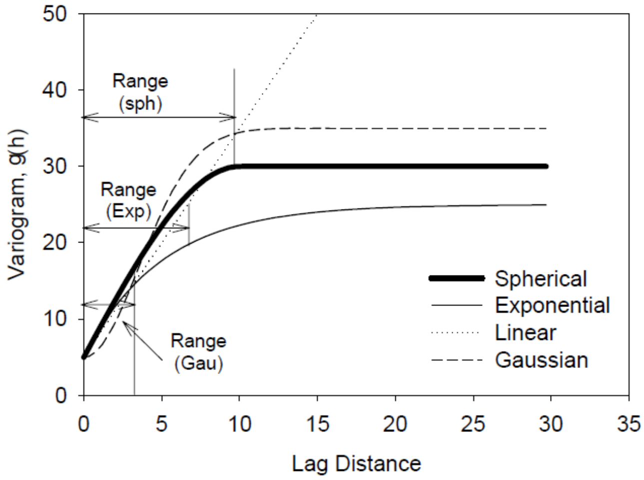


Figure 3.3. Comparison of commonly used theoretical variogram models (White et al. 2007)

The theoretical variogram models are based on three parameters: the *nugget*, the *sill*, and the *range* (Figure 3.2).

Linear Model:

$$\gamma(0)=0, \tag{3.21}$$

$$\gamma(h)=p.h \qquad\qquad if \quad h>0. \tag{3.22}$$

where

h = Lag distance,
p = Slope of the line.

Spherical Model:

Spherical model is the simplest and most commonly used model (Smith, 2013). The model is defined by a range $-a-$ and a contribution $-c-$ as:

$$\gamma(0) = 0, \tag{3.23}$$

$$\gamma(h) = c\left[1.5\frac{h}{a} - 0.5\left(\frac{h}{a}\right)^3\right] \qquad if \quad \frac{h}{a} \le a, \tag{3.24}$$

$$\gamma(h) = c \qquad if \quad \frac{h}{a} > a, \tag{3.25}$$

where

a = Range of influence,
c = Sill,
h = Lag distance.

Spherical models provide a better fit when spatial autocorrelation decreases to a point after which it becomes zero (Cressie, 1993).

Exponential Model:

The *exponential model* is defined by a parameter - a - and a contribution - c - as:

$$\gamma(0) = 0, \tag{3.26}$$

$$\gamma(h) = c\left[1 - \exp\left(-3.\frac{h}{a}\right)\right] \qquad if \quad h > 0. \tag{3.27}$$

where

a = Range of influence,
c = Sill,
h = Lag distance.

Cressie (1993) noted that exponential models fit best when the spatial autocorrelation decreases exponentially with increasing distance.

The *exponential* and *spherical models* exhibit linear behavior at the origin, appropriate for representing properties with a higher level of short-range variability (Bohling, 2005). Kalkhan (2011) noted that the linear model assumes a constant increase of the variance with the distance and hence there is neither range nor sill; it is the case of the spatial gradient.

Gaussian Model:

The *Gaussian model* has a parabolic nature at the origin and used to represent very smoothly varying features (Bohling, 2005), and is defined by a parameter - a - and a contribution - c - as:

$$\gamma(0) = 0, \tag{3.28}$$

$$\gamma(h) = c\left[1 - \exp\left(-\frac{3h^2}{a^2}\right)\right] \qquad if \qquad h > 0, \tag{3.29}$$

where

a = Range of influence,
p = Slope of the line,
c = Sill.

CHAPTER 4

PROPOSED METHODOLOGICAL FRAMEWORK

This chapter discusses the proposed *Markov chain Monte Carlo (MCMC)* simulation methodology for investigating digital elevation model (DEM) uncertainties in *GIS-based solar radiation models*.

The methodological framework proposed in the thesis is given in Figure 4.1. It consists of four phases. Phase I: Identification and classification of solar radiation model inputs, Phase II: Stochastic (*MCMC*) simulation and convergence analysis for DEM, Phase III: Check for variogram reproduction, and Phase IV: Execution of solar radiation models and uncertainty assessment. Sections 4.2.1 - 4.2.4 discusses these phases in details. This study considers only two *GIS-based solar radiation models*, that is, *Solar Analyst* (Rich et al. 1995) - a commercial software developed for ArcGIS, and *r.sun* model (Hofierka and Šúri, 2002), an Open Source GIS software developed for GRASS. The versions of the software used in this research are; ArcGIS 9.3 and GRASSS 6.4.0.

4.1 Phase I: Collection and Classification of Solar Radiation Models Inputs

The first stage in this phase involves the collection of *Solar Analyst* and *r.sun* models inputs (Figure 4: PI-1). These inputs are given in Tables 2.1 and 2.2 for *Solar Analyst* and *r.sun*, respectively. Then followed by pre-processing which deals with processing of both digital and analog data collected to required data format for the respective models, e.g., conversion of analogue to digital data format, conversion from one file format to another and spatial formats, transformation from one projection to another (Figure 4: PI-2). The third stage is the classification of model inputs into probabilistic and deterministic (Figure 4.1: PI-3, PI-4, PI-5). Probabilistic (stochastic) inputs are those in which random events and effects play an important role (Edwards and Hamson, 1989). In contrast, the deterministic inputs are those variables that whenever given to a model the outputs will be same (World Health Organization, 2005). DEM, slope, aspect, linke turbidity factor (T_{LK}), and ground albedo has the characteristics of being considered as probabilistic inputs. However, in this study T_{LK} and ground albedo are considered as deterministic inputs. The remaining inputs exhibit deterministic characteristics (Tables 2.1 and 2.2). As the aim of this study is to investigate the DEM uncertainty on *GIS-based solar radiation models*, therefore, Shuttle Radar Topography Mission (SRTM) DEM of the study area is obtained (Figure 4: PI-6) for utilization in the analysis. However, it is important to note that the proposed methodological framework is applicable to any kind of DEM data available.

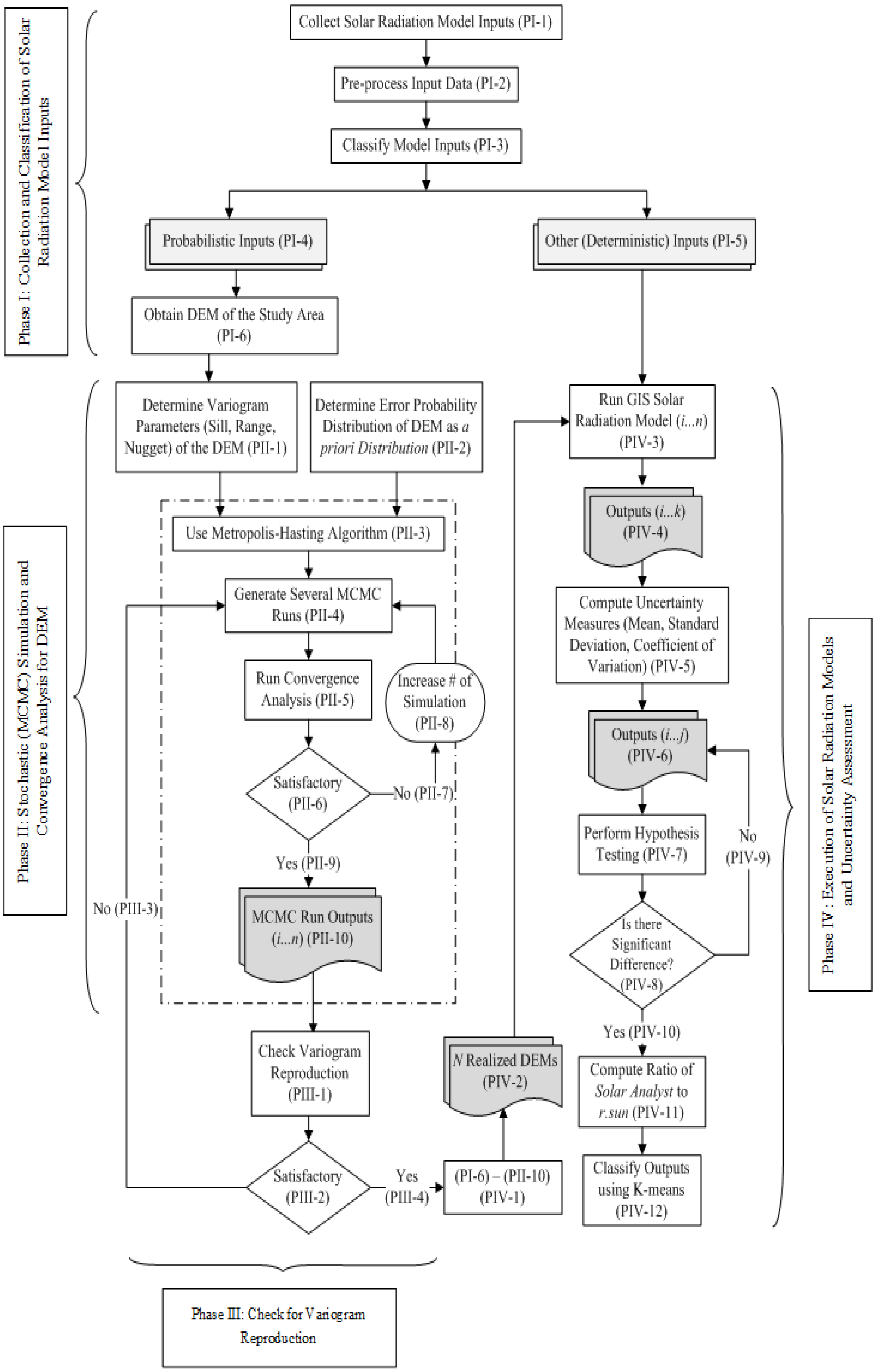


Figure 4.1. Flow chart of the proposed methodological framework

4.2 Phase II: Stochastic (MCMC) Simulation and Convergence Analysis for DEM

This phase deals with formulating and implementing the *MCMC* simulation, and then assessing its convergence (Figure 4). The first stage (Figure 4: PII-1 and PII-2) provides the necessary inputs for the *MCMC* simulation by determining the; (1) DEM variogram characteristics or spatial dependence (autocorrelation) of the study area. (2) DEM error distribution.

Spatial autocorrelation can best be explained by the first law of geography which is expressed by variogram/semivariogram of the DEM. Equation (4.1) (Journel and Huijbregts, 1978) is used to compute the semivariogram. The obtained variogram model and its parameters are then used in conditioning the *MCMC* simulation such that it generates several equiprobable realizations that provide the same spatial dependence (autocorrelation) structure of the original DEM.

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left[z(x_i) - z(x_i + h) \right]^2, \tag{4.1}$$

where

$z(x_i)$ is the elevation value at location x_i,
$N(h)$ is the number of elevation pairs $\left[z(x_i), z(x_i + h) \right]$ separated by a lag distance of h.

The DEM error distribution (Figure 4: PII-2) deals with identification of error probability distribution function (pdf) of the SRTM DEM data and its parameters (mean and standard deviation). Generally, this can either be obtained from referenced data set with higher accuracy (Bolstad and Stowe, 1994; Fisher and Tate, 2006) or from the literature. This information serves as *a priori* pdf when formulating and implementing the *MCMC* simulation. In this study, SRTM DEM is utilized and as such DEM error distribution is extracted from Rodríguez et al. (2006) and then analyzed using distribution fitting software. Rodríguez et al. (2006) reported that the SRTM performance observed by comparing against the ground-truth (kinematic GPS transects) met/exceeded its performance requirements, often by a factor of 2.

The second stage deals with stochastic simulation, i.e., *MCMC* based on *Metropolis-Hasting (MH) algorithm* (Figure 4: PII-3). There are several *MCMC algorithms*: *Metropolis-Hastings algorithm* (Metropolis et al. 1953), *Gibbs sampler* (Geman and Geman, 1984), *Slice sampler* (Neal, 2003) as explained in Chapter 3. However, in this study, the *MH algorithm* is adopted because of its generality, simplicity and powerfulness (Robert and Casella, 2010). To formulate the proposed stochastic *MCMC algorithm* for this study, the DEM error pdf obtained in the previous stage is utilized as *a priori* pdf while the variogram model is used to generate the spatial autocorrelation or covariance matrix.

The *MH algorithm* designed and implemented for this study is based on Navarro and Perfors (2011) and described below:

(1) Generation of a candidate, denoted as x^*: The value of x^* is generated from the proposal distribution, denoted by $Q\left(x^* | x_n\right)$, which depends on the current state of the Markov chain, x_n. There are a few minor technical constraints on what one can use as a proposal distribution, but generally it can be selected arbitrarily. A typical way to do this is to use a normal distribution centered on the current state x_n. That is

$$x^* \Big| x_n \approx Normal\left(x_n, \partial^2\right), \tag{4.2}$$

for variance of ∂^2 specified by the user.

(2) Accept-reject step: First, the acceptance probability $A\left(x_n \to x^*\right)$, in Equation (4.3) is calculated:

$$A\left(x_n \to x^*\right) = \min\left(\frac{p\left(x^*\right)}{p\left(x_n\right)} \cdot \frac{Q\left(x_n \middle| x^*\right)}{Q\left(x^* \middle| x_n\right)}, 1\right), \tag{4.3}$$

The probability of both the candidate (x^*) and current state (x_n) is calculated using the *prior* pdf determined in (Figure 4: PII-2). In this study, a lognormal pdf is found as the most appropriate distribution that models the SRTM DEM error. This conclusion is reached after analyzing the data extracted from Rodríguez et al. (2006) who assessed the performance of SRTM by comparing against the ground-truth (kinematic GPS transects) see subsection 5.2.3 for details. In this regard, the multivariate log-normal distribution (Tarmast, 2001) given in Equation (4.4) is used:

$$f(y_{ij}) = (2\pi)^{-P/2} |D|^{-1/2} |y_1, y_2, ..., y_p|^{-1} \exp\left[-\left(\log_{ij} - v_{ij}\right)' D^{-1} \left(\log_{ij} - v_{ij}\right)/2\right] \quad (0 < y_i < \alpha). \tag{4.4}$$

where

$\log = [\log_1,\ \log_2,\ ...,\ \log_p]$ is a *p*-component column vector,

$y_i = \exp\left(x_i\right)$,

v = mean,

D = covariance matrix calculated from the exponential variogram model.

As lognormal distribution can be transformed to normal distribution instead of using Equation (4.4) it is decided to use a multivariate normal pdf as given in Equation (4.5) (Do, 2008). After all the calculations the final samples are back transformed to lognormal.

$$p(x) = \frac{1}{(2\pi)^{\frac{k}{2}} |\Sigma|^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2}(x-\mu)' \Sigma^{-1} (x-\mu)\right), \tag{4.5}$$

where

x denotes candidate elevation error,
k-dimension (size of candidate elevation),
Σ covariance matrix,
μ is the mean.

Equation (4.6) (Tarantola, 2008) is used to calculate covariance from an exponential model:

$$C(h) = v^2 \exp\left(\frac{-h}{a}\right) \qquad (h > 0), \tag{4.6}$$

where

v^2 = variance,
a = range,
h = lag-distance.

Having proposed the candidate and calculated acceptance probability, $A\left(x_n \to x^*\right)$.

To make the decision of acceptance or rejection of a candidate, uniformly distributed random number (u) between 0 and 1 are generated. Then the acceptance and rejection decision is made based on:

$$x_{n+1} = \begin{cases} x^*, \ if \ u \le A\left(x_n \to x^*\right), \\ x_n, \ if \ u \ge A\left(x_n \to x^*\right). \end{cases} \tag{4.7}$$

Based on the above algorithm, the *MCMC* simulation is then run a finite number of (n) times for m parallel chains using different starting values that are over dispersed with respect to the target distribution (Figure 4: PII-4).

The Markov chain generated by *MH algorithm* eventually converges to the distribution used in the calculation of the acceptance criteria for any form of the proposal distribution, given that the Markov chain is ergodic (Gilks et al. 1996). There exist several methods for testing *MCMC* convergence, for instance, Gelman and Rubin (1992); Raftery and Lewis (1992); Geweke (1992). This study utilizes the Brooks and Gelman (1998) approach, which is a generalized version of Gelman and Rubin (1992) that is simple and generally applied to the output of any iterative simulation (Figure 4: PII-5, PII-6, PII-7, PII-8, PII-9). The approach is based on analyzing multiple simulated *MCMC* chains (m parallel chains) by comparing the variances within each chain and the variance between chains. A large deviation between these two variances indicates non-convergence and vice-versa. The method involves two steps as summarized below:

(1) Stage I is performed before commencing the sampling. It involves obtaining an over dispersed estimate of the target distribution and then use it to generate the starting points for the required number of independent chains.
(2) Stage II is performed for each scalar quantity of interest after running the *MH sampler* chains for the required number of iterations (n). This involves using the last n-iterations to re-estimate the target distribution of the scalar quantity as a conservative Student t

distribution, the scale parameter involves both the between- and within-chain variance. The details of the procedure can be found from Gelman and Rubin (1992).

According to Brooks and Gelman (1998) when the estimation of a vector parameter θ is based upon observations $\theta_{jt}^{(i)}$, denoting the ith element of the parameter vector in chain j at time t. With regards to higher dimensions, this means the estimation of the posterior variance-covariance matrix by using Equations (4.8), (4.9) and (4.10):

$$\hat{V} = \frac{n-1}{n}W + \left(1+\frac{1}{m}\right)B/n, \tag{4.8}$$

where

$$W = \frac{1}{m(n-1)}\sum_{j=1}^{m}\sum_{t=1}^{n}\left(\theta_{jt} - \bar{\theta}_{j.}\right)\left(\theta_{jt} - \bar{\theta}_{j.}\right)'. \tag{4.9}$$

and

$$B/n = \frac{1}{m-1}\sum_{j=1}^{m}\left(\bar{\theta}_{j.} - \bar{\theta}_{..}\right)\left(\bar{\theta}_{j.} - \bar{\theta}_{..}\right)'. \tag{4.10}$$

Equations (4.9) and (4.10) represent the (p-dimensional) within- and between-sequence covariance matrix estimates of the p-variate functional θ, respectively. Therefore, both $\hat{V}$ and W, can be monitored by determining convergence when any rotationally invariant distance measure between the two matrices shows "sufficient closeness".

The multivariate potential scale reduction factor (MPSRF) denoted as $\hat{R}$ is then computed as follows:

$$\hat{R} = \sqrt{\frac{n-1}{n} + \frac{m+1}{m}\lambda_1}, \tag{4.11}$$

where

λ represent the positive definite matrix $(1/n)W^{-1}B$ of the largest eigenvalue.

The value of $\hat{R}$ in Equation (4.11) declines towards 1 as the simulation converges. Gelman et al. (2004) suggest that values for $\hat{R}$ less than 1.1 may be regarded as an indication that the *MCMC sampler* has converged. The achievement of convergence gives several equiprobable outputs of DEM error (Figure 4: PII-10).

4.3 Phase III: Check for Variogram Reproduction

Since the *MCMC* simulated realizations in this study are required to adequately reproduce not only the *prior* distribution but also the input variogram model, a further check is also performed on the simulation output after achieving convergence by comparing the resulting variograms over multiple realizations with the reference (theoretical) variogram model (Figure 4: PIII-1, PIII-2, PIII-3, PIII-4). This aids in suggesting the acceptance or rejection of the computed realizations as correct numerical representations of the phenomenon under study, i.e., whether the realizations obtained after convergence possess spatial autocorrelation characteristic similar to the DEM of the study area, or not.

A multivariate hypothesis test developed by Ortiz and Leuangthong (2007) is adopted for this study. The method is based on Hotelling's T^2-statistic and is used to measure whether the fit is acceptable within a 95% confidence level ($\alpha = 0.05$).

Lags or distances (p) are chosen for verification, and the resultant n realization variograms form the sample variogram values for each lag. Considering all relevant lags taken together requires a joint test of how well the variogram model is being reproduced. A hypothesis test is constructed based on the sample mean variogram values and compared against the model variogram value for each lag distance. The null hypothesis H_0 and alternative hypothesis H_1 are constructed as:

$$H_0: \ \mu = \mu_0,$$
$$H_1: \ \mu \neq \mu_0,$$

where

μ is the p x 1 mean vector obtained from the variogram realizations for p different lags,
μ_0 is the p x 1 mean vector based on the reference or input variogram for the same p lags.

The test is constructed to measure the squared distance of the sample mean from the reference mean, standardized by the sample covariance. For this multivariate context, the Hotelling's T^2-statistic is distributed as

$$\frac{(n-1)p}{n-p} F_{p,n-p}, \tag{4.12}$$

where,

$F_{p,n-p}$ represents an F-distribution with p and n-p degrees of freedom, respectively,
p is the number of lag distances to check,
n is the number of sample variogram values available at each lag and corresponds to the number of realizations generated.

Under this multivariate context, the null hypothesis is rejected if

$$T^2 = n\left(\bar{X} - \mu_0\right)' S^{-1}\left(\bar{X} - \mu_0\right) > \frac{(n-1)p}{n-p} F_{p,n-p(\alpha)}, \tag{4.13}$$

where,

S is the sample covariance matrix and calculated as

$$\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(X_i - \bar{X}\right)'}{n-1}. \tag{4.14}$$

4.4 Phase IV: Execution of Solar Radiation Models and Uncertainty Assessment

In this study, it is assumed that the DEM uncertainty can be modeled by subtracting from every grid elevation error term for a random field spatially dependent on neighbors generated using *MCMC* simulation. Let $Z(x, y)$ be a DEM elevation at the position (x, y), then:

$$Z(x, y) = z(x, y) - \mathbf{R}, \tag{4.15}$$

where,

$z(x, y)$ is the elevation from DEM,
R is a random error generated based on the proposed MCMC simulation.

Thus, this phase (Figure 4) involves subtracting each (n) simulated DEM error (Figure 4: PII-10) from the study area DEM (Figure 4: PI-6), which is assumed to consist of the true elevation at that point plus an unknown amount of error to produce N different but equally probable realizations of the topography with errors in the study area (Figure 4: PIV-2). Since the topographic realizations take into account measured elevation errors, each different but equally probable realization is in theory a more accurate representation of the real topography than the original DEM (Haneberg, 2006). In other words, these alternative equiprobable realizations, which span the range of possible attribute values given the model of uncertainty at any location, are used as alternative inputs for the solar radiation models, thereby generating possible results. Therefore, the study area DEM (Figure 4: PI-6), the N realized DEMs (Figure 4: PIV-2), and other inputs are used to execute the solar radiation models (Figure 4: PIV-3) for the purpose of generating the direct, diffuse and global radiation, and direct radiation duration maps in *Solar Analyst* (Figure 4: PIV-4). Similarly, beam, diffuse, ground reflected and global (total) irradiation, and insolation time maps are obtained in *r.sun* model (Figure 4: PIV-4). Then, mean, standard deviation, and coefficient of variation (CoV) are computed to provide measures of uncertainty (Figure 4: PIV-5, PIV-6). The mean of the perturbed DEMs or realized DEMs is calculated using Equation (4.16), whereas, Equations (4.17), (4.18) and (4.19) are used for calculating standard deviation, and CoV, respectively.

$$\bar{x}_{(i,j)} = \frac{\sum_{i=1}^{n} x_{(i,j)}}{n}, \tag{4.16}$$

where

$\bar{x}_{(i,j)}$ is the mean of a cell at row i and column j,

$x_{1,(i,j)}$, $x_{2,(i,j)}$, ..., $x_{n,(i,j)}$ are cell values for each of the realized DEMs,

n is the total number of simulations.

$$\sigma_{(i,j)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_{(i,j)} - \bar{x}_{(i,j)}\right)^2}, \tag{4.17}$$

where

$\sigma_{(i,j)}$ is the standard deviation of a cell at row *i* and column *j*,

$x_{1,(i,j)}$, $x_{2,(i,j)}$, ..., $x_{n,(i,j)}$ are cell values,

$\bar{x}_{1,(i,j)}$, $\bar{x}_{2,(i,j)}$, ..., $\bar{x}_{n,(i,j)}$ are the means of individual cells.

$$CoV_{(i,j)} = \frac{\sigma_{(i,j)}}{\bar{x}_{(i,j)}}.100, \tag{4.18}$$

where

$CoV_{(i,j)}$ denotes the CoV of a cell at row *i* and column *j*,

$\sigma_{1,(i,j)}$, $\sigma_{2,(i,j)}$, ..., $\sigma_{n,(i,j)}$ is the standard deviation of individual cells,

$\bar{x}_{1,(i,j)}$, $\bar{x}_{2,(i,j)}$, ..., $\bar{x}_{n,(i,j)}$ are the means of the individual cells.

A relative comparison based on statistical analysis is carried out with a view of determining the significant difference between the radiation estimates of the two models. The outputs considered are both generated by the two models and comprised of direct/beam radiation, diffuse radiation, global radiation, and direct duration. The reflected radiation generated only by the *r.sun* is not considered.

For statistical testing of difference in the radiation estimates of the two models, a hypothesis test is performed (Figure 4: PIV-7). If a significant statistical difference is found in the estimated distributions of the models (Figure 4: PIV-8, PIV-10), a ratio of *Solar Analyst* to *r. sun* is calculated (Figure 4: PIV-11) and the results are classified into two or three classes depending on the variations using *K-means clustering method* (Figure 4: PIV-12). If otherwise (Figure 4: PIV-8, PIV-9), the process is ended at (Figure 4: PIV-6). The K-means clustering is a partitioning procedure where the data are grouped into K group by the user. The routine tries to find the best positioning of the K centers and then assigns each point to the center that is the nearest (Akpinar, 2005). In other words, the K-means is an algorithm for clustering $\boldsymbol{N}$ data points into K disjoint subsets S_j to minimize the criterion N_j (Equation 4.19) in which x_n is a vector representing the n^{th} data point and μ_j is the geometric centroid of the data points in S_j (Math World, 2013).

$$J = \sum_{j=1}^{K} \sum_{n \in S_j} \left| x_n - \mu_j \right|. \tag{4.19}$$

CHAPTER 5

IMPLEMENTATION

The proposed stochastic methodology is illustrated with a case study from Abuja, which is the administrative capital of the Federal Republic of Nigeria. The implementation is done using MaTLAB 2012a®, R software, EasyFit Professional software 5.5®, ArcGIS 9.3®, GRASS GIS 6.4 and SPSS 15®. Further analyses are also conducted using different tools like Microsoft Excel. The implementation process is divided into three, namely; (1) Proposed *Markov chain Monte Carlo (MCMC)* simulation, (2) Simulation of *GIS-based solar radiation models*, and (2) Assessment of digital elevation model (DEM) uncertainties on *GIS-based solar radiation models* outputs.

5.1 The Study Area

The case study area is 9.225 km^2 (3.15 km by 3.15 km) and located in the northeastern part of Abuja (Figure 5.1). The area lies between latitude 9^o 7' 28" and 9^o 5' 42" North, and longitude 7^o 27' 39" and 7^o 29' 25" East.

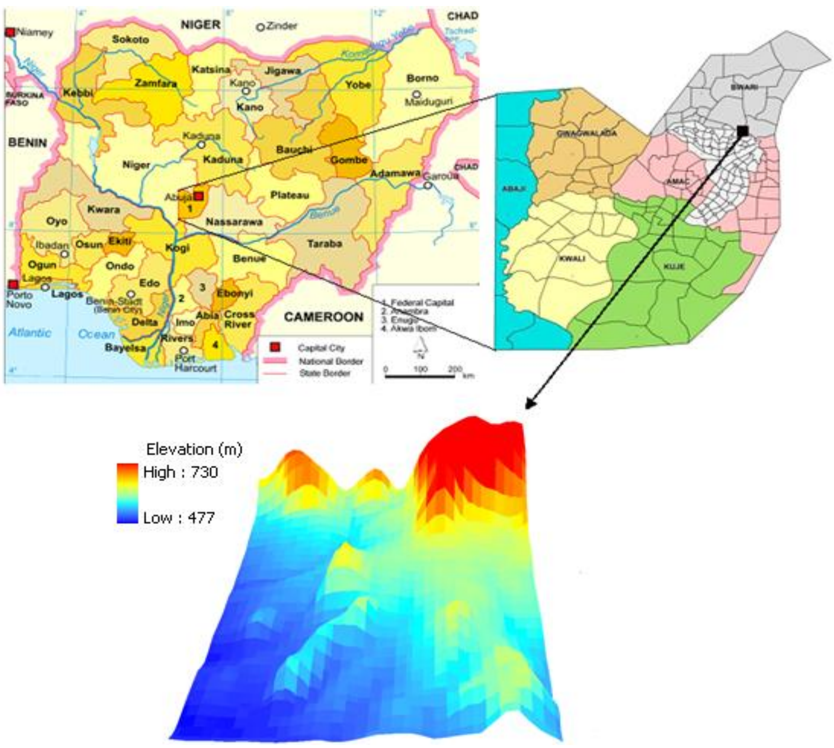


Figure 5.1. The Study area

A free downloadable Shuttle Radar Topography Mission (SRTM) DEM with spatial resolution of 90 m is acquired from the Consortium for Spatial Information (CGIAR-CSI) website and utilized for this study. The tile (srtm_3811) covering the study area is projected to WGS84 UTM Zone 32 N and then study area is extracted. The area has potential for solar radiation resource that will enable the establishment of solar power plants because it is situated in a tropical climate region where the weather is generally hot (average monthly temperature ranges between 29^{o}C to 40^{o}C) and sunny throughout the year. It is also characterized with highlands, lowlands, mountains, and relatively flat terrain. This will enable the assessment of DEM uncertainties with respect to these different landforms. In addition, if the solar radiation potential is utilized, it will contribute immensely in addressing the electricity problems, which the region and the country are currently experiencing.

5.1.1 Energy Outlook of Nigeria

Nigeria, the most populous country in Africa with 168.8 million people (World Bank, 2012), is the biggest energy basin in the continent with 37.14 billion barrels proven oil reserves, 5,118 billion m^3 proven natural gas reserves (OPEC, 2013), and 209 million short tons recoverable coal reserves (EIA, 2013). On the other hand, the Nigeria's renewable energy potential include 10,000 megawatts (MW) of large scale hydropower, 734 MW small scale hydropower, fuel wood (13,071,464 hectares) of forest land, animal waste (61 million tons/year), crop residue (83 million tons/year), solar radiation (3.5-7.0 kWh/m^2/day), and 2-4 m/s (annual average) wind (ECN and UNDP, 2005). Figure 5.2 shows the electricity production portfolio of Nigeria's main resources.

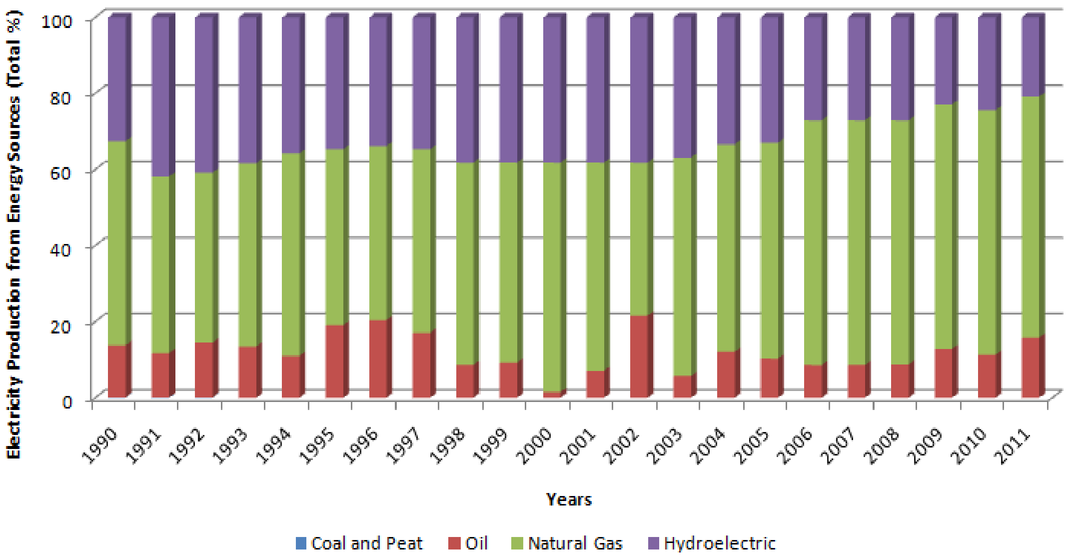


Figure 5.2. Electricity production portfolio of Nigeria showing the main sources (IEA, 2013)

However, in reality these endowed natural resources are not adequately harnessed to meet the Nigerian electricity demand due to lack of sound policies and commitments. As such, Nigeria is currently facing serious electricity supply shortages. An increasing trend in electricity demand is very common in this developing country. Therefore, a non increasing and under-utilized generating capacity (Figure 5.3) adversely affects the small, medium and large-scale enterprises, living standards of Nigerians, inflow of foreign investment, and the balance of payment. As evidence, only about 40% of Nigeria’s population have access to grid-based electricity and less than 20% of the rural population is connected to the national grid (PHCN, 2005; ICEED, 2006). In addition, more than 60% of the factories depend on generators for their main source of power supply, which increases the cost of manufactured goods.

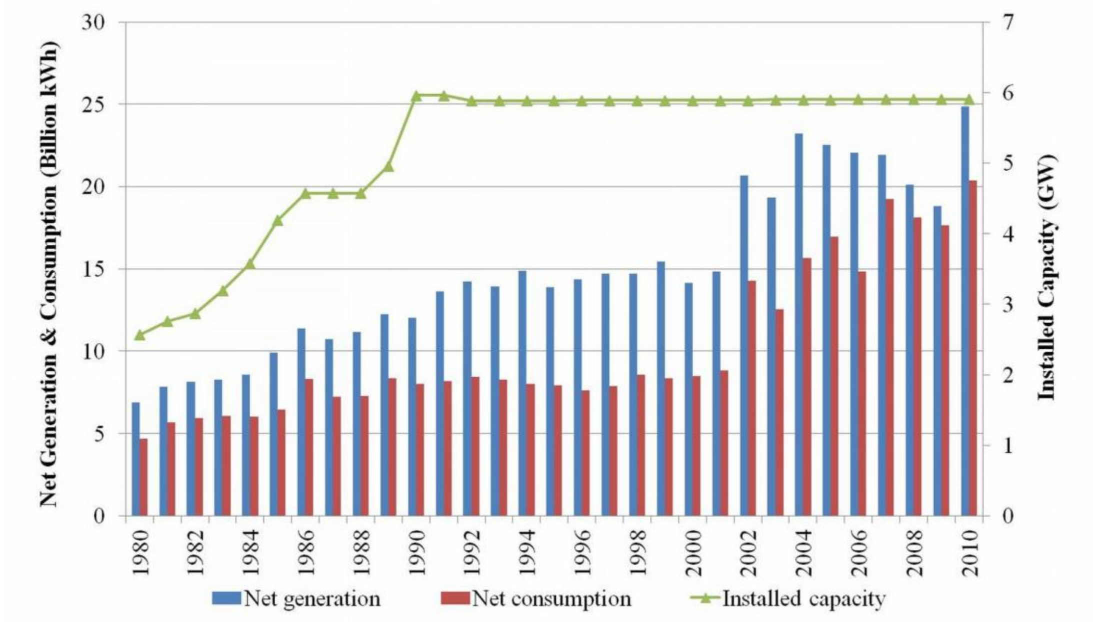


Figure 5.3. The Nigerian electricity market in the past 30 years (EIA, 2012)

5.2 Data Description for MCMC simulation

In this section, the data necessary for the formulation and implementation of the proposed *MCMC* simulation is described, and then the *MCMC* simulation implementation is demonstrated. They include DEM, DEM error probability distribution function (pdf), spatial autocorrelation model of the study area DEM, linke turbidity factor (T_{LK}) and ground albedo.

5.2.1 Digital Elevation Model (DEM) of the Case Study Area

DEM is a major input for both the *r.sun* and *Solar Analyst* models. There are several free downloadable DEMs with different spatial resolutions for the entire World on World Wide Web, for instance, SRTM: 90 m, Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) - 30 m, GTOPO30 (30 arc second, approximately 1 km). However, in this study, SRTM DEM is utilized and in particular the updated SRTM version 4.1 with 90 m resolution. The utilized DEM represents a raster of 35 cells by 35 cells with minimum elevation of 477 m and maximum of 730 m. The mean topographical elevation is 553.29 m having a standard deviation of 56.54 m (Table 5.1 and Figures 5.4 and 5.5).

Table 5.1. Summary statistics for SRTM DEM of the study area

Statistic	Values
Count	1,225
Minimum elevation (m)	477
Maximum elevation (m)	730
Mean elevation (m)	553.29
Standard deviation (m)	56.54
Skewness	1.38

Table 5.1 Summary statistics for SRTM DEM of the study area (continued)

Statistic	Values
Kurtosis	4.66
1st quartile	514.75
Median	543
3rd quartile	570

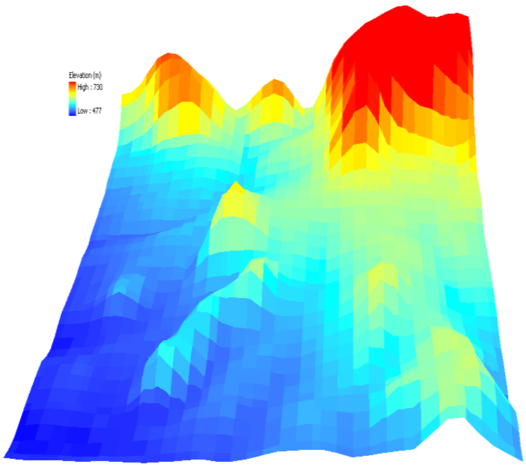


Figure 5.4. SRTM DEM of the study area

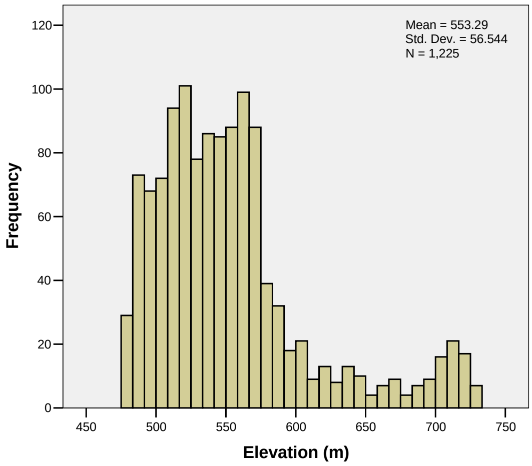


Figure 5.5. Histogram of SRTM DEM of the study area

The SRTM mission was flown using single-pass synthetic aperture radar (SAR) interferometry in February 2000 by NASA and provides a topography covering continental areas from 60° N to 56° S latitudes with a 1 and 3 arc sec spatial resolution (about 30 and 90 m), respectively. The 1 arc sec data set is publicly available for the United States while the 3 arc dataset is available for the remaining part of the world. The mission covered nearly 80% of the Earth's surface and aimed to achieve an absolute horizontal and vertical accuracy of 20 and 16 m, respectively (NASA, 2005). The vertical accuracy of ±16 m linear error at 90% confidence level is fulfilled (Sun et al. 2003; Rodríguez et al. 2005). For instance, by comparing SRTM (1 arc sec) with ICESat for the western United States, Carabajal and Harding (2006) found a mean and standard deviation of elevation differences of -0.60±3.46 m over area with low relief and -5.61±15.68 m for higher relief. In the case of 3 arc sec data set, Rodríguez et al.(2005) reported an absolute vertical error of 5.6 m for Africa, which is the lowest value among the six continents studied (Africa, Australia, Eurasia, Islands, North America, and South America). However, according to Jarvis et al. (2008) the SRTM version 4.1 represents a significant improvement from the previous versions, using new interpolation algorithms and better auxiliary DEMs. These enhancements include:

- an improved ocean mask which includes some small islands previously been lost in the cut data,
- fixed single no-data line of pixels along meridians,
- all GeoTiffs with 6000 x 6000 pixels,
- for ASCII format files the projection definition included in .prj files, and

- for GeoTiff format files the projection definition included in the .tfw (ESRI TIFF World) and a .hdr file that reports PROJ.4 equivalent projection definitions.

In addition, it is distributed in ARC GRID, ARC ASCII and Geotiff format, in decimal degrees; and projected in a geographic (Latitude/Longitude) projection, with the WGS84 horizontal datum and the EGM96 vertical datum. The CGIAR has processed this data to provide seamless continuous topography surfaces. Areas with regions of no data in the original SRTM data have been filled using interpolation methods.

The SRTM DEM data in this study is used for two purposes. First, the original data is used as an input to both the *r.sun* and *Solar Analysts* models. Second, the distribution and variogram characteristics are extracted and used in the formulation and implementation of the proposed *MCMC* simulation with a view of generating *N* realized equiprobable DEMs.

5.2.2 Probability Distribution of SRTM

In order to assess the distribution of the case study area SRTM DEM data, first, the basic descriptive parameters of the distributions are derived, including mean, median, minimum, maximum, standard deviation, and inter-quartile range (Table 5.1). Second, the degree of normality of the distribution is determined using skewness and kurtosis (Table 5.1). The skewness (1.38) is positive; indicating that the data is positively skewed or right skewed, which means the right tail of the distribution is longer than the left. On the other hand, the kurtosis (4.66) is less than 3; this indicates a *platykurtic distribution*, i.e., a distribution with a wider peak and flatter, shorter and thinner tails than a normal distribution (Table 5.1). Moreover, the probability for extreme values is less than for a normal distribution, and the values are widely spread around the mean. Third, histogram and normal Q-Q plot (Figures 5.5 and 5.6) are plotted to compare the distribution of the data to a standard normal distribution, providing yet another measure of the normality of the data. From the histogram and normal Q-Q plot it is important to note that the data is right skewed, meaning it deviates from normality.

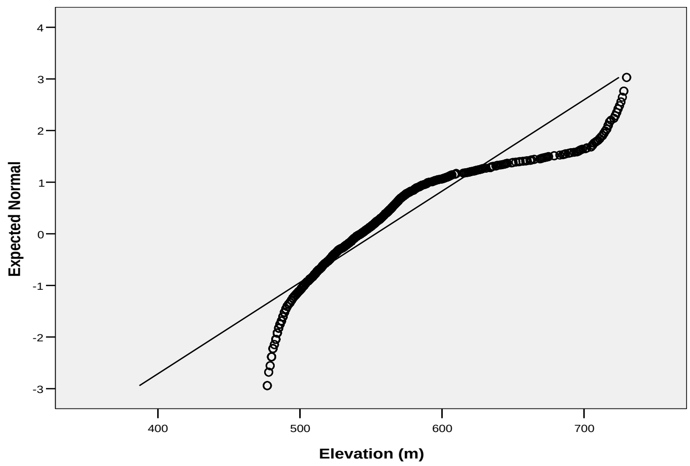


Figure 5.6. The Normal Q-Q Plot of Study Area SRTM DEM

5.2.3 Error Distribution of SRTM

Table 5.2 shows summary performance of the SRTM, while Figure 5.7 shows the histograms and cumulative distribution function for the height error magnitude from Rodríguez et al.

(2006) analysis of SRTM DEM performance. To find out the DEM error distribution, data is extracted from Rodríguez et al. (2006) analysis result for the height error (m), specifically Figure 5.7(a). The extracted data are statistically analyzed using the distribution fitting software (EasyFit Professional software 5.5®). The software tests and ranks several pdfs using the goodness-of-fit tests of Kolmogorov-Smirnov, Anderson-Darling, and Chi-Square. Based on the result of this analysis, a lognormal pdf is found to be the most appropriate distribution that models the SRTM DEM error distribution (Table 5.3 and Figure 5.8) with mean (μ) = 4.209 and standard deviation (σ) = 0.054. Thus, lognormal pdf is used as *a priori* input distribution function in addition to other parameters to generate equiprobable DEMs using the *MCMC* simulation.

Table 5.2. Summary of kinematic GPS GCP comparison with SRTM data (Rodríguez et al. 2006)

Continent	Mean (m)	Standard Deviation (m)	90% Absolute Error (m)
Africa	1.3	3.8	6.0
Australia	1.8	3.5	6.0
Eurasia	-0.7	3.7	6.6
North America	0.1	4.0	6.5
New Zealand	1.4	5.9	10.0
South America	1.7	4.1	7.5

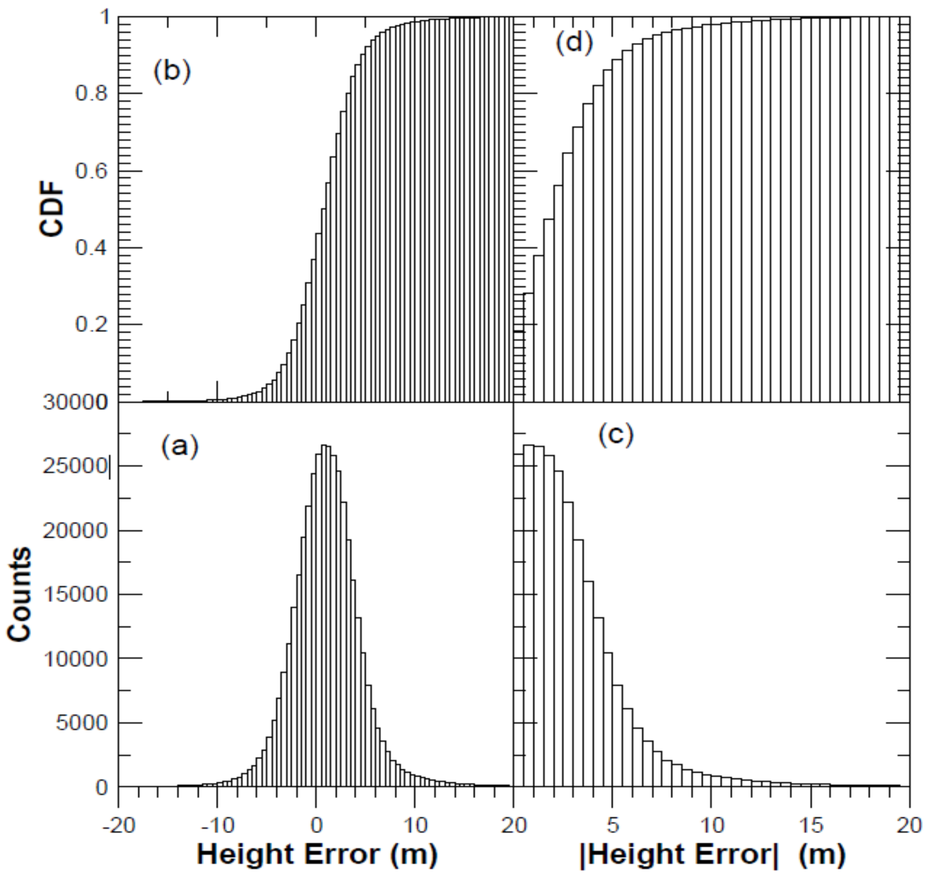


Figure 5.7. Africa kinematic GPS height comparison. Panel (a) shows the distribution of the signed error; panel (b) the corresponding cumulative distribution function; panel (c) is the distribution of the error magnitude; and panel (d) is the cumulative distribution of the error magnitude (Rodríguez et al. 2006).

Table 5.3. Goodness of fit summary

Lognormal (3P)										
	Kolmogorov-Smirnov					**Anderson-Darling**				
Sample Size	69					69				
Statistic	0.05973					57.757				
P-Value	0.95399									
Rank	1					3				
alpha	0.2	0.1	0.05	0.02	0.01	0.2	0.1	0.05	0.02	0.01
Critical Value	0.12675	0.14483	0.16088	0.1799	0.19303	1.3749	1.9286	2.5018	3.2892	3.9074
Reject?	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes

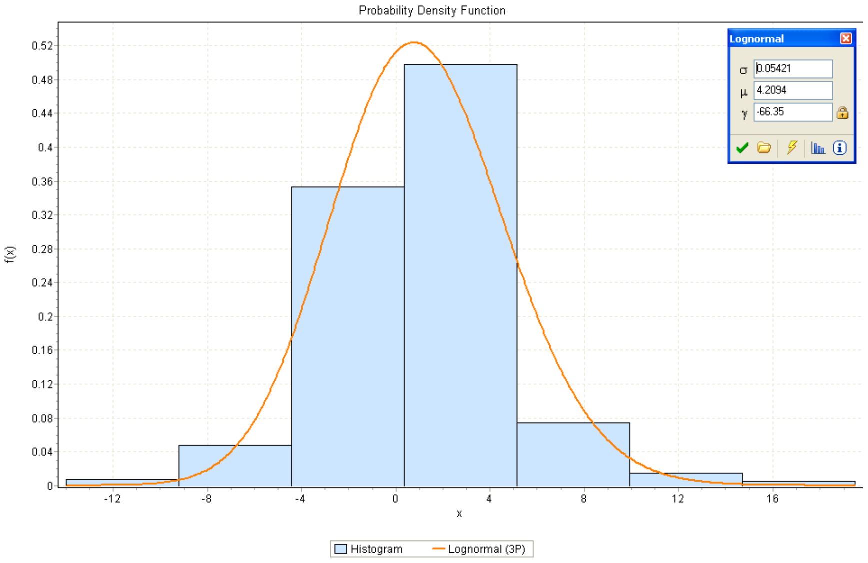


Figure 5.8. Lognormal probability density function

5.2.4 Variogram Model of SRTM for Autocorrelation Assessment

The traditional way of finding a suitable variogram model is produced by fitting a theoretical variogram model to experimental variogram model (Bivand et al. 2008; Minasny et al. 2011). There are two basic fitting approaches: manual and automated. In the manual fitting, a theoretical semivariogram model is selected based on visual inspection of the empirical semivariogram, for example, Hohn (1988); Olea (1999). The automated approach involves performing model fitting in an automated manner. For this task one can use methods such as least squares, maximum likelihood, and robust methods (Cressie, 1993). The main disadvantages of the manual fitting method according to Li and Lu (2010) include, (1) laborious and time-demanding, (2) lacks uniform and objective format, and (3) affects the automation process of the entire geostatistics computations. In this regard, this study adopts the automated fitting method since experimental variogram values are best represented by the method (Li and Lu, 2010). Therefore, the spatial autocorrelation of the study area SRTM DEM was assessed using variogram based on a code written and executed in R software which automatically fits the variogram model (Appendix A). The outcome of the analysis showed that the exponential model is the best model that fit the data (Table 5.4). The exponential model is also found to be appropriate in previous studies like Holmes et al.

(2000); Oksanen and Sarjakoski (2005). Thus, the result indicates that spatial autocorrelation decreases exponentially with increasing distance. The exponential model is linear at very short distances near the origin; however, it rises more steeply and then flattens out more gradually (Alves et al. 2009).

Table 5.4. Variogram parameters of the study area SRTM DEM

Variogram parameters	Values
Nugget	0
Partial sill	1.008936e+07
Range	20499218
Fitted variogram model	Exponential
SSError	12875384

5.3 Implementation of the Proposed MCMC Simulation

The parameters of the lognormal pdf, i.e., μ = 4.209 and σ = 0.054, and exponential variogram model (*sill*: 125.74, *range*: 1,556.85, *nugget*: 0) are used as *a prior* information in which the proposed *MCMC* simulation based on *MH algorithm* is coded and executed in a MaTLAB® programming environment (Appendix B). A total of 1,080 simulations is run to achieve convergence using $2m$ chains. Whereas, the initial *burn-in* period of eighty which represent 7.41% of the simulation is discarded. See Appendix C for some selected outputs of the simulation.

After running a total of 1,080 *MCMC* simulations, the results show that the multivariate potential scale reduction factor (MPSRF) is 0.99 and this indicates that the *MCMC sampler* has converged to a stationary distribution, since it is less than 1.1 as recommended by Gelman et al. (2004). As such, the results assumed to be drawn from the target probability distribution (lognormal pdf). On the other hand, the check for spatial dependence (i.e., variogram reproduction) based on 95% confidence level shows that T^2 = 1.751, whereas, its corresponding *F*-statistic value is 23.189. Since T^2 is less than the *F*-statistic, the null hypothesis can be safely rejected and it can be concluded that the variogram simulation for the study area remains valid.

Finally, the 1,000 simulated errors are subtracted from the original SRTM DEM of the study area to generate 1,000 realized equiprobable DEMs which serves as input to the solar radiation models. Appendix D presents examples of these realizations. The total computing time for executing the code is five hours and twenty minutes on an Intel (R) Core (TM) 2 CPU, T5600 @ 1.83 GHz, 987 MHz, and 2 GB of RAM.

The following sections utilized the realized DEMs as inputs to estimate solar irradiation raster maps and then DEM uncertainties on these outputs are computed.

5.4 GIS-Based Solar Radiation Model Simulation

In this section, the data used for both the *Solar Analyst* and *r.sun* models are discussed, and the models implementation using the 1,000 realized equiprobable DEMs from the previous section and other inputs is demonstrated.

DEM is a mandatory input data for both the *Solar Analyst* and *r.sun* solar radiation models. In this research, two types of DEMs are used as input to the models namely: (1) original

SRTM DEM of the study area, and (2) the 1,000 realized equiprobable DEMs from the *MCMC* simulation.

In the case of *Solar Analyst*, slope and aspect maps are generated from the input DEM automatically by the model and then utilized in calculating the solar radiation. On the other hand, the *r.sun* model requires separate slope and aspect maps as inputs. Therefore, these maps are calculated for both the original SRTM DEM of the study area and the 1,000 realized equiprobable DEMs using *r.slope.aspect* module available in GRASS GIS. Thus, for the original SRTM DEM of the study area, there is one raster map for both aspect and slope (Appendix E). Whereas, for the 1,000 realized equiprobable DEMs there are; 1,000 slope maps (Appendix F) and 1,000 aspect maps (Appendix G) generated. All these are used as inputs to the model.

The T_{LK}, is a major input of *r.sun* model. It "often normalized at an air mass = 2 to reduce its dependence on air mass and refers to the overall spectrally integrated attenuation, which includes the presence of gaseous water vapor and aerosols" (Nguyen and Pearce, 2010). It expresses the atmospheric turbidity or equivalently the attenuation of the direct solar radiation flux (Djafer and Irbah, 2013). The larger the T_{LK}, the lower the sky transparency and thus the higher the attenuation of the solar radiation (Nguyen and Pearce, 2010). Typical values of the T_{LK} varies between 1 and 10. High values of the T_{LK} mean that the solar radiations are more attenuated in a clear sky atmosphere (Djafer and Irbah, 2013). A worldwide database of the T_{LK} can be obtained from SoDa database. The SoDa's T_{LK} data is derived from radiation or aerosol field measurements (AERONET), and satellite sources, such as global clear sky radiation, perceptible water vapor, and aerosol optical depth (pathfinder). T_{LK} has been calculated with beam or global radiation measurements at the ground with the help of European Solar Radiation Atlas clear sky radiation model. Satellite and ground information have been fitted together (SoDa, 2012a) and the root mean square error (RMSE) is 0.73 T_{LK} units (Remund et al. 2003). In this study, the average monthly values of the T_{LK} used for the *r.sun* model is obtained from the SoDa database by SoDa (2012), see Table 5.5.

Table 5.5. Linke turbidity factor (SoDA, 2012)

Month	**Monthly average Linke turbidity factor (AM^2)**
January	6.7
February	6.2
March	6.8
April	4.8
May	6.8
June	6.8
July	5.4
August	4.7
September	5.8
October	5.8
November	6.6
December	6.7

Ground albedo or ground reflectance is also an important input of *r.sun* model. The term “ground albedo” is defined as the coefficient of reflection found in visible range of the spectrum, while “reflectance” connotes the reflected fraction of short-wave energy (Muneer and Tham, 2013). The ground albedo data is obtained from the NASA (2013) as point data.

It is a monthly average over the 22-year time frame (July 1983 - June 2005). Twelve raster layers for ground albedo are generated for 12 months by interpolation using the *s.vol.rst* module in GRASS GIS containing a tri-variate version of Regularized Spline with Tension (Neteler and Mitasova, 2004). To extract only the study area, each of the twelve rasters is clipped with the study area boundary polygon. However, examining the raster layers show that ground albedo in the study area is not varied; therefore, a constant value for each respective month is utilized in the calculation (Table 5.6).

Table 5.6. Ground albedo (Neteler and Mitasova, 2004)

Month	Monthly average ground albedo
January	0.180687
February	0.190578
March	0.205051
April	0.205992
May	0.209048
June	0.216052
July	0.223974
August	0.224589
September	0.218602
October	0.204639
November	0.194905
December	0.185265

Another important input to the *r.sun* model is the clear sky index (Kc). The Kc is calculated using different methods as follows: (1) ratio of global horizontal irradiation under clear sky condition to global horizontal irradiation under cloudy conditions is available (Hofierka and Šúri, 2002); (2) linear regression (Kasten and Czeplak, 1979) and (3) based on a cloud cover index derived from remote sensing data (Martins et al. 2007). However, in this study, the Kc data are obtained from the NASA (2013) as point data. It is a monthly average over the 22-year period (July 1983 - June 2005). Similar method used in generating the ground albedo data is also adopted here.

To avoid simulating 365 times for each day of the year (365 days), Klein's definition of mean day (KDMD) is utilized. The KDMD "defines the mean day of each month to be the day for which daily horizontal extraterrestrial irradiance is approximately the same with the mean monthly averages" (Nguyen and Pearce, 2010). Based on KDMD, only 12 simulations are carried out. Table 5.7 in Duffie and Beckman (1991) gives the specific day of the month and year, and sun declination (δ) values to supply as input to the models during the simulation (Nguyen and Pearce, 2010). Thus, both the *Solar Analyst* and *r.sun* models are executed based on the dates or day of year from Table 5.7.

Table 5.7. Monthly average days, dates and declinations (Duffie and Beckman, 1991)

Month	Date	Day of year (*n*)	Sun's declination (δ)
January	17	17	-20.9
February	16	47	-13.0
March	16	75	-2.4
April	15	105	9.4
May	15	135	18.8
June	11	162	23.1
July	17	198	21.2
August	16	228	13.5

Table 5.7. Monthly average days, dates and declinations (Duffie and Beckman, 1991) (continued)

Month	Date	Day of year (n)	Sun's declination (δ)
September	15	258	2.2
October	15	288	-9.6
November	14	318	-18.9
December	10	344	-23.0

The simulation of *GIS-based solar radiation models* is done for both the two models considered in this study. These models are *Solar Analyst* and *r.sun.*

The *Solar Analyst* model in ArcGIS is implemented using the original SRTM DEM of the study area, and each of the 1,000 realized equiprobable DEMs from the *MCMC* simulation. The diffuse model type of standard overcast sky is used together with other inputs, for example, elevation, slope and aspect computed from the input DEM, and latitude. Default values of 0.3 and 0.5 are used for the diffuse proportion and transmissivity, respectively. In terms of time configuration, the model is run for the whole year of 2012 based on Klein's mean day (Table 5.7). Clear sky conditions are applied in these calculations which, for the selected site, are a reasonable approximation. In total, 1,225 data points are used in the study area. The raster outputs generated from this model comprised of four different types, global radiation (Wh/m^2), direct radiation (Wh/m^2), diffuse radiation (Wh/m^2), and direct duration (hours). In the case of original SRTM DEM of the study area, there is a total of four (4) outputs only (Figures 6.23, 6.31, 6.39, 6.47), whereas, for the 1,000 realized equiprobable DEMs from the *MCMC* simulation, there are four thousand (4,000) output, i.e., 1,000 for each respective output type (Appendix H). Similarly, these outputs are imported to MaTLAB® for further analysis.

From the 2 operational modes available in *r.sun*, Mode 2 computes the sum of solar radiation for a specific day and for selected components of solar radiation which comprised of, beam, diffuse, reflected, and global radiations, and direct beam duration. Thus, all these components are selected when executing the model. The first step is specifying the DEM (elevation raster map), aspect and slope raster maps. The DEMs used for analysis are; original SRTM DEM of the study area and each of the 1,000 realized equiprobable DEMs from the *MCMC* simulation. On the other hand, the aspect and slope maps calculated previously are used. Second, the Klein's mean day (Table 5.7) is used for each respective month (i.e., January - December). Third, the values of T_{LK} (Table 5.5) and ground albedo (Table 5.6) is defined for each month accordingly. Fourth, the shadowing effect of terrain is incorporated. In addition, default options for some of the optional parameters such as time step (0.5) and sampling distance step coefficient (1.0) are maintained. Moreover, the analysis is done for clear-sky condition. Finally, the names of the output raster maps are specified. These outputs comprised of beam irradiation (Wh/m^2), diffuse irradiation (Wh/m^2), insolation time (hour), ground reflected irradiation (Wh/m^2), and ground (total) irradiation (Wh/m^2). All the outputs are produced in floating point, and as such both export and import followed the same format. In total, for the original SRTM DEM of the study area there are sixty output raster maps. However, this is reduced to five outputs by using the raster calculator to sum the monthly outputs of (January - December) to generate one single year's output for each of the five solar radiation components. Four of these outputs are presented in Figures 6.55, 6.63, 6.71, 6.79. On the other hand, the outputs for the 1,000 realized DEMs is 60,000 raster maps, i.e., 12,000 raster maps for each of the respective five outputs. Similarly, these outputs are reduced to 5,000 raster maps using the map algebraic tool (Appendix I). The 5,000 raster maps are exported as tiff format for further analysis in MaTLAB®.

Finally, using the annually aggregated outputs of both the *Solar Analyst* and *r.sun* models generated above, the mean, standard deviation, coefficient of variation (CoV), and ratios are computed to provide a measure of uncertainty. The mean for each of the solar radiation component outputs is calculated using Equation (4.16), whereas, Equations (4.17) and (4.18) are used for computing standard deviation, and CoV, respectively. The results of these analyses are discussed extensively in Chapter 6.

CHAPTER 6

RESULTS AND DISCUSSION

In this Chapter, the implementation results of the proposed methodology are presented and discussed. The Chapter begins with assessment of uncertainty propagation in the digital elevation model (DEM), slope and aspect. The second part covers DEM uncertainty propagation on the *Geographic Information System (GIS)-based solar radiation model* outputs. Finally, the Chapter concludes with a comparison of *Solar Analyst* and *r.sun* models outputs.

6.1 Uncertainty Propagation in DEM

Figure 6.1 represents the mean of the realized DEMs in which the lowest and highest elevations are 472.5 m and 725.9 m, respectively. In comparison to Figures 5.4 and 5.5, it can be noted that similar landform patterns: a mountain in the North-East and some hills in the North, North-West, central part and South-East, while in the South-West a low land is observed. However, the major important differences observed is that the minimum elevation in the mean realized DEMs is lower than the corresponding minimum elevation from Shuttle Radar Topography Mission (SRTM) DEM of the study area by 4.5 m, whereas, the maximum elevation is lower by 4.1 m. Figure 6.2 shows the mean of 549.02 m and standard deviation of 56.55 m.

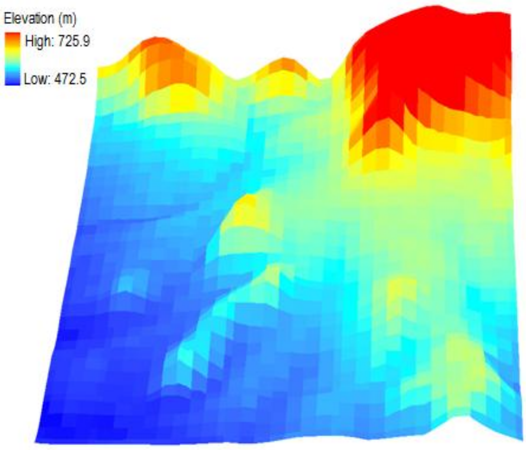


Figure 6.1. Mean of realized DEMs

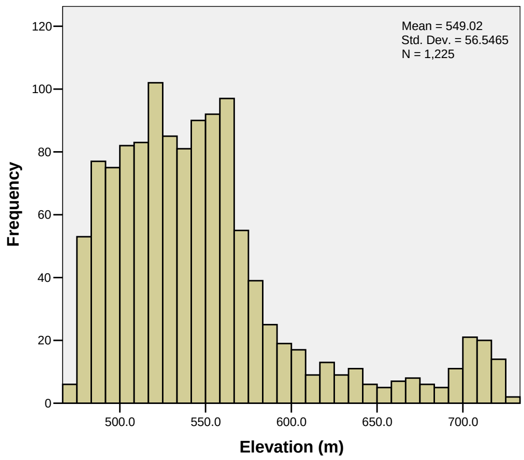


Figure 6.2. Histogram for mean of realized DEMs

The standard deviation map gives an idea of the local uncertainty associated with the computation of disparity (Senegas et al. 2006). The standard deviation of elevation error increases from 0.05 to 0.56 and mean of 0.15 (Figures 6.3 and 6.4). Figure 6.5 shows the coefficient of variation (CoV) map which is the ratio of the standard deviation to the mean of the realized DEMs multiplied by 100. It can be noted that the highest variability reaches up to 0.12% while the lowest is 0.01%. Pixels with higher CoV indicate greater elevation variability and vice-versa. The corresponding histogram (Figure 6.6) indicates the mean of 0.03%. As expected, the elevation error would increase as the surface slope increases based on fundamental research in topographic mapping (Maling, 1989). The findings of this study indicate that majority of the higher standard deviation of the elevation errors can be observed in steeper slopes. Thus, the results are consistent with the previous findings, for instance, Wechsler and Kroll (2006) who noted that elevation error is manifested in steeper slopes. However, the CoV also indicates that high standard deviation do not only occur in steep slopes but also in the valleys. This is very important because knowing error magnitudes and their spatial distribution helps identification of places requiring more consideration during the decision making (Gonga-Saholiariliva et al. 2011). Moreover, the finding shows that the local variability of errors is far better than the global values reported in the form of root mean squared error (RMSE) which dominates DEM uncertainty literature in the past decades.

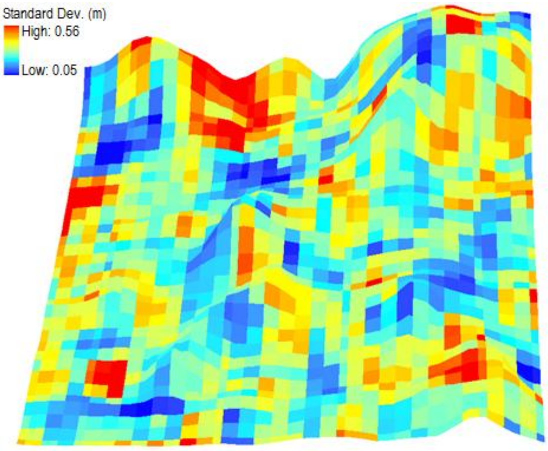


Figure 6.3. Standard Deviation for Mean of realized DEMs

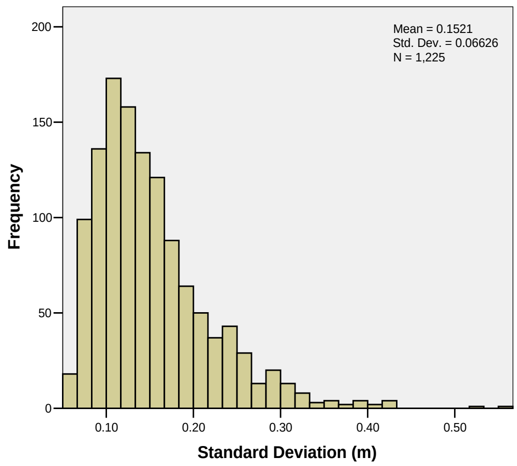


Figure 6.4. Histogram of Standard Deviation for Mean of realized DEMs

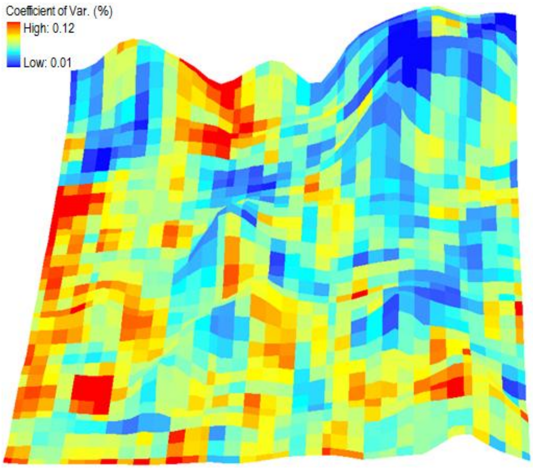


Figure 6.5. CoV for Mean of realized DEMs

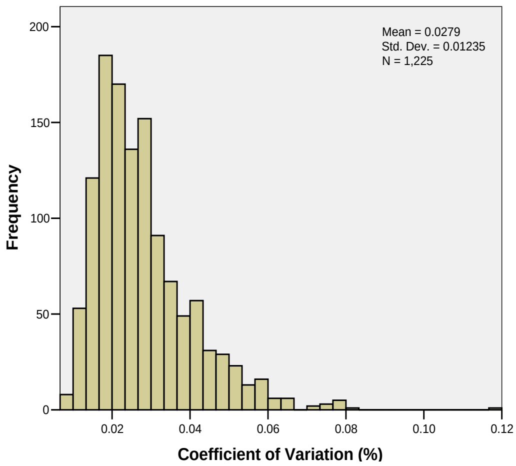


Figure 6.6. Histogram of CoV for Mean of realized DEMs

6.2 Uncertainty Propagation in Slope

The slope maps are derived using ArcGIS 9.3® (Appendix J). Figure 6.7 shows the slope derived from the SRTM DEM of the study area. The slope values range between 0° and 26.6°. The mean slope is 5.51° while 4.95° represent the standard deviation (Figure 6.8). On the other hand, Figure 6.9 presents the mean slope of the realized DEMs. The minimum slope in Figure 6.9 is 0.12° and the maximum is 26.58°. As observed from Figure 6.10, the mean slope of the realized DEMs has a mean slope value of 5.52° and standard deviation of 4.95°. The standard deviation map and its corresponding histogram (Figures 6.11 and 6.12) show that a slope error standard deviation ranges from 0.23° in steeper slopes to 16.87° in flat surfaces. This is more noticeable from the CoV map and its corresponding histogram (Figure 6.13 and 6.14) in which the CoV lies between 0.09% in steeper areas to 6.78% in flat lands. Thus, on the spatial distribution of slope errors, the results of this study revealed that slope errors are more pronounced in flat terrain than steeper areas. The outcome of this study aligns with those of Carter (1992); Vieux (1993); and Zhou et al. (2006) who reported that slope errors generally are more prominent in flat surfaces and not consistent with Chang and Tsai (1991) who stated that slope errors are mainly concentrated in areas of steep slopes. The results can be explained due the fact that as the more gentle slopes are eliminated from consideration, the frequency distribution tended to be more uniform (Carter, 1992).

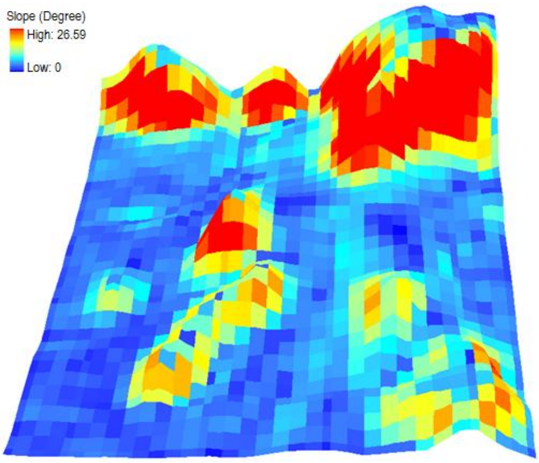


Figure 6.7. Slope for SRTM DEM of the study area

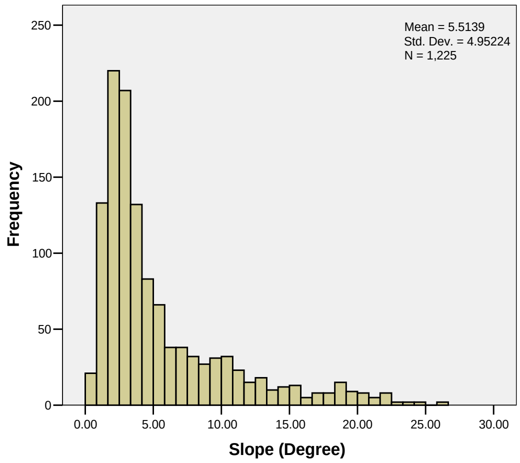


Figure 6.8. Histogram of slope for SRTM DEM of the study area

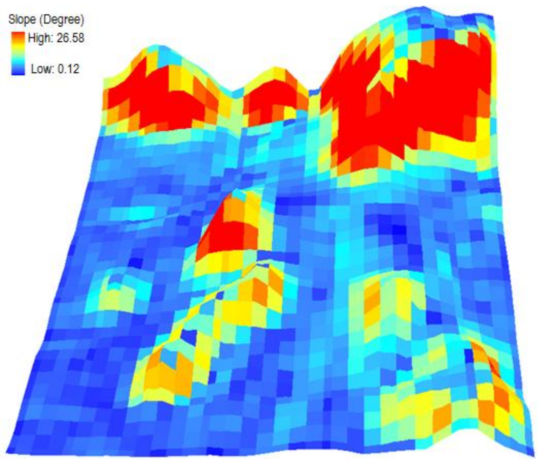


Figure 6.9. Mean slope of realized DEMs

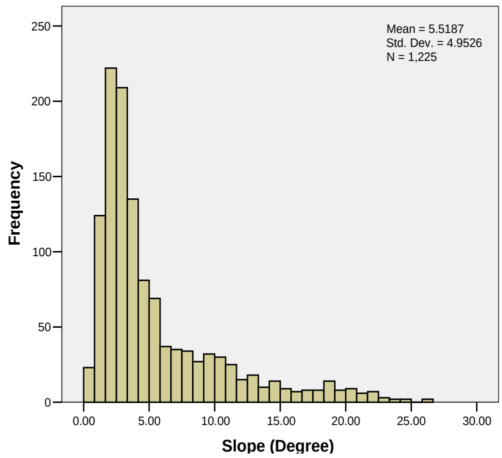


Figure 6.10. Histogram for mean slope of realized DEMs

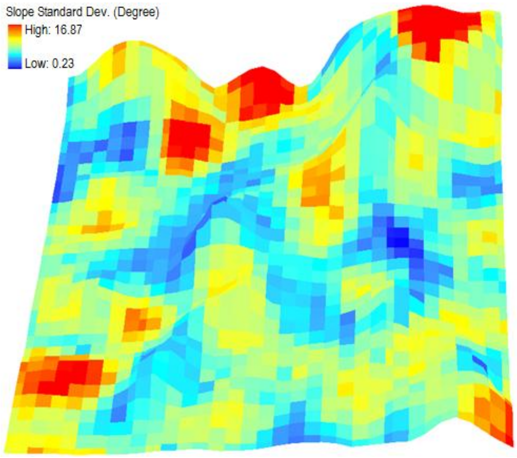


Figure 6.11. Standard deviation for mean slope of realized DEMs

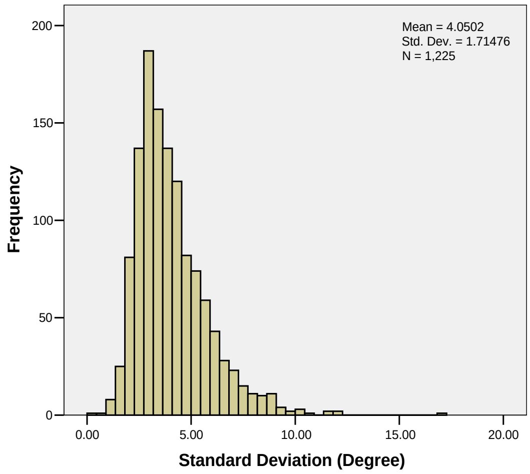


Figure 6.12. Histogram of standard deviation for mean slope of realized DEMs

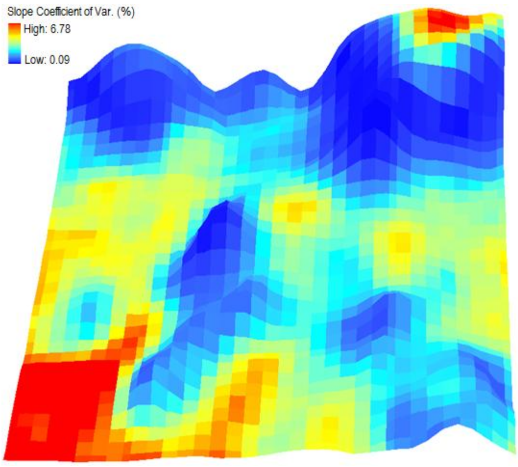


Figure 6.13. CoV for mean slope of realized DEMs

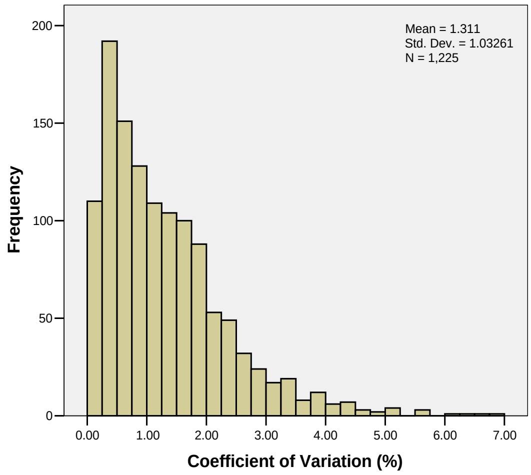


Figure 6.14. Histogram of CoV for mean slope of realized DEMs

6.3 Uncertainty Propagation in Aspect

The aspect maps are derived using ArcGIS 9.3® (Appendix K). The aspect map derived from SRTM DEM of the study area has minimum and maximum value of -1^{o} and 358.26^{o}, respectively (Figure 6.15). It can be observed from Figure 6.16 that the mean aspect for the original SRTM DEM of the study area is 206.52^{o}, while 69.33^{o} represent the standard deviation. Figure 6.17 which have a minimum of 2.6^{o} and maximum of 357^{o}, represents the mean aspect of the realized DEMs. Based on Figure 6.18, 206.95^{o} and 68.91^{o} represent the mean and standard deviation, respectively for the mean aspect of the realized DEMs. Figure 6.19 gives the spatial distribution of standard deviation of aspect derived from the realized DEMs. In Figure 6.20, 7.96^{o} and 8.54^{o} represent mean and standard deviation, respectively. It can be observed that there are higher standard deviations on flat terrains than undulating one. The highest CoV value reaches up to 7.6% on flat terrains as compared to 0% on undulating terrains (Figure 6.21). In the corresponding histogram (Figure 6.22), mean (0.46) and a standard deviation (0.61) are observed. The findings of this study confirm the assertion published in the works of Carter (1992); Chang and Tsai (1991); and Zhou et al. (2006) who noted that the concentration of aspect error is quite high at flat terrains. From the corresponding histogram (Figure 22), a mean of about 0.56% and a standard deviation of 0.61 is observed.

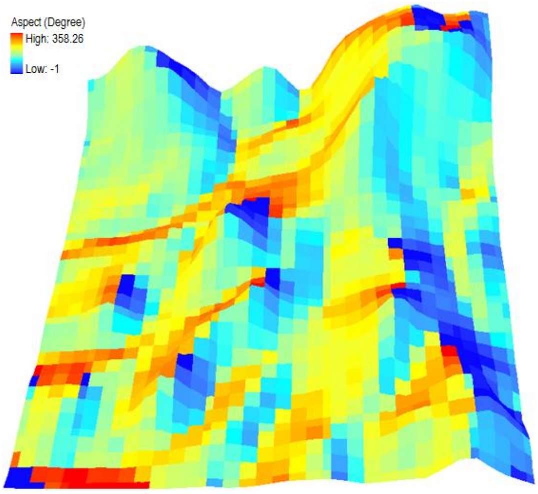


Figure 6.15. Aspect for SRTM DEM of the study area

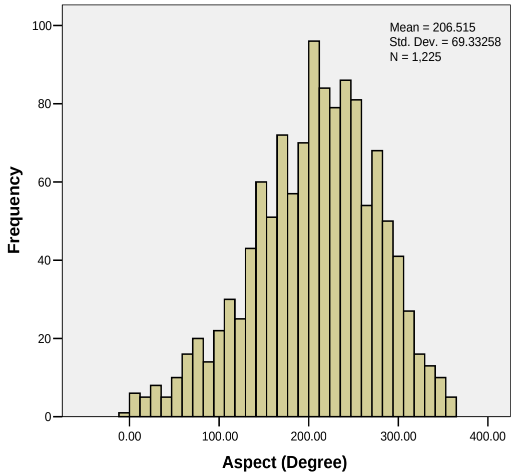


Figure 6.16. Histogram of aspect for SRTM DEM of the study area

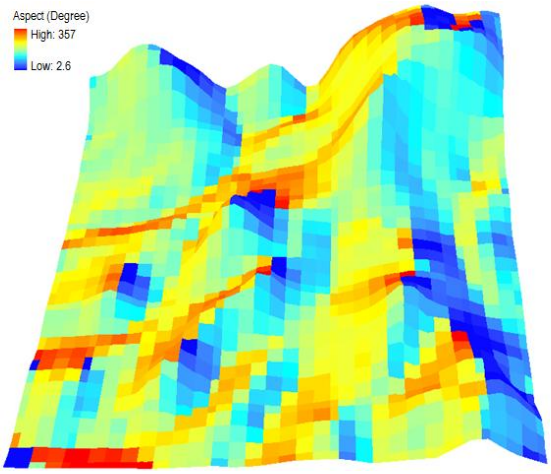


Figure 6.17. Mean aspect of realized DEMs

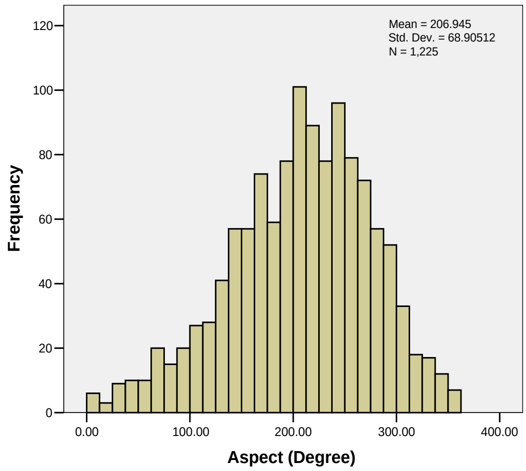


Figure 6.18. Histogram for mean aspect of realized DEMs

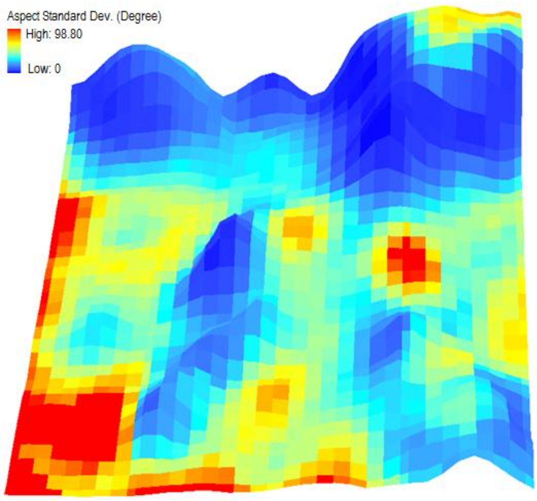


Figure 6.19. Standard deviation for aspect of realized DEMs

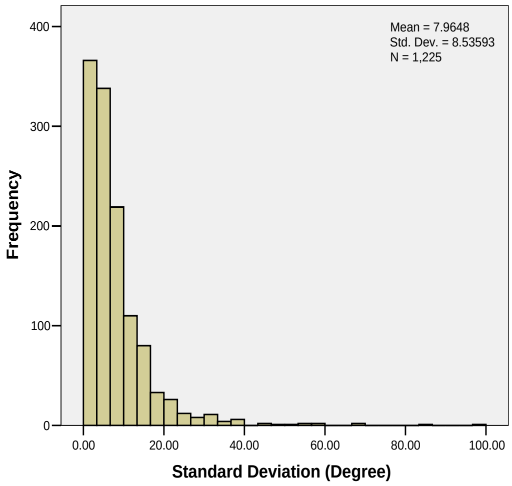


Figure 6.20. Histogram of standard deviation for aspect of realized DEMs

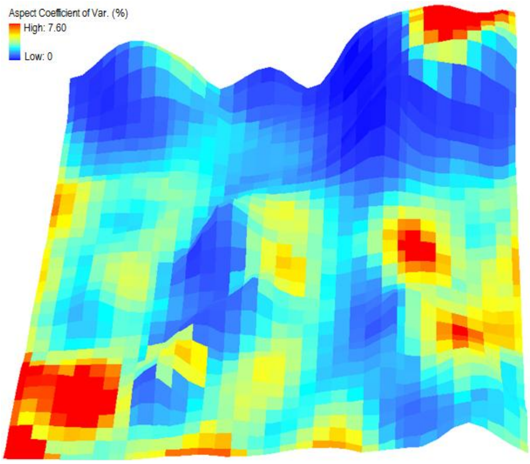


Figure 6.21. CoV for aspect of the realized DEMs

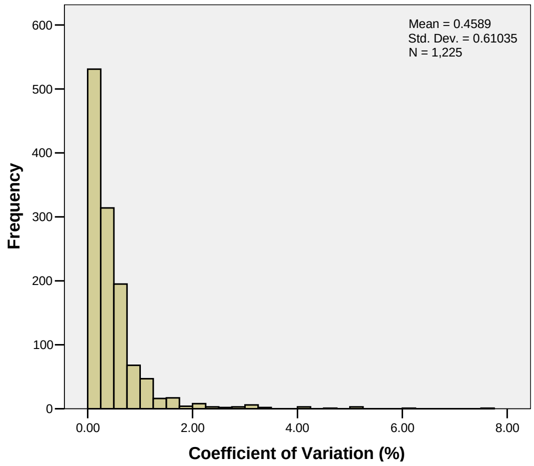


Figure 6.22. Histogram for CoV of realized DEMs

6.4 Uncertainty Proportion in *Solar Analyst* Outputs

The direct radiation derived from the SRTM DEM (Figure 6.23) and mean direct radiation of realized DEMs (Figure 6.25) shows similar spatial distribution patterns. Higher terrains that are relatively flat produces the highest amount of direct radiation estimates than lower areas, valleys and steep slopes. However, it is observed that the original SRTM DEM of the study area overestimated the direct radiation in which the highest value reaches up to 49,922.0 Wh/m^2/year while the lowest value is 43,921.3 Wh/m^2/year. The mean direct radiation of the realized DEMs has 49,898.0 Wh/m^2/year and 43,891.4 Wh/m^2/year as maximum and minimum values, respectively. Figure 6.24 indicates that the direct radiation derived from the SRTM DEM has a mean of 48,160.68 Wh/m^2/year and standard deviation of 598.34 Wh/m^2/year. On the other hand, Figure 6.26 reveals that 48,138.46 Wh/m^2/year represents the mean, while standard deviation is 599.03 Wh/m^2/year for the mean of the realized DEMs. Figures 6.27 and 6.29 shows similar patterns, where hill tops that are relatively flat exhibit high error variations. However, an exception is noticed from the south-western part of the maps which is a low land but shows a similar pattern to the highlands.

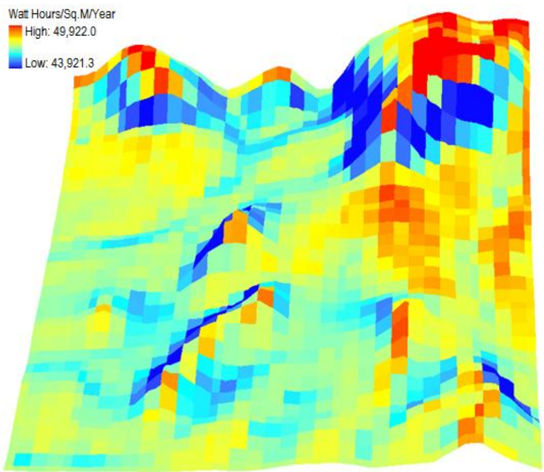


Figure 6.23. Direct radiation from the original SRTM DEM of the study area

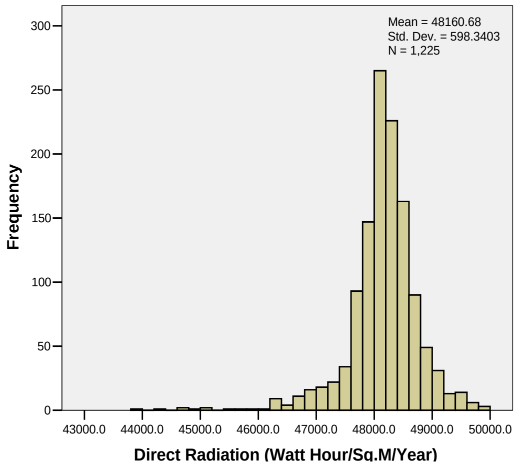


Figure 6.24. Histogram for direct radiation of SRTM DEM of the study area

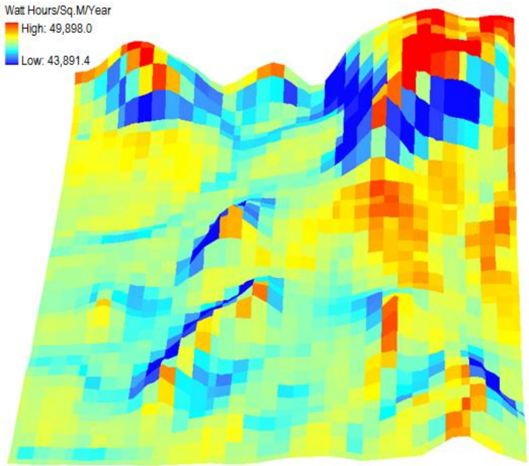


Figure 6.25. Mean direct radiation of realized DEMs

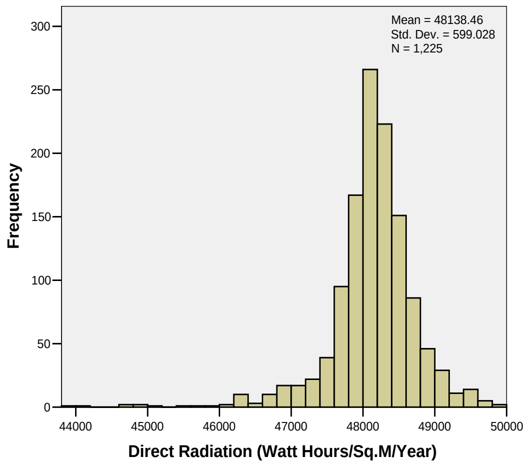


Figure 6.26. Histogram for mean direct radiation of realized DEMs

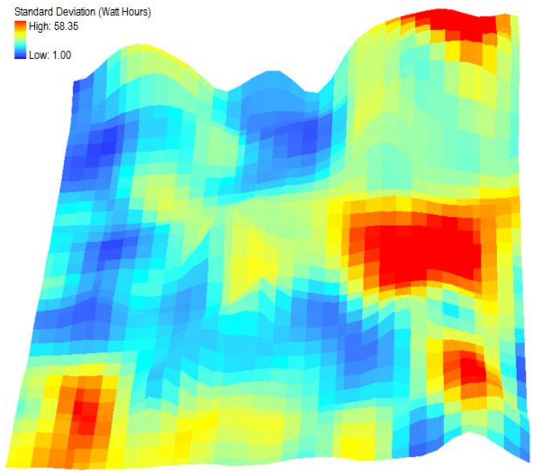


Figure 6.27. Standard deviation for mean direct radiation of realized DEMs

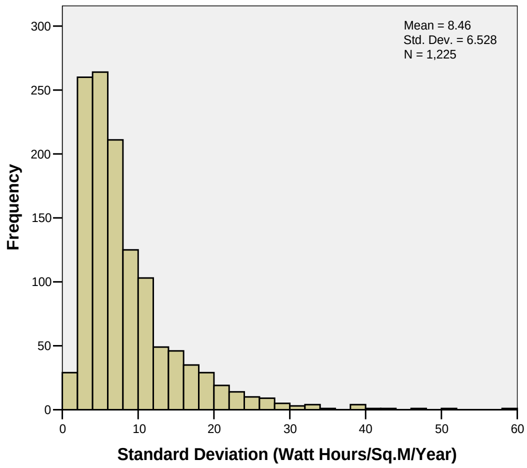


Figure 6.28. Histogram for standard deviation for mean direct radiation of realized DEMs

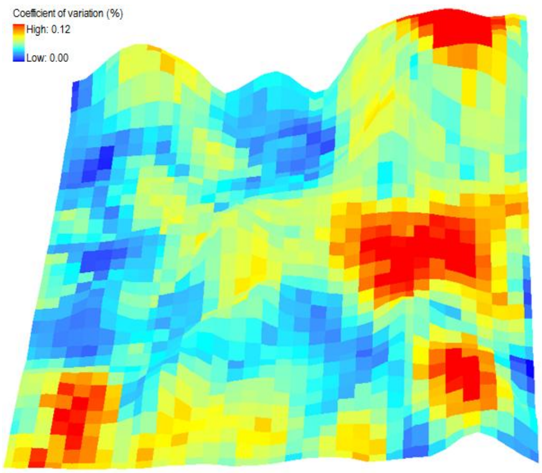


Figure 6.29. CoV for mean direct radiation of realized DEMs

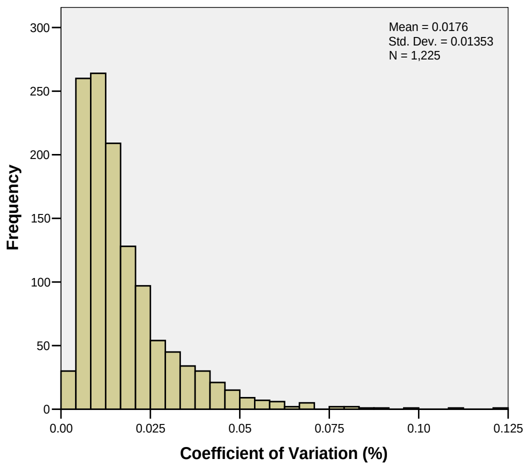


Figure 6.30. Histogram of CoV for mean direct radiation of realized DEMs

The diffuse radiation derived from the SRTM DEM (Figure 6.31) and mean diffuse radiation of realized DEMs (Figure 6.33) also show similar spatial distribution patterns in which the estimated values of diffuse radiation shows a significant linear increase with elevation. In other word, higher terrain surfaces produced higher amount of diffuse radiation when compared to lower elevation terrains, valleys and steep slopes. Comparing the two diffuse radiation maps, one can observe that the original SRTM DEM of the study area slightly overestimated the diffuse radiation as the highest value is 14,189.7 Wh/m^2/year and the lowest value is 13,245.2 Wh/m^2/year. Whereas, the mean direct radiation of the realized DEMs has 14,182.9 Wh/m^2/year and 13,241.6 Wh/m^2/year as maximum and minimum values, respectively. The corresponding histogram of diffuse radiation maps shown in Figure 6.32 indicates that the diffuse radiation generated from the SRTM DEM has a mean of 13,756.07 Wh/m^2/year and standard deviation of 126.30 Wh/m^2/year. For mean of the realized DEMs (Figure 6.34), it is observed that 13,749.08 Wh/m^2/year represent the mean and 126.25 Wh/m^2/year constitute the standard deviation. The results of Figures 6.35 and 6.37 reveal that higher standard deviation of diffuse radiation error mainly occur on steep slopes. The corresponding standard deviation and CoV histograms are shown in Figures 6.36 and 6.38, respectively.

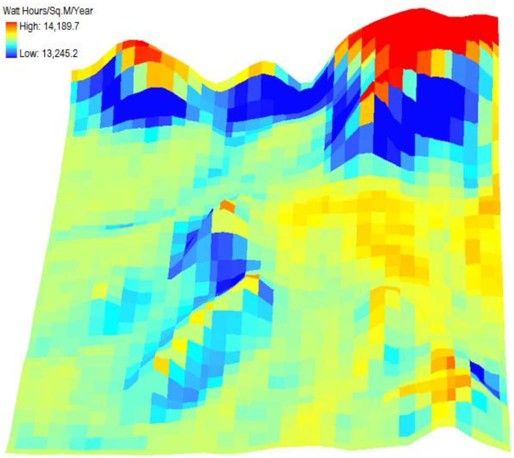


Figure 6.31. Diffuse radiation from the original SRTM DEM of the study area

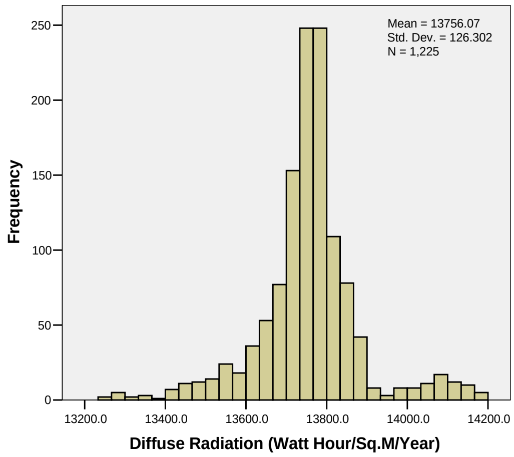


Figure 6.32. Histogram for diffuse radiation from the original SRTM DEM of the study areas

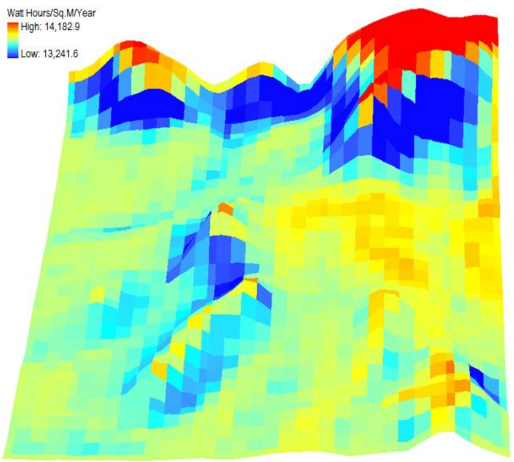


Figure 6.33. Mean diffuse radiation of realized DEMs

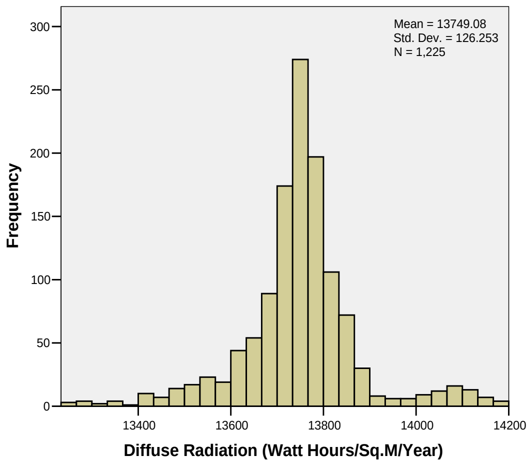


Figure 6.34. Histogram for mean diffuse radiation of realized DEMs

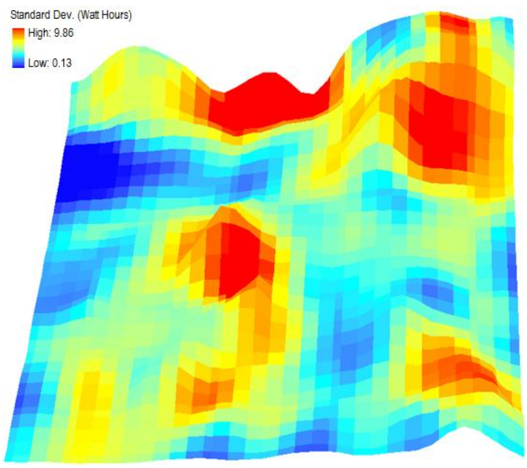


Figure 6.35. Standard deviation for mean diffuse radiation of realized DEMs

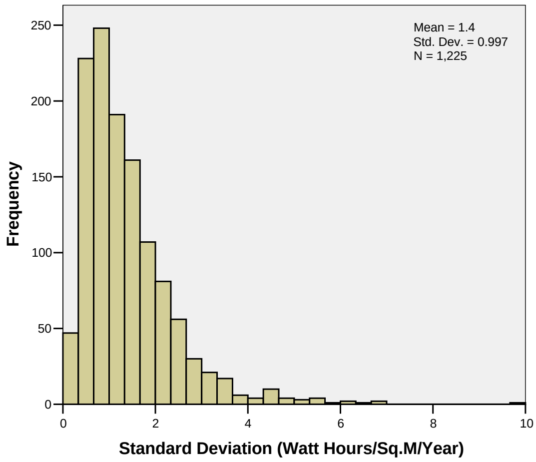


Figure 6.36. Histogram for standard deviation for mean diffuse radiation of realized DEMs

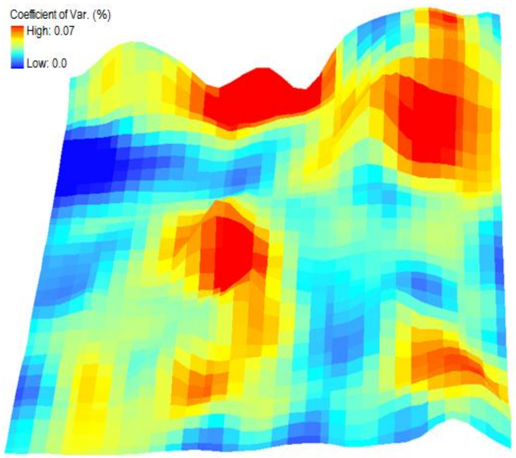


Figure 6.37. CoV for mean diffuse radiation of realized DEMs

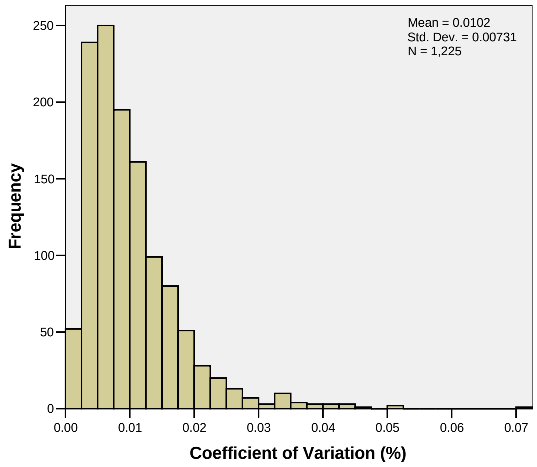


Figure 6.38. Histogram of CoV for mean diffuse radiation of realized DEMs

The spatial distribution patterns observed in direct and diffuse radiation maps is also noticeable in the global radiation maps (Figures 6.39 and 6.41). Higher elevation terrains produced higher values of global radiation than the lower elevation surfaces, valleys and steep slopes. The similarities of the results may be attributed to the fact that global radiation is the summation of direct and diffuse radiations. Assessment of the two global radiation maps reveals that the original SRTM DEM of the study area slightly overestimated the global radiation in which the highest value reaches 64,083.1 Wh/m^2/year and 57,209.5 Wh/m^2/year represent the lowest value. On the other hand, the mean global radiation of the realized DEMs has 64,052.6 Wh/m^2/year and 57,172.7 Wh/m^2/year as maximum and minimum values, respectively. The corresponding histograms of global radiation maps indicated in Figure 6.40 reveal that the global radiation derived from the SRTM DEM has a mean of 61,916.75 Wh/m^2/year and standard deviation of 698.98 Wh/m^2/year, while Figure 6.42 shows that the mean and standard deviation of the mean of the realized DEMs are 61,887.54 Wh/m^2/year and 699.69 Wh/m^2/year, respectively. Figures 6.43 and 6.45 shows similar spatial patterns where hill tops that are relatively flat exhibit high variations of error. However, the south-western part of the maps which is a low land but exhibited a similar pattern to the highlands is an exception. Figures 6.44 and 6.46 represent the global radiation corresponding histograms for standard deviation and CoV, respectively. In Figure 6.44, the mean is 9.55 Wh/m^2/year and standard deviation is 7.16 Wh/m^2/year. The mean of the CoV is 0.02, whereas, 0.01 represent the standard deviation (Figure 6.46).

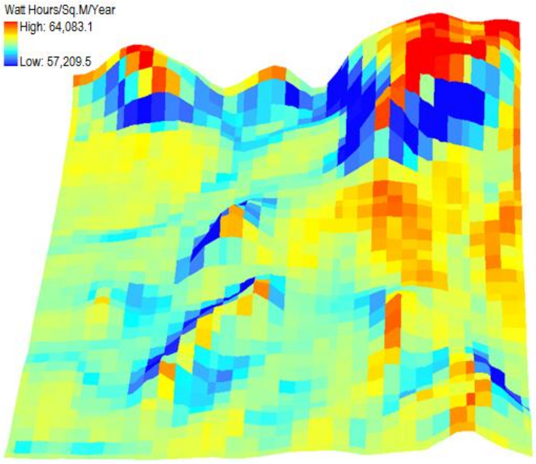


Figure 6.39. Global radiation from the original SRTM DEM of the study area

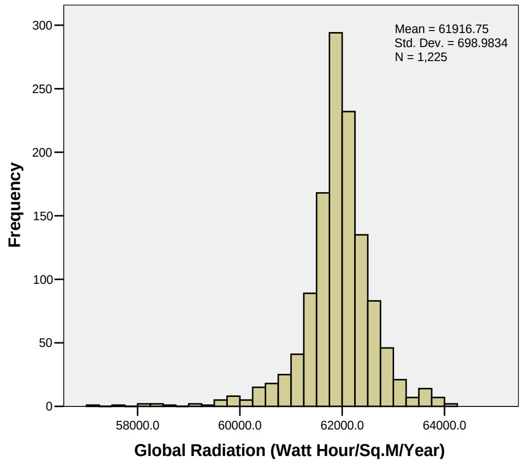


Figure 6.40. Histogram for global radiation from the original SRTM DEM of the study area

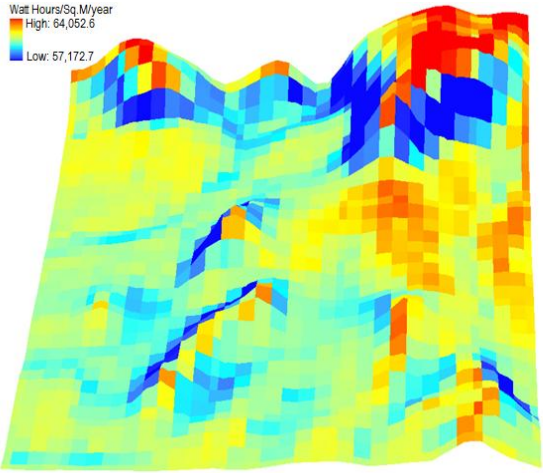


Figure 6.41. Mean global radiation of realized DEMs

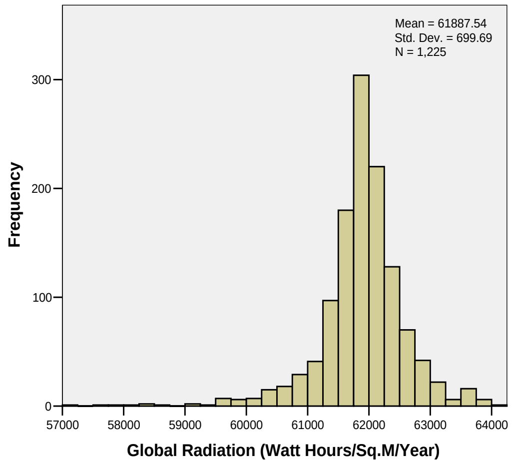


Figure 6.42. Histogram for mean global radiation of realized DEMs

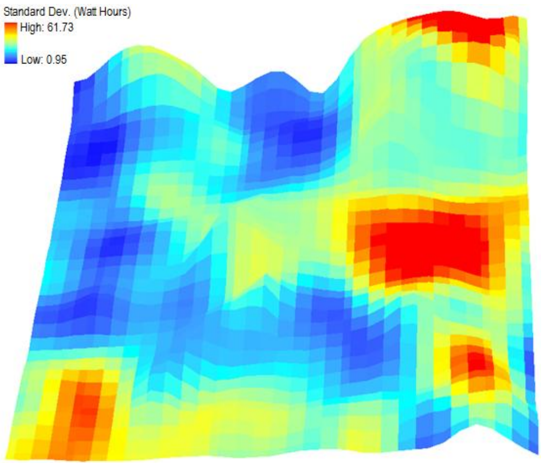


Figure 6.43. Standard deviation for mean global radiation of realized DEMs

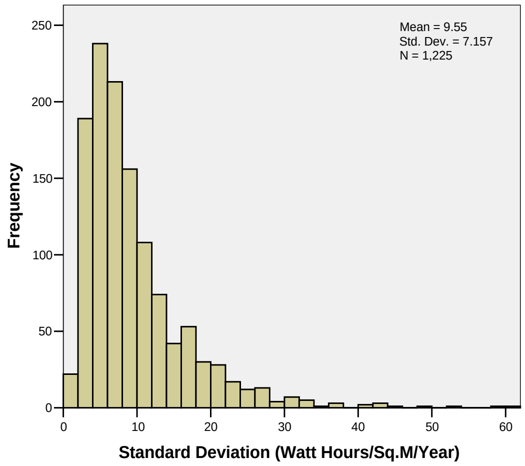


Figure 6.44. Histogram for standard deviation formean global radiation of realized DEMs

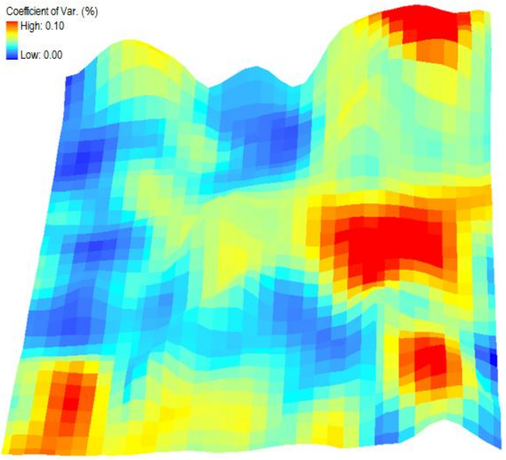


Figure 6.45. CoV for mean global radiation of realized DEMs

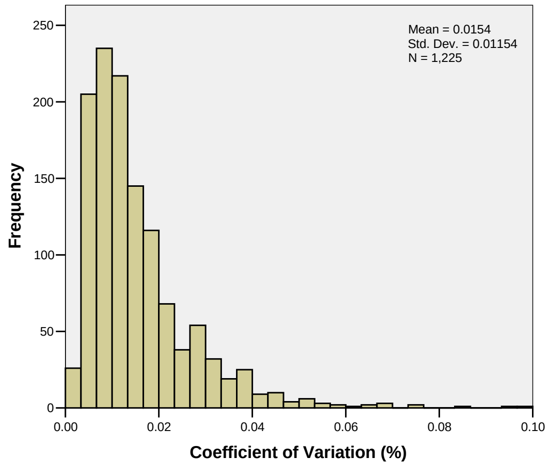


Figure 6.46. Histogram of CoV for mean global radiation of realized DEMs

Figures 6.47 and 6.49 shows the direct duration derived from the SRTM DEM and mean direct radiation of realized DEMs, respectively. These maps show similar spatial distribution patterns in both highlands and lowlands that are relatively flat and produced the highest amount of direct duration when compared to the valleys and steep slopes. In Figure 6.47, it can be observe that 143.30 hours/m^2/year and 121.00 hours/m^2/year represent the maximum and minimum estimates, respectively. The result of Figure 6.49 indicates that 120.92 hours/m^2/year (highest) and 143.13 hours/m^2/year (lowest) represent the mean direct duration of the realized DEMs. The corresponding histograms of direct duration maps in Figure 6.48 show that the direct duration derived from the SRTM DEM has a mean value of 138.16 hours/m^2/year and standard deviation of 3.84 hours/m^2/year, while Figure 6.50 shows that the mean (138.01 hour/m^2/year) and standard deviation (3.84 hour/m^2/year) represent the mean of the realized DEMs. The standard deviation and CoV are presented in Figures 6.51 and 6.53, respectively. The Figures reveal that standard deviations in higher elevation surfaces and some steep slopes are more pronounced than the lowlands and valleys.

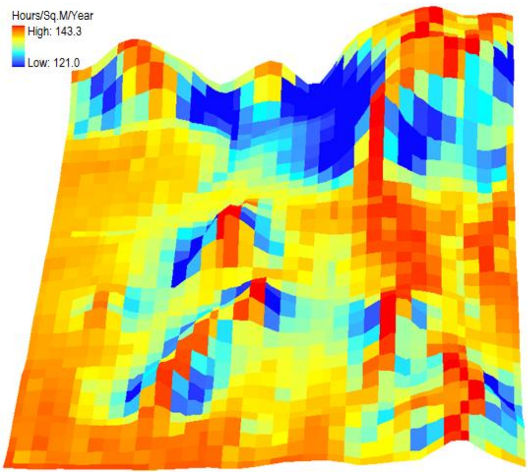


Figure 6.47. Direct duration from SRTM DEM of the study area

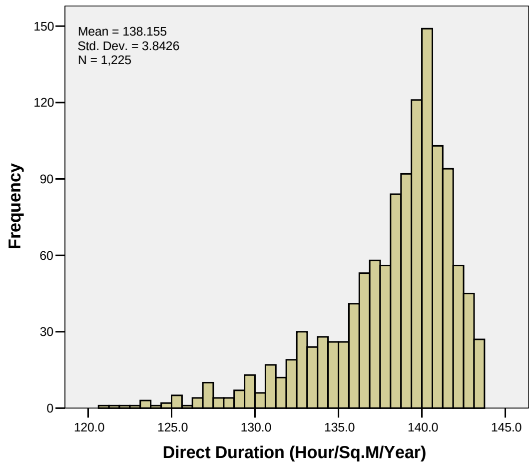


Figure 6.48. Histogram for direct duration from SRTM DEM of the study area

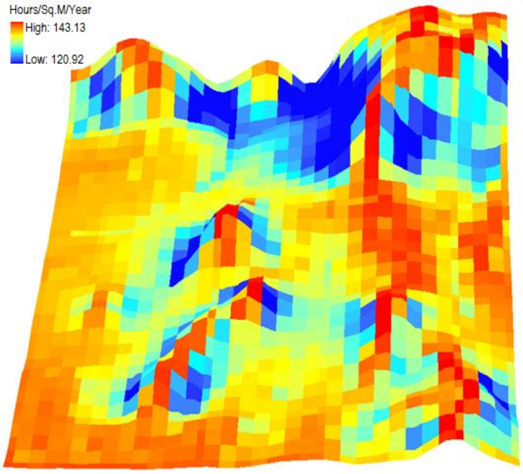


Figure 6.49. Mean direct duration of realized DEMs

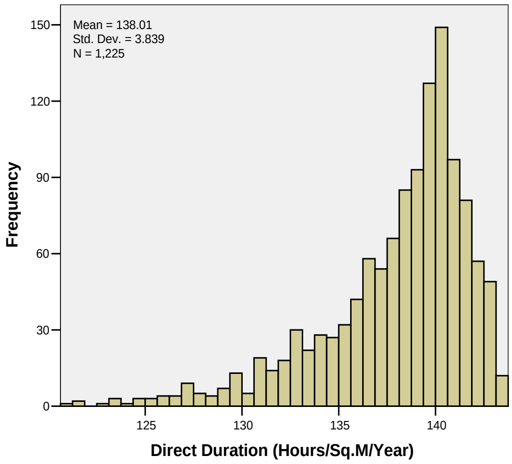


Figure 6.50. Histogram for mean direct duration of realized DEMs

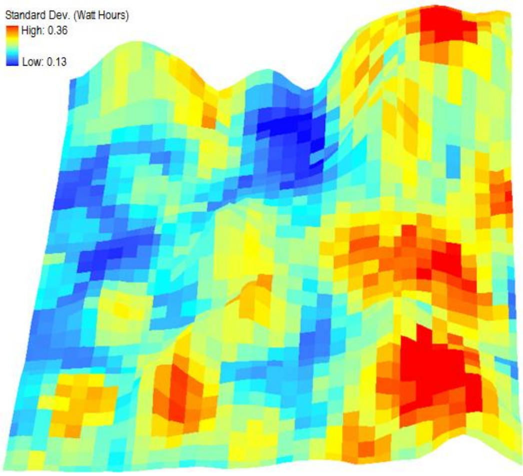


Figure 6.51. Standard deviation for mean direct duration of realized DEMs

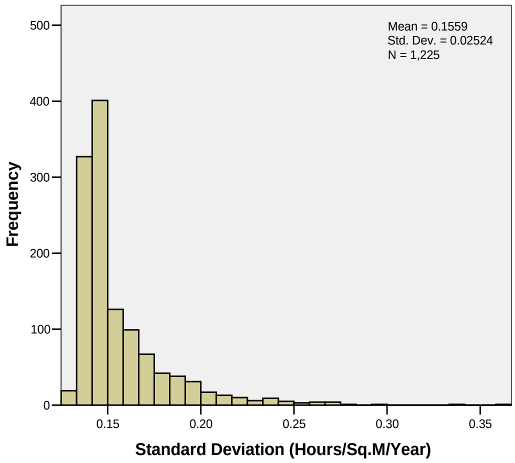


Figure 6.52. Histogram of standard deviation for mean direct duration of realized DEMs

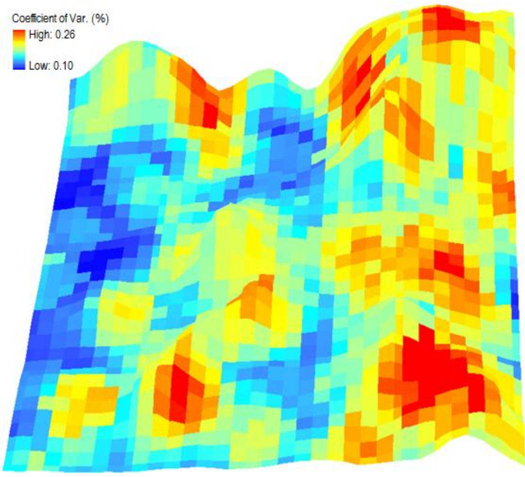


Figure 6.53. CoV mean direct duration of realized DEMs

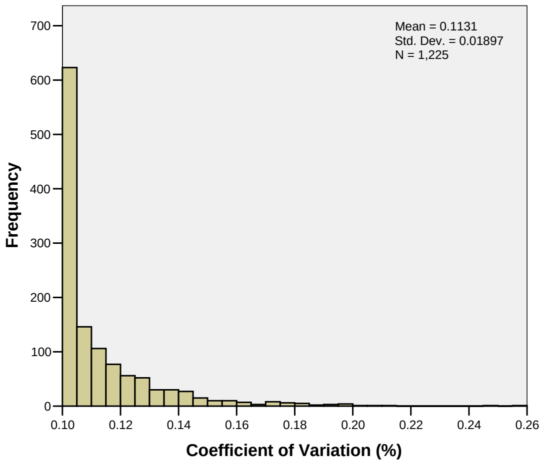


Figure 6.54. Histogram of CoV of mean direct duration of realized DEMs

6.5 Uncertainty Proportion in *r.sun* Outputs

Figures 6.55 and 6.57 report the *r.sun's* direct radiation outputs derived from the study area SRTM DEM and the mean of realized DEMs, respectively. By examining these two maps, one may observe that they both display similar spatial pattern in which relatively flat terrains account for higher values than the steep slopes. However, 49,305.50 Wh/m^2/year and 55,161.00 Wh/m^2/year represent the minimum and maximum values in Figure 6.55. While in Figure 6.57, minimum value is 49,320.10 Wh/m^2/year and the maximum value reaches up to 55,146.50 Wh/m^2/year. The mean direct radiation as observed from Figure 6.56 is 54,001.06 Wh/m^2/year while 604.27 Wh/m^2/year represents the standard deviation. In the case of mean realized DEMs (Figure 6.58), 53,987.04 Wh/m^2/year and 604.79 Wh/m^2/year stand for the mean and standard deviation, respectively. The results of Figure 6.59 representing the standard deviation map, and Figure 6.61 as CoV reveals that the direct radiation errors occur in areas that are relatively flat in nature and some west-facing slopes. The corresponding histograms of standard deviation and CoV are shown in Figures 6.60 and 6.62, respectively.

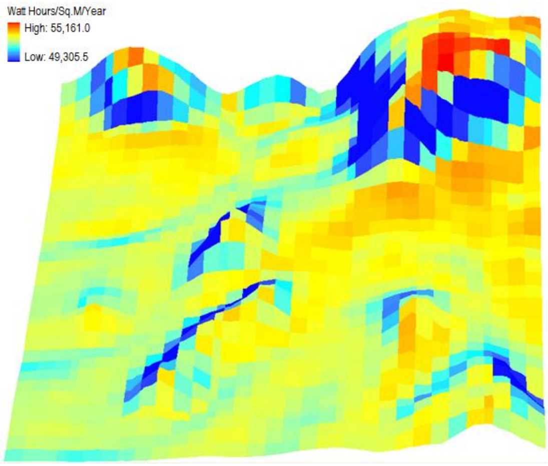


Figure 6.55. Direct radiation from the original SRTM DEM of the study area

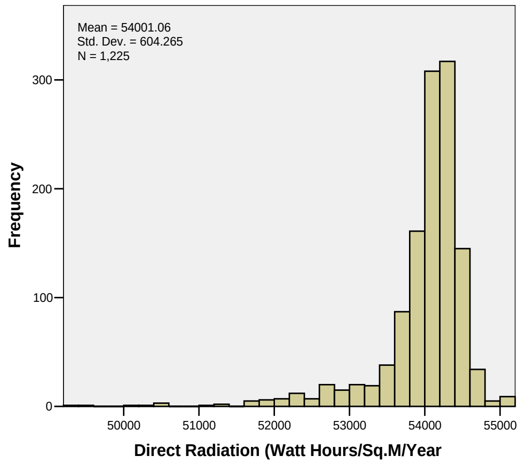


Figure 6.56. Histogram for direct radiation from the original SRTM DEM of the study area

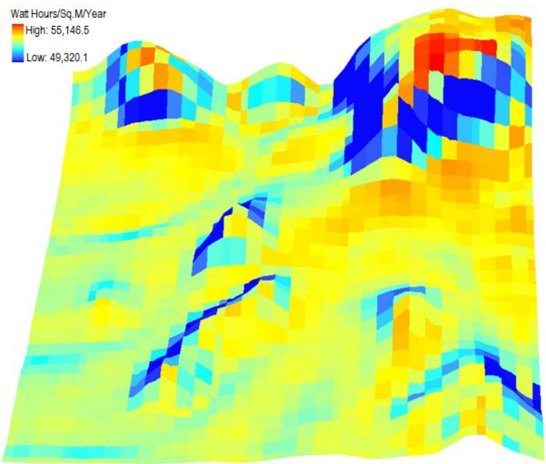


Figure 6.57. Mean direct radiation of realized DEMs

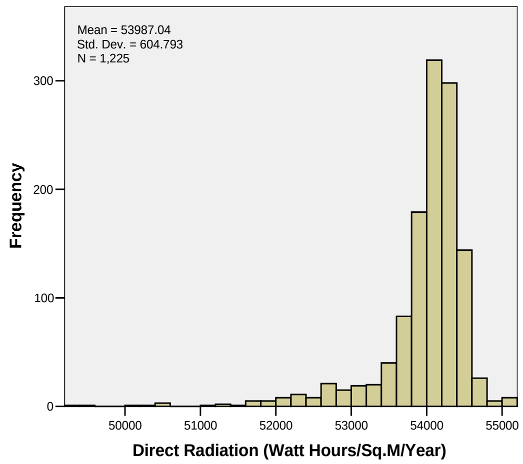


Figure 6.58. Histogram of mean direct radiation of realized DEMs

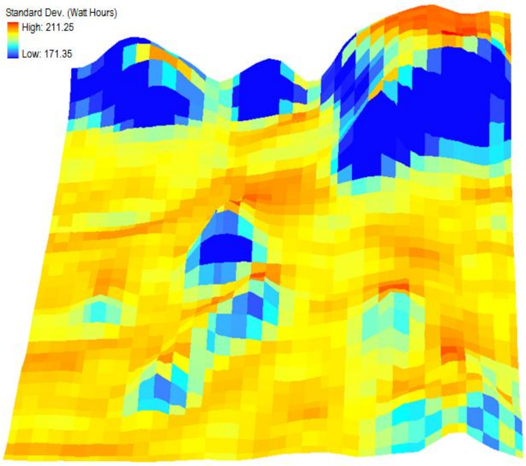


Figure 6.59. Standard deviation for mean direct radiation of realized DEMs

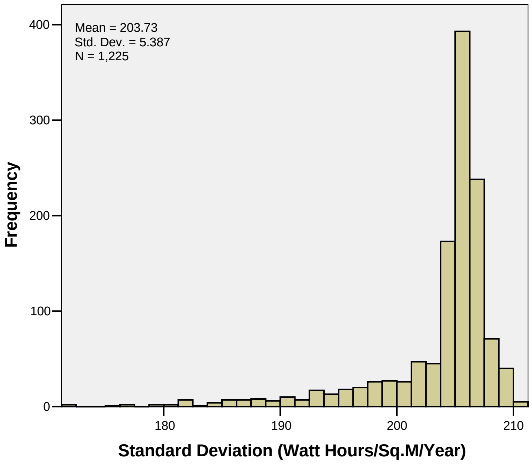


Figure 6.60. Histogram of standard deviation for mean direct radiation of realized DEMs

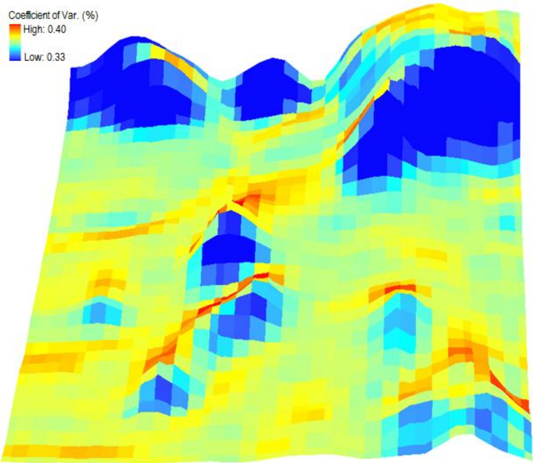


Figure 6.61. CoV for mean direct radiation of realized DEMs

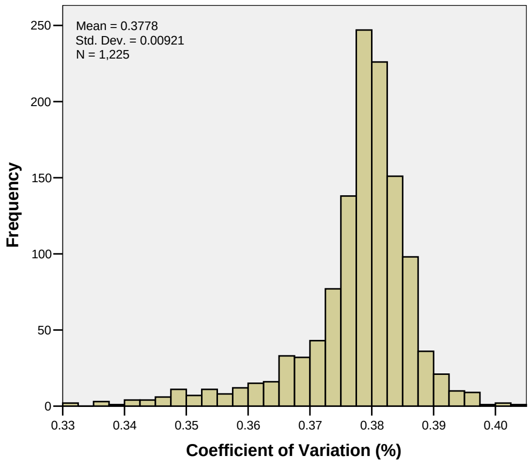


Figure 6.62. Histogram for CoV for mean direct radiation of realized DEMs

The outputs of diffuse radiation derived from the SRTM DEM and mean diffuse radiation of realized DEMs are presented in Figures 6.63 and 6.65, respectively. These figures show an identical spatial pattern where lowlands and gentle slopes located at the bottom of the major hills exhibit higher values, whereas the steep slopes have low values. Furthermore, the original SRTM DEM of the study area slightly overestimated the diffuse radiation as the highest value reaches 26,497.50 Wh/m^2/year and the lowest value is 25,455.40 Wh/m^2/year. On the hand, the mean diffuse radiation of the realized DEMs has 26,482.20 Wh/m^2/year and 25,451.40 Wh/m^2/year as maximum and minimum values, respectively. The corresponding histogram of the diffuse radiation map indicated in Figure 6.64 reveals that the diffuse radiation derived from the SRTM DEM has a mean value of 26,340.67 Wh/m^2/year and standard deviation of 113.00 Wh/m^2/year. Figure 6.66 shows the result obtained from the mean of the realized DEMs with a mean of 26,338.68 Wh/m^2/year and standard deviation of 112.73 Wh/m^2/year. Results presented in Figures 6.67 and 6.69 indicates that the diffuse radiation error occurs in relatively flat terrains irrespective of the differences in elevation/height. Figures 6.68 and 6.70 represent the corresponding histograms of standard deviation and CoV, respectively.

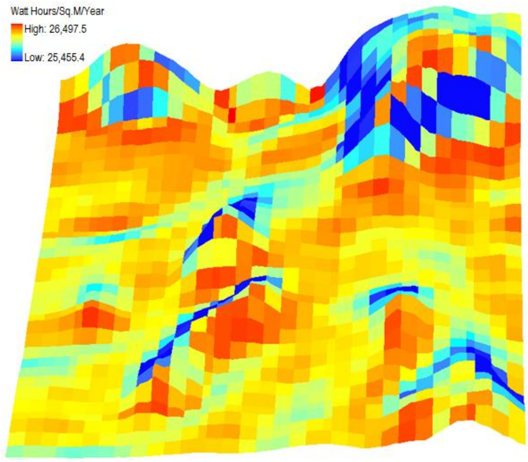


Figure 6.63. Diffuse radiation from SRTM DEM of the study area

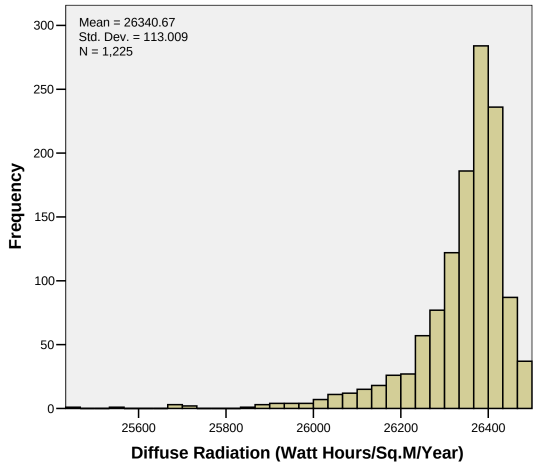


Figure 6.64. Histogram for diffuse radiation from SRTM DEM of the study area

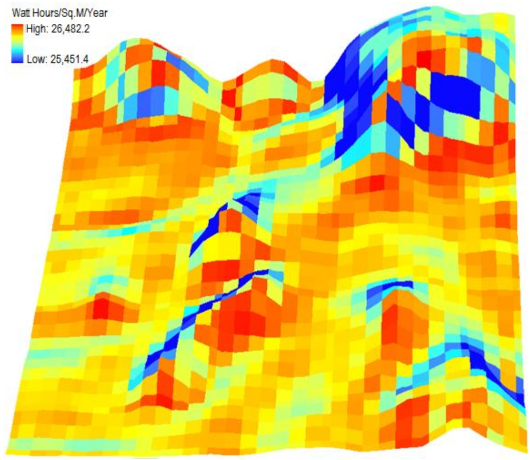


Figure 6.65. Mean diffuse radiation of realized DEMs

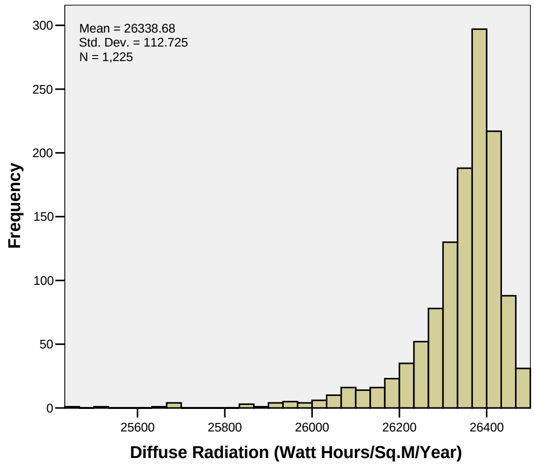


Figure 6.66. Histogram of mean diffuse radiation of realized DEMs

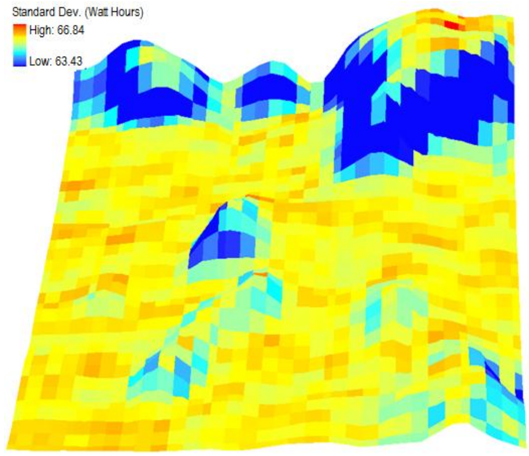


Figure 6.67. Standard deviation for mean diffuse radiation of realized DEMs

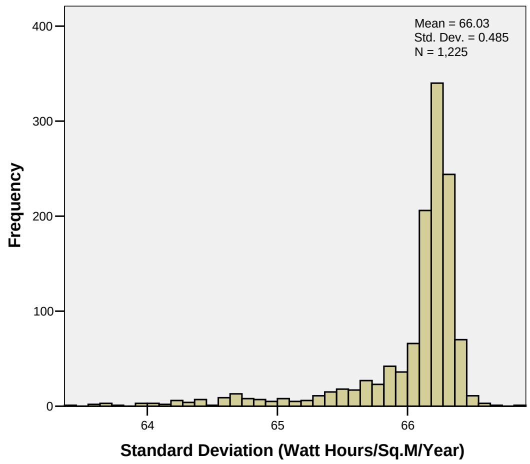


Figure 6.68. Histogram of standard deviation for mean diffuse radiation of realized DEMs

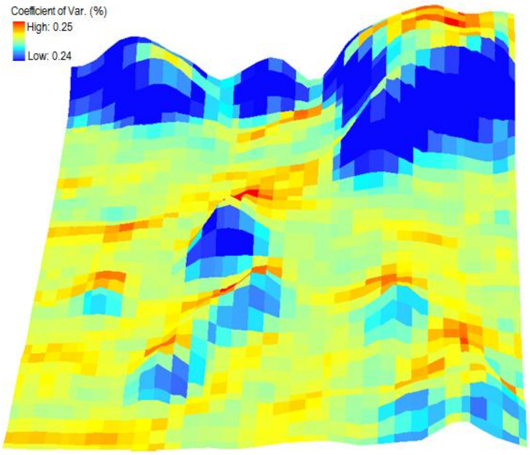


Figure 6.69. CoV for mean diffuse radiation of realized DEMs

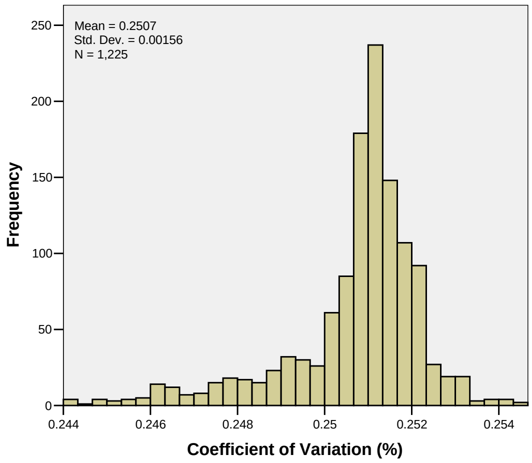


Figure 6.70. Histogram of CoV for mean diffuse radiation of realized DEMs

The global radiation outputs of *r.sun* derived from the SRTM DEM of the study area is represented in Figure 6.71 and the mean global radiation of realized DEMs in Figure 6.73. From Figure 6.71, the highest value of the global radiation estimates reaches 81,727.60 Wh/m^2/year and the lowest value is 75,550.80 Wh/m^2/year. In the case of Figure 6.73, the maximum value is 81,660.70 Wh/m^2/year and 75,463.40 Wh/m^2/year represents the minimum value. These maps indicate slight differences but similar pattern in which higher terrains indicate higher values while lowlands, valleys and steep slopes shows low estimated values. Figure 6.72 reveal a mean value of 80,415.71 Wh/m^2/year and standard deviation of 627.33 Wh/m^2/year. On the other hand, 80,350.99 Wh/m^2/year and 630.02 Wh/m^2/year represents the mean and standard deviation values in Figure 6.74, respectively. Based on the results presented in Figure 6.75, the error deviations of global radiation estimates are more prominent in the bottom and top edges of hills. The corresponding histogram (Figure 6.76) indicates a mean standard deviation of 666.33. Whereas, the CoV map (Figure 6.77) shows higher occurrence of deviations in relatively flat surfaces irrespective of elevation differences and the east-ward facing steep slope. Figure 6.78 shows the corresponding histogram of CoV.

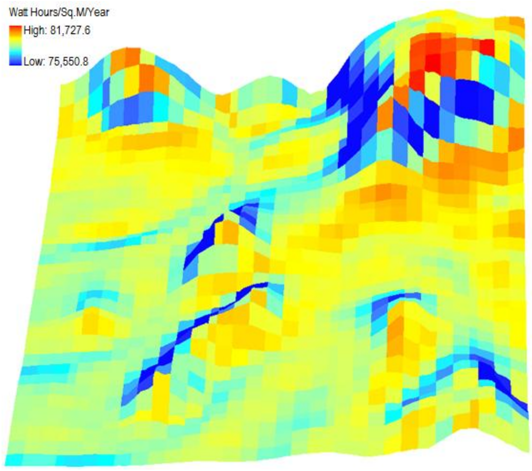


Figure 6.71. Global radiation from the SRTM DEM of the study area

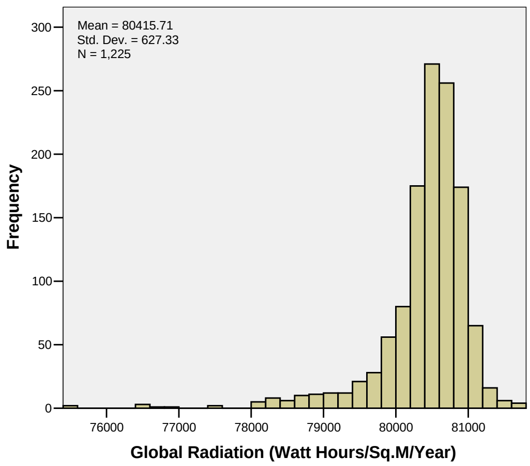


Figure 6.72. Histogram for global radiation from the original SRTM DEM of the study area

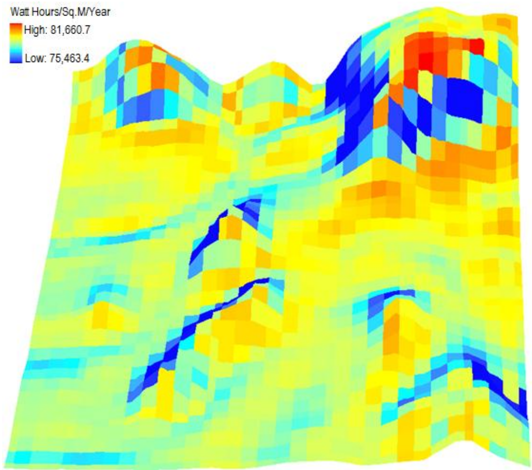


Figure 6.73. Mean global radiation of realized DEMs

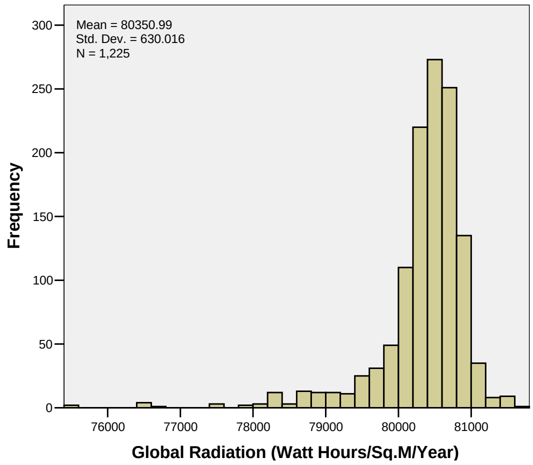


Figure 6.74. Histogram of mean global radiation of realized DEMs

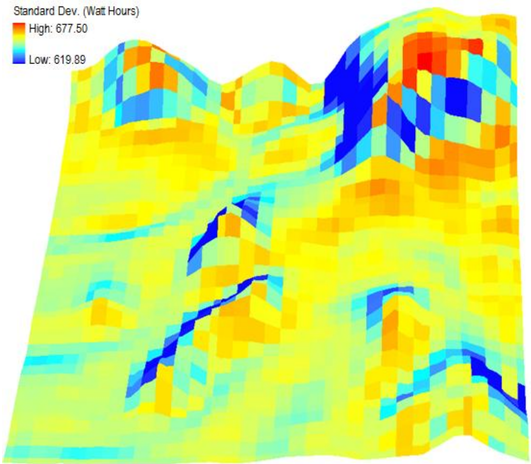


Figure 6.75. Standard deviation for mean global radiation of realized DEMs

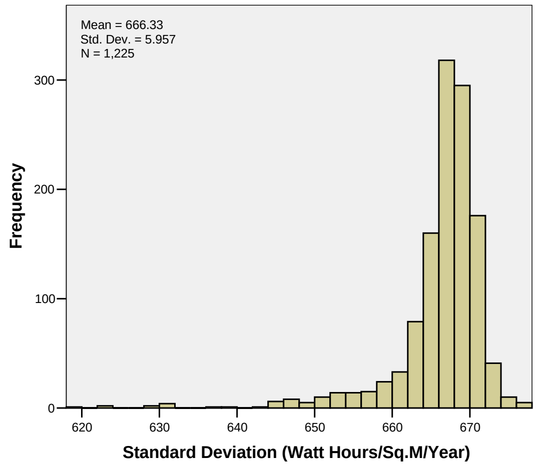


Figure 6.76. Histogram of standard deviation for mean global radiation of realized DEMs

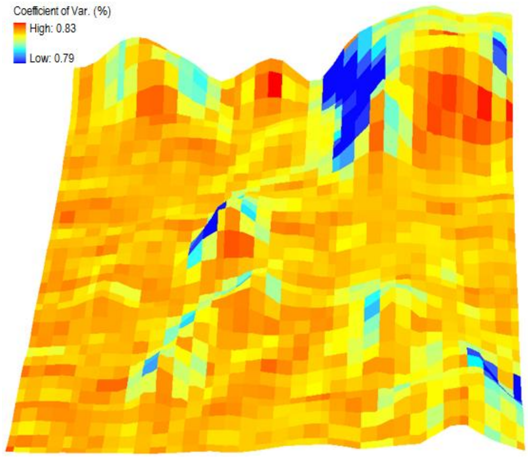


Figure 6.77. CoV for mean global radiation of realized DEMs

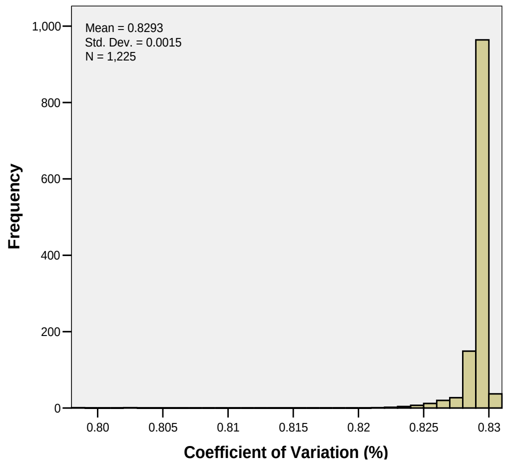


Figure 6.78. Histogram for CoV for mean global radiation of realized DEMs

As observed from Figures 6.79 and 80 and 6.81 and 82, the characterization of direct duration by *r.* sun model in the study area SRTM DEM and Mean realized DEMs are similar in terms of spatial distribution. However, results of Figure 6.79 show that areas located in the valleys exhibit lower estimates of direct duration than what is obtainable in Figure 6.81. In respect to SRTM DEM of the study area, it can be observed that 143.50 hours/m^2/year and 122.00 hours/m^2/year represent the maximum and minimum values, respectively, while 143.97 hours/m^2/year (highest) and 126.00 hours/m^2/year (lowest) represent the direct radiation mean of the realized DEMs. The corresponding histograms of direct duration maps indicated in Figure 6.80 reveal that the direct duration derived from the SRTM DEM has a mean of 140.83 Wh/m^2/year and standard deviation of 2.99 Wh/m^2/year, while Figure 6.82 shows that the mean (140.82 Wh/m^2/year) and standard deviation (3.01 Wh/m^2/year) are obtained from the direct duration estimates derived from the mean of realized DEMs. Figures 6.83 and 6.85 shows a similar spatial distribution pattern in which the west-facing slopes exhibit higher standard deviation of errors and then followed by some pockets of relatively higher standard deviation that are sparsely distributed irrespective of the nature of the terrain. The corresponding standard deviation and CoV are shown in Figures 6.84 and 6.86, respectively.

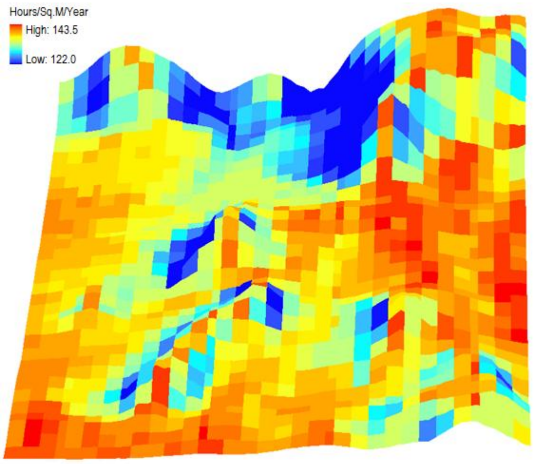


Figure 6.79. Direct duration from the study area SRTM DEM

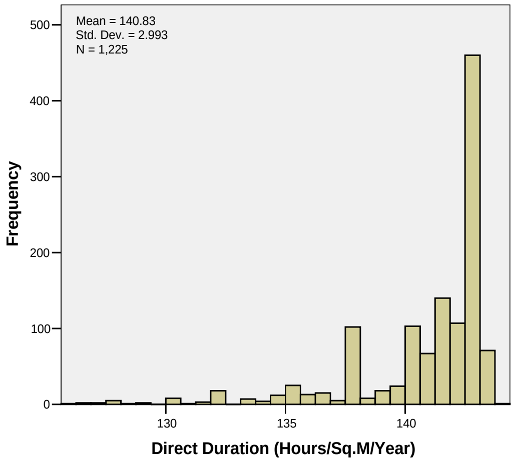


Figure 6.80. Histogram for direct duration from the study area SRTM DEM

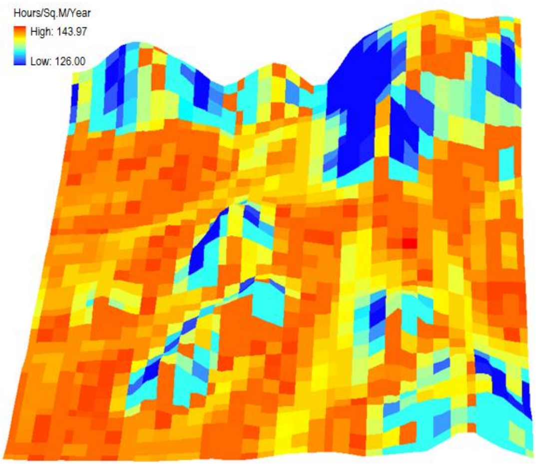


Figure 6.81. Mean direct duration of realized DEMs

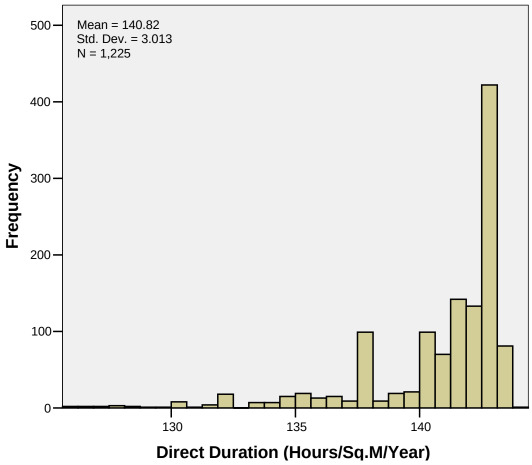


Figure 6.82. Histogram of mean direct duration of realized DEMs

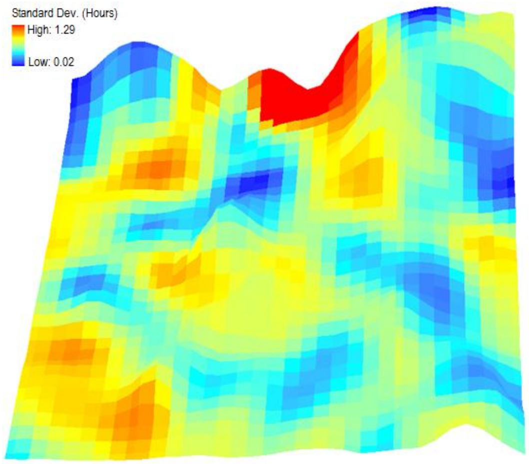


Figure 6.83. Standard deviation for mean direct duration of realized DEMs

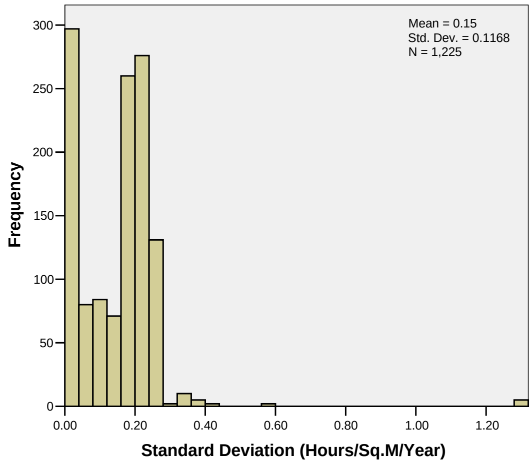


Figure 6.84. Histogram of standard deviation for mean direct duration of realized DEMs

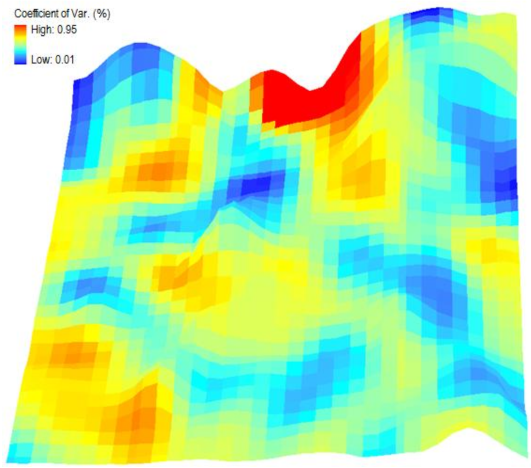


Figure 6.85. CoV for mean direct duration of realized DEMs

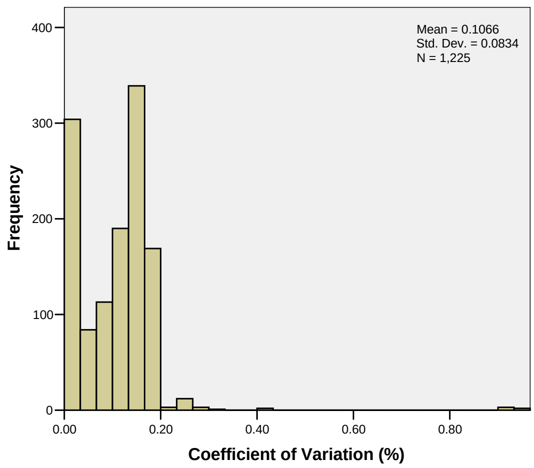


Figure 6.86. Histogram of CoV for mean direct duration of realized DEMs

6.6 Comparison of *Solar Analyst* and *r.sun* Outputs

There is no meteorological station that observes radiation data in the study area. Furthermore, an attempt is made to validate the outputs of the two models (S*olar Analyst* and *r.sun*) obtained in this study with the NASA radiation data and/or SoDA radiation data but unfortunately the spatial resolution of these data sets are very low. For instance, the spatial resolution of NASA's radiation data is 1^o x 1^o (111.12 km x 111.12 km). In this regard, a relative comparison based on statistical analysis is carried out with a view of determining the significant difference between the radiation estimates of the two models. The outputs considered are both generated by the two models and comprised of direct/beam radiation, diffuse radiation, global radiation, and direct duration. The reflected radiation generated only by the *r.sun* is not considered.

First, test of normality is conducted for all the outputs and the result indicates that all the model distribution are not normally distributed (Appendix L). Hence, the *Mann-Whitney U test* which is a non-parametric test is used (Appendix M). The results show that at 95% confidence level, there exists a significant statistical difference in the models (S*olar Analyst* and *r.sun*) estimates. This may be related to the different approaches that these models apply to obtain solar radiation estimates. However, it is important to note that both models compute solar radiation estimates based on the terrain information (elevation, slope, and aspect) available in a DEM (Ruiz-Arias et al. 2009).

Second, since there exists a significant statistical difference in the estimated distributions of the models, the ratio of *Solar Analyst* to *r. sun* is calculated and the results are classified into two or three classes depending on the variations using K-means clustering method (Equation 4.19).

6.6.1 Ratio of *Solar Analyst* to *r.sun*

In this section, the *Solar Analyst* estimates of direct radiation, diffuse radiation, global radiation and direct duration are compared with respect to that of *r.sun* by computing their ratio. The aim here, is to find areas of similarity and dissimilarity in the estimates of the two models.

For direct radiation, it can be observed from Figure 6.87 that the highest estimate differences between the two models occur on steep slopes and lowlands. Whereas, ridges represent low differences. In contrast, valleys give similar values. Based on this analysis, it can be concluded that the *Solar Analyst* gives greater differences than *r.sun*.

The estimates of diffuse radiation by the two models are similar in high elevation areas that are relatively flat. However, *Solar Analyst* exhibit higher estimates than *r.sun* model with respect to steep slopes, valleys, and lowlands (Figure 6.88).

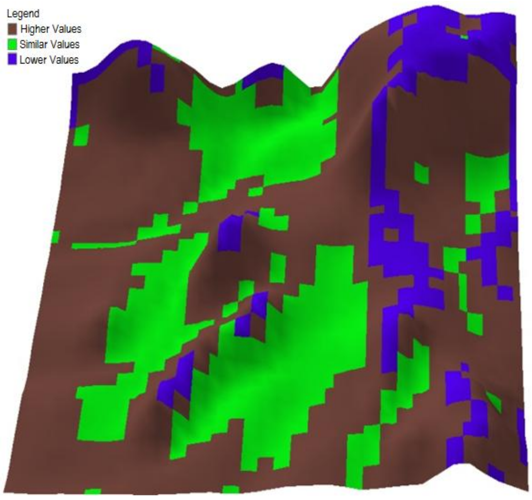


Figure 6.87. Ratio of *Solar Analyst* to *r*.sun mean direct radiation

Figure 6.88. Ratio of *Solar Analyst* to *r.sun* mean diffuse radiation

The result presented in Figure 6.89 revealed that differences in global radiation estimated values of *Solar Analyst* is similar to that of *r. sun* in the highlands. On the other hand, *Solar Analyst* model has higher differences in estimates than *r.sun* model in lowlands. However, the *Solar Analyst* performs better than *r. sun* in steep slopes and valleys.

The *Solar Analyst's* greatest magnitude of the estimated values of direct duration over *r. sun* model occurs in highlands, lowlands, ridges, and east-west facing slopes. Whereas, it is noted that the two models have similar estimates of direct duration estimated values mostly in valleys (Figure 6.90).

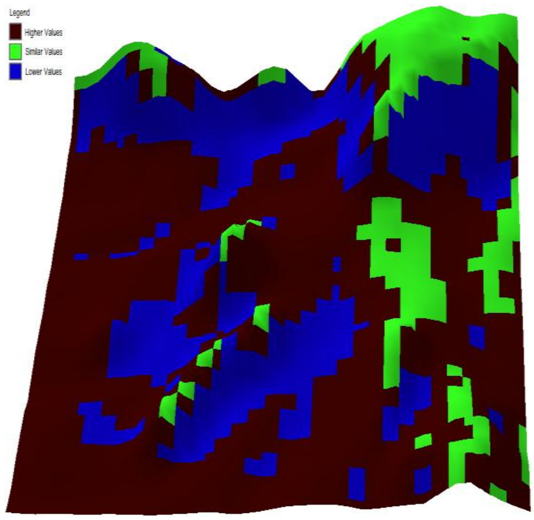


Figure 6.89. Ratio of *Solar Analyst* to *r.sun* mean global radiation

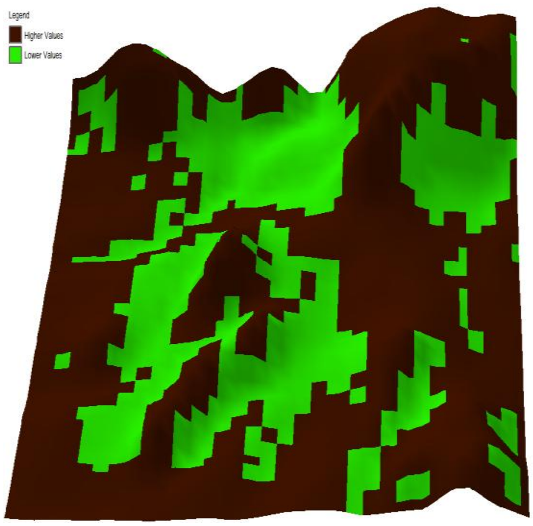


Figure 6.90. Ratio of *Solar Analyst* to *r.sun* mean direct duration

6.6.2 Ratio of Uncertainties in *Solar Analyst* to *r.sun*

Figure 6.91 reveals that the uncertainties of the estimated values of direct radiation by *Solar Analyst* model are greater than *r.sun* model in areas comprises of lowlands, valleys, and south-facing steep slopes. On the other hand, similar uncertainty values for both models are observed on low lands and east-facing slopes. *Solar Analyst* has lower uncertainty than *r.sun* in the highlands which are relatively flat in nature (Figure 6.91). Though, an exception is observed in the south-western part of the map where lowland exhibited low uncertainty. In general, *Solar Analyst* model exhibits higher uncertainty than *r.sun* model.

Based on the pattern shown in Figure 6.92 it can be noted that *Solar Analyst* exhibit higher uncertainty than *r.sun* model for the estimated values of diffuse radiation. The areas that exhibited the characteristics of high uncertainty include ridges and lowlands. On the other hand, the south-facing slopes produce similar uncertainty with respect to both models.

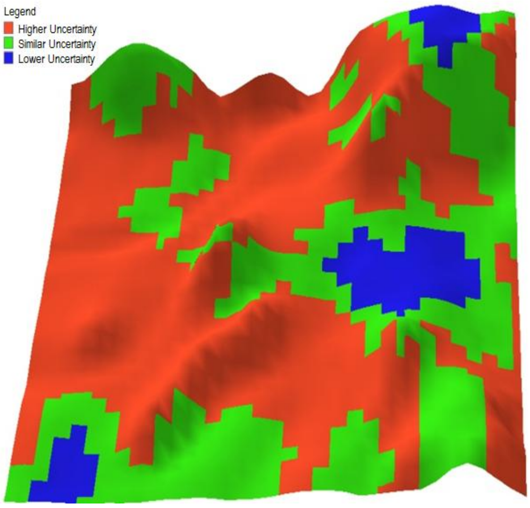


Figure 6.91. Ratio of *Solar Analyst* to *r.sun* direct radiation's standard deviation uncertainty

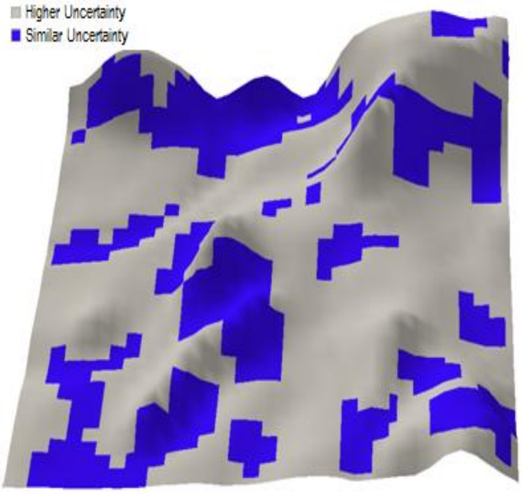


Figure 6.92. Ratio of *Solar Analyst* to *r.sun* diffuse radiation's standard deviation uncertainty

Figure 6.93 shows that the uncertainties in global radiation estimated by *Solar Analyst* for the valleys and lowland that are relatively flat in nature are higher than the *r.sun* model's results. Both models produce similar uncertainty mostly in steep slopes. Whereas, uncertainties of *Solar Analyst* results are lower than *r. sun* model in relatively flat surfaces irrespective of elevation height.

Based on the results of Figure 6.94, it can be noted that similar uncertainties with respect to hours of direct radiation of both models are randomly distributed in nature and found on relatively flat terrains, valleys and steep slopes. On the other hand, the *Solar Analyst* shows higher uncertainty than the *r.sun* model for direct duration.

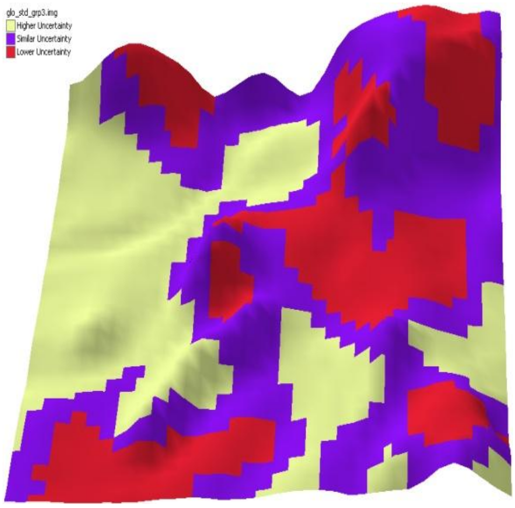


Figure 6.93. Ratio of *Solar Analyst* to *r*.sun global radiation's standard deviation uncertainty

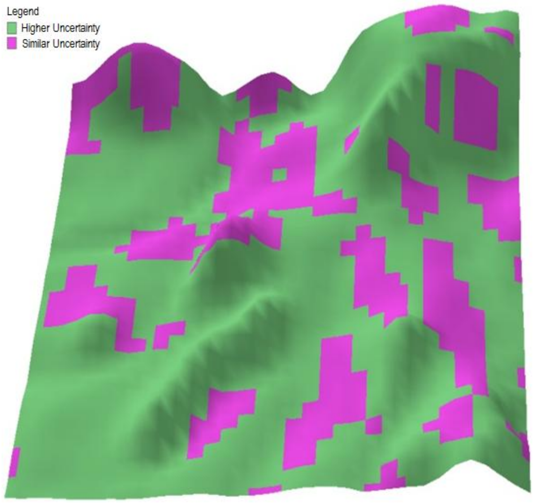


Figure 6.94. Ratio of *Solar Analyst* to *r.sun* direct duration's uncertainty

6.6.3 Ratio of *Solar Analyst* to *r.sun* Coefficient of Variation (CoV) Uncertainty

Figure 6.95 illustrates that the two models have similar CoV for direct radiation in areas that are relatively flat irrespective of elevation height differences. On the other hand, a high percentage of lowlands and steep slopes indicate greater variability in *Solar Analyst* model.

The CoV ratio of *Solar Analyst* to *r.sun* model with respect to the diffuse radiation is shown in Figure 6.96. Valleys and steep slopes exhibit similar values for both models, while gentle slopes and relatively flat terrains indicate that *Solar Analyst* has higher CoV than the *r.sun* model.

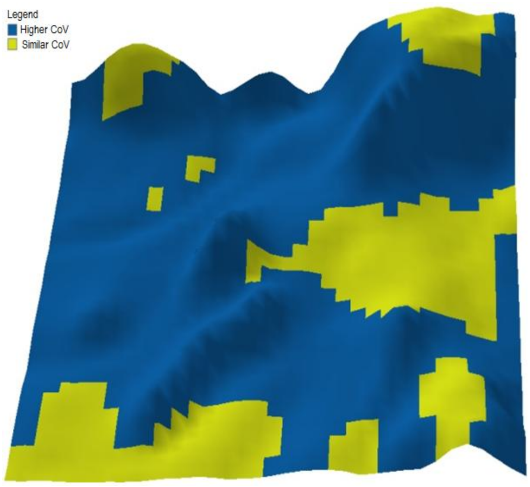


Figure 6.95. Ratio of *Solar Analyst* to *r.sun* Direct Radiation's CoV uncertainty

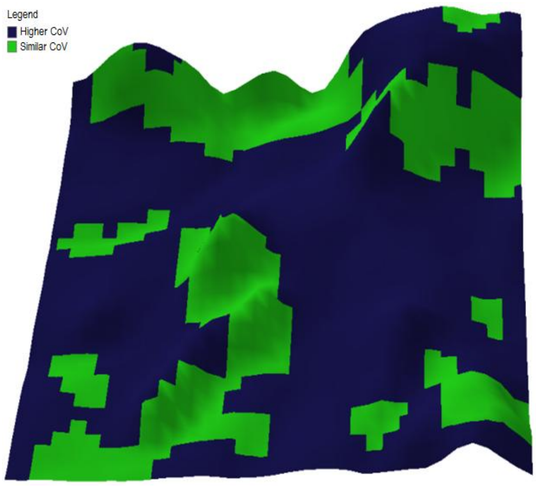


Figure 6.96. Ratio of *Solar Analyst* to *r.sun* Diffuse Radiation's CoV uncertainty

The global radiation's CoV ratio of the two models shows a similar result obtained in direct radiation where areas characterized as relatively flat, irrespective of variations in elevations, reveals similar values. However, greater variability of *Solar Analyst* over *r.sun* model estimates are noticed in low lands and steep slopes (Figure 6.97).

Based on the results shown in Figure 6.98, higher variations of *Solar Analyst* over *r.sun* model estimates of direct duration occur in low lands and steep slopes. On the other hand, similar variation values for both models are randomly observed irrespective of variation in elevation. The lower variations for both models are also randomly distributed.

Figure 6.97. Ratio of *Solar Analyst* to *r*.sun Global Radiation's CoV uncertainty

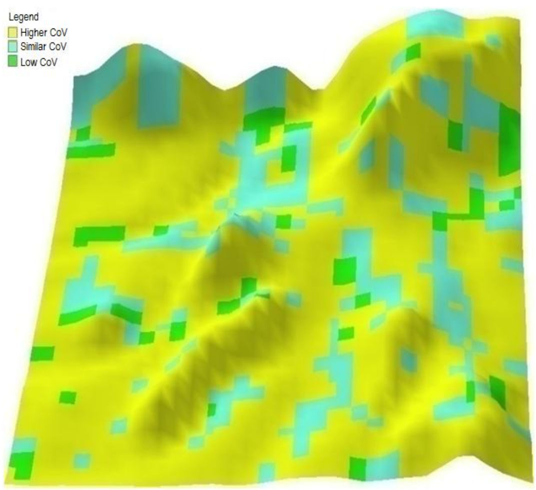


Figure 6.98. Ratio of *Solar Analyst* to *r.sun* Direct Duration's CoV uncertainty

CHAPTER 7

CONCLUSIONS AND OUTLOOK

7. 1 Conclusions

Solar radiation resource maps are vital for many areas of human endeavor especially for the purpose of decision-making regarding resource management and scientific research. Therefore, decision makers or scientists are often presented with solar radiation resource maps derived from Geographic Information System (GIS) for use in the process of discharging their legitimate responsibilities. However, in most cases, adequate information concerning the accuracy of those maps is not provided (also common to many GIS analysis products). As such, there is a great potential for decision-makers or scientists to err by over- or under-estimating the solar radiation map accuracy level. Thus, it is important that GIS community provides uncertainty information and educate decision-makers so as to have a better understanding of the possible sources of errors related to solar radiation maps and how to minimize them. As decision-makers become more aware of data accuracy and uncertainty propagation with respect to the products of *GIS-based solar radiation models*, they will be able to make a better decision which reassures both the decision makers and related stakeholders. On the other hand, decisions which are based on inaccurate data may lead to an increased probability of implementing wrong actions. This will ultimately bring about erroneous resource management actions which can have devastating repercussions on the limited resources. For instance, resource degradation, negative impacts on ecosystem(s), potential detrimental human health impacts, and economical impacts. As products of *GIS-based solar radiation models* are increasingly utilized as the basis of decision-making on resource management and regulatory issues, there is a great possibility for an increased number of litigation cases.

In this study, an *MCMC* simulation method that incorporates spatial autocorrelation in the digital elevation model (DEM) is developed with the aim of assessing the impact of DEM uncertainty in *GIS-based solar radiation models*. The method utilizes SRTM DEM error probability distribution function (pdf), and variogram model parameters of the study area SRTM DEM. Statistical analysis of the data extracted from Rodríguez et al. (2006) indicates that SRTM DEM error shows the lognormal distribution with a mean of 4.209 m and standard deviation of 0.054 m. On the other hand, exponential variogram model with *sill* of 125.74, *range* of 1,556.85, and *nugget* of 0 represent the spatial autocorrelation of the study area's DEM and is incorporated into the *Metropolis-Hasting algorithm* as *a priori* information. A total of 1,080 simulations is executed using $2m$ chains. However, the initial *burn-in* period of 80, representing 7.41% of the simulation is discarded. A multivariate potential scale reduction factor (MPSRF) of 0.99 is obtained after executing 1,080 MCMC simulations which indicate that the *MCMC sampler* has converged to a stationary distribution, since it is less than 1.1 as recommended by Gelman et al. (2004). Therefore, it is assumed that the results are drawn from the target probability distribution (lognormal pdf).

On the other hand, the check for variogram reproduction based on 95% confidence level indicates that the variogram simulation remains valid since T^2=1.751 is less than the corresponding *F*-statistic of 23.19. The above process is coded and executed in MaTLAB. The total computing time for executing the codes is five hours and twenty minutes (5:20) on a laptop which has an Intel (R) Core (TM) 2 CPU, T5600 @ 1.83 GHz, 987 MHz, and 2 GB of RAM.

Thus, the proposed framework allows analysis and management of uncertainty in the DEM data. The 1,000 simulated errors are deducted from the original SRTM DEM of the study area to generate 1,000 realized equiprobable DEMs which serves as input to the solar radiation models. Therefore, each of the 1,000 realized DEMs, the original SRTM DEM of the study area, and other inputs of the models (*Solar Analyst* and *r.sun*) are used to estimate the solar radiation products in the study area. The processing time of these models indicates that *r.sun* is faster than the *Solar Analyst.* However, one major advantage of *Solar Analyst* over the *r.sun* is that it is structured in such a way that one can either compute solar radiation for a specific day, month or year. Unlike *Solar Analyst*, the *r.sun* model is structured to compute solar radiation on daily basis which is time demanding. DEM uncertainty propagation in slope, aspect, and solar radiation maps estimated using *Solar Analyst* and *r.sun* are investigated by computing summary statistics such as mean, standard deviation, and coefficient of variation (CoV) (the last two are a measure of output uncertainty). Finally, the outputs of the two solar radiation models are compared. The proposed framework is implemented in a case study area of Abuja, Nigeria.

DEM quality generally influences the results produced by *GIS-based solar radiation models* (Pons and Ninyerola, 2008). Terrain nature is a major factor that determines the variance of the solar radiation spatial distribution. Thus, variation in elevation, slope, aspect and shadowing effects of terrain features all affect the solar radiation values obtained at any specific geographic location (Fu and Rich, 1999; Huang and Fu, 2009). Therefore, DEM uncertainty propagates in *GIS-based solar radiation models* through elevation and its derivatives; slope and aspect. Larger differences in elevation, slope and aspect are found to be correlated with steep slopes and rugged terrain.

Based on the implementation outcomes of this study, the following concluding remarks are obtained;

- Widespread DEMs, for example, SRTM are distributed with accuracy figures that only give global measures such as root mean square error (RMSE) lacking any information regarding the spatial distribution of error (Hebeler, 2008). However, in this study the *MCMC* simulation is used to generate multiple realizations of the DEM error surface that reproduce the error measurements at their original locations. These DEM errors had a significant impact on terrain attributes which compound elevation values of many grid cells (for instance, slope and aspect). The findings show that local variability of errors are far better than the global values reported in the form of RMSE which dominates DEM uncertainty literature in the past decades.
- Spatial autocorrelation is essential for spatial uncertainty analysis. Therefore, this present study reproduced the spatial autocorrelation structure of SRTM DEM for the study area.

- As expected, the elevation error would increase as the surface slope increases based on fundamental research in topographic mapping (Maling, 1989). The findings of this study indicate that majority of large deviations of elevation errors are manifested in steeper slopes. Therefore, the results are consistent with the previous findings of Wechsler and Kroll (2006). In addition to steep slopes, the CoV reveals that larger deviations also occur in the valleys. This is very important because knowing error magnitudes and their spatial distribution helps identification of places requiring more consideration during the decision making process (Gonga-Saholiariliva et al. 2011).
- It is observed that the spatial distribution of slope errors is more pronounced in flat terrain than steeper areas. The results of this study are aligned with those of Carter (1992); Vieux (1993); Zhou et al. (2006) but not consistent with Chang and Tsai (1991) who stated that slope errors are mainly concentrated in areas of steep slopes. The results can be explained due the fact that as the more gentle slopes were eliminated from consideration, the frequency distribution tended to be more uniform (Carter, 1992).
- The findings of this study support the assertion published in the works of Carter (1992); Chang and Tsai (1991); Zhou et al. (2006) who noted that the concentration of aspect error is quite high at flat terrains.
- Generally, the two models (*Solar Analyst* and *r.sun*) overestimate the incoming radiation with respect to the original SRTM DEM of the study area. Note that this overestimation indicates lower reliability of SRTM DEM.
- The results of both models showed a better performance using the realized DEMs from the proposed *MCMC* simulation than the original SRTM DEM.
- For *Solar Analyst* model, DEM uncertainty has greater effect on diffuse radiation and direct duration. Whereas, direct and global radiations are less affected.
- For *r.sun* model, the DEM uncertainty has less influence on solar radiation outputs.
- Comparison of the two models shows that *Solar Analyst* has higher uncertainty than the *r.sun*. In other words, *r.sun* is more robust.
- Interestingly, the study reveals that relatively flat terrains where DEM uncertainty seems to be low exhibit higher uncertainty of solar radiation products. This indicates that DEM may not be the only input associated with uncertainty. Hence, there is a need for further investigation with respect to the other inputs.
- Solar radiation maps, for example, direct radiation maps are important for concentrated solar photovoltaic (CSP) or concentrated photovoltaic (CPV), and photovoltaic technologies; diffuse radiation maps are needed for building climatology – day lighting, photovoltaic technologies; and global radiation are required for solar collectors that are flat in nature, heat and agricultural purposes. Therefore, using these maps without considering uncertainty will have negative consequences on the decision making, for instance, waste of resources, low output/productivity, and low agricultural harvest.

The implications of the results obtained from this study is discussed on two major issues related to the current challenges facing the Nigerian electricity sector, namely, (1) increase in electricity generation for domestic and industrial use, and (2) a resource for assisting decision-makers in allocating sites for CSP or CPV projects.

The study reveals that there are vast potential of solar energy resources in Nigeria even though the area coverage of study is very small compared to the size of the country.

However, previous studies such as Sambo (2005); Fadare (2009) and the like have reported the Nigeria's endowment of such a huge solar resource. Moreover, the location of Nigeria near the Equator is another indicator. Therefore, harnessing this resource through the establishment of CSP or CPV plants will contribute immensely to the electricity generation capacity for Nigeria. For instance, Alstom (2013) noted that CSP is an ideal approach for both harvesting the solar energy potential for large scale grid connected power generation, and remote industrial applications. Currently, there are various capacities of CSP plants ranging from 1 to 64 megawatts (MW). However, it is important to note that the sitting of CSP and CPV projects are generally based on exclusion criteria, for instance, protected areas, and land use such as agricultural and forestry areas, grasslands, or any area assumed unsuitable due to socio-geographical reasons.

The importance of using solar energy in Nigeria will not only be confined to improving the electricity generation which will be distributed through the national grid with the aim of helping to meet the shortfall of peak-load demand that the country is presently experiencing, but also contribute to meeting the demands of both the remote sites and urban centers. Photovoltaic (PV) technology is a good potential for such locations because it has shown an increased promise in terms of efficiency improvements and costs. The estimated lifetime of PV modules based on manufacturer's specifications are 25 and 30 years making them exceptionally attractive for investors. Today, all the PV modules in the Nigerian market are imported. Solar PV systems can be extensively used for a wide range of electrical energy requirements, including; solar home systems, water pumping systems, refrigeration and telecommunications that will reduce the load curve of electricity demand. These applications have positive social and economic impact on the lives of individual users, businesses and communities. Therefore, since solar radiation in Nigeria is fairly well distributed, a rural electrification drive based on PV power systems should be pursued for supplying energy to homes, schools, clinics, small and medium scale farms, and small businesses. In addition, urban areas are not excluded when considering rooftop applications.

Therefore, investing in these technologies will definitely reduce the Nigeria's heavy reliance on fossil fuel sources which by all evidences could not provide sufficient and sustainable energy for Nigeria's developmental objectives and the reduction of environmental pollution. Moreover, substituting fossil fuels with renewable energy sources is regarded as a significant measure for cutting global carbon emissions (IPCC, 2001). Full use of these resources can help mitigate global warming in environmental terms, meet energy needs in economic terms, and provide employment in rural areas in socioeconomic terms (UNCSD, 2002).

The results of this study will also be a very useful resource in assisting decision-makers of such agencies like the Energy Commission of Nigeria, Nigerian Electricity Regulatory Commission, among others in allocating appropriate sites for establishing CSP or CPV projects. Thereby preventing them from unnecessary litigations and waste of resources.

7. 2. Future Outlook

Spatially varying inputs such as slope, aspect, linke turbidity factor, and ground albedo also have uncertainty associated with them. Therefore, the challenge of incorporating all the spatially varying inputs needs to be addressed and their impact shall be assessed.

There is also the need to apply the methodology using different kind of DEMs, for instance, Advanced Space borne Thermal Emission and Reflection Radiometer (ASTER), GTOPO30, and hypothetical DEMs. The application on hypothetical DEMs may be very useful because one knows what is going on during its creation which will assist in explaining the outcome of the analysis.

REFERENCES

Abbaspour, R.A., Delavar, M.R. and Batouli, R. (2003). The issue of uncertainty propagation in spatial decision making. http://www.scangis.org/scangis2003/papers/13.pdf, last accessed on 10/07/2011.

Ackermann, F. (1996). Airborne laser scanning for elevation models, *GIM International*, 10(10): 24-25.

Ades, A.E. and Lu, G. (2003). Correlations between parameters in risk models: estimation and propagation of uncertainty by Markov chain Monte Carlo. *Risk Analysis*, 23(6): 1165-1172.

Ahmed, E.A. and Adam, M.E. (2013). Estimate of global solar radiation by using artificial neural network in Qena, Upper Egypt. *Journal of Clean Energy Technologies,* 1(2): 148-150.

Akpinar, E. (2005). Using geographic information systems in analysing the pattern of crime incidents and the relationship between landuse and incidents. Msc. Thesis, Department of Geodetic and Geographic Information Technologies, Middle East Technical University, Ankara - Turkey.

Alai, J. (1993). Spatial uncertainty in a GIS. MSc. Thesis, Department of Geomatic Engineering, University of Calgary, Canada.

Alawi, S. and Hinai, H. (1998). An ANN based approach for predicting global radiation in locations with no direct measurement instrumentation. *Renewable Energy,* 14(1-4): 199-204.

Alsamamra, H., Ruiz-Arias, J.A., Pozo-Vazquez, D. and Tovar-Pescador, J. (2009). A comparative study of ordinary and residual kriging techniques for mapping global solar radiation over southern Spain. *Agricultural and Forest Meteorology*, 149: 1343–1357.

Alstom (2013). Concentrated solar power solutions. http://www.alstom.com/Global/Power/Resources/Documents/Brochures/concentrated-solar-power-solutions.pdf, last accessed on 10/09/2013.

Alves, M., Moreira da Silva, F., Pozza, E.A. and Silva de Oliveira, M. (2009). Modeling spatial variability and pattern of rust and brown eye spot in coffee agroecosystem. *Journal of Pest Science,* 82(2): 137-148.

Angstrom, A. (1924). Solar and terrestrial radiation. *Journal of the Royal Meteorological Society,* 50: 121-125.

Ardiansyah, P.O.D. and Yokoyama, R. (2002). DEM generation method from contour lines based on the steepest slope segment chain and a monotone interpolation function. *ISPRS Journal of Photogrammetry and Remote Sensing,* 57: 86-101.

Atkinson, P.M. (1999). Geographical information science: geostatistics and uncertainty. *Progress in Physical Geography,* 23(1): 134.

Austin, G.E., Thomas, C.J., Houston, D.C. and Thompson, D.B.A. (1996). Predicting the spatial distribution of Buzzard buteo buteo nesting areas using a geographical information system and remote sensing. *Journal of Applied Ecology,* 33(6): 1541-1550.

Azadeh, A., Maghsoudi, A. and Sohrabkhani, S. (2009). An integrated artificial neural networks approach for predicting global radiation. *Energy Conversion and Management,* 50(6): 1497–1505.

Barker, A.A. (1965). Monte Carlo calculations of the radial distribution functions for a proton-electron plasma. *Australian Journal of Physics,* 18: 119-133.

Barnes, R. (2011). Variogram tutorial. Golden Software, Inc., http://www.goldensoftware.com/variogramTutorial.pdf, last accessed on 20/04/2012.

Besag, J.E. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society,* 36(2): 192-236.

Billera, L.J., and Diaconis, P. (2001). A geometric interpretation of the Metropolis-Hastings algorithm. *Statistical Science,* 16: 335-339.

Bivand, R.S., Pebesma, E.J. and Gómez-Rubio, V. (2008). Applied spatial data analysis with R. New York, USA. *Springer*.

Blomgren, S. (1999). A digital elevation model for estimating flooding scenarios at the Falsterbo Peninsula. *Environmental Modelling & Software,* 14(6): 579-587.

Bohling, G. (2005). Introduction to geostatistics and variogram analysis, http://people.ku.edu/~gbohling/cpe940/Variograms.pdf, last accessed on 25/05/2011.

Bolstad, P.V. and Stowe, T. (1994). An evaluation of DEM accuracy: elevation, slope, and aspect. *Photogrammetric. Engineering and Remote Sensing,* 60(11): 1327-1332.

Borcard, D., Legendre, P. and Drapeau, P. (1992). Partialling out the spatial component of ecological variation. *Ecology,* 73(3): 1045-1055.

Borisov, M., Banković, R. and Drobnjak, S. (2009). Modelling the uncertainty of digital elevation models with geostatistical simulation and application on spatially distribution of soil nitrogen. *Scientific Technical Review,* 3(4): 54-59.

Brimicombe, A. (2010). GIS, environmental modeling and engineering (2nd edition). Boca Raton: *CRS Press*.

Brooks, S.P. (1996). Quantitative convergence diagnosis for MCMC via CUSUMS. Technical Report.*University of Bristol*, UK.

Brooks, S.P. (1998). Markov chain Monte Carlo method and its application. *The Statistician,* 47(1): 69-100.

Brooks, S.P. and Gelman, A. (1997). Alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics,* 7: 434-455.

Brooks, S.P. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics,* 7(4): 434-455.

Brooks, S.P. and Roberts, G.O. (1996). Disussion on convergence of Markov chain Monte Carlo algorithms. In Bernardo, J.M. Berger,J.O., Dawid, A.P. and Smith, A.F.M. (eds): Bayesian Statistics 5. Oxford, UK. *Oxford University Press*.

Burrough, P.A. (1986). Principles of geographical information systems for land resources assessment: Oxford, UK, *Clarendon Press*. Page 194.

Burrough, P.A. and McDonnell, R.A. (1998). Principles of geographical information systems: New York, *Oxford University Press*. Page 333.

Campbell, J. (1981). Spatial autocorrelation effects upon the accuracy of supervised classification of land cover. *Photogrammetric. Engineering and Remote Sensing,* 47: 355-363.

Cano, D., Morget, J.M. Albuisson, M., Guillard, H., Regas, N. and Wald, L. (1986). A method for determination of the global solar radiation from meteorological satellite data. *Solar Energy*, 37: 31-39.

Carabajal, C.C. and Harding, D.J. (2006), SRTM C-band and ICESat laser altimetry elevation comparisons as a function of tree cover and relief. *Photogrammetric. Engineering and Remote Sensing,* 72(3): 287-298.

Carlisle, B.H. (2005). Modelling the spatial distribution of DEM error. *Transactions in GIS,* 9(4): 521-540.

Caruso, V.M. (1987). Standards for digital elevation models. Processings of the ACSM/ASPRS Annual Convention, Baltimore. Pages 159-166.

Chang, H.C., Li, X. and Linlin, G. (2010). Assessment of SRTM, ACE2 and ASTER-GDEM using RTK-GPS, http://www.gmat.unsw.edu.au/snap/publications/chang_etal2010a.pdf, last accessed on 14/06/2012.

Chang, K. and Tsai, B. (1991). The effect of DEM resolution on slope and aspect mapping. *Cartography and Geographic Information Science,* 18(1): 69-77.

Charabi, Y. and Gastli, A. (2010). GIS assessment of large CSP plant in Duqum, Oman. *Renewable and Sustainable Energy Reviews,* 14: 835-841.

Chaubey, I., Cotter, A.S., Costello, T.A. and Soerens, T. (2005). Effect of DEM data resolution on SWAT output uncertainty. *Hydrological Processes*, 19(3): 621-628.

Chib, S. and Greenberg, E. (1995). Markov chain Monte Carlo simulation methods in econometrics, *Economic Theory,* 12: 409-431.

Clark, I. and Harper, W. (2002). Practical geostatistics 2000, 3rd reprint, Ecosse North America Llc, Columbus, Oh., USA.

Clifton, J. and Boruff, B.J. (2010). Assessing the potential for concentrated solar power development in rural Australia. *Energy Policy,* 38: 5272-5280.

Cowles, M.K. and Carlin, B.P. (1996). Markov chain Monte Carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association,* 91(434): 883-904.

Cressie, N.A.C. (1993). Statistics for spatial data. New York, USA, *John Wiley & Sons.*

Crosetto, M. and Tarantola, S. (2001). Uncertainty and sensitivity analysis: tools for GIS based model implementation. *International Journal of Geographical Information Science,* 15: 415-437.

Crosetto, M., Tarantola, S. and Saltelli, A. (2000). Sensitivity and uncertainty analysis in spatial modelling based on GIS. *Agriculture Ecosystems and Environment,* 81: 71-79.

Dagpunar, J.S. (2007). Simulation and Monte Carlo. West Sussex, England. *John Wiley & Sons Ltd*. Page 1.

de Swart, E.O.A.M., van der Valk, A.G., Koehler, K.J. and Barendregt, A. (1994). Experimental evaluation of realized niche models for predicting responses of plant species to a change in environmental conditions. *Journal of Vegetation Science,* 5: 541-552.

Devillers, R., Gervais, M., Bedard, Y. and Jeansoulin, R. (2002). Spatial data quality: from metadata to quality indicators and contextual end-user manual. *OEEPE/ISPRS Joint Workshop on Spatial Data Quality Management,* March 21-22, Istanbul - Turkey.

Djafer, D. and Irbah, A. (2013). Estimation of atmospheric turbidity over Ghardaia city. *Atmospheric Research,* 128: 76-84.

Dogniaux, R. (1984). De l'influence de l'estimation du facteur total de trouble atmosphérique sur l'évaluation du rayonnement solaire direct par ciel clair. Application aux données radio-métriques de L'IRM à Uccle. in Institut Royal Météorologique de Belgique (IRM), Miscellanea - Serie C 20.

Donatelli, M., Bellocchi, G. and Fontana, F. (2003). RadEst3.00: Software to estimate daily radiation data from commonly available meteorological variables. *European Journal of Agronomy*, 18: 363-367.

Dubayah, R. (1992). Estimating net solar radiation using landsat thematic mapper and digital elevation data. *Water Resources Research,* 28(9): 2469-2484.

Dubayah, R. and Rich, P.M. (1995). Topographic solar radiation models for GIS. *International Journal of Geographic Information Science,* 9(4): 405-419.

Dubayah, R. and Rich, P.M. (1996). GIS-based solar radiation modeling in GIS and environmental modeling: progress and research issues. GIS World Books. Fort Collins, CO. Pages 129-135.

Dubayah, R.C. (1994). Modeling a solar radiation topoclimatology for the Rio Grande River Basin. *Journal of Vegetation Science,* 5: 627-640.

Duffie, J.A., Beckman, W.A. (1991). Solar engineering of thermal processes, 2nd edition. New Jersey, USA. *John Wiley & Sons.*

Eastman, J.R., Kyem, P.A.K., Toledano, J. and Jin, W. (1993). GIS and decision making, explorations in geographical information systems technology, Vol. 4, Worcester, MA, Unitar, Clark Labs.

ECN and UNDP (2005), Renewable energy master plan (REMP). Energy commission of Nigeria, Federal Ministry of Energy Resources, Nigeria.

Edwards, D. and Hamson, M. (1989). Guide to Mathematical modeling. London: Macmillan.

Ehlschlaeger, C.R. and Shortridge, A. (1997). Modelling elevation uncertainty in geographical analyses. In Kraak, M.J. and Molenaar, M. (eds), Advances in GIS Research: Proceedings of the Seventh International Symposium on Spatial Data Handling. London, *Taylor and Francis*. Pages 585-95.

Ehlschlaeger, C.R., Shortridge, A.M. and Goodchild, M.F. (1997). Visualizing spatial data uncertainty using animation. *Computers & Geosciences,* 23(4):387-395.

Energy Information Administration (EIA) (2012). Nigeria. http://www.eia.gov/countries/country-data.cfm?fips=NI, last accessed on 10/11/2012.

Energy Information Administration (EIA) (2013). International energy outlook 2013. http://www.eia.gov/forecasts/ieo/pdf/0484%282013%29.pdf, last accessed 10/10/2013.

Ertekin, C. and Yaldiz, O. (2000). Comparison of some existing models for estimating global solar radiation for Antalya (Turkey). *Energy Conversion & Management,* 41: 311-330.

ESRI (2008). ArcGIS Desktop Help. Environmental Systems Research Institute (ESRI), Redlands, California.

Fadare, D. A. (2009). Modelling of solar energy potential in Nigeria using an artificial neural network model, *Applied Energy*, 86(9): 1410-1422.

Farah, A., Talaat, A. and Farrag, F.A. (2008). Accuracy assessment of digital elevation models using GPS. *Artificial Satellites,* 43(4): 151-161.

Fisher, P. (1998). Improved modeling of elevation error with geostatistics: *GeoInformatica,* 2: 215-233.

Fisher, P.F. (1991). "First experiments in viewshed uncertainty: the accuracy of the viewable area," *Photogrammetric Engineering and Remote Sensing*, 57: 1321-1327.

Fisher, P.F. (1999). Models of uncertainty in spatial data – In: Longley P.A., Goodchild M.F., Maguire D.J., Rhind D.W.: Geographical Information Systems, Vol 1. *John Wiley & Sons*. Pages 191-205.

Fisher, P.F. and Tate, N.J. (2006): Causes and consequences of error in digital elevation models. *Progress in Physical Geography,* 30(4): 467-489.

Florinsky, I.V. (1998). Accuracy of local topographic variables derived from digital elevation models. *International Journal of Geographical Information Science,* 12 (1): 47-61.

Fu, P. and Rich, P.M. (1999). Design and implementation of the Solar Analyst: an ArcView extension for modeling solar radiation at landscape scales. *Proceedings of the Nineteenth Annual ESRI User Conference,* http://proceedings.esri.com/library/userconf/proc99/proceed/papers/pap867/p867.htm, last accessed on 12/02/2012.

Fu, P. and Rich, P.M. (2000). The solar analyst 1.0 user manual. *Helios Environmental Modeling Institute, LLC,* KS, USA.

Fu, P. and Rich, P.M. (2002). A geometric solar radiation model with applications in agriculture and forestry. *Computers and Electronics in Agriculture,* 37: 25-35.

Gafarain, A.V., Ancker, C.J. and Morisaku, T. (1978). Evaluation of commonly used rules for detecting “steady state” in computer simulation. *Naval Research Logistics*, 25: 511-529.

Gamerman, D. (1997). Markov chain Monte Carlo: Stochastic simulation for Bayesian inference. London: *Chapman and Hall.*

Garbrecht, J. and Starks, P. (1995). Note on the use of USGS level 1 7.5-minute DEM coverages for landscape drainage analyses. *Photogrammetric Engineering and Remote Sensing*, 61(5): 519-522.

Garren, S.T. and Smith, R.L. (1993). Convergence diagnostics for Markov chain samplers. Technical report, University of North Carolina, Dept. of Statistics.

Gastli, A. and Charabi, Y. (2010). Solar electricity prospects in Oman using GIS-based solar radiation maps. *Renewable and Sustainable Energy Reviews*, 14: 790-797.

Gelfand, A.E. and Smith, A.F.M. (1990). Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85: 398-409.

Gelfand, A.E., Hills, S.E., Racine-Poon, A. and Smith, A.F.M. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association*, 85: 972-985.

Gelman, A. and Rubin, D. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7: 457-511.

Gelman, A. and Rubin, D.B. (1996). Markov chain Monte Carlo methods in biostatistics. *Statistical Methods in Medical Research*, 5(4): 339-355.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. (2004), Bayesian data analysis, 2^{nd} edition. Boca Raton, *Chapman & Hall/CRC.*

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6: 721-741.

Genz, R. and van Genderen, J.L. (1996). SAR interferometry - issues, techniques, applications, *International Journal of Remote Sensing*, 17(10): 1803-1835.

Gessler, P.E., Moore, I.D., McKenzie, N.J. and Ryan, P.J. (1993). Soil-landscape modeling in southeastern Australia. In Goodchild, M.F., Parks, B.O. and Steyaert, L.T.

(eds.): Environmental Modeling with GIS. New York, USA, *Oxford Press*. Pages 53-58.

Geweke, C.J. (1992). Practical Markov chain Monte Carlo. *Statistical Science,* 7: 473-511.

Geyer, C.J. (1992). Markov chain Monte Carlo maximum likelihood, http://users.stat.umn.edu, last accessed on 06/05/2012.

Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. (1996). Markov chain Monte Carlo in practice. London: Chapman and Hall.

Gonga-Saholiariliva, N., Gunnell, Y., Petit, C. and Mering, C. (2011). Techniques for quantifying the accuracy of gridded elevation models and for mapping uncertainty in digital terrain analysis. *Progress in Physical Geography,* 35(6): 739-764.

Goodchild, M.F. (1992). Closing report NCGIA research initiative I: accuracy of spatial databases I. University of California, National Center for Geographic Information and Analysis.

Goodchild, M.F. and Gopal, S. (1989). Accuracy of spatial databases. London, UK,*Taylor & Francis.*

Goodchild, M.F., Guoqing, S. and Shiren, Y. (1992). Development and test of an error model for categorical data. *International Journal of GIS,* 6: 87-104.

Grass Development Team (2006). http://www.grass.itc.it, last accessed on 11/06/2012.

Grenander, U. (1983). Tutorial in pattern theory. Report: Division of Applied Mathematics, Brown University.

Greve, C. (1996). Digital photogrammetry: an addendum to the manual of photogrammetry. *Falls Church, Va: American Society for Photogrammetry and Remote Sensing*.

Guth, P.L. (1992). Spatial analysis of DEM error. Pp 187-196 in ASPRS/ACSM/RT Technical Papers 2, Washington, DC, USA.

Guth, P.L. (1999). Contour line ghosts in USGS level 2 DEMs. *Photogrammetric Engineering and Remote Sensing,* 65: 289-96.

Hammersley, J.M. and Clifford, P. (1971). Markov fields on finite graphs and lattices. Unpublished manuscript.

Hammersley, J.M. and Handscomb, D.C. (1979). Monte Carlo methods. London, *Chapman & Hall*.

Hansen, J.W. (1999). Stochastic daily solar irradiance for biological modeling applications. *Agricultural and Forest Meteorology*, 94: 53-63.

Harrison, K. Kumar, S., Peters-Lidard, C. and Santanello, J. (2011). Uncertainty analysis in the decadal survey era: a hydrologic application using the land information system (LIS), http://esto.nasa.gov/conferences/estf2011/papers/Harrison_ESTF2011.pdf, last accessed on 28/03/2012.

Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika,* 57: 97-109.

Haxeltine, A. and Prentice, I.C. (1996). BIOME3: An equilibrium terrestrial biosphere model based on ecophysiological constraints, resources availability, and competition among plant function types. *Global Biogeochemical Cycles,* 10: 693-709.

Heath, M.T. (2002). Scientific computing: an introductory survey (2nd edition). *McGrawHill*. Page 511.

Hebeler, F. (2008). Modelling topographic uncertainty: impacts on large scale environmental modeling. PhD Thesis, University of Zurich, http://www.geo.uzh.ch/fileadmin/files/content/abteilungen/gis/research/phd_theses/thesis_FelixHebeler_2008.pdf, last accessed 06/06/2012.

Hebeler, F. and Purves, R. (2009). Modelling DEM data uncertainties for Monte Carlo simulations of Ice Sheet models, Stein, A., Shi, J. and Bijker, W., eds., Quality Aspects in spatial Data Mining. Boca Raton,*Taylor & Francis Group, LLC*. Pages 175-196.

Heidelberger, P. and Welch, P.D. (1983). Simulation run length control in the presence of an initial transient. *Operations Research,* 31: 1109-1144.

Henderson, B. and Bui, E. (2005). Determining uncertainty in sediment & nutrient transport models for ecological risk assessment. Report No 2. LWA/MDBC Project UMO43: Risk-based approaches to managing contaminants in catchments, http://www.waterscience.com.au/pdf/Risk_based_Approaches_Report_2.pdf, last accessed on 05/06/2012.

Her, Y.G. (2011). HYSTAR: Hydrology and sediment transport simulation using time-area method. PhD Thesis. Virginia Polytechnic Institute and State University, USA.

Hetrick, W.A., Rich, P.M., Barnes, F.J., Weiss, S.B. (1993). GIS-based solar radiation flux models. American Society for Photogrammetry and Remote Sensing Technical papers. *GIS, Photogrammetry and Modeling,* 3: 132-143.

Heuvelink, G.B., Burrough, P.A. and Stein, A. (1989). Propagation of erosion spatial modelling with GIS. *International Journal of Geographical Information Systems,* 3(4): 303-322.

Heuvelink, G.B.M. (1998a). Uncertainty analysis in environmental modelling under a change of spatial scale. *Nutrient Cycling in Agroecosystems,* 50: 255–264.

Heuvelink, G.B.M. (1998). Error propagation in environmental modelling with GIS. *Taylor and Francis,* London.

Heuvelink, G.B.M. (2002). Analysing uncertainty propagation in GIS: why it is not that simple? In Foody, G.M. and Atkinson, P.M., editors, Uncertainty in Remote Sensing and GIS. Chichester, England. *John Wiley& Sons, Ltd.*

Heuvelink, G.B.M. and Burrough, P. (2002). Developments in statistical approaches to spatial uncertainty and its propagation. *International Journal of Geographic Information Science,* 16: 111-113.

Heywood, I., Cornelius, S., Carver, S. (2002). An introduction to geographical information Systems. Harlow, *Prentice Hall*.

Higdon, D. (1994). Spatial applications of Markov chain Monte Carlo for bayesian inference. PhD thesis, University of Washington, USA.

Hodges, T., Botner, D., Sakamoto, C. and Hays Haug, J. (1987). Using the CERES-Maize model to estimate production for the U.S. Cornbelt. *Agricultural and Forest Meteorology,* 40(4): 293-303.

Hofierka, J. (1997). Direct solar radiation modelling within an open GIS environment. *Proceedings of the Joint European GI Conference*, Vienna, 575-584.

Hofierka, J. and Šúri, M. (2002). The solar radiation model for Open source GIS: implementation and applications, Proceedings of the Open source GIS - GRASS users conference - Trento, Italy, 11-13 September 2002. Pages 1-19.

Hofierka, J. and Zlocha, M. (2012). A new 3-D solar radiation model for 3-D city models. *Transactions in GIS,* 16(5): 681–690.

Hofierka, J., Cebecauer, T. (2008): Spatially distributed assessment of solar resources for energy applications in Slovakia. Acta Facultatis Studiorum Humanitatis et Naturae Universitatis Prešoviensis. Prírodné vedy, Folia Geographica 12. Pages 97-114.

Hofierka, J., Parajka, P., Mitasova, M. and Mitas, L. (2002). Multivariate interpolation of precipitation using regularized spline with tension. *Transactions in GIS,* 6: 135-50.

Hohn, M. (1988), *Geostatistics and Petroleum Geology*, New York, USA. *Van Nostrand Reinhold*.

Holmes, K. W., Chadwick, O. A. and Kyriakidis, P. C. (2000). Error in a USGS 30-meter digital elevation model and its impact on terrain modeling. *Journal of Hydrology,* 233: 154-173.

Houlding, S.W. (2000). *Practical geostatistics: modeling and spatial analysis*. Springer-Verlag, Germany.

Hoyer-Klick, C., Beyer, H., Dumortier, D., Schroedter-Homscheidt, M., Wald, L., Martinoli, M., Schillings, C., Gschwind, B., Menard, L., Gaboardi, E., Polo, J., Cebecauer, T., Huld, T., Šúri, M., de Blas, M., Lorenz, E., Kurz, C., Remund, J., Ineichen, P., Tsvetkov, A., Hofierka, J. (2009). MESoR - management and exploitation of solar resource knowledge. SolarPACES 2009. Berlin, Germany.

Huang, S. and Fu, P. (2009). Modeling small areas is a big challenge. ArcUser Spring 2009, http://www.esri.com/news/arcuser/0309/files/solar.pdf, last accessed on 19/01/2011.

Hunter, G. and Goodchild M. (1995). Dealing with error in spatial databases: a simple case study. *Photogrammetric Engineering and Remote Sensing,* 61(5): 529-537.

Hunter, G.J. (1993). Handling uncertainty in spatial database. Ph.D. Thesis, Department of Surveying and Land Information, University of Melbourne, Australia.

Hunter, G.J. (1999). New tools for handling spatial data quality: moving from academic concepts to practical reality. *URISA Journal,* 17: 261-268.

Hunter, G.J. and Goodchild, M.F. (1997). Modeling the uncertainty of slope and aspect estimates derived from spatial databases. *Geographical Analysis,* 29(1): 35-49.

Hunter, G.J., Caetano, M. and Goodchild, M.F. (1995). A methodology for reporting uncertainty in spatial database products. *Journal of the Urban and Regional Information Systems Association,* 7(2): 11-21.

Huss, R.E. and Pumar, M.A. (1997). Effect of databaseerrors on intervisibility estimation. *Photogrammetric Engineering and Remote Sensing*, 63: 415-24.

Hwang, D., Karimi, H.A. and Byun, D.W. (1998). Uncertainty analysis of environmental models within gis environments. *Computers & Geosciences,* 24(2): 119-130.

IEA International Energy Agency (2013). International energy statistics, http://www.iea.org/, last accessed on 06/10/2013.

Intergovernmental Panel on Climate Change - IPCC, (1996). Climate Change 1995: The Science of Climate Change. Cambridge, UK. *Cambridge University Press.*

IPCC (2001). Climate change 2001: mitigation, A report of Working Group III of the Intergovernmental Panel on Climate Change, http://www.grida.no/publications/ other/ipcc_tar/?src=/climate/ipcc_tar/wg3/index.htm, last accessed on 10/09/2013.

Isaaks, E.H., and R.M. Srivastava. (1989). An introduction to applied geostatistics. New York, *Oxford University Press*.

Isikwue, B.C., Amah, A.N., Agada, P. O. (2012). Empirical model for the estimation of global solar radiation in Makurdi, Nigeria. *Global Journal of Science Frontier Research Physics & Space Science,* 12(1): 59-61.

Iwahashi, J., Watanabe, S. and Furuya, T. (2001). Landform analysis of slope movements using DEM in Higashikubiki area, Japan. *Computers & Geosciences,* 27: 851-865.

Jahanpeyma, M., Delavar, M. R., Malek, M. R. and Kamalian, N. (2007). Analytical evaluation of uncertainty propagation in seismic vulnerability assessment of Tehran using GIS. Proc. ISSDQ 2007, Enschede, The Netherlands. Pages 13-15.

Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E. (2008). Hole-filled SRTM for the globe Version 4, http://srtm.csi.cgiar.org, last accessed on 7/08/2012.

Jones, C.B., Kidner, D.B. and Ware, J.M. (1994). The implicit triangulated irregular network and multiscale spatial databases. *The Computer Journal*, 37(1): 43-57.

Jones, J.W., Hoogenboomb, G., Portera, C.H., Bootea, K.J., Batchelorc, W.D., Huntd, L.A., Wilkense, P.W., Singhe, U., Gijsmana, A.J. and Ritchie, J.T. (2003). The DSSAT cropping system model. *European Journal of Agronomy,* 18(3-4): 235-265.

Kalkhan, M.A. (2011). Spatial statistics - geospatial information modelling and thematic mapping. *CRC Press*.

Karssenberg, D. and De Jong, K. (2005). Dynamic environmental modelling in GIS: 2. Modelling error propagation. *International Journal of Geographical Information Science*, 19(6): 623-637.

Kasten, F., Czeplak, G. (1979). Solar and terrestrial radiation dependent on the amount and type of cloud. *Solar Energy,* 24(2): 177-189.

Kimbler, D.L. and Knight, B.D. (1987). A survey of current methods for the elimination of initialization bias in digital simulation. In *Proc. 20th A. Simulation Symposium.* Pages 133-152.

Kirkpatrick, S., Gelatt, C.D. and Vecchi, M.P. (1983). Optimization by simulated annealing. *Science*, 220: 671-680.

Kryza, M., Szymanowski, M., Migała, K. and Pietras, M. (2010). Spatial information on total solar radiation - application and evaluation of *r.sun* model for the Wedel Jarlsberg Land, Svalbard. *Polish Polar Research,* 31(1):17–32.

Kumar, L. (2011). Reliability of GIS-based solar radiation models and their utilisation in agro-meteorological research, http://www.isprs.org/proceedings/2011/isrse34/2111 04015final00289.pdf, last accessed on 14/06/2012.

Kumler, M.P.(1994). An intensive comparison of triangulated ırregular networks (TIN) and digital elevation models (DEM). *Cartographica*, 31(2): 1-99.

Kyriakidis, P.C., Shortridge, A. M., and Goodchild, M.F. (1999). Geostatistics for conflation and accuracy assessment of digital elevation models. *International Journal of Geographical Information Science,* 13: 677-707.

Lam, N.S. (1983). Spatial interpolation methods: a review. *The American Cartographer* 10(2): 129-149.

Lawal, D.U., Matori, A., Hashim, A.M., Yusof, K.W. and Chandio, I.A. (2012). Detecting flood susceptible areas using GIS-based analytic hierarchy process. *International Conference on Future Environment and Energy,* IPCBEE vol.28, IACSIT Press, Singapore.

Legendre, P. and Fortin, M. (1989). Spatial pattern and ecological analysis. *Vegetatio,* 80: 107-138.

Lei, S., Daniels, John L., Bian, Z. and Wainaina, N. (2010). Improved soil temperature modeling. *Environmental Earth Sciences,* 62(6): 1123-1130.

Li, J. (2008). Spatial distribution of incoming potential solar radiation based on solar analyst model and DEM in Xinjiang, China", *Proc. SPIE* 7146, Geoinformatics 2008 and Joint Conference on GIS and Built Environment: Advanced Spatial Data Models and Analyses, 71462M (November 11, 2008); doi:10.1117/12.813190.

Li, S. and Lu, W. (2010). Automatic fit of the variogram. 2010 Third International Conference on Information and Computing, http://ieeexplore.ieee.org/stamp/ stamp.jsp? arnumber=05514035, last accessed on 21/07/2012.

Liang, F., Liu, C. And Carroll, R.J. (2010). Advanced Markov chain Monte Carlo methods: Learning from past samples. UK, *John Wiley & Sons, Ltd*.

Liu, X. and Bian, L. (2008). Accuracy assessment of DEM slope algorithms related to spatial autocorrelation of DEM errors. In Zhou, Q., Lees, B. and Tang, G. (Eds.) Advances in digital terrain analysis. Berlin, Germany. *Springer-Verlag*. Pages 307-322.

Liu, X.F. (2009). Quantum simulation of Markov Chains, http://arxiv.org/pdf/0909.4623. pdf, last accessed on 21/11/2011.

Maling, D.H. (1989). Measurements from maps. New York, *Pergamon Press*.

Mariscal, M.J., Orgaz, F., Villalobos, F.J. (2000). Modelling and measurement of radiation interception by olive canopies. *Agricultural and Forest Meteorology*, 100: 183-197.

Martínez-Durbán, M., Zarzalejo, L.F., Bosch, J.L., Rosiek, S., Polo, J., Batlles, F.J. (2009). Estimation of global daily irradiation in complex topography zones using digital elevation models and meteosat images: comparison of the results. *Energy Conversion Management*, 50(9): 2233–2238.

Martins, F.R., Pereira, E.B., Ebreu, S.L. (2007). Satellite-derived solar resource maps for Brazil under SWERA project. *Solar Energy*, 81: 517-528.

Math World, (2013). K-means clustering algorithm, http://mathworld.wolfram.com/K-MeansClusteringAlgorithm.html, last accessed on 7/07/2013.

Mckenney, D.W., Mackey, B.G. and Zavitz, B.L. (1999). Calibration and sensitivity analysis of a spatially−distributed solar radiation model. *International Journal of Geographic Information Science*, 13 (1): 49-65.

Mészároš, I., Miklánek, P. and Parajka, J. (2002). Solar energy income modelling in mountainous areas. In L. Holko, P. Miklánek, J. Parajka and Z. Kostka (eds), ERB and NEFRIEND Proj.5 Conf. Interdisciplinary Approaches in Small Catchment Hydrology: Monitoring and Research. Bratislava: Slovak NC IHP UNESCO/UH SAV. Pages 127-135.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21: 1087-1091.

Meyn, S.P. and Tweedie, R.L. (1993). Markov chains and stochastic stability. New York, USA, *Springer*.

Mikhail, E.M. and Ackermann, F. (1976). Observations and least squares. Lanham, MD. *University Press of America*. Page 497.

Mikhail, E.M., McGlone, J.C. and Bethel, J.S. (2001). Introduction to modern photogrammetry, New York, USA. *John Wiley & Sons Ltd*.

Miklanek, P. (1993). The estimation of energy incoming grid points over the basin using simple digital elevation model. *Annales Geophysicae*, 11: 296.

Miklánek, P. and Mészároš, I. (1993): Modelling of insolation characteristics in mountainous environment. In: *Advances in water sciences, Vol.II: Hydrological processes and transport phenomena in environment*. UH SAV, Stara Lesna, Slovakia. Pages 32 - 36.

Miliaresis, G.C. (2008). Quantification of terrain processes. In Zhou, Q., Lees, B. and Tang, G. (Eds.) Advances in digital terrain analysis. Berlin, Germany, *Springer-Verlag*. Pages 13-28.

Minasny, B., Vrugt, J.A. and McBratney, A.B. (2011). Confronting uncertainty in model-based geostatistics using Markov chain Monte Carlo simulation. *Geoderma*, 163: 150-162.

Monogan, J. (2010). Markov Chain Monte Carlo, http://stats.wustl.edu/~jmonogan/teaching/statComp/mcmc.pdf, last accessed on 03/02/2011.

Moore, I.D., Grayson, R.B. and Lfadson, A.R. (1991). Digital terrain modeling - a review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, 5(1): 3-30.

Mowrer, H.T. (1997). Propagating uncertainty through spatial estimation processes for old-growth subalpine forests using sequential Gaussian simulation in GIS. *Ecological Modelling*, 98: 73-86.

Muneer, T. and Yieng Wei Tham, Y.W. (2013). Solar energy measurements. In Kutz, M. (ed.), A multidisciplinary reference of engineering measurement tools, techniques, and applications-Volume 1. *John Wiley & Sons, Inc.*

Murray, C.J., Chien, Y.-J. and Thorne, P.D. (2004). A geostatistical analysis of historical field data on tritium, technetium-99, iodine-129, and uranium. *Pacific Northwest National Laboratory Richland, Washington*. http://www.pnl.gov/main/publications/external/technical_reports/pnnl14618rev0.pdf, last accessed on 12/06/2012.

NASA (2005). Shuttle radar topography mission, http://www2.jpl.nasa.gov/srtm/mission.htm, last accessed on 20/08/2011.

NASA (2009). Solar: monthly and annual average direct normal irradiance GIS data at 11 resolution of the world from NASA/SSE, http:eosweb.larc.nasa.gov/sse, last accessed on 10/01/2012.

NASA (2103). Surface meteorology and solar energy, http://eosweb.larc.nasa.gov/cgi-bin/sse/sse.cgi/, last accessed on 10/06/2013.

Navarro, D. and Perfors, A. (2011). The Metropolis-Hastings algorithm, http://www.psychology.adelaide.edu.au/personalpages/staff/danielnavarro/ccs2011/technote_metropolishastings.pdf,last accessed on 24/07/2012.

Neal, R.M. (2003). Slice sampling. *The Annals of Statistics*, 31(3): 705-767.

Neteler, M. and Mitasova, H. (2004). Open source GIS: a GRASS GIS approach. Boston, USA, *Kluwer Academic Publishers*.

Nguyen H.T. and Pearce J.M. (2010). Estimating potential photovoltaic yield with *r.sun* and the open source Geographical-Resources Analysis Support System. *Solar Energy*, 84: 831-843.

NIST (1992). Spatial data transfer standard (FIPS 173). Washington DC, USA, National Institute of Standards and Technology, US Department of Commerce.

Nummelin, E. (1984). General irreducible Markov chains and non-negative operators. Cambridge, UK, *Cambridge University Press*.

O'Callaghan, J.F. and Mark, D.M. (1984). The extraction of drainage networks from digital elevation data. Computer Vision, *Graphics and Image Processing*, 28: 323-344.

Oksanen, J. and Sarjakoski, T. (2005a). Error propagation of DEM-based surface derivatives. *Computers & Geoscience*, 31: 1015-1027.

Oksanen, J. and Sarjakoski, T. (2005b). Error propagation analysis of DEM-based drainage basin delineation. *International Journal of Remote Sensing*, 26(14): 3085-3102.

Olea, R.A. (1999). Geostatistics for engineers and earth scientists, Boston, USA, *Kluwer Academic Press*.

Oliver, D.S., Cunha, L.B. and Reynolds, A.C. (1997). Markov chain Monte Carlo methods for conditioning a permeability field to pressure data. *Mathematical Geology,* 29: 61-91.

Oliver, M., Webster, R. and Gerrard, J. (1989). Geostatistics in physical geography. Part II: applications. *Transactions of the Institute of British Geography,* 14: 270-286.

Onstad, C.A. and Brakensiek, D.L. (1968). Watershed simulation by stream path analogy. *Water Resources Research,* 4(5): 965-971.

OPEC (2013). OPEC annual statistical bulletin. http://www.opec.org/opec_web/static_files_project/media/downloads/publications/ASB2013.pdf, last accessed on 06/10/2013.

Ortiz, J.M. and Leuangthong, O. (2007). A practical approach to validate the variogram reproduction from geostatistical simulation, http://www.captura.uchile.cl/jspui/bitstream/2250/10330/1/124905_C2_2007-06-APCOMOrtizLeuangthong.pdf, last accessed on 20/07/2012.

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R. and Vignola, F. (2002). A new operational model for satellite-derived irradiances: description and validation. *Solar Energy,* 73(5): 307-317.

Pons, X. and Ninyerola, M. (2008). Mapping a topographic global solar radiation model implemented in a GIS and refined with ground data, *International Journal of Climatology,* 28: 1821-1834.

Power Holding Company of Nigeria (PHCN) (2005) Generation and transmission annual reports. Osogbo-Nigeria.

Prodan, A. and Prodan, R. (2001). Stochastic simulation and modeling, http://epp.eurostat.ec.europa.eu/cache/ITY_PUBLIC/NTTS2001/46.pdf, last accessed on 11/05/2012.

Raaflaub, L.D. and Collins, M.J. (2006). The effect of error in gridded digital elevation models on the estimation of topographic parameters. *Environmental Modelling & Software*, 21(5): 710-732.

Raftery, A.E. and Lewis, S.M. (1992), How many iterations in the Gibbs sampler? In *Bayesian Statistics* 5 (Bernardo, J.M., Berger, J.O., Dawid, A.P. and Smith, A.F.M. (eds.), Proceedings of the Fourth Valencia International Conference on Bayesian Statistics. New York, USA. *Oxford University Press.* Pages 763-774.

Refsgaard, J.C., van der Sluijs, J.P., Højberg, A.L. and Vanrolleghem, P.A. (2007). Uncertainty in the environmental modelling process: a framework and guidance. *Environmental Modelling & Software,* 22: 1543-1556.

Remund J., Wald L., Lefevre M., Ranchin T. and Page J. (2003). Worldwide linke turbidity information. Proceedings of ISES Solar World Congress, 16-19 June, G.teborg, Sweden, CD-ROM. *International Solar Energy Society*.

Rich, P.M., Hetrick, W.A. and Saving, S.C. (1995). Modeling topographic influences on solar radiation: a manual for the Solarflux model. Los Alamos National Laboratory Report LA-12989-M.

Ritter, C. and Tanner, M.A. (1992), The Gibbs stopper and the griddy Gibbs sampler. *Journal of the American Statistical Association,* 87: 861-868.

Robert, C. (1996). Convergence assessments for Markov chain Monte Carlo methods. *Stastistical Science,* 10: 231-253.

Robert, C.P. and Casella, G. (1999). Monte Carlo statistical methods. New York, USA, *Springer-Verlag*.

Robert, C.P. and Casella, G. (2010). Introductory to Monte Carlo methods with R. New York, USA, *Springer*. Page 167.

Roberts, G.O. (1992). "Convergence diagnostics of the Gibbs sampler," in *Bayesian Statistics 4*, eds. J. M. Bernardo, J. Berger, A . P. Dawid, and A. F. M. Smith, Oxford, UK, *Oxford University Press.* Pages 775-782.

Rodríguez, E. Morris, C.S., Belz, J.E., Chapin, E.C., Martin, J.M., Daffer, W. and Hensley, S. (2005). An assessment of the SRTM topographic products. *Jet Propulsion Laboratory*. Pasadena, California -USA.

Rodríguez, E., Morris, C.S. and Eric Belz, J. (2006). A global assessment of the SRTM performance. *Photogrammetric Engineering & Remote Sensing.* 72(3): 249-260.

Rubinstein, R.Y. and Kroese, D.P. (2007). Simulation and the Monte Carlo method. New Jersy, USA. *John Wiley & Sons Ltd*. Page 175.

Ruiz-Arias, J.A., Alsamamra, H., Tovar-Pescador, J. and Pozo-Vazquez, D. (2010). Proposal of a regressive model for the hourly diffuse solar radiation under all sky condition. *Energy Conversion and Management*, 51: 881-893.

Ruiz-Arias, J.A., Tovar-Pescador, J., Pozo-Vazquez, D. and Alsamamra (2009). A comparative analysis of DEM-based models to estimate the solar radiation in mountainous terrain. *International Journal of Geographic Information Science,* 23(8): 1049-1076.

Sambo, A. S. (2005). Renewable energy for rural development: the Nigerian perspective, ISESCO Science and Technology Vision, 1 (May 2005): 12-22.

Sandborn, A.N., Griffiths, T.L. and Shiffrin, R.M. (2010). Uncovering mental representations with Markov chain Monte Carlo. *Cognitive Pyschology,* 60(2): 63-106.

Scharmer K. and Greif J. (eds) (2000). The European solar radiation atlas, Volume 2: Database and Exploitation Software. Les Presses de l'École des Mines, Paris, http://www.helioclim.net/esra/index.html, last accessed on 10/07/2011.

Senegas, J., Schmitt, M. and Nonin, P. (2006). Conditional simulation applied to uncertainty assessment in DTMs. In Foody, G.M. and Atkinson, P.M., editors, Uncertainty in Remote Sensing and GIS. Chichester, England. *John Wiley & Sons, Ltd.*

SFITZ, Swiss Federal Institute of Technology Zurich (2006). Advanced Monte Carlo methods I & II: An ETHZ/SAM course, http://www.cs.fsu.edu/~mascagni/ Advanced _Monte_Carlo_Methods.html, last accessed on 5/05/2012.

Shekhar, S. and Xiong, H. (2008). Encyclopedia of GIS. *Springer-Verlag*.

Sheng, J., Wilson, J.P. and Lee, S. (2009). Comparison of land surface temperature (LST) modeled with a spatially-distributed solar radiation model (SRAD) and remote sensing data. *Environmental Modelling & Software*, 24: 436-443.

Shortridge, A.M. (2001). Characterizing uncertainty in digital elevation models. In Spatial uncertainty in ecology: implications for remote sensing and GIS applications. Hunsaker, C.T., Goodchild, M.F., Friedl, M.A. and Case, T.J. (Editors). New York, USA, *Springer-Verlag*. Pages 238-257.

Smith, T.E. (2013). Notebook for spatial data analysis: Part II. Continuous Spatial Data Analysis, http://www.seas.upenn.edu/~ese502/NOTEBOOK/Part_II/4_Variograms .pdf, last accessed on 25/6/2013.

SoDA (2012): Services for professionals in solar energy and radiation, http://www.soda-is.com/eng/services/climat_free_eng.php#lin, last accessed on 7/09/2012.

SoDA (2012a): Linke turbidity. http://www.soda-is.org/eng/services/Linke_Turbidity_info. html?&lang=en_us, last accessed on 7/08/2012.

Sorapipatana, C. (2010). An assessment of solar energy potential in Kampuchea. *Renewable and Sustainable Energy Reviews,* 14: 2174-2178.

Stockle, C.O., Donatelli, M. and Nelson, R. (2003). CropSyst, a cropping systems simulation model. *European Journal of Agronomy*, 18: 289-307.

Sun, G., Ranson, K., Khairuk, V. and Kovacs, K. (2003). Validation of surface height from shuttle radar topography mission using shuttle laser altimeter, *Remote Sensing of Environment*, 88(4):401-411.

Šúri M., Huld T.A., Dunlop E.D. and Ossenbrink H.A. (2007). Potential of solar electricity generation in the European Union member states and candidate countries. *Solar Energy,* 81: 1295-1305.

Šúri, M. (2002). Modelovanie a kartografické zobrazovanie globálneho žiarenia. In Feranec, J. Pravda J., eds. *Proceeding of the workshop "Aktivity v kartografii 2002"*, Bratislava (Slovak Cartographic Society and Institute of Geography SAS). Pages 126-133.

Šúri, M. (2007). Solar resource data and tools for an assessment of photovoltaic systems. In Jäger-Waldau A. (ed), Status Report 2006, Office for Official Publications of the European Communities, Luxembourg. Pages 93-98. http://iamest.jrc.it/pvgis/doc/report/solar_inventory_2006.pdf.

Šúri, M. and Hofierka, J. (2004). A new GIS-based solar radiation model and its application to photovoltaic assessment. *Transactions in GIS,* 8(2): 175-190.

Šúri, M., Huld T.A. and Dunlop E.D. (2005). PVGIS: a web-based solar radiation database for the calculation of PV potential in Europe. *International Journal of Sustainable Energy,* 24: 55-67.

Šúri, M., Jones, A.R., Dunlop, E.D. (2002). Mapping the potential of photovoltaic systems in urban areas of Slovakia. In: Brebbia, C.A., Martin-Dunque, J.F., Wadhwa, L.C. (eds), The Sustainable City II, Urban Regeneration and Sustainability. Southampton, UK, *WIT press*. Pages 1007-1016.

Tajchman. S.J. (1981). On computing topographic characteristics of a mountain catchment. *Canadian Journal of Forest Research,* 11: 768-774.

Tate, E. and Maidment, D. (1999). Floodplain mapping using HEC-RAS and ArcView GIS,http://www.crwr.utexas.edu/gis/gishydro04/introduction/digitallibrary/tate/rpt99-1.pdf, last accessed on 27/06/2012.

Thapa, K. and Bossler, J. (1992). Accuracy of spatial data used in geographic information systems. *Photogrammetric Engineering & Remote Sensing,* 58: 835-841.

The International Research Center for Energy and Economic Development (ICEED) (2006): Renewable electricity action program (REAP). Federal Ministry of Power and Steel, Federal Republic of Nigeria

Theobald, D.M.(1989). Accuracy and bias issues in surface representation. In: Goodchild, M.F., Gopal, S. (Eds.), The accuracy of spatial database. New York, USA, *Taylor and Francis*. Pages 99-106.

Thompson, G. (2003). Effects of DEM resolution on GIS-based solar radiation model output: a comparison with the national solar radiation database. An MA thesis submitted to the Graduate School of the University of Cincinnati, USA.

Tobler, W.R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography,* 46: 234-40.

Tovar-Pescador, J., Pozo-Vazquez, D., Ruiz-Arias, J.A., Batlles, J., Lopez, G. and Bosch, J.L. (2006). On the use of the digital elevation model to estimate the solar radiation in areas of complex topography. *Meteorological Applications,* 13(3): 279-287.

Tsai, V. J. D. (1993). Delaunay triangulations in TIN creation: An overview and a linear-time algorithm. *International Journal of Geographic Information Science,* 7(6): 501-524.

UNCSD (UN Commission on Sustainable Development) (2002). Plan of implementation of the world summit on sustainable development. http://www.un.org/esa/sustdev/documents/WSSD_POI_PD/English/WSSD_PlanImpl.pdf

University Consortium for Geographic Information Science [UCGIS] (1998). Uncertaınty in geographic data and GIS-based analyses, http://www.ucgis.org/priorities/research/research_white/1998%20Papers/uncert.html, last accessed on 5/06/2011.

van Zyl, J. and Labadie, J.W. (2011). GIS analysis for multicriteria reservoir site selection: NISP project, Colorado, http://proceedings.esri.com/library/userconf/proc11/papers/3758_75.pdf, last accessed on 28/06/2012.

Vann, J., Bertoli, O. and Jackson, S. (2002). An overview of geostatistical simulation for quantifying risk. Paper originally published at a Geostatistical Association of Australasia symposium “Quantifying Risk and Error” March 2002.

Veregin, H. (1999). Data quality parameters. In: Longley, P.A., Goodchild, M.F., Maguire, D.J., and Rhind, D.W. (eds.): Geographical information systems, Volume 1 - Principles and technical issues: New York, *John Wiley and Sons*. Pages 177-189.

Vivoni, E.R., Ivanov, V.Y., Bras, R.L. and Entekhabi, D. (2004). Generation of triangulated irregular networks based on hydrological similarity. *Journal of Hydrologic Engineering*, 9(4): 288-302.

Wang, G. and Chen, S. (2012). A review on parameterization and uncertainty in modeling greenhouse gas emissions from soil. *Geoderma*, 170: 206-216.

Wang, P., Du, J., Feng, X. and Hu, S. (2006). Effect of DEM uncertainty on the distributed hydrological model TOPMODEL. *Chinese Geographical Science*, 16(4): 320-326.

Watson, D. F., and Philip, G. M. (1984). Systematic triangulations. *Comput. Vis. Graph. Image Process.*, 2: 217-223.

Wechsler, S. (2007). Uncertainty associated with digital elevation models for hydrologic applications: a review. *Hydrology and Earth System Sciences Discussions,* 3: 2343-2384.

Wechsler, S.P. (1999). Digital elevation model (DEM) uncertainty: evaluation and effect on topographic parameters, http://proceedings.esri.com/library/userconf/proc99/proceed/papers/pap262/p262.htm#References, last accessed on 28/05/2012.

Wechsler, S.P. (2000). Effect of DEM uncertainty on topographic parameters, DEM scale and terrain evaluation. Unpublished PhD dissertation, State University of New York, USA.

Wechsler, S.P. and Kroll, C.N. (2006). Quantifying DEM uncertainty and its effect on topographic parameters. *Photogrammetric Engineering & Remote Sensing,* 72(9): 1081-1090.

Weibel, R. and Brändli, M. (1995). Adaptive methods for the refinement of digital terrain models for geomorphometric applications. *Zeitschrift für Geomorphologie*, 11: 13-30.

Wendy, L.M and Angel, R.M. (2007). Computational statistics handbook with MaTLAB. 2nd edition. Boca Raton, *Chapman & Hall*.

Weng, Q. (2002). Quantifying uncertainty of digital elevation models derived from topographic maps.In Richardson, D. and van Oosterom, P. (eds): Advances in spatial data handling. New York, USA, *Springer-Verlag*. Pages 403-418.

White, D.J., Thompson, M.J. and Vennapusa, P.K. (2007). Field validation of intelligent compaction monitoring technology for unbound materials (Final report). *Minnesota Department of Transportation*, Minnesota, USA.

Wilson, J.P. and Gallant, J.C. (2000). Secondary topographic attributes. In: Wilson, J.P., Gallant, J.C. (Eds.), Terrain analysis: principles and applications. New York, USA. *John Willey & Sons.* Pages 51-85.

Winslow, J.C., Hunt Jr, E.R. and Piper, S.C. (2001). A globally applicable model of daily solar irradiance estimated from air temperature and precipitation data. *Ecological Modelling*, 143: 227-243.

World Bank (2012). Nigeria, http://data.worldbank.org/country/nigeria, last accessed on 01/12/2012.

World Health Organization (2005). Principles of characterizing and applying human exposure models. IPCS harmonization project document no. 3. WHO Document Production Services, Geneva, Switzerland.

Yu, B. (1994), Monitoring the convergence of Markov samplers based on estimated L^1 error. Technical Report 409, University of California at Berkeley, Dept. of Statistics.

Yu, S., van Kreveld, M., and Snoeyink, J. (1997) Drainage queries in TINs: from local to global and back again. In Kraak, M.J. and Molenaar, M. (eds) *Advances in GIS Research II: Proceedings of the Seventh International Symposium on Spatial Data Handling.* London, Taylor and Francis. Page 829-42.

Zellner, A. and Min, C.K. (1995). Gibbs sampler convergence criteria. *Journal of the American Statistical Association,* 90: 921-927.

Zhang, J. and Goodchild, M.F. (2002). Uncertainty in geographic information, *Taylor and Francis,* London, UK.

Zhao, W.D. and Tang, G.A. (2008). A new data structure for refined digital elevation models. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B4, Beijing, http://www.isprs. org/proceedings/XXXVII/congress/4_pdf/107.pdf.

Zhou, J., Wu, Y.Z. and Yan, G.(2005). General formula for estimation of monthly average daily global solar radiation in China. *Energy Conversion & Management,* 46: 257-268.

Zhou, Q. and Liu, X. (2004). Analysis of errors of derived slope and aspect related to DEM data properties. *Computers & Geosciences,* 30(4): 369-378.

Zhou, Q. and Liu, X. (2004). Error analysis on grid-based slope and aspect algorithms. *Photogrammetric Engineering and Remote Sensing*, 70(8): 957-962.

APPENDIX A

R CODES

```
# load library
library(rgdal)
library(maptools)
library(spdep)
library(spgwr)
library(gstat)
library(automap)
library(spatstat)
library(lattice)
library(sp)
library(RColorBrewer)
library(mgcv)
library(rpart)
library(kernlab)
library(StatDA)
library(raster)
library(ape)

station<-readShapePoints("C:\\abdurrahman\\dem_sa.shp")
variogram = autofitVariogram(dem_sa,elev)

plot(variogram)
variogram
```

APPENDIX B

MATLAB CODES FOR THE PROPOSED MCMC SIMULATION

I. Main

```
clc; % clear command window
clear; % clear variables and functions from memory and removes all variables from the
workspace

global range mu_normal var_normal spat_autoco length_of_iteration array_size
mparalel_chain nb_of_files original_dem % global variables

sill=125.74 % Variogram parameters (depends on array_size - study area)
range=1556.85 % Variogram parameters (depends on array_size - study area)

lognormal_mean=4.2094;lognormal_var=0.05421; % lognormal pdf parameters
mu_normal=log(lognormal_mean)-0.5*log(1+(lognormal_var/(lognormal_mean^2))); %
calculates normal mean from lognormal mean
var_normal=log(1+(lognormal_var/(lognormal_mean^2))); % calculates normal variance
from lognormal variance
length_of_iteration=1080 % number of iteration
array_size=1225 % study area size (must be a square area)
spat_autoco=spatial_autoco(array_size); % calls spat_autoco function
normal_mean=0;normal_sigma=var_normal; % parameters of normal distribution for
generating candidates (arbitrary)
mparalel_chain=2 % number of different runs
mean_matrix=[]; % initialize mean_matrix with empty matrix
threed_storage_matrix=zeros(array_size,length_of_iteration+1,mparalel_chain); %
initialization of 3D storage matrix
nb_of_files=length_of_iteration % number of image files to be created (last columns to
be selected)
original_dem = double((imread('dem_35x35.tif'))); % study area dem data

for k=1:mparalel_chain

    initial_array=rand(array_size,1)*0.1+1.4; % random array for mparalel chains
    dispersed starting values
    current_array=initial_array; % current sample
    storage_matrix=initial_array; % stores the initial array
```

```
    for i=1:length_of_iteration

    candidate_array=normrnd(normal_mean,normal_sigma,[array_size,1])+current_array; % generate candidate vector of random numbers from normal distribution
    p_candidate=accept_prob(candidate_array); % calculate probability of candidate
    p_current=accept_prob(current_array); % calculate probability of current state
    acceptance_criteria=min([1 p_candidate/p_current]); % calculates the acceptance criteria

    if acceptance_criteria-rand>0 % subtracts acceptance criteria from a random value (scalar) and compare with zero (0)
            current_array=candidate_array; % if the previous process is greater than zero then current array is equal to candidate array
            storage_matrix=[storage_matrix candidate_array]; % appends the storage matrix  with the accepted candidate
    else
            storage_matrix=[storage_matrix current_array]; % appends the storage matrix with the current_array
end

end

if (k==1)
    threed_storage_matrix(:,:,1)=storage_matrix; % stores the storage matrix as the first item in the 3D storage matrix drawer when k==1
else
    threed_storage_matrix=cat(3,threed_storage_matrix,storage_matrix); % drawer gets deeper with the new storage matrix
end

    mean_matrix=[mean_matrix mean(storage_matrix,2)]; % row-wise mean of each storage matrix in each chain

end

% Brooks and Gelman multivariate potential scale reduction factor (MPSRF)convergence diagnosis
w_chainvariance=within_chain_var(mean_matrix,threed_storage_matrix); % calls the within chain variance function
btw_chain_var=between_chain_var(mean_matrix); %calls the between chain variance function
v_hat=(length_of_iteration-1)/length_of_iteration*w_chainvariance*(1+1/mparalel_chain)*btw_chain_var % calculates the variance estimate
multivariate_potential_scale_reduction_factor=sqrt(v_hat) % calculates the multivariate potential scale reduction factor (MPSRF)
r_hata=sqrt(v_hat/w_chainvariance)
zz_mat=(1/length_of_iteration)*inv(w_chainvariance)*btw_chain_var;
```

```
r_hat=sqrt((length_of_iteration-1)/length_of_iteration+(mparalel_chain+1)/mparalel_chain*max(eig(zz_mat)))

% variogram reproduction check
ref_variogram_model=reference_variogram(sill,range) % calls reference var function (reference variogram)

% selects last n-samples of the iteration from the storage matrix
svar_of_some_dist_in_last_nsamples_mat=[]; % initialize an emty matrix for semivariance of specific distances
mean_array_of_svar_of_some_dist_in_last_nsamples=[]; % initialize an emty matrix to store row-wise mean of semivariances for specific distances of n-samples

for n=length_of_iteration:-1:length_of_iteration-99 % creates a dummy variable for last n-samples (n - pointer and iteration)

    n;
    svar_mat_of_last_nsamples=semivariance_cal(storage_matrix(:,n),array_size); % calculates the semivariance for each of the last 500 samples using semivariance_cal
    mean_lag_dist_array=mean_lag_dist_cal(svar_mat_of_last_nsamples); % calculates mean semivariance of all distances using mean_lag_dist_cal function
    %important for defining the corresponding distances to reference variogramdistance
    svar_of_some_dist_in_last_nsamples=mean_lag_dist_array(2:sqrt(array_size):array_size); % mean semivariance of distances corresponding to those of reference semivariance model
    svar_of_some_dist_in_last_nsamples_mat=[svar_of_some_dist_in_last_nsamples_mat svar_of_some_dist_in_last_nsamples]; % creates a matrix that contains semivariances for specific lag distances of last n-samples

end

mean_array_of_svar_of_some_dist_in_last_nsamples=[mean_array_of_svar_of_some_dist_in_last_nsamples mean(svar_of_some_dist_in_last_nsamples_mat,2)]
diff_btw_mean_array_n_ref_var=mean_array_of_svar_of_some_dist_in_last_nsamples-ref_variogram_model; % difference between mean of lag distances and reference variogram
sample_cov_mat=sample_cov_mat_cal(svar_of_some_dist_in_last_nsamples_mat,mean_array_of_svar_of_some_dist_in_last_nsamples); % computes the covariance matrix using svar_of_some_dist_in_last_nsamples_mat
svar_of_some_dist_in_last_nsamples_mat_size=size(svar_of_some_dist_in_last_nsamples_mat); %calculates the number of last n-samples
T_square=svar_of_some_dist_in_last_nsamples_mat_size(2)*diff_btw_mean_array_n_ref_var'*sample_cov_mat*inv(sample_cov_mat)*diff_btw_mean_array_n_ref_var % calculates T-squares
fdist_1=(svar_of_some_dist_in_last_nsamples_mat_size(2)-1)*svar_of_some_dist_in_last_nsamples_mat_size(1)/(svar_of_some_dist_in_last_nsamples_mat_size(2)-svar_of_some_dist_in_last_nsamples_mat_size(1))
```

```
f_tabulated =
finv(0.95,svar_of_some_dist_in_last_nsamples_mat_size(1),svar_of_some_dist_in_last_
nsamples_mat_size(2)-svar_of_some_dist_in_last_nsamples_mat_size(1))
f_dist_value=fdist_1*f_tabulated % f-distribution value for alpha=0.05

% Transformation of storage matrix from array to matrix

storage_matrix_size=size(storage_matrix); % finds the size of storage matrix
lastmatrices_lognorm_3d_mat=zeros(sqrt(array_size),sqrt(array_size),nb_of_files); %
initialization of 3D storage matrix
new_dems_3d_mat=zeros(sqrt(array_size),sqrt(array_size),nb_of_files); % initialization
of 3D storage matrix

for c=1:nb_of_files

    column=storage_matrix(:,storage_matrix_size(2)-c); % finds the last n columns of
    the storage matrix
    lastmatrices=reshape(column,sqrt(array_size),sqrt(array_size)); % reshape the last n
    columns of the storage matrix to a matrix
    lastmatrices_lognorm=exp(lastmatrices); % transform the last n-matrices to
    lognormal
    lastmatrices_lognorm_3d_mat(:,:,c)=lastmatrices_lognorm; % the 3d matrix is filled
    with lastmatrices_lognorm
    lastmatrices_lognorm_32=uint16(lastmatrices_lognorm.*1000); % convert float to
    integer
    imwrite(lastmatrices_lognorm_32,strcat('sim_error_',num2str(c),'.tif')); % Transform
    the matrix to image files

end

% calculates mean, standard deviation and coefficient of variation using simulated errors
mean_error=error_mean(lastmatrices_lognorm_3d_mat) % calls the function of mean
error
std_dev_of_error=error_std_dev(lastmatrices_lognorm_3d_mat,mean_error) % calls
function standard deviation error
coeff_of_var=(std_dev_of_error./mean_error)*100 % calculate the coefficient of
variations

% converts mean error to integer and write as tif file
mean_error_32=uint16(mean_error.*1000);
imwrite(mean_error_32,strcat('mean_dem','.tif'));

% convert standard deviation of error to integer and write as tif file
std_dev_of_error_32=uint16(std_dev_of_error.*1000);
imwrite(std_dev_of_error_32,strcat('stddev_of_error','.tif'));
```

II. Functions

a. Acceptance Probability

```
function acceptance_prob=accept_prob(candidate_array) % declares the function
acceptprob

global mu_normal spat_autoco % declares the global variables

size_of_cand_array=size(candidate_array);

m=1/(2*3.14)^(size_of_cand_array(1)/2);
n=1/(sqrt(norm(spat_autoco)));
q=exp(-0.5*(candidate_array-mu_normal)'*inv(spat_autoco)*(candidate_array-
mu_normal));

acceptance_prob=m*n*q;
```

b. Spatial Autocorrelation

```
function spat_autoco=spatial_autoco(array_size) % declares function spatial_autoco
which calculate spatial autocorrelation

global range var_normal % global variables

for i=1:array_size
        for j=1:array_size
                distmat(i,j)=abs(j-i);  % calculates distance matrix
                auto_covariance(i,j)=var_normal*(exp(-distmat(i,j)*range/
                array_size/range/0.3)); % autocovariance matrix of normal
                distribution
        end
end
spat_autoco=auto_covariance; % autocovariance matrix of normal r.v.s
end
```

c. Within Chain Variance

```
function w_chainvariance=within_chain_var(mean_matrix,threed_storage_matrix)
% calculates within chain variance

global length_of_iteration mparalel_chain % declaration of global variables
sum=0;

for g=1:mparalel_chain
        for h=1:length_of_iteration
                a=threed_storage_matrix(:,h,g)-mean_matrix(:,g);
                sum=sum+a'*a;
end
```

```
end
w_chainvariance=1/(mparalel_chain*(length_of_iteration-1))*sum;
```

d. Between Chain Variance

```
function btw_chain_var=between_chain_var(mean_matrix)

global mu_normal mparalel_chain % declaration of global variables

for i=1:mparalel_chain
        b=mean_matrix(:,i)-mu_normal;
        sum=b'*b;
end
btw_chain_var=1/(mparalel_chain-1)*sum;
```

e. Reference Variogram

```
function ref_variogram_model=reference_variogram(sill,range) % calculates
reference variogram model

global array_size % defines the global variable

for lag_dist=1:sqrt(array_size) % defines the distances
        lag_dist_values(lag_dist)=sill*exp(-3*(lag_dist*sqrt(array_size))/range); %
        calculates values for different lag distances
end
ref_variogram_model=lag_dist_values';
```

f. Semivariance Calculation

```
function semivariance_mat_of_samples=semivariance_cal(array,array_size) %
declares function

for i=1:array_size % assign values to the dummy variable for iteration with step size
        for j=1:array_size %
                semivariance_mat(i,j)=0.5*(array(i)-array(j))^2;  % calculates the
                semivarince of ith and jth of array
        end
end
semivariance_mat_of_samples=semivariance_mat; % calculates the semivariance
matrix of the sample
end
```

g. Mean Lag Distance

```
function mean_lag_dist_array=mean_lag_dist_cal(semivariance_mat_of_samples)

global array_size % declares global variables

for i=1:array_size
```

```
            diagonal_means(i)=mean(diag(semivariance_mat_of_samples,i-1));
end
mean_lag_dist_array=diagonal_means';
```

h. Sample Covariance Matrix

```
function
sample_cov_mat=sample_cov_mat_cal(svar_of_some_dist_in_last_nsamples_mat,m
ean_array_of_svar_of_some_dist_in_last_nsamples);

sum=0;
svar_of_some_dist_in_last_nsamples_mat_size=size(svar_of_some_dist_in_last_nsa
mples_mat); % calculates the number of culumns in semivariancefor specific
distances of last n-samples matrix (i.e. number of selected last iterations)

for i=1:svar_of_some_dist_in_last_nsamples_mat_size(2)
            z=svar_of_some_dist_in_last_nsamples_mat(:,i)-
            mean_array_of_svar_of_some_dist_in_last_nsamples;
            sum=sum+z*z';
end
sample_cov_mat=sum/(svar_of_some_dist_in_last_nsamples_mat_size(2)-1);
```

i. Mean Error

```
function [mean_error]=error_mean(lastmatrices_lognorm_3d_mat)

global nb_of_files % global variable

size_matrix = size(lastmatrices_lognorm_3d_mat);
sum = zeros(size_matrix(1),size_matrix(2));

for i=1:size_matrix(1)
            for j=1:size_matrix(2)
                        for k=1:size_matrix(3)
                                    temp = lastmatrices_lognorm_3d_mat(i,j,k);
                                    sum(i,j)= [sum(i,j)+temp];
                        end
                        temp =[];
            end
end
sum;
mean_error=sum./(nb_of_files);
```

j. Standard Deviation of Error

```
function [std_dev_map]=stddev_map(new_dems_3d_mat,mean_dem_map)

global array_size nb_of_files % global variables
sum=zeros(sqrt(array_size),sqrt(array_size));
```

```
for i=1:sqrt(array_size) % row function
          for j=1:sqrt(array_size) % column function
                    for k=1:nb_of_files
                     deviations =(new_dems_3d_mat(i,j,k)-mean_dem_map(i,j))^2;
                    sum(i,j)=[sum(i,j)+deviations];
          end
          deviations =[];
end
          sum;
          std_dev=sqrt(sum./nb_of_files);
end
std_dev_map=std_dev;
```

APPENDIX C

SIMULATED DEM ERRORS

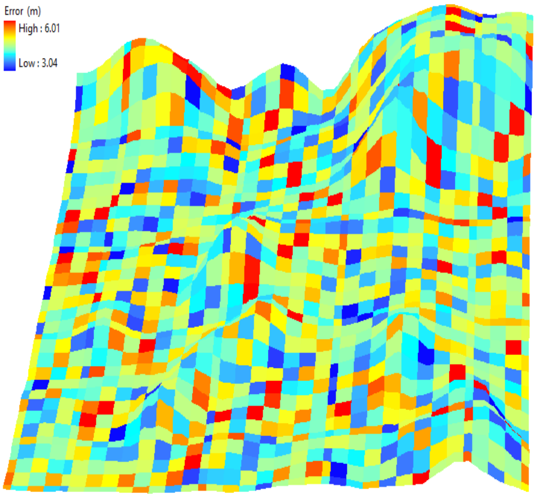


Figure C1. Error Realization #1

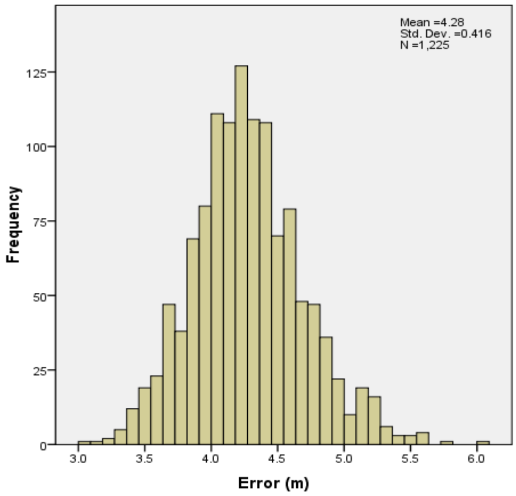


Figure C2. Histogram of Error Realization #1

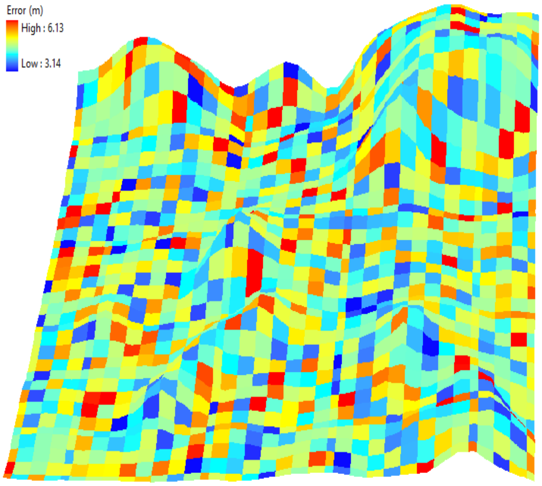


Figure C3. Error Realization #100

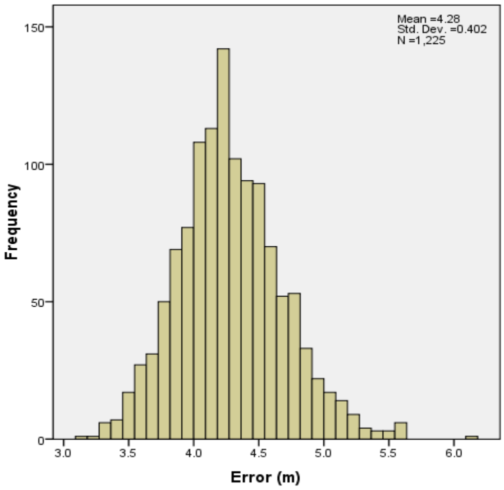


Figure C4. Histogram of Error Realization #100

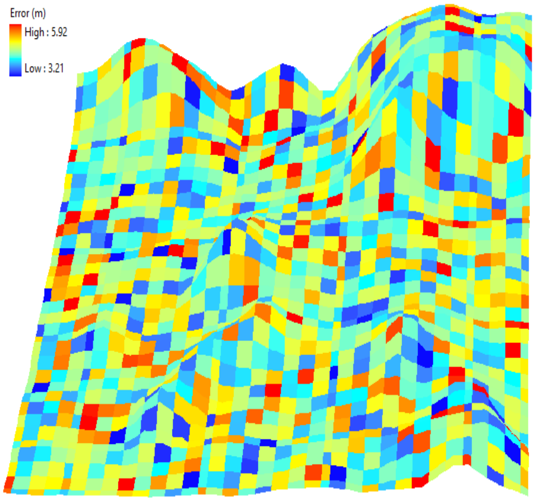


Figure C5. Error Realization #200

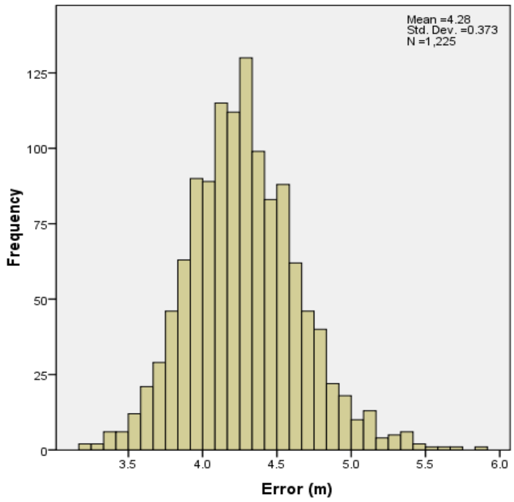


Figure C6. Histogram of Error Realization #200

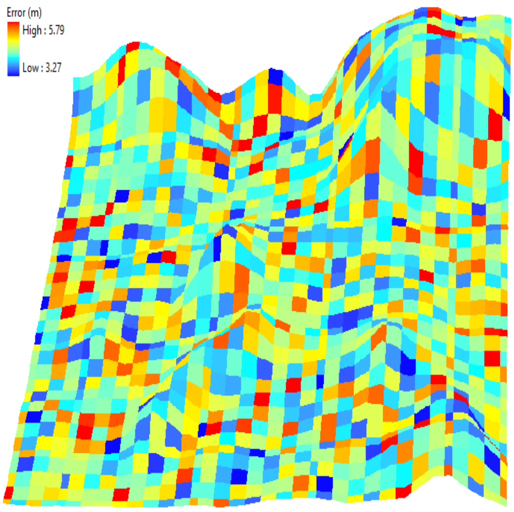


Figure C7. Error Realization #300

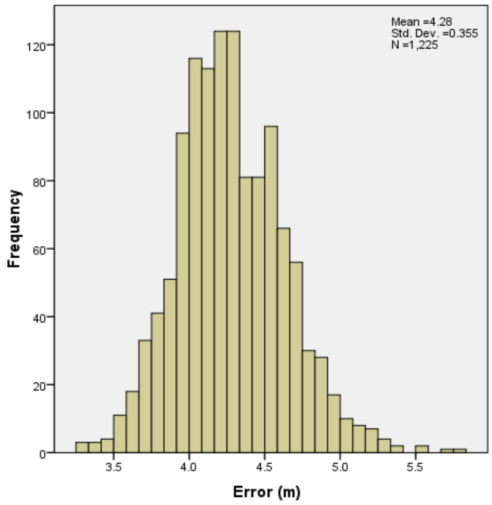


Figure C8. Histogram of Error Realization #300

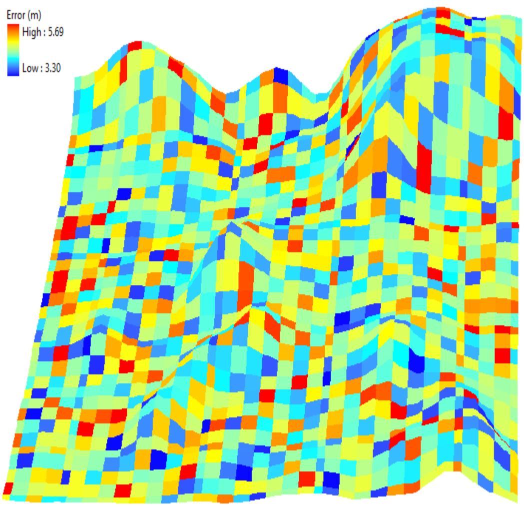


Figure C9. Error Realization #400

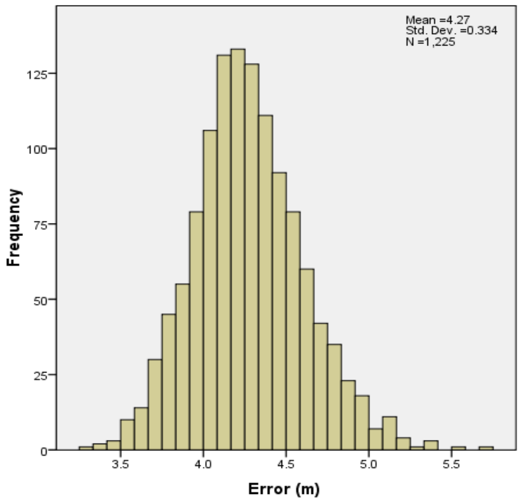


Figure C10. Histogram of Error Realization #400

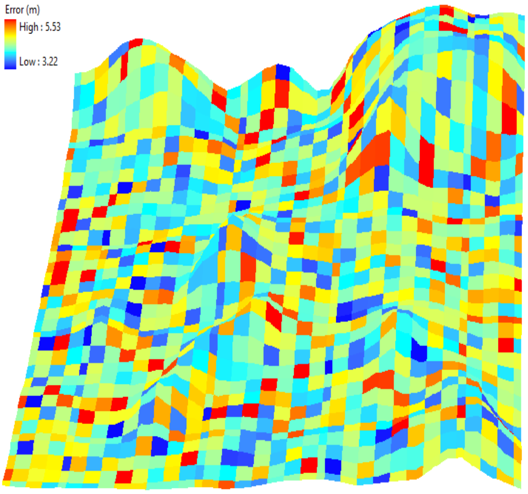


Figure C11. Error Realization #500

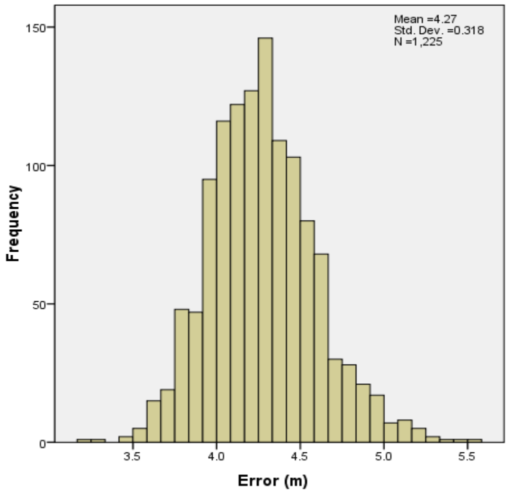


Figure C12. Histogram of Error Realization #500

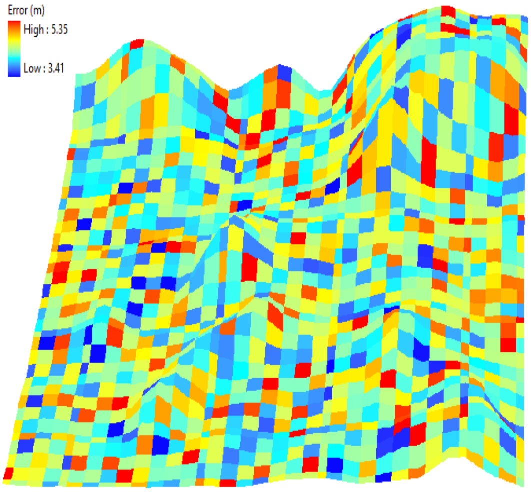


Figure C13. Eror Realization #600

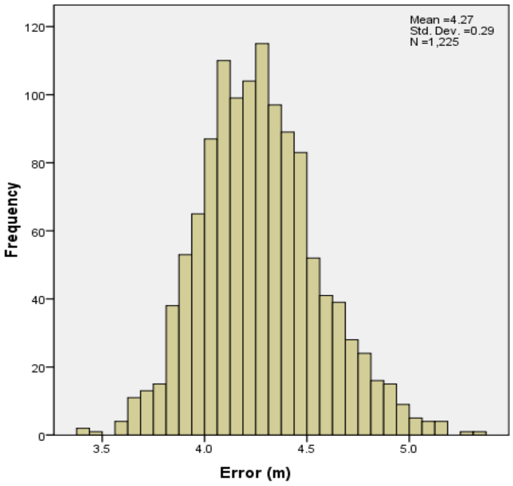


Figure C14. Histogram of Error Realization #600

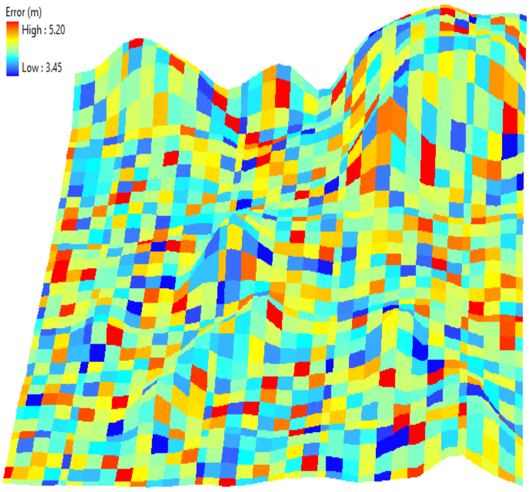


Figure C15. Error Realization #700

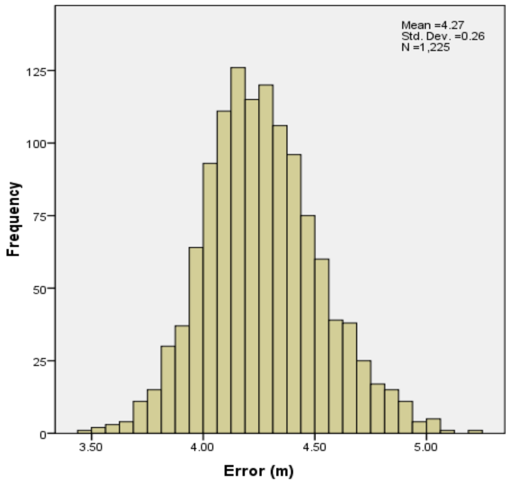


Figure C16. Histogram of Error Realization #700

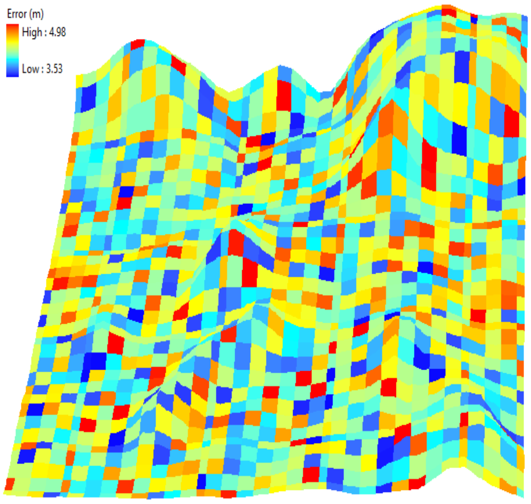


Figure C17. Error Realization #800

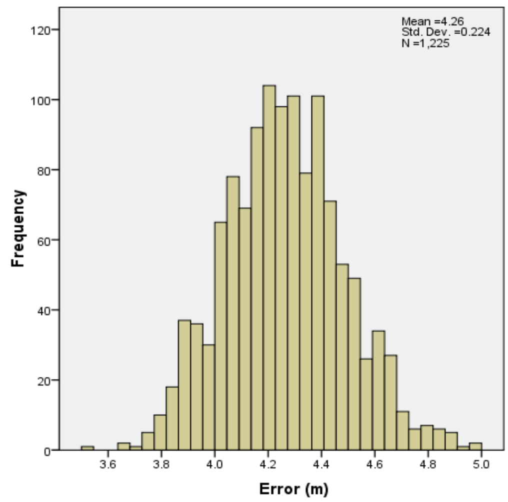


Figure C18. Histogram of Error Realization #800

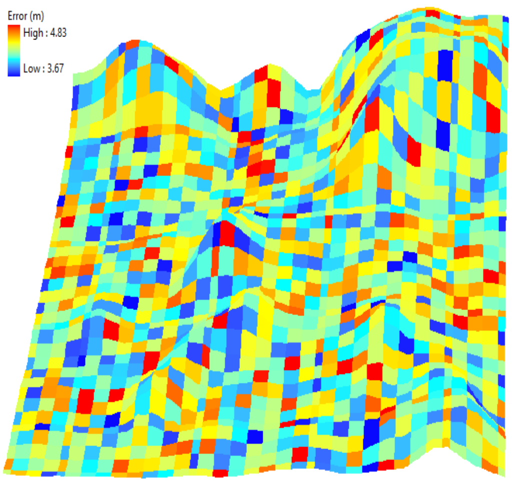


Figure C19. Error Realization #900

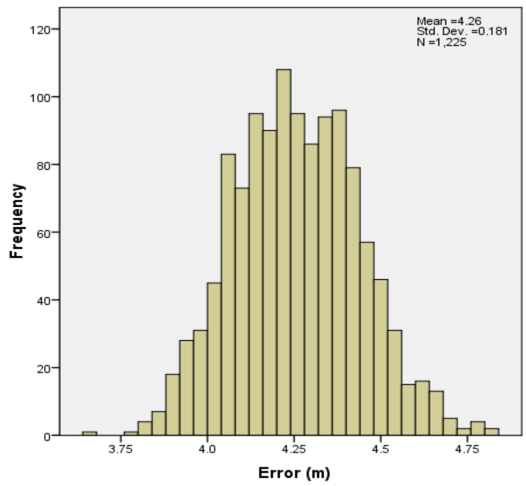


Figure C20. Histogram of Error Realization #900

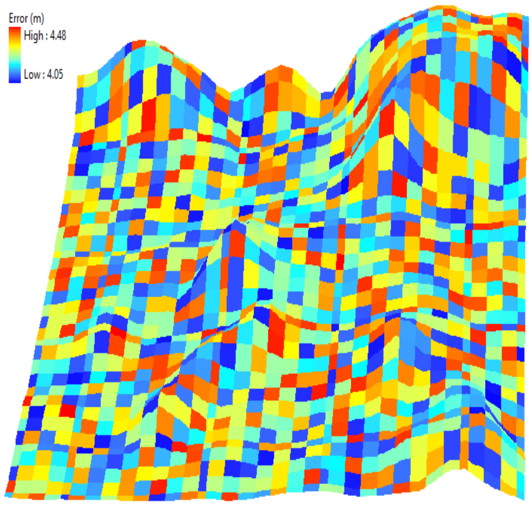


Figure C21. Error Realization #1000

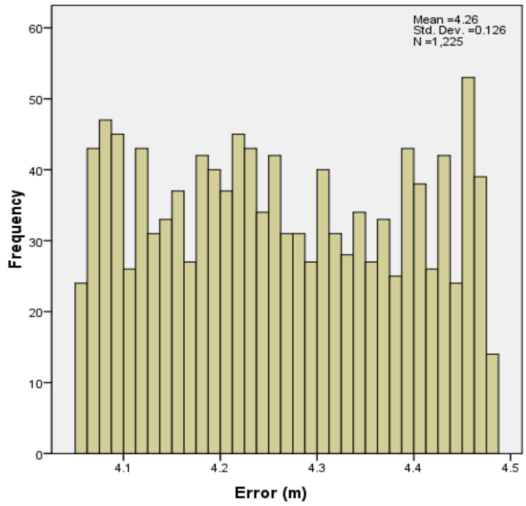


Figure C22. Histogram of Error Realization #1000

APPENDIX D

REALIZED DEMS

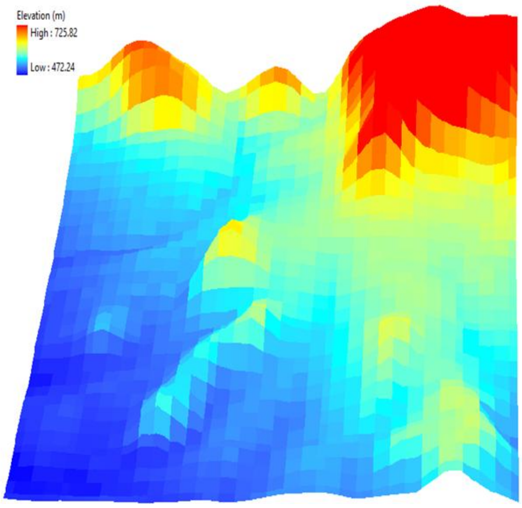


Figure D1. DEM Realization #1

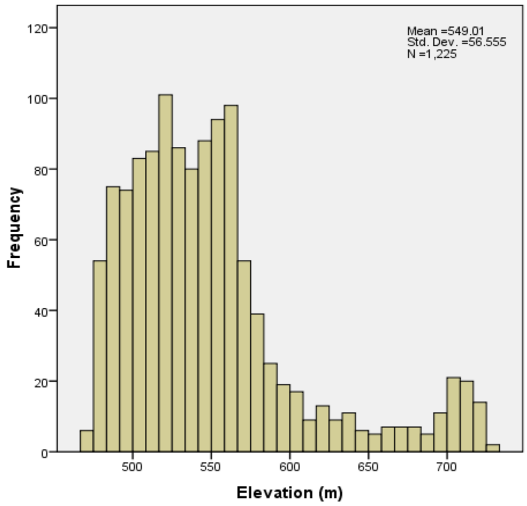


Figure D2. Histogram of DEM Realization #1

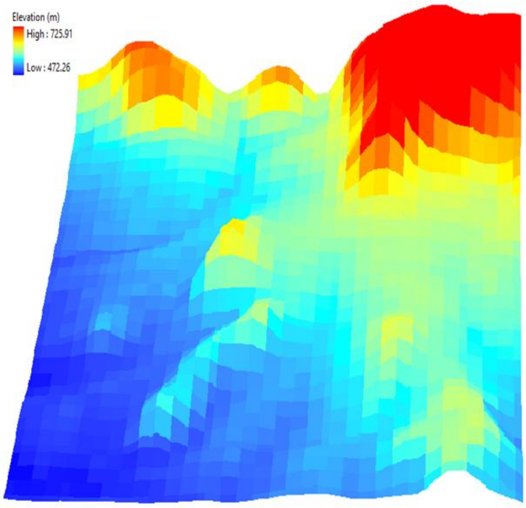


Figure D3. DEM Realization #100

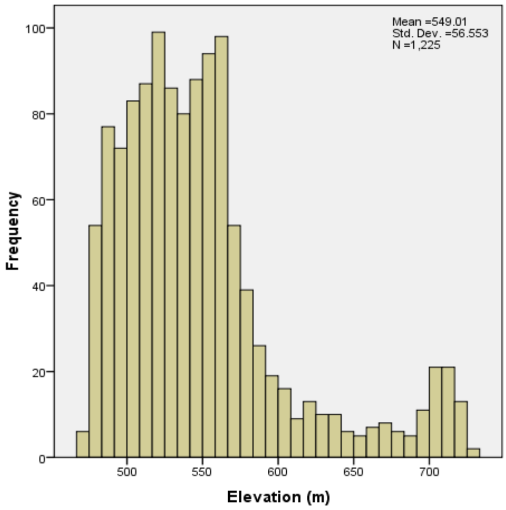


Figure D4. Histogram of DEM Realization #100

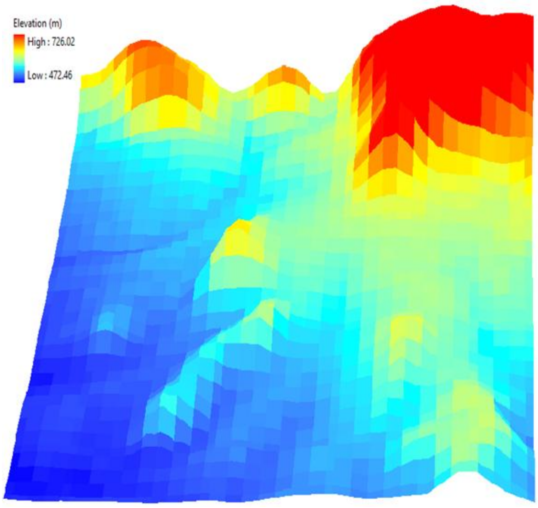


Figure D5. DEM Realization #200

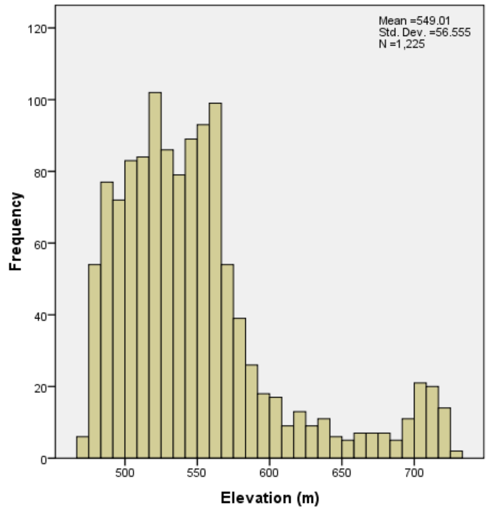


Figure D6. Histogram of DEM Realization #200

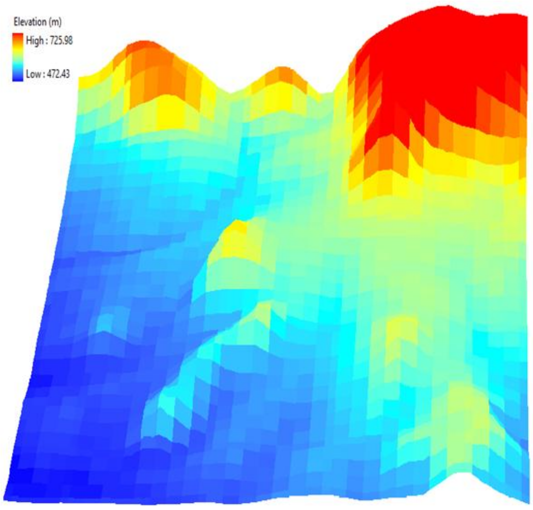


Figure D7. DEM Realization #300

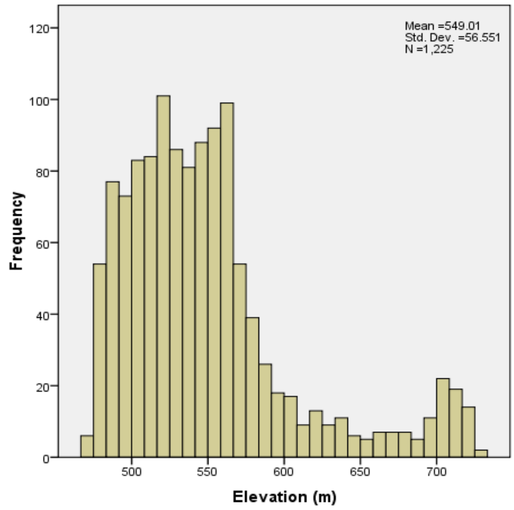


Figure D8. Histogram of DEM Realization #300

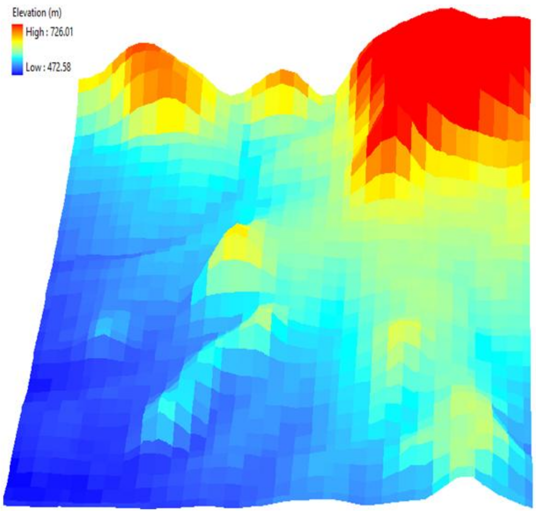


Figure D9. DEM Realization #400

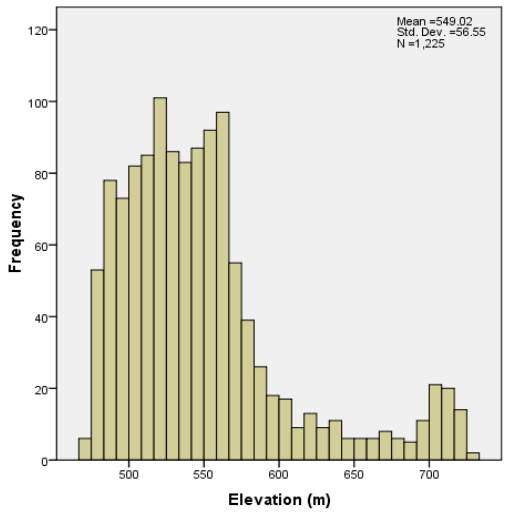


Figure D10. Histogram of DEM Realization #400

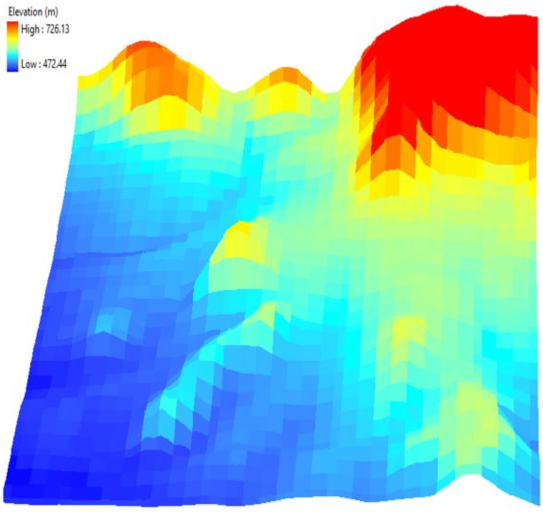


Figure D11. DEM Realization #500

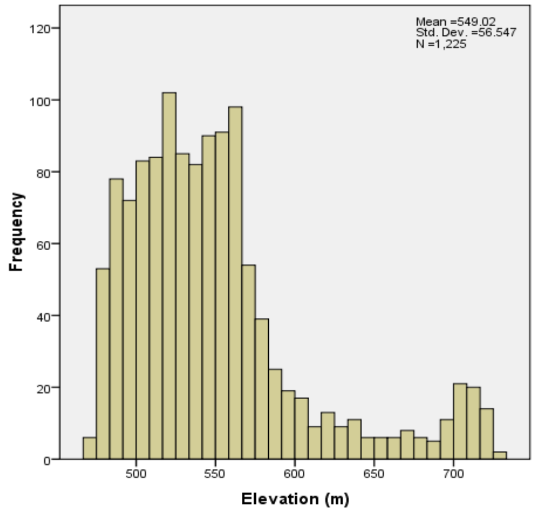


Figure D12. Histogram of DEM Realization #500

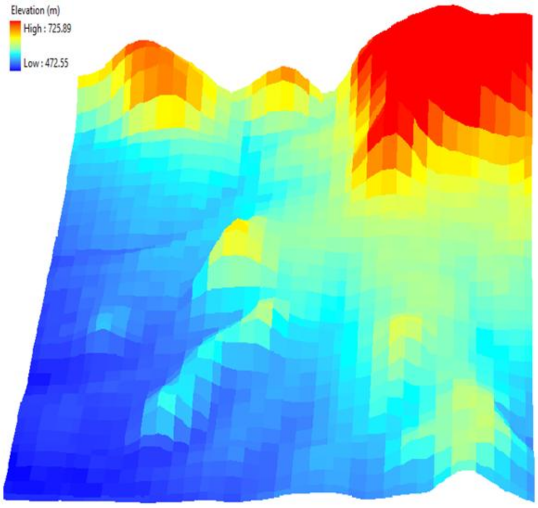


Figure D13. DEM Realization #600

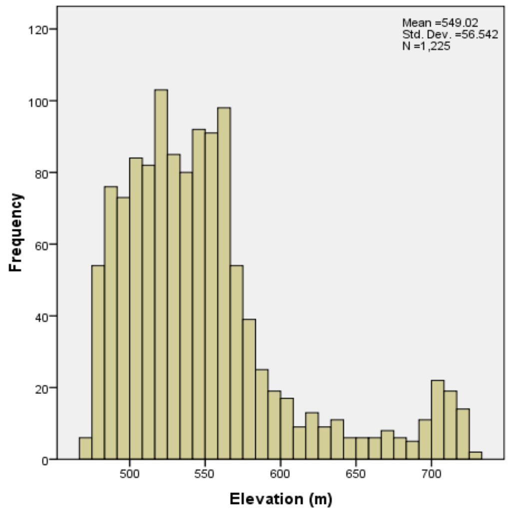


Figure D14. Histogram of DEM Realization #600

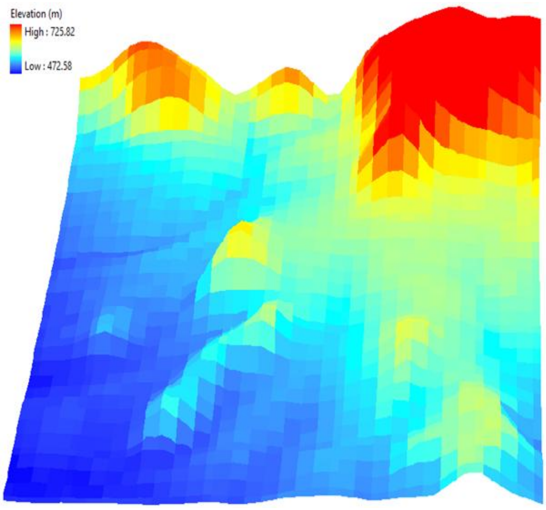


Figure D15. DEM Realization #700

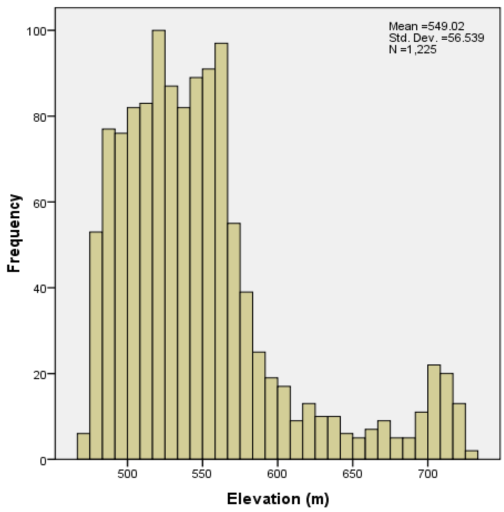


Figure D16. Histogram of DEM Realization #700

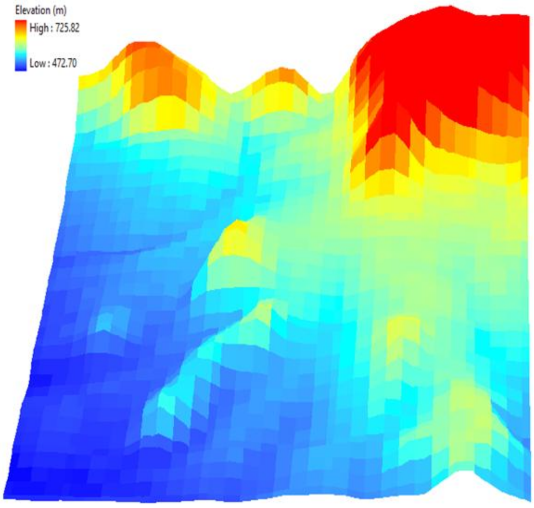


Figure D17. DEM Realization #800

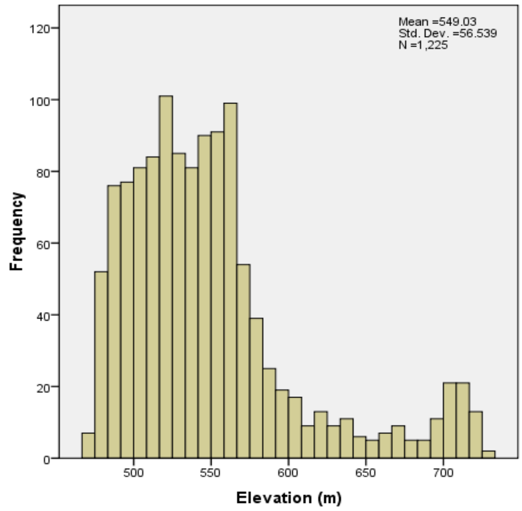


Figure D18. Histogram of DEM Realization #800

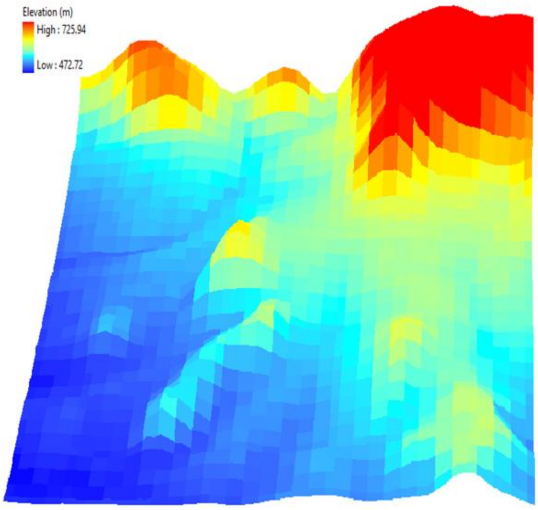


Figure D19. DEM Realization #900

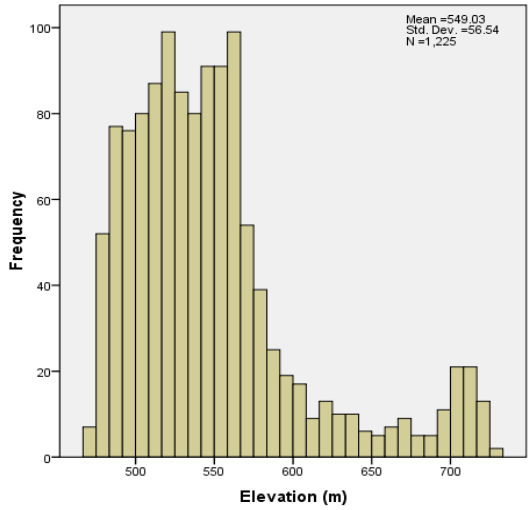


Figure D20. Histogram of DEM Realization #900

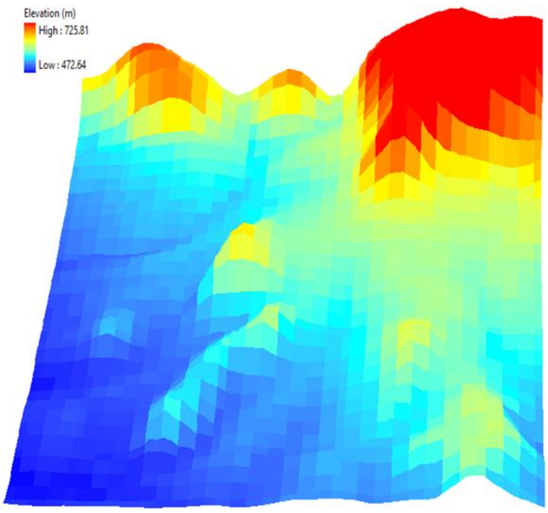


Figure D21. DEM Realization #1000

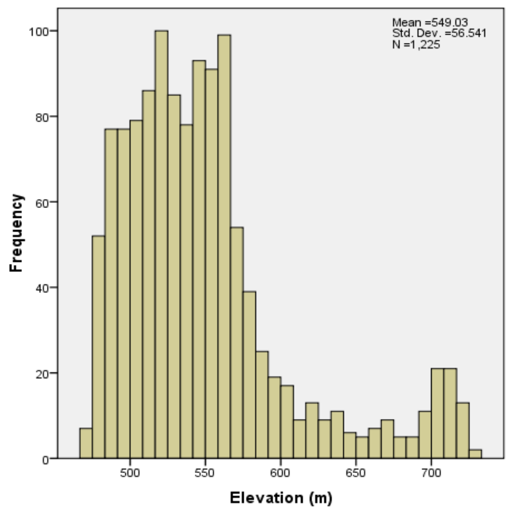


Figure D22. Histogram of DEM Realization #1000

APPENDIX E

SLOPE AND ASPECT MAPS GENERATED USING *R.SLOPE.ASPECT* MODULE IN GRASS GIS FROM THE SRTM DEM

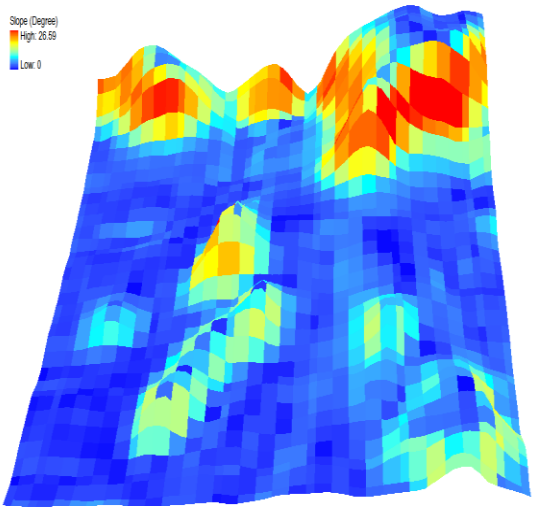


Figure E1. Slope of Study Area SRTM using *r.slope.aspect* module

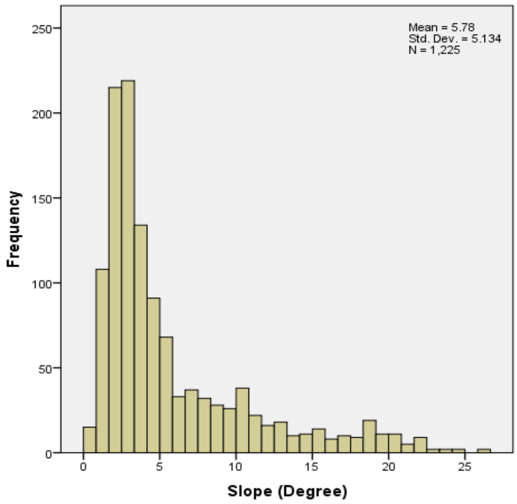


Figure E2. Histogram for Study Area SRTM using *r.slope.aspect* module

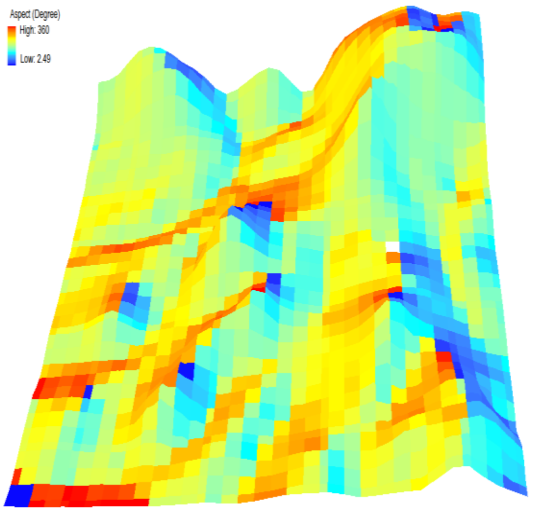


Figure E3. Aspect of Study Area SRTM using *r.slope.aspect* module

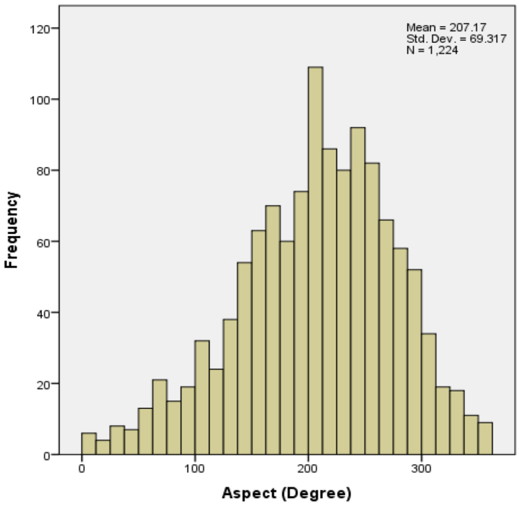


Figure E4. Histogram for Study Area SRTM using *r.slope.aspect* module

APPENDIX F

SLOPE MAPS GENERATED USING *R.SLOPE.ASPECT* MODULE IN GRASS GIS FROM THE REALIZED DEMS

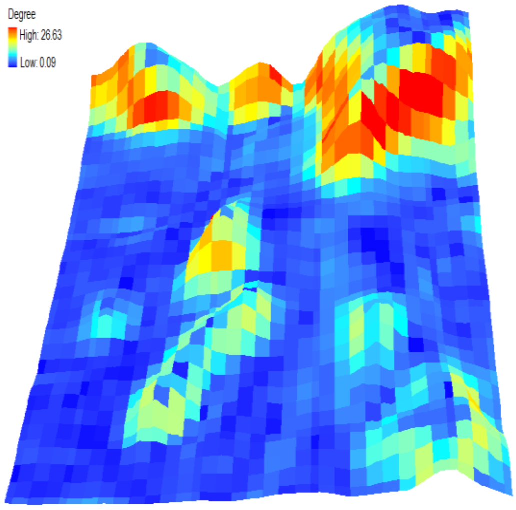


Figure F1. Slope #1 using *r.slope.aspect* module

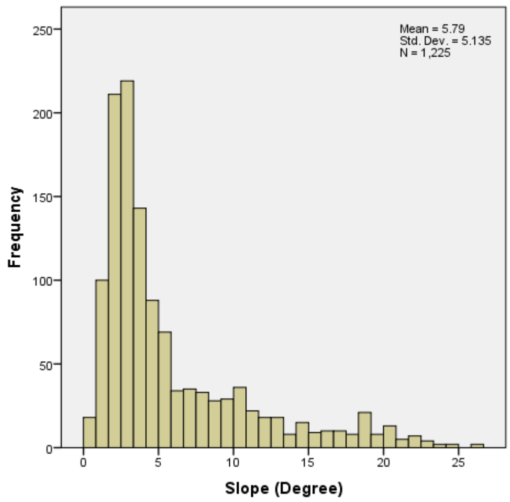


Figure F2. Histogram for Slope #1 using *r.slope.aspect* module

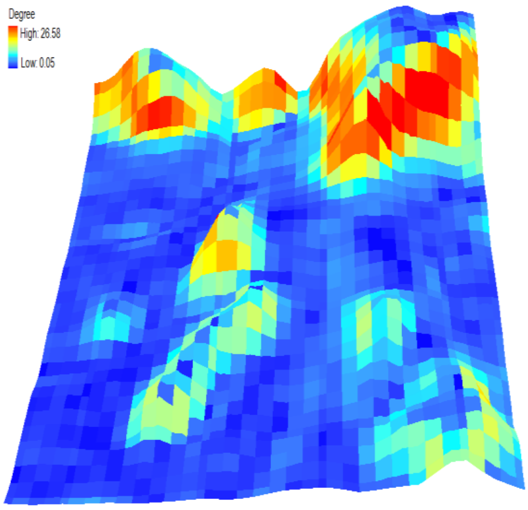


Figure F3. Slope #100 using *r.slope.aspect* module

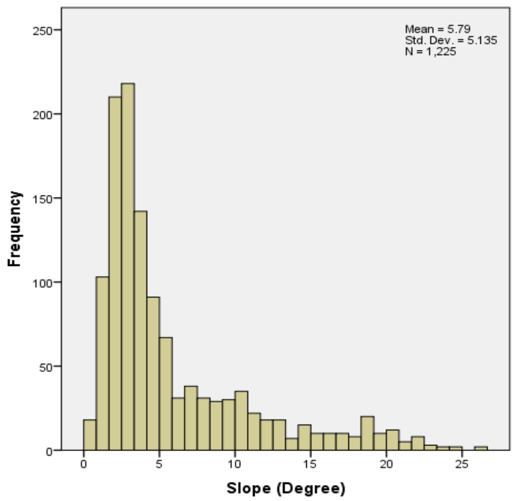


Figure F4. Histogram for Slope #100 using *r.slope.aspect* module

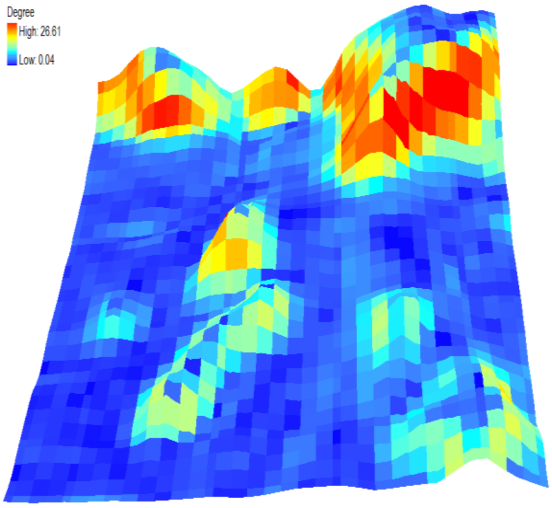


Figure F5. Slope #200 using *r.slope.aspect* module

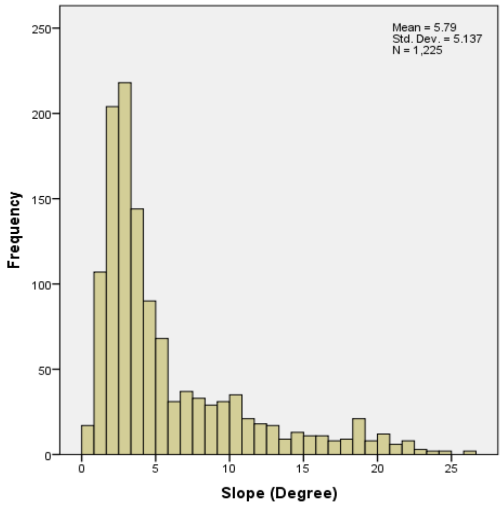


Figure F6. Histogram for Slope #200 using *r.slope.aspect* module

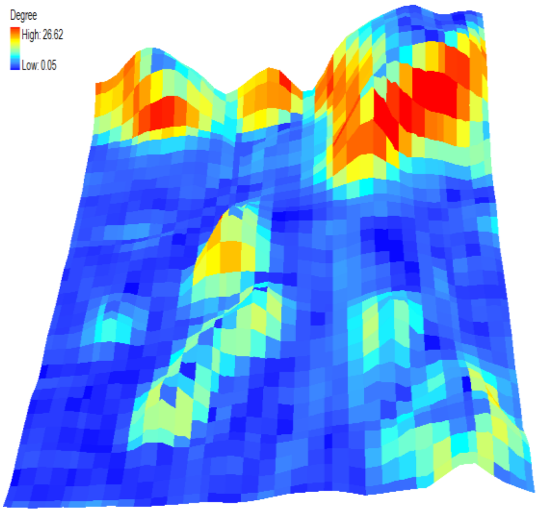


Figure F7. Slope #300 using *r.slope.aspect* module

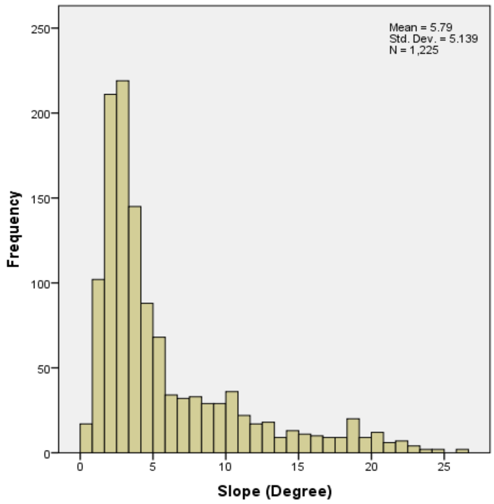


Figure F8. Histogram for Slope #300 using *r.slope.aspect* module

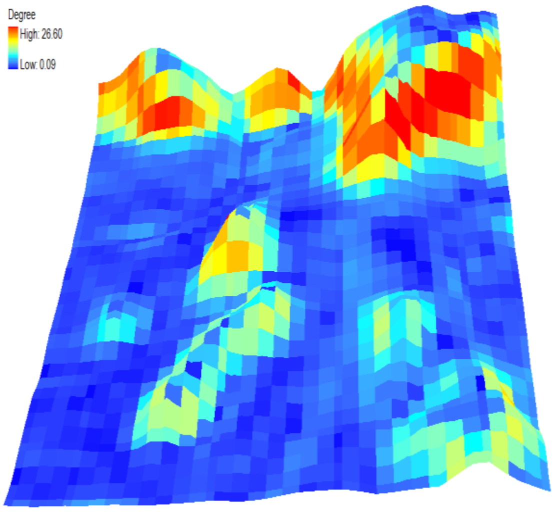


Figure F9. Slope #400 using *r.slope.aspect* module

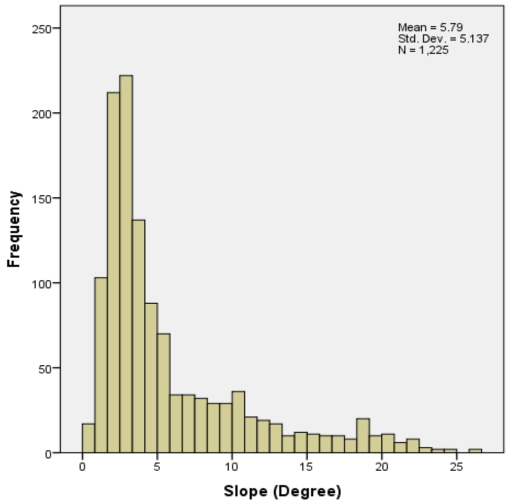


Figure F10. Histogram for Slope #400 using *r.slope.aspect* module

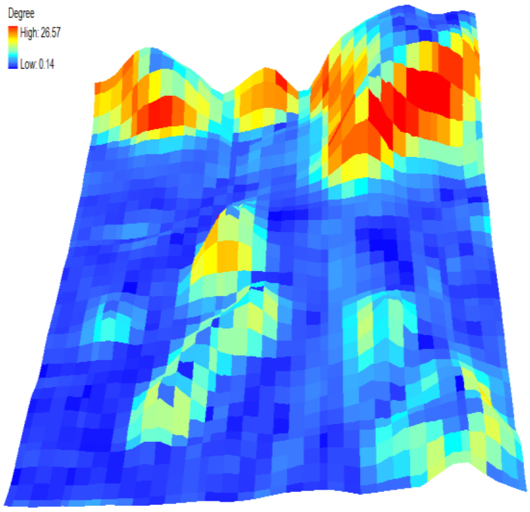


Figure F11. Slope #500 using *r.slope.aspect* module

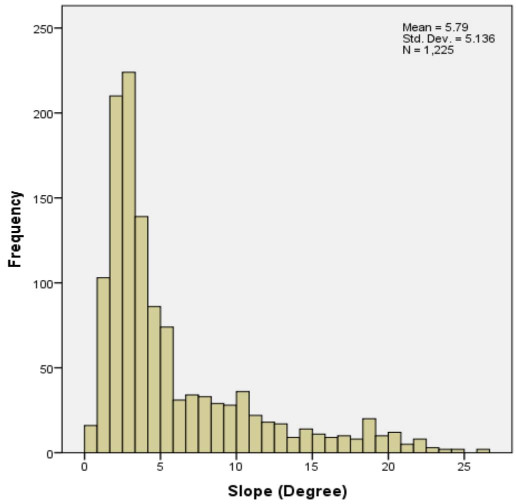


Figure F12. Histogram for Slope #500 using *r.slope.aspect* module

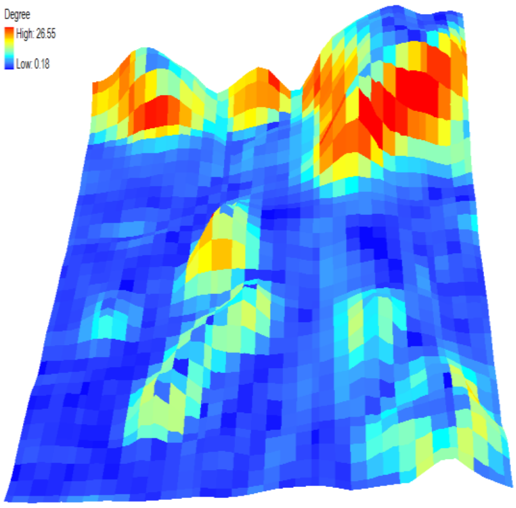


Figure F13. Slope #600 using *r.slope.aspect* module

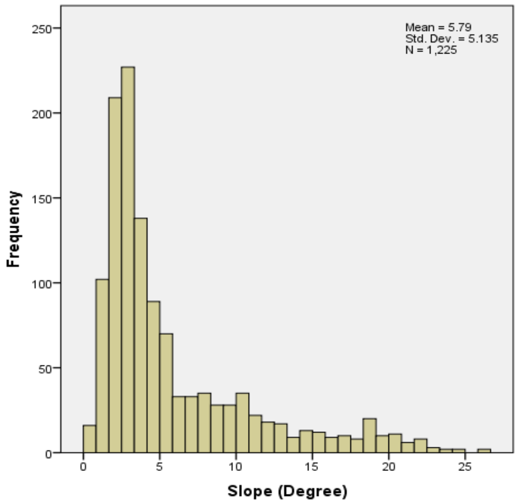


Figure F14. Histogram for Slope #600 using *r.slope.aspect* module

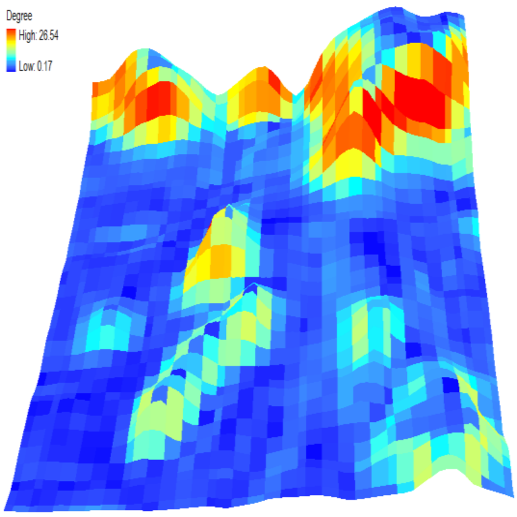


Figure F15. Slope #700 using *r.slope.aspect* module

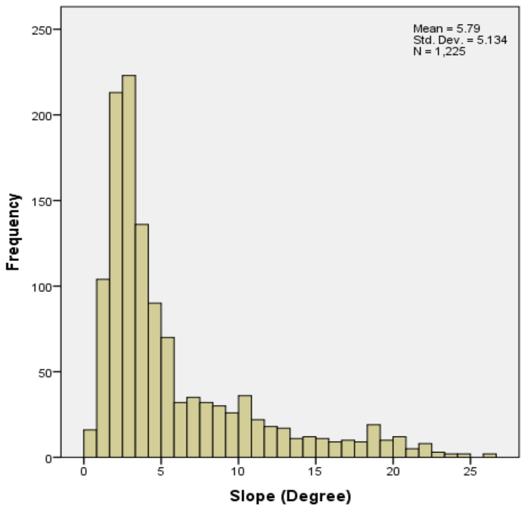


Figure F16. Histogram for Slope #700 using *r.slope.aspect* module

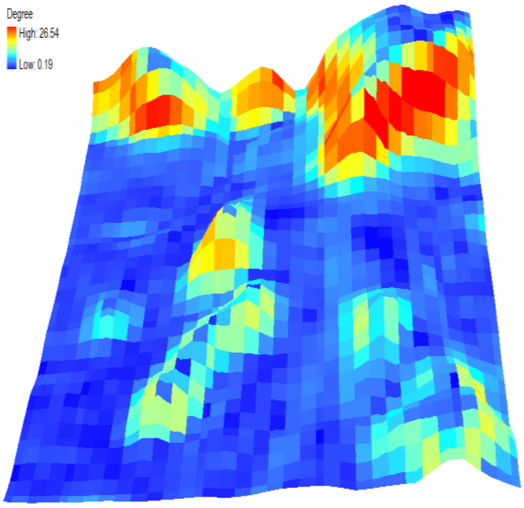


Figure F17. Slope #800 using *r.slope.aspect* module

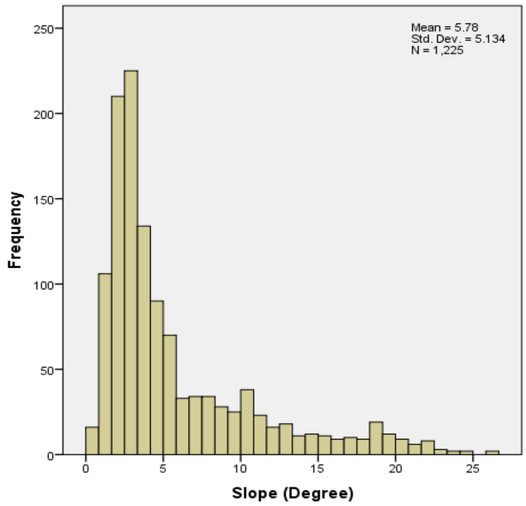


Figure F18. Histogram for Slope #800 using *r.slope.aspect* module

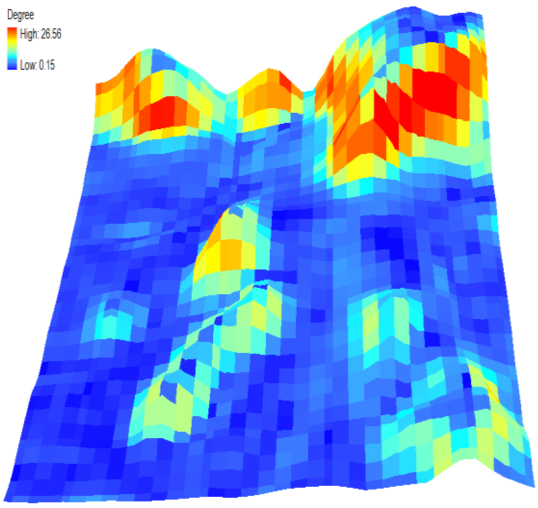


Figure F19. Slope #900 using *r.slope.aspect* module

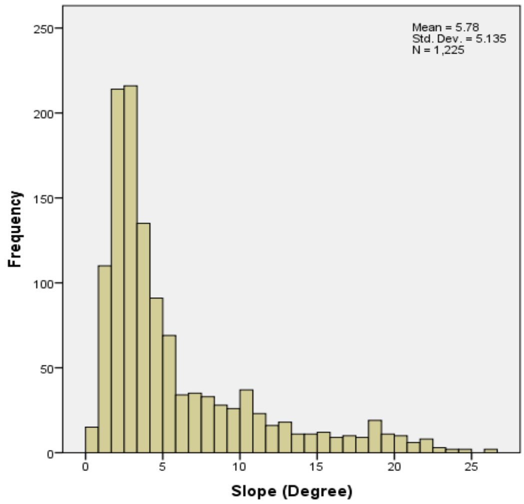


Figure F20. Histogram for Slope #900 using *r.slope.aspect* module

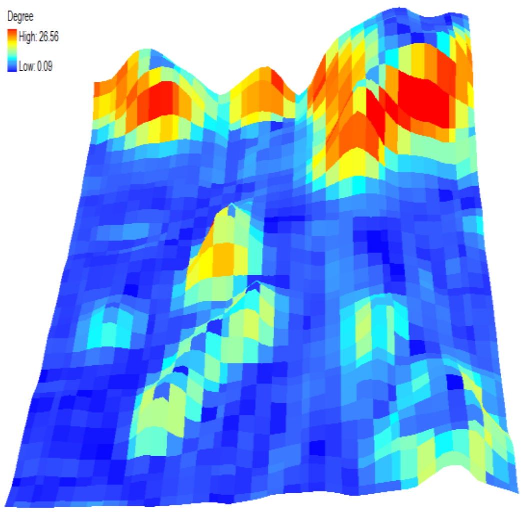


Figure F21. Slope #1000 using *r.slope.aspect* module

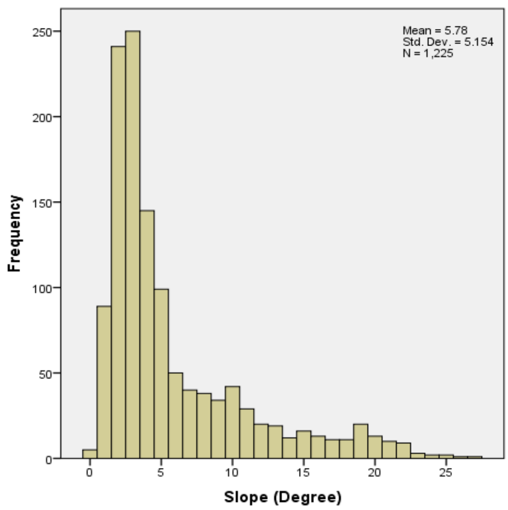


Figure F22. Histogram for Slope #1000 using *r.slope.aspect* module

APPENDIX G

ASPECT MAPS GENERATED USING *R.SLOPE.ASPECT* MODULE IN GRASS GIS FROM THE REALIZED DEMS

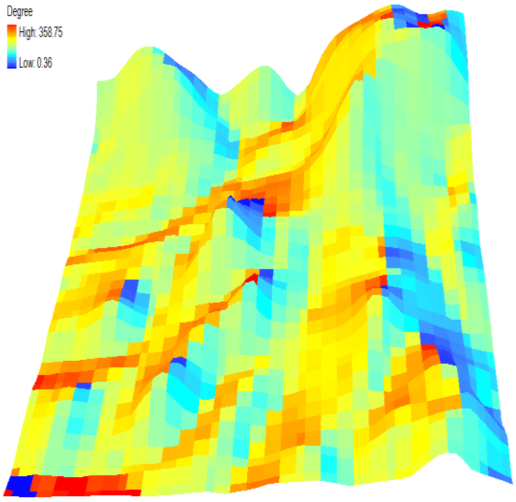


Figure G1. Aspect #1 using *r.slope.aspect* module

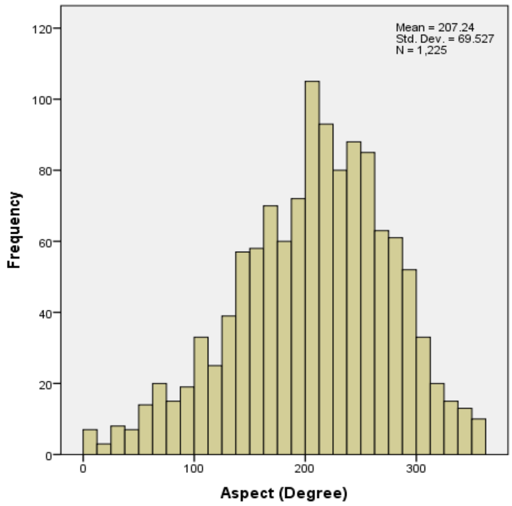


Figure G2. Histogram for Aspect #1 using *r.slope.aspect* module

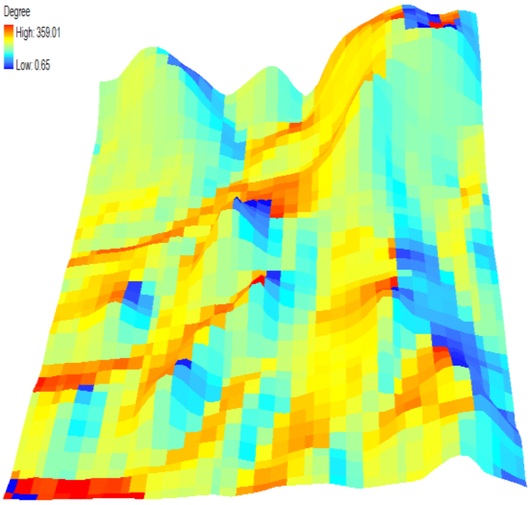


Figure G3. Aspect #100 using *r.slope.aspect* module

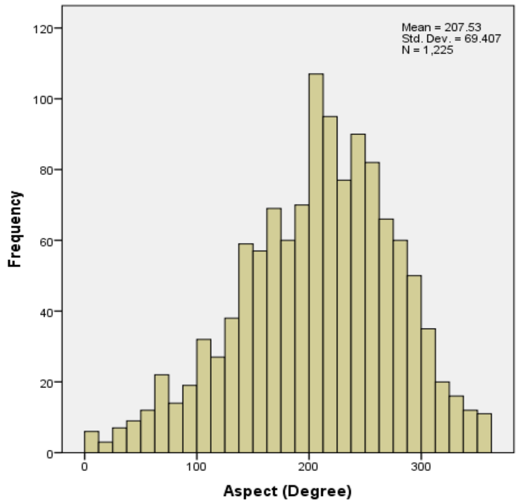


Figure G4. Histogram for Aspect #100 using *r.slope.aspect* module

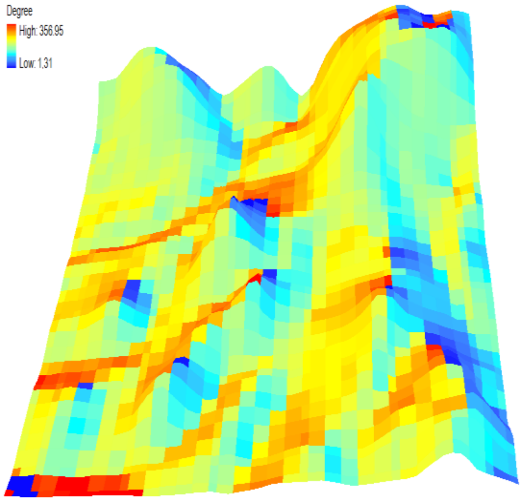


Figure G5. Aspect #200 using *r.slope.aspect* module

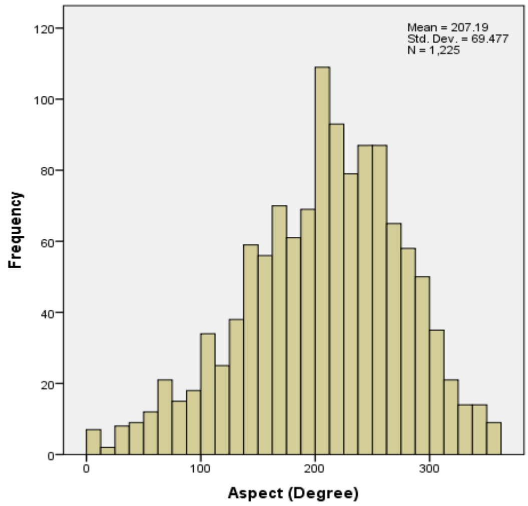


Figure G6. Histogram for Aspect #200 using *r.slope.aspect* module

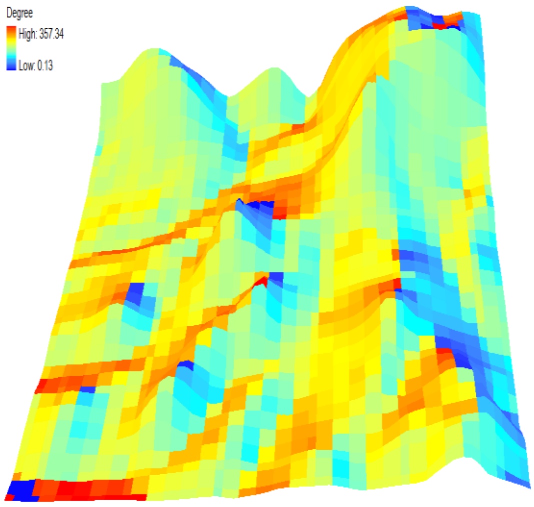


Figure G7. Aspect #300 using *r.slope.aspect* module

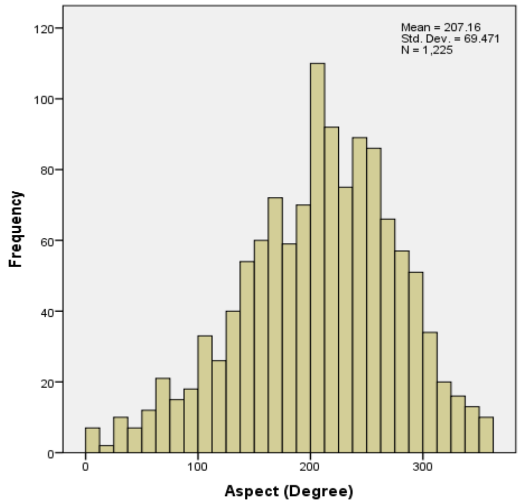


Figure G8. Histogram for Aspect #300 using *r.slope.aspect* module

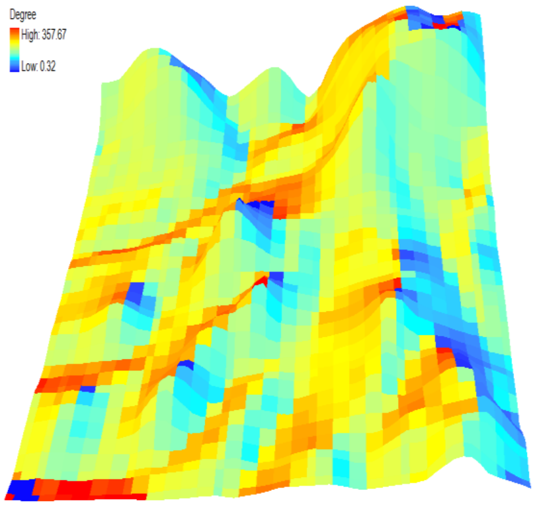


Figure G9. Aspect #400 using *r.slope.aspect* module

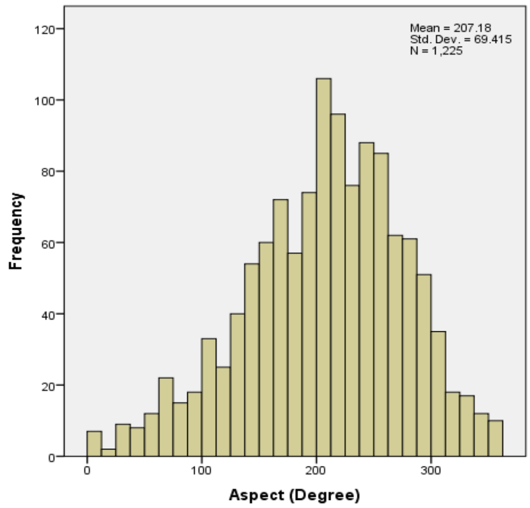


Figure G10. Histogram for Aspect #400 using *r.slope.aspect* module

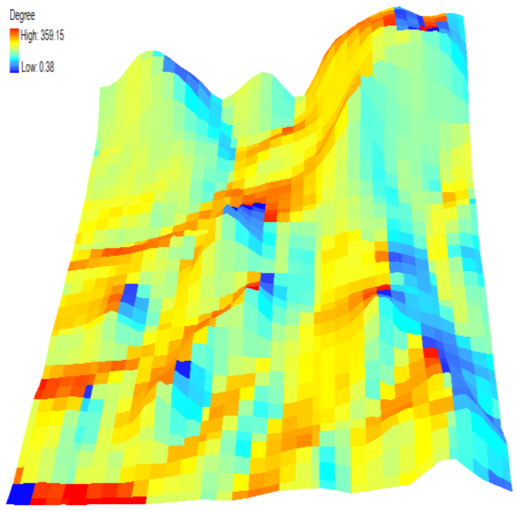


Figure G11. Aspect #500 using *r.slope.aspect* module

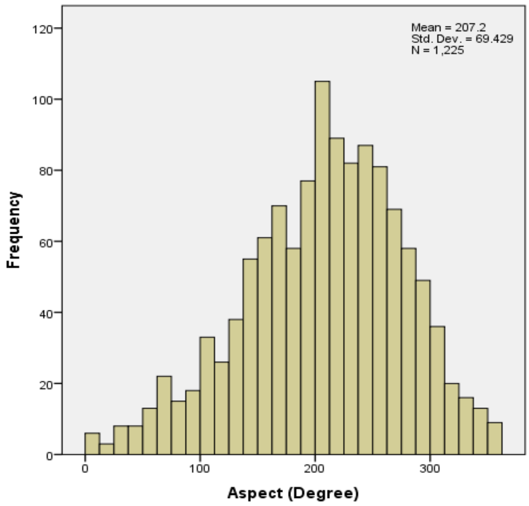


Figure G12. Histogram for Aspect #500 using *r.slope.aspect* module

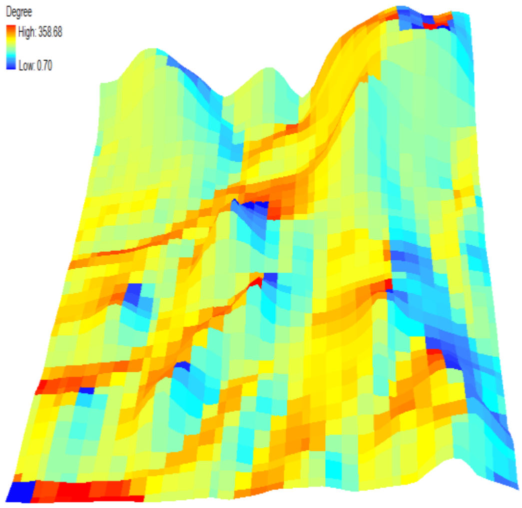


Figure G13. Aspect #600 using *r.slope.aspect* module

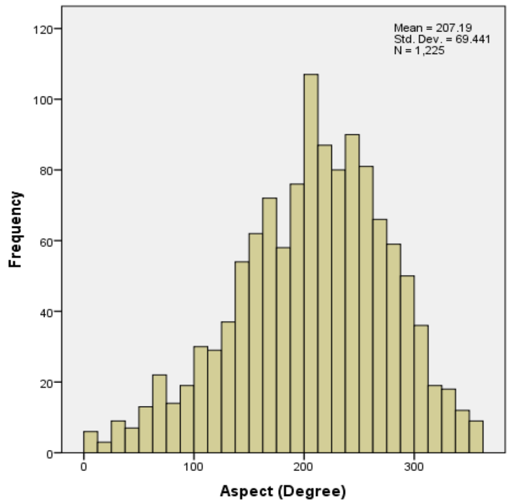


Figure G14. Histogram for Aspect #600 using *r.slope.aspect* module

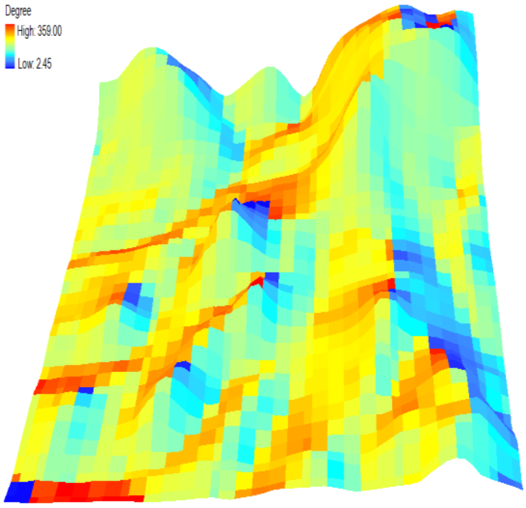


Figure G15. Aspect #700 using *r.slope.aspect* module

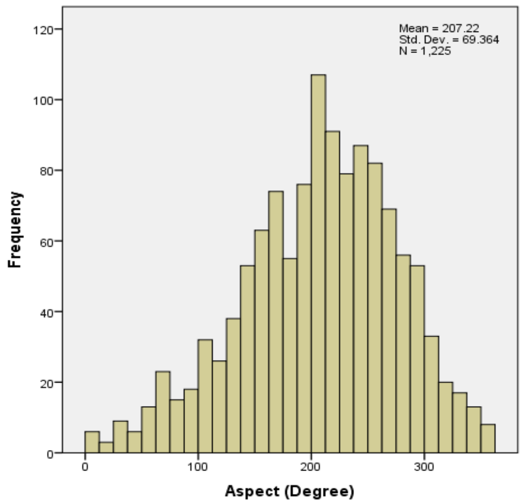


Figure G16. Histogram for Aspect #700 using *r.slope.aspect* module

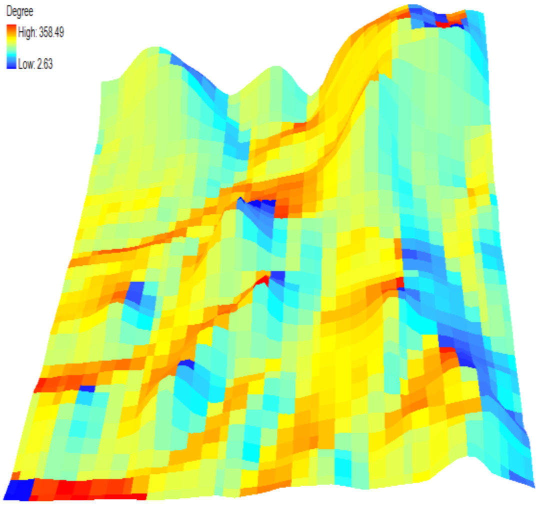


Figure G17. Aspect #800 using *r.slope.aspect* module

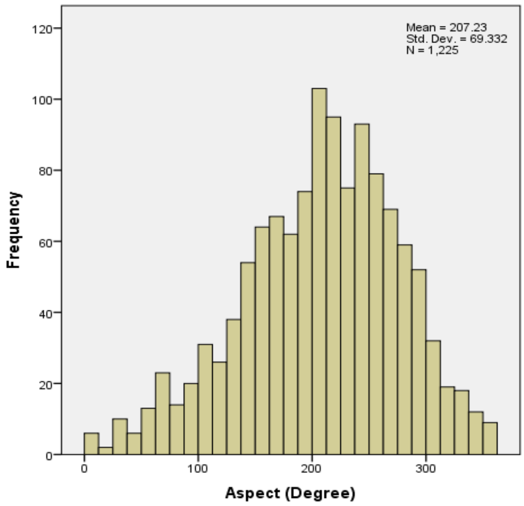


Figure G18. Histogram for Aspect #800 using *r.slope.aspect* module

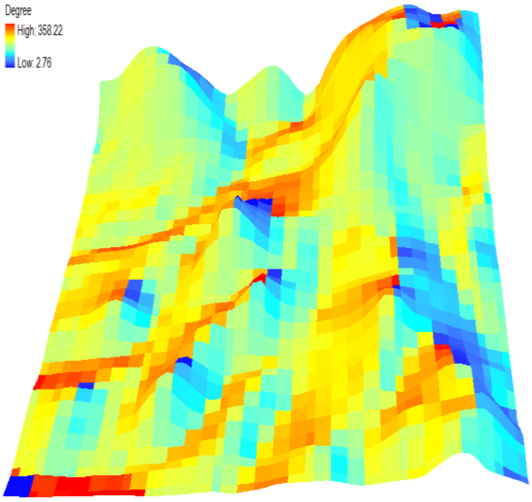


Figure G19. Aspect #900 using *r.slope.aspect* module

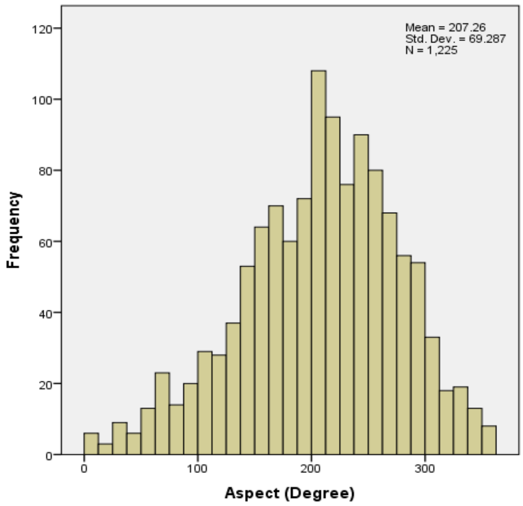


Figure G20. Histogram for Aspect #900 using *r.slope.aspect* module

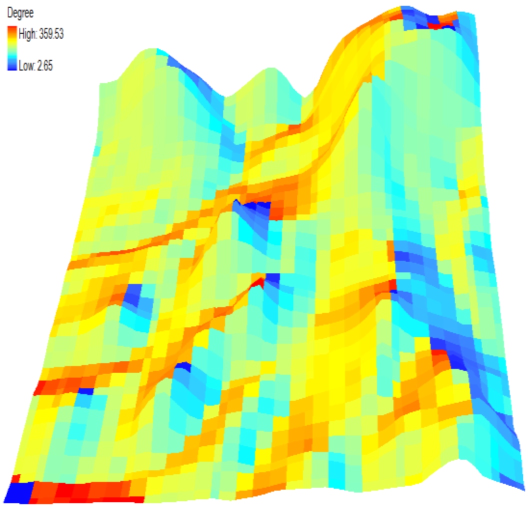


Figure G21. Aspect #1000 using *r.slope.aspect* module

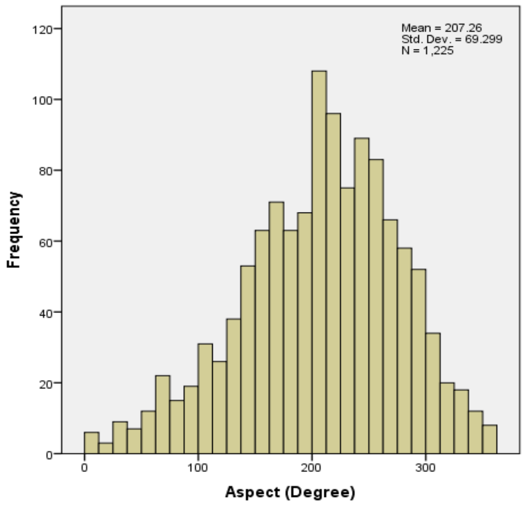


Figure G22. Histogram for Aspect #1000 using *r.slope.aspect* module

APPENDIX H

SELECTED OUTPUTS OF *SOLAR ANALYST* (ARCGIS) FROM THE REALIZED DEMS

Direct Radiation

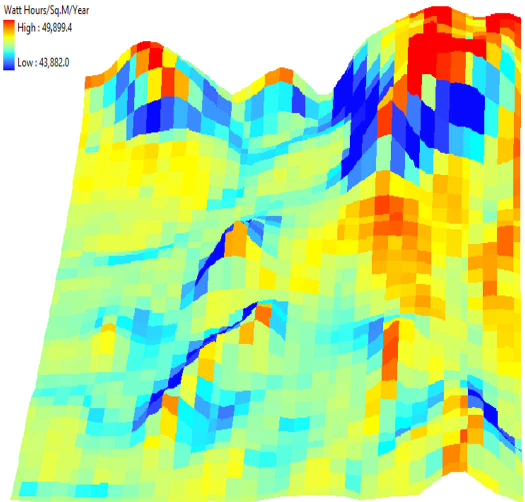


Figure H1. Direct Radiation #1 of *Solar Analyst*

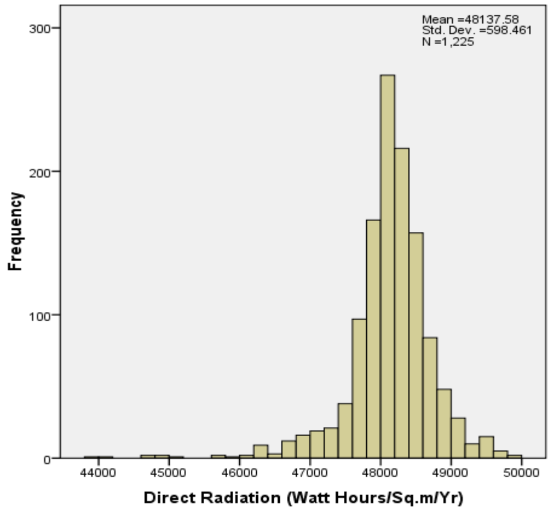


Figure H2. Histogram for Direct Radiation #1 of *Solar Analyst*

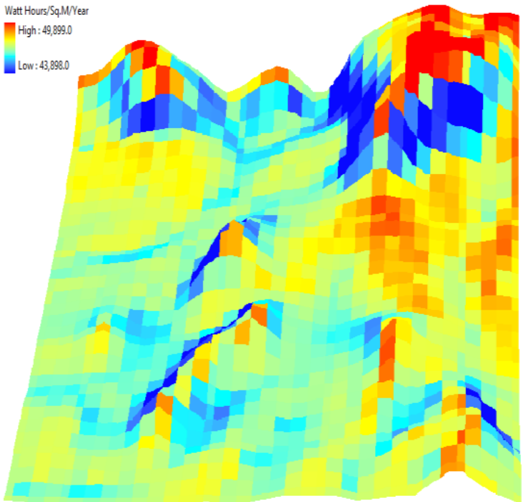


Figure H3. Direct Radiation #100 of *Solar Analyst*

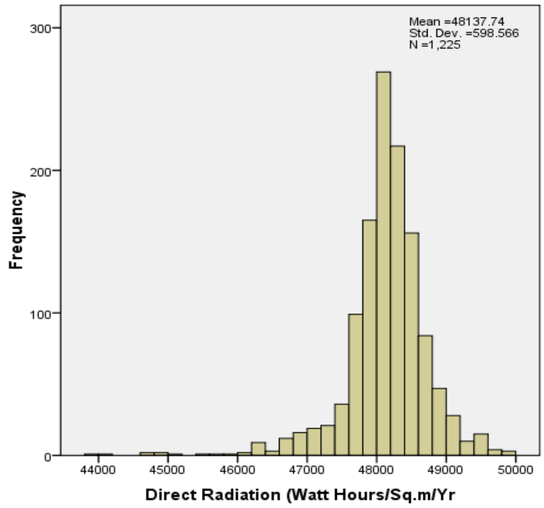


Figure H4. Histogram for Direct Radiation #100 of *Solar Analyst*

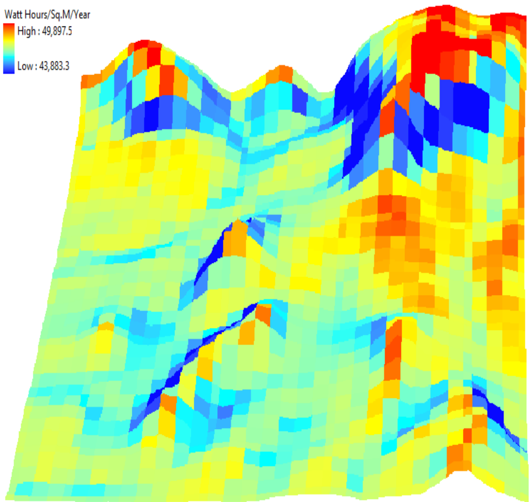


Figure H5. Direct Radiation #200 of *Solar Analyst*

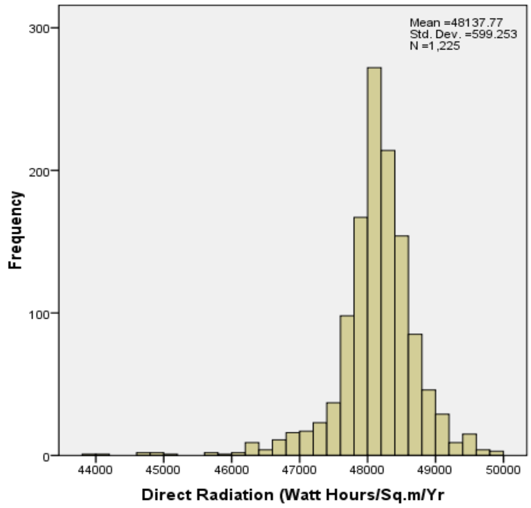


Figure H6. Histogram for Direct Radiation #200 of *Solar Analyst*

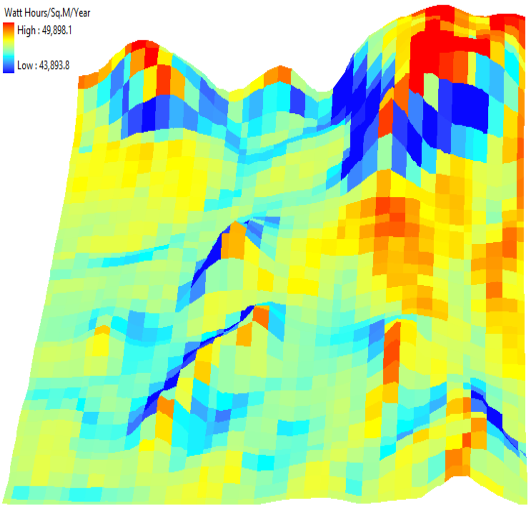


Figure H7. Direct Radiation #300 of *Solar Analyst*

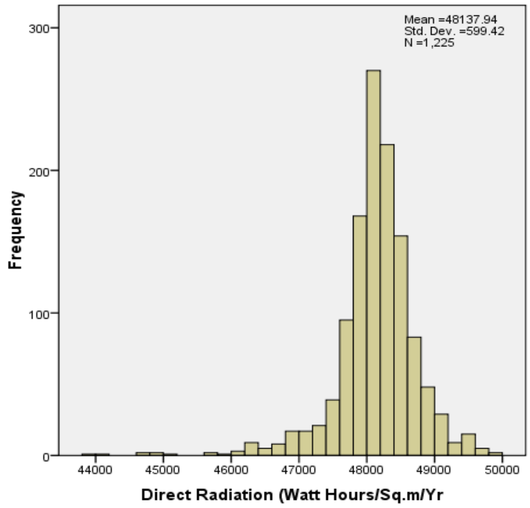


Figure H8. Histogram for Direct Radiation #300 of *Solar Analyst*

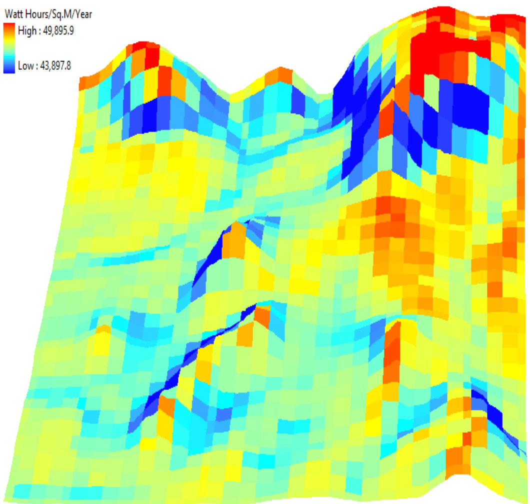


Figure H9. Direct Radiation #400 of *Solar Analyst*

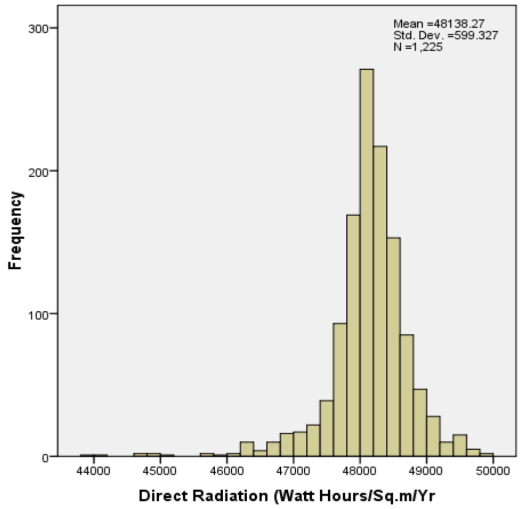


Figure H10. Histogram for Direct Radiation #400 of *Solar Analyst*

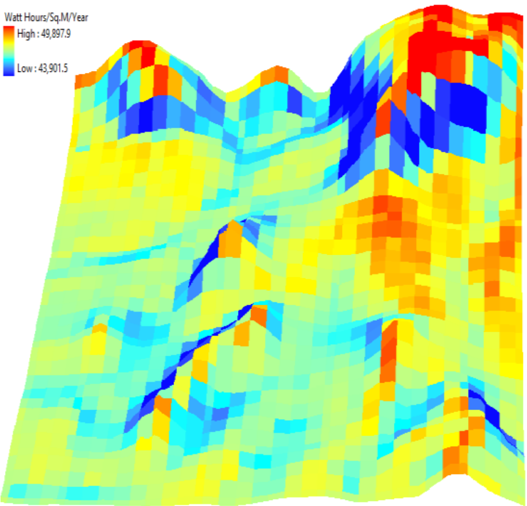


Figure H11. Direct Radiation #500 of *Solar Analyst*

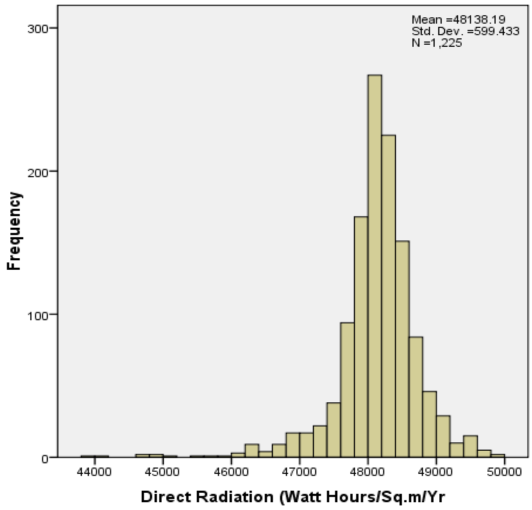


Figure H12. Histogram for Direct Radiation #500 of *Solar Analyst*

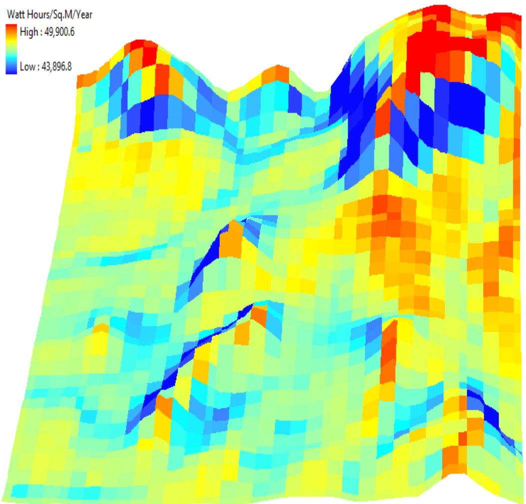


Figure H13. Direct Radiation #600 of *Solar Analyst*

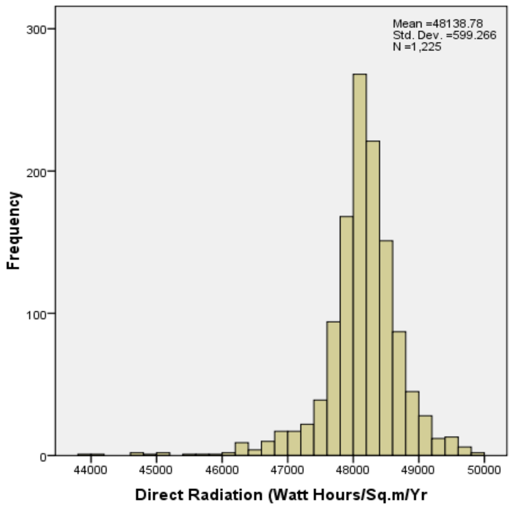


Figure H14. Histogram for Direct Radiation #600 of *Solar Analyst*

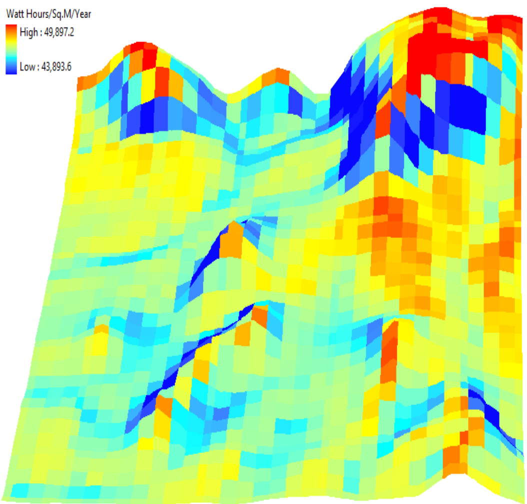


Figure H15. Direct Radiation #700 of *Solar Analyst*

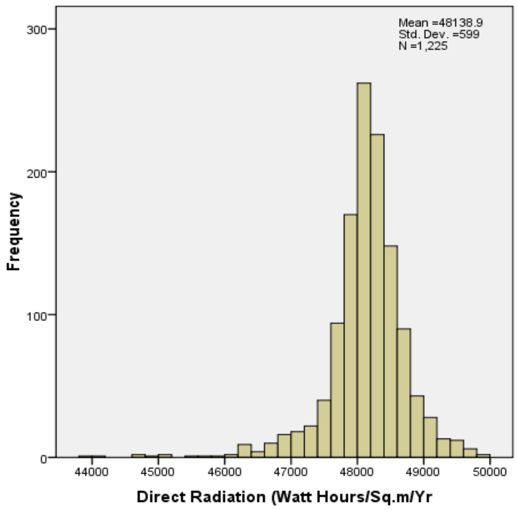


Figure H16. Histogram for Direct Radiation #700 of *Solar Analyst*

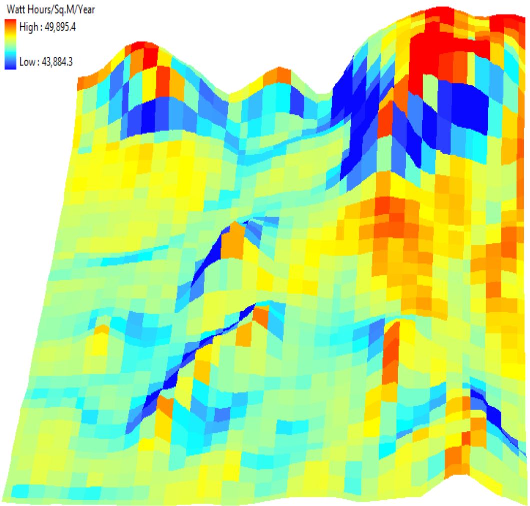


Figure H17. Direct Radiation #800 of *Solar Analyst*

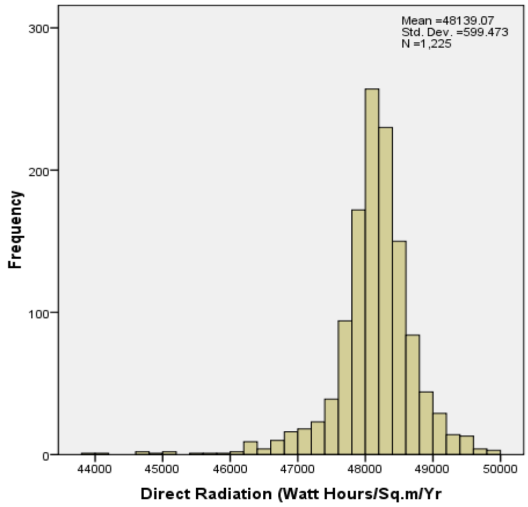


Figure H18. Histogram for Direct Radiation #800 of *Solar Analyst*

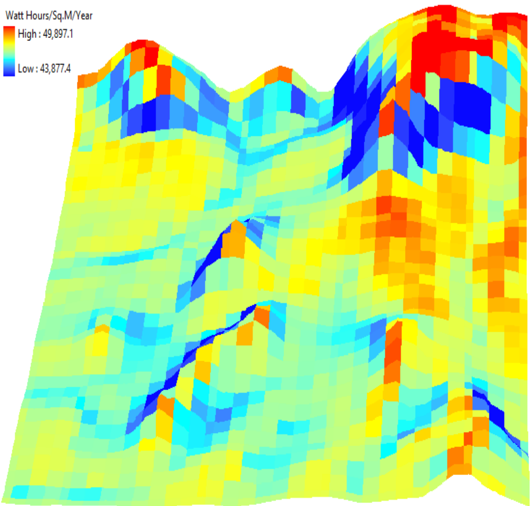


Figure H19. Direct Radiation #900 of *Solar Analyst*

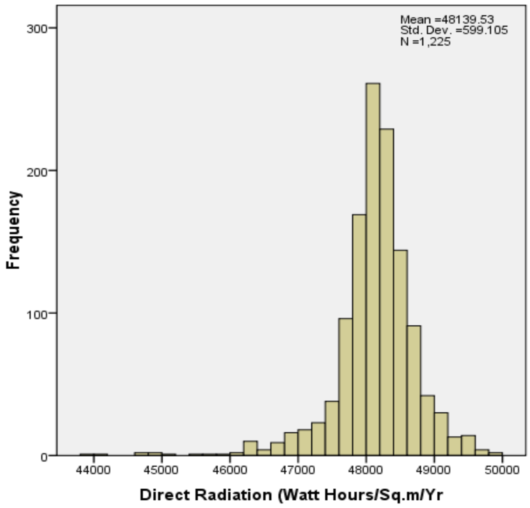


Figure H20. Histogram for Direct Radiation #900 of *Solar Analyst*

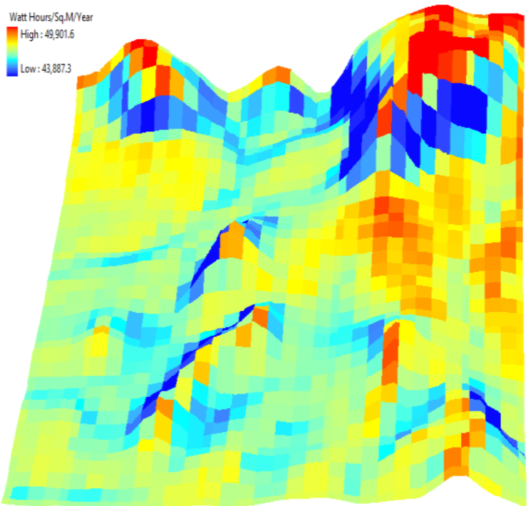


Figure H21. Direct Radiation #1000 of *Solar Analyst*

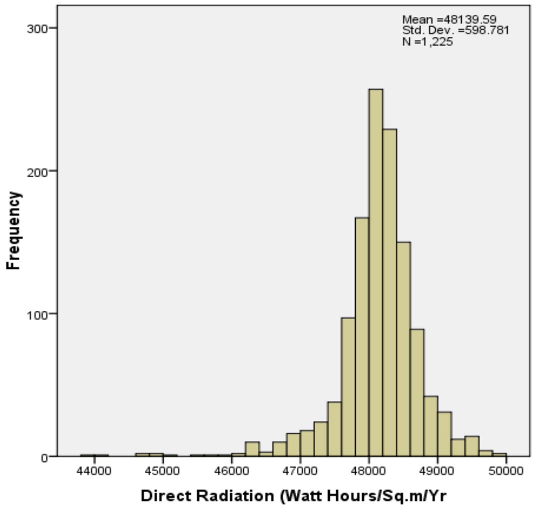


Figure H22. Histogram for Direct Radiation #1000 of *Solar Analyst*

Diffuse Radiation

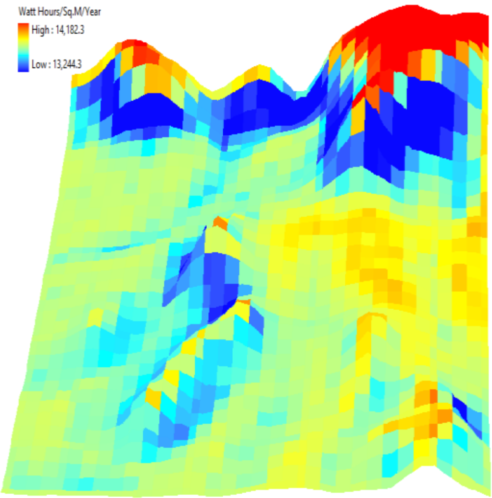


Figure H23. Diffuse Radiation #1 of *Solar Analyst*

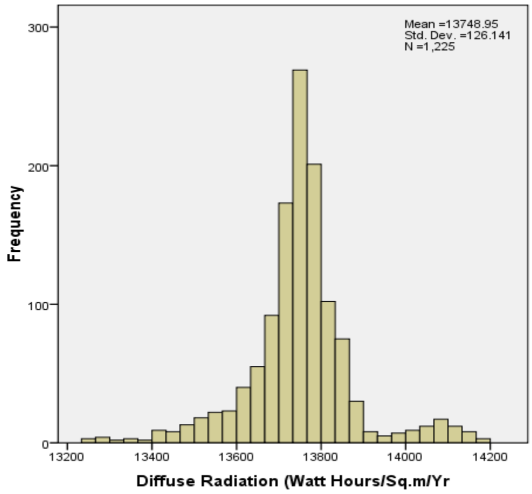


Figure H24. Histogram for Diffuse Radiation #1 of *Solar Analyst*

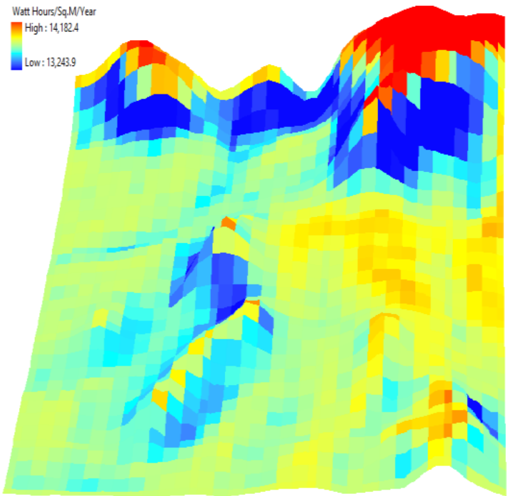


Figure H25. Diffuse Radiation #100 of *Solar Analyst*

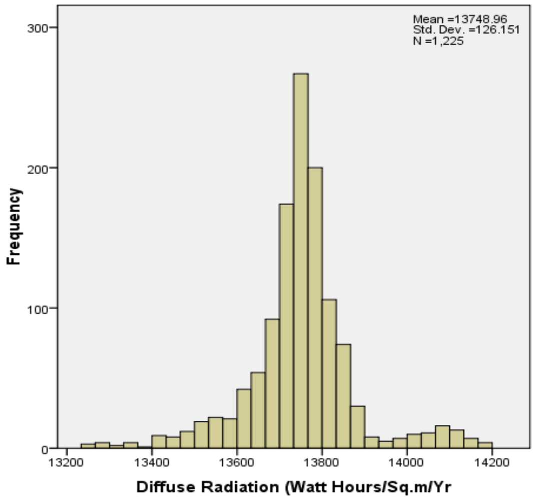


Figure H26. Histogram for Diffuse Radiation #100 of *Solar Analyst*

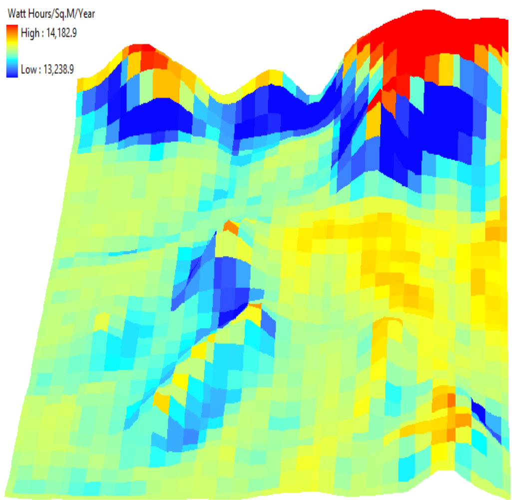


Figure H27. Diffuse Radiation #200 of *Solar Analyst*

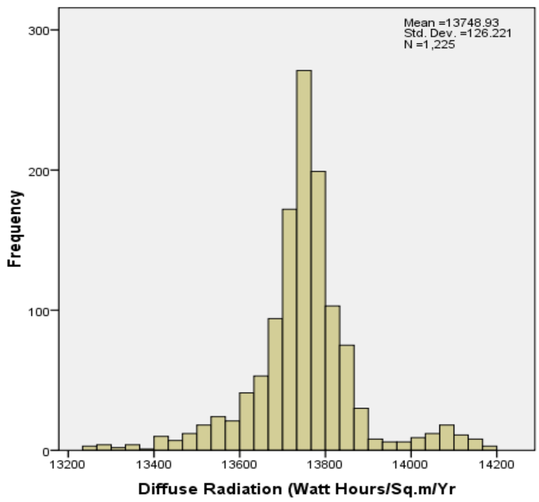


Figure H28. Histogram for Diffuse Radiation #200 of *Solar Analyst*

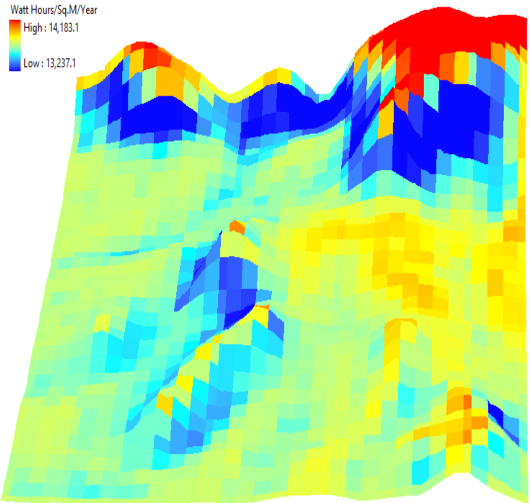


Figure H29. Diffuse Radiation #300 of *Solar Analyst*

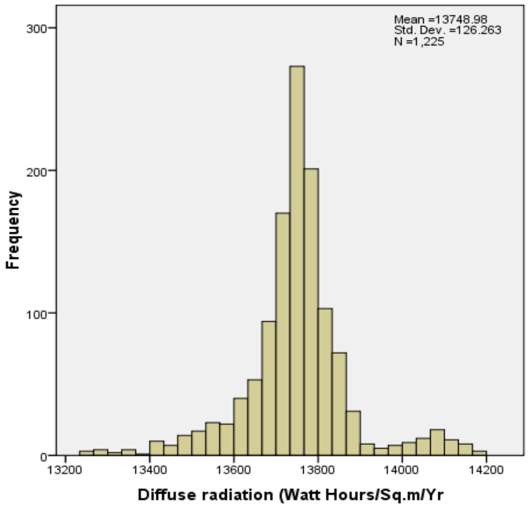


Figure H30. Histogram for Direct Radiation #300 of *Solar Analyst*

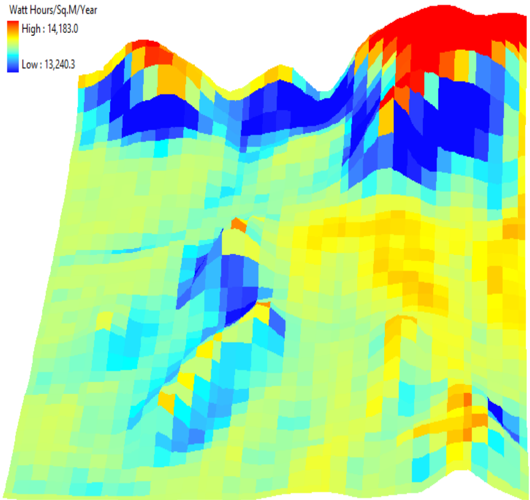


Figure H31. Diffuse Radiation #400 of *Solar Analyst*

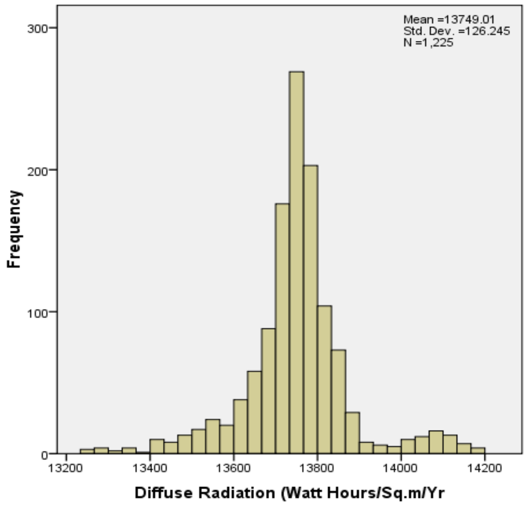


Figure H32. Histogram for Diffuse Radiation #400 of *Solar Analyst*

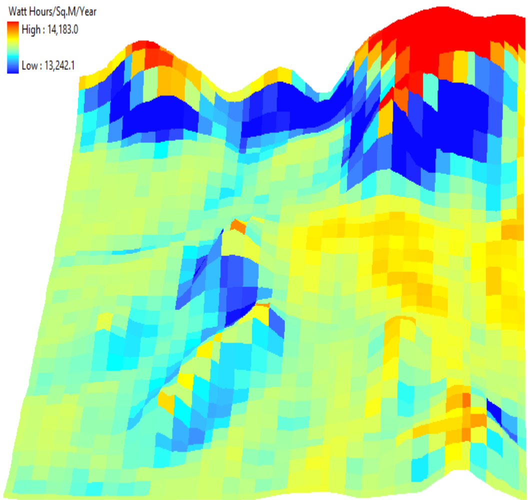


Figure H33. Diffuse Radiation #500 of *Solar Analyst*

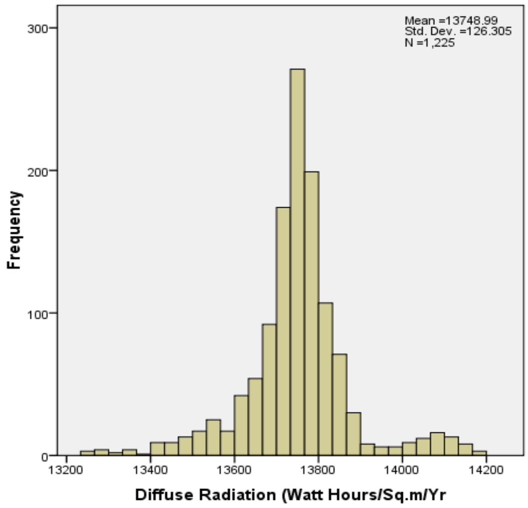


Figure H34. Histogram for Diffuse Radiation #500 of *Solar Analyst*

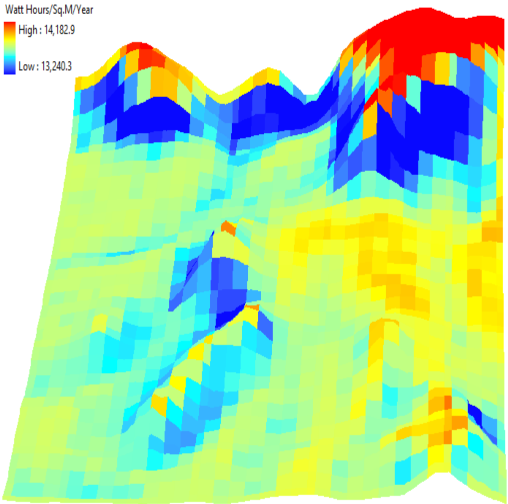


Figure H35. Diffuse Radiation #600 of *Solar Analyst*

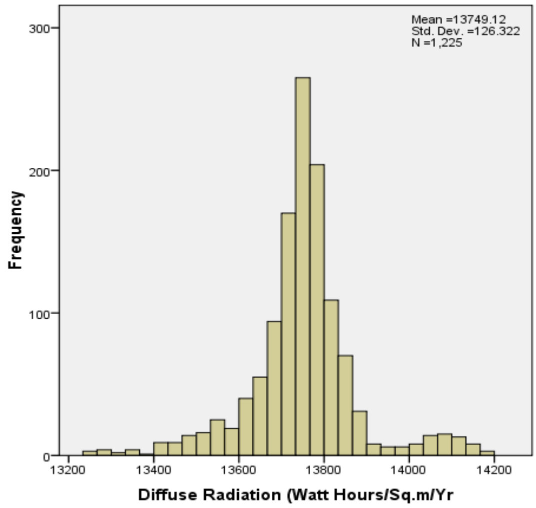


Figure H36. Histogram for Diffuse Radiation #600 of *Solar Analyst*

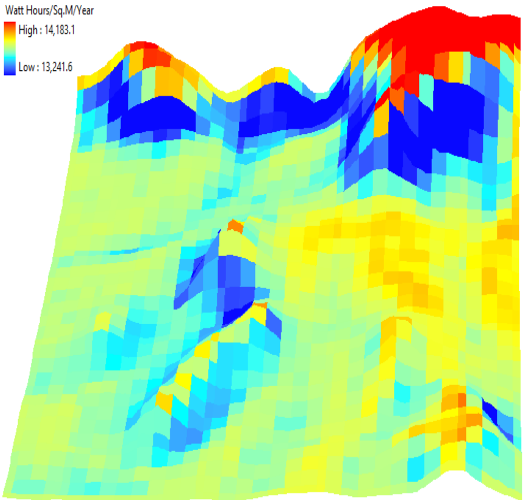


Figure H37. Diffuse Radiation #700 of *Solar Analyst*

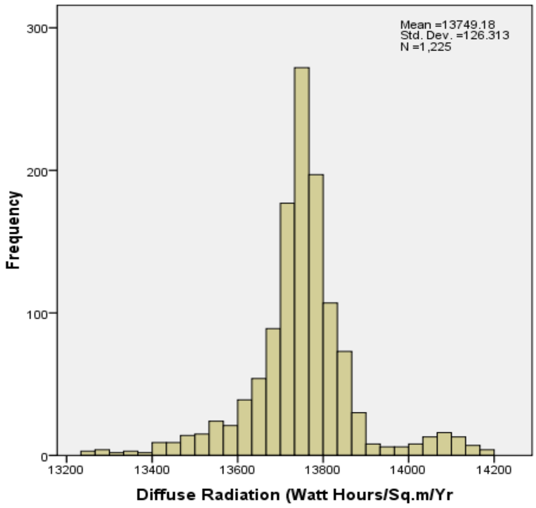


Figure H38. Histogram for Diffuse Radiation #700 of *Solar Analyst*

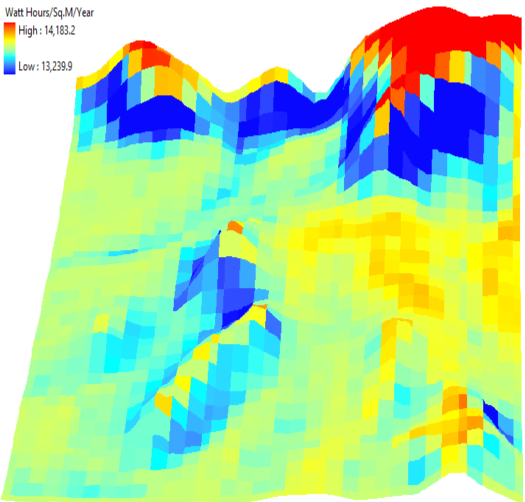


Figure H39. Diffuse Radiation #800 of *Solar Analyst*

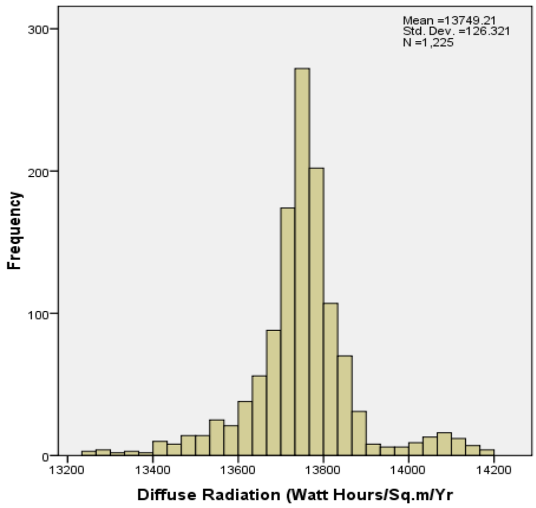


Figure H40. Histogram for Direct Radiation #800 of *Solar Analyst*

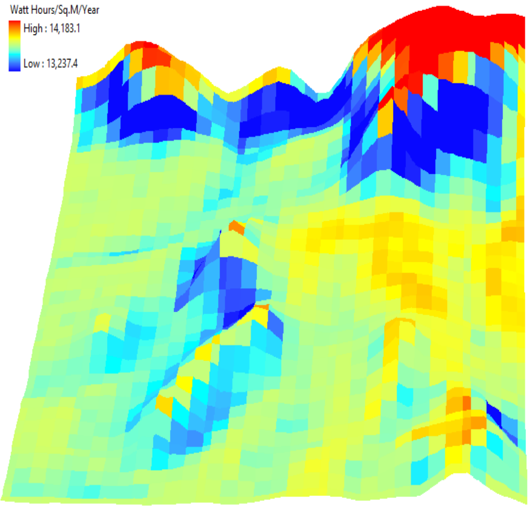


Figure H41. Diffuse Radiation #900 of *Solar Analyst*

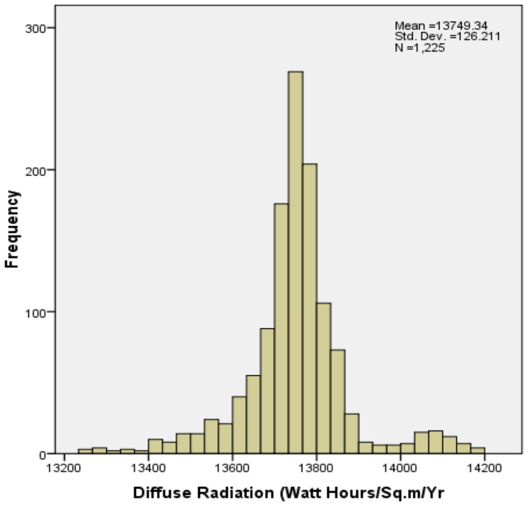


Figure H42. Histogram for Diffuse Radiation #900 of *Solar Analyst*

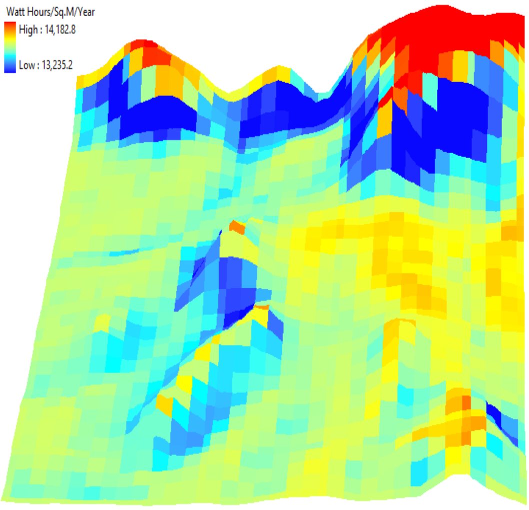


Figure H43. Diffuse Radiation #1000 of *Solar Analyst*

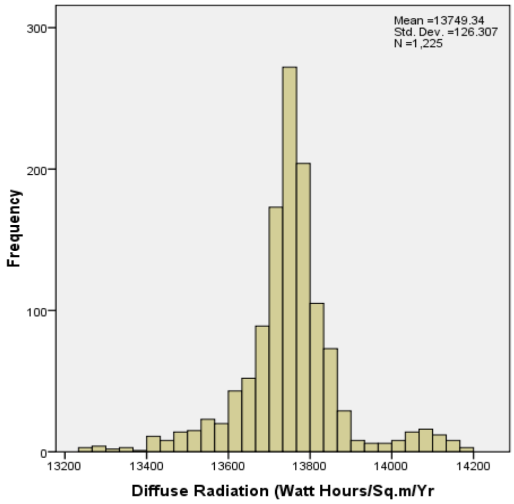


Figure H44. Histogram for Diffuse Radiation #1000 of *Solar Analyst*

Global Radiation

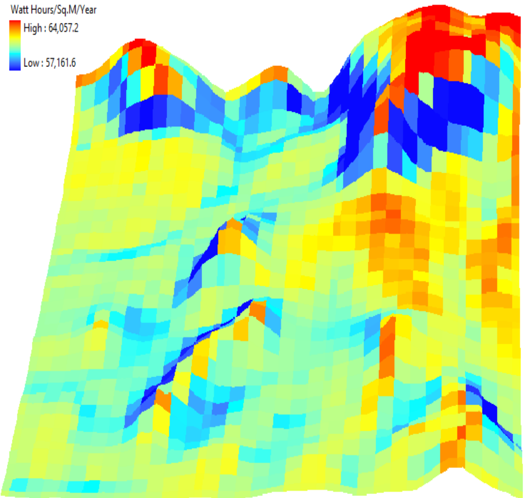


Figure H45. Global Radiation #1 of *Solar Analyst*

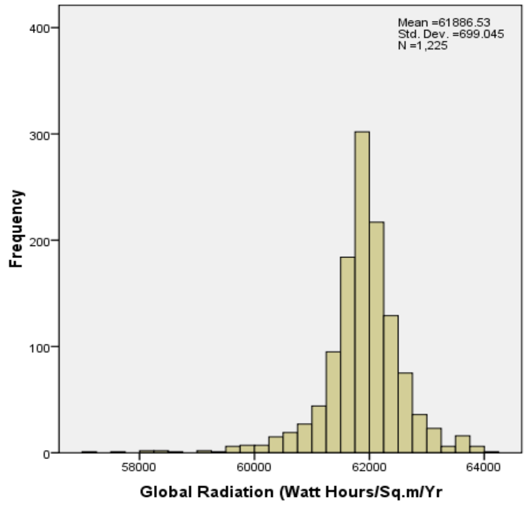


Figure H46. Histogram for Global Radiation #1 of *Solar Analyst*

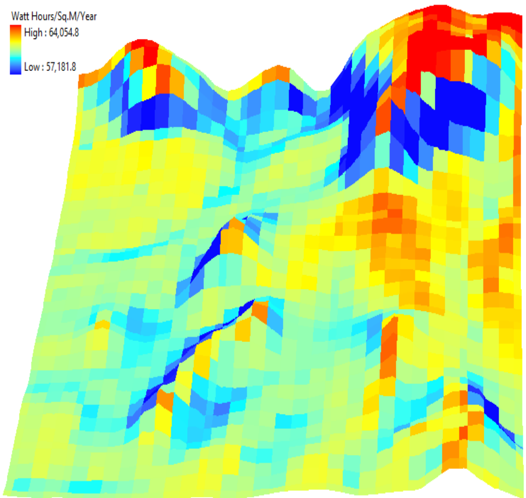


Figure H47. Global Radiation #100 of *Solar Analyst*

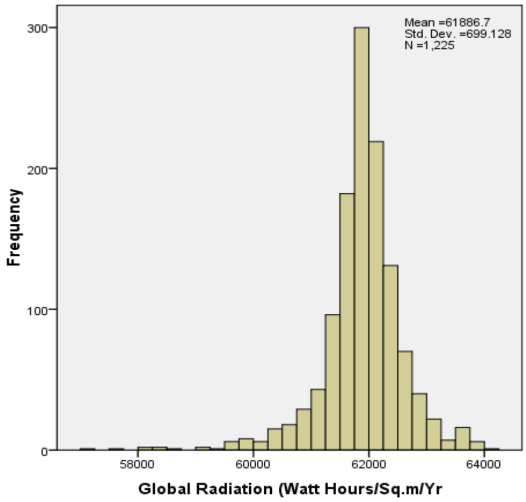


Figure H48. Histogram for Global Radiation #100 of *Solar Analyst*

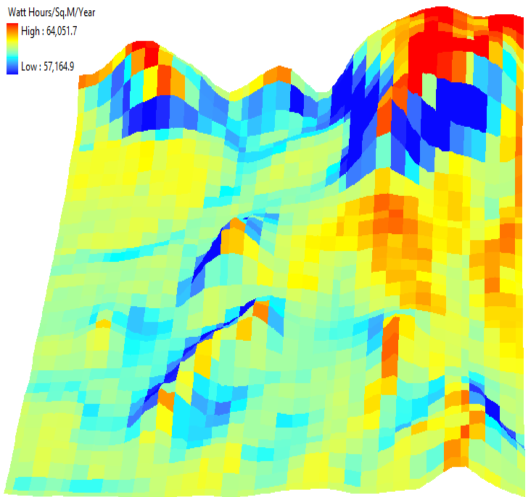


Figure H49. Global Radiation #200 of *Solar Analyst*

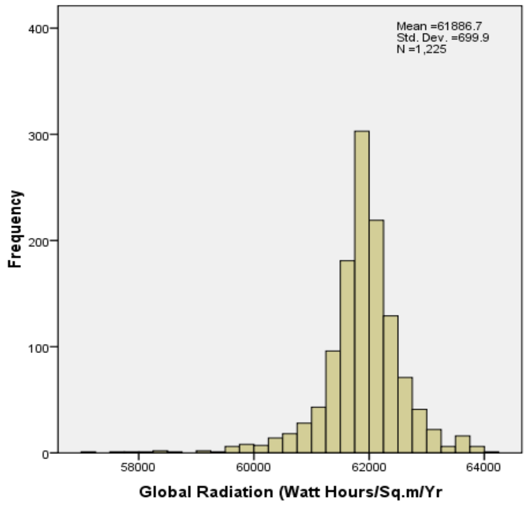


Figure H50. Histogram for Global Radiation #200 of *Solar Analyst*

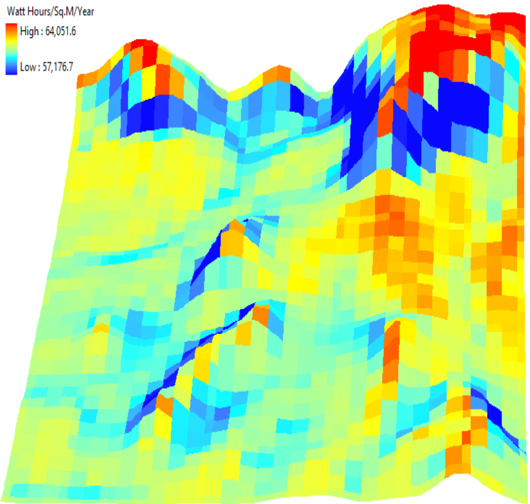


Figure H51. Global Radiation #300 of *Solar Analyst*

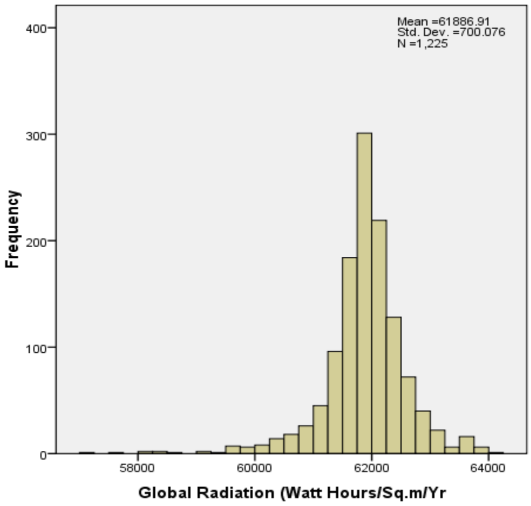


Figure H52. Histogram for Global Radiation #300 of *Solar Analyst*

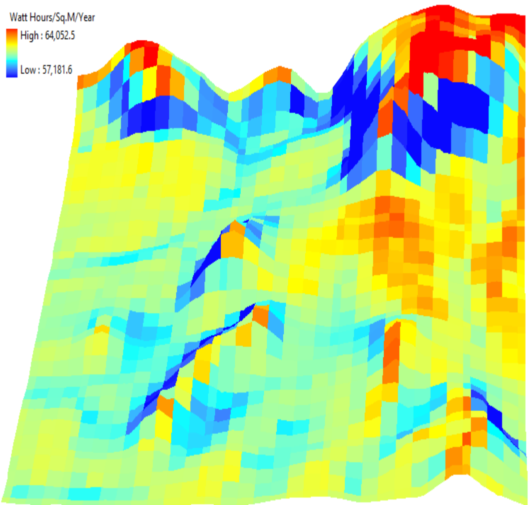


Figure H53. Global Radiation #400 of *Solar Analyst*

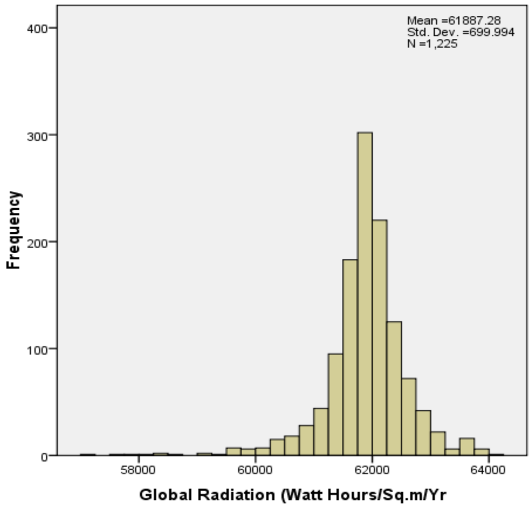


Figure H54. Histogram for Global Radiation #400 of *Solar Analyst*

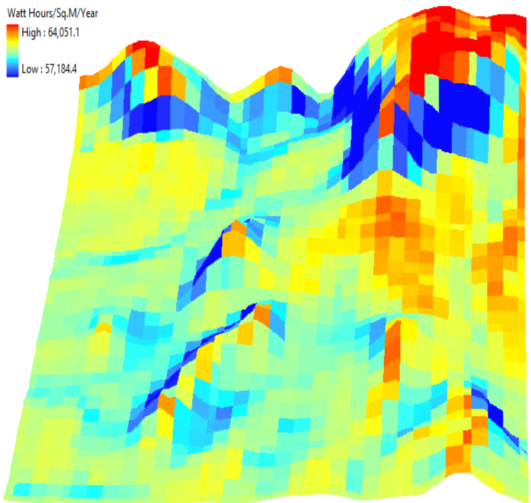


Figure H55. Global Radiation #500 of *Solar Analyst*

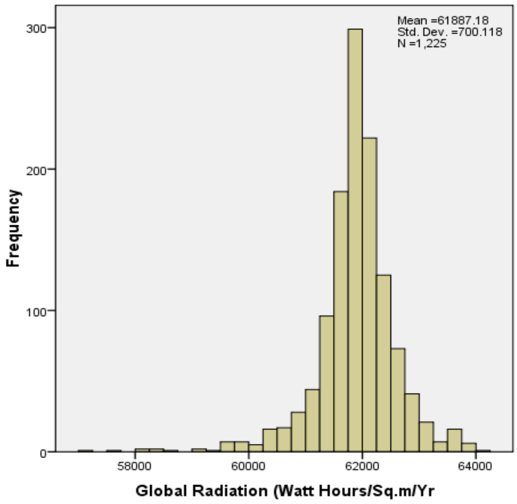


Figure H56. Histogram for Global Radiation #500 of *Solar Analyst*

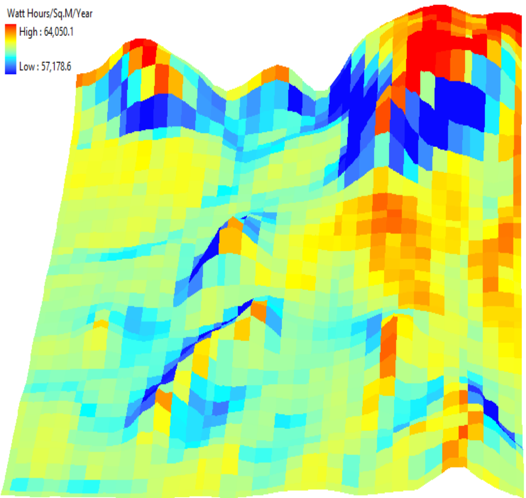


Figure H57. Global Radiation #600 of *Solar Analyst*

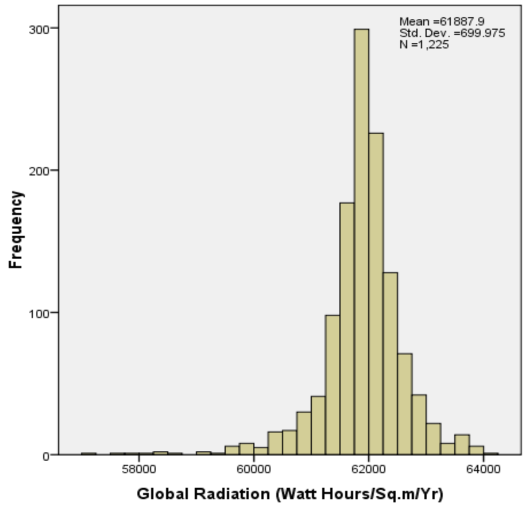


Figure H58. Histogram for Global Radiation #600 of *Solar Analyst*

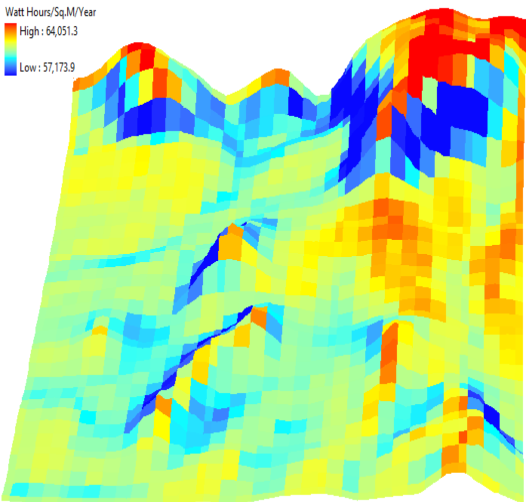


Figure H59. Global Radiation #700 of *Solar Analyst*

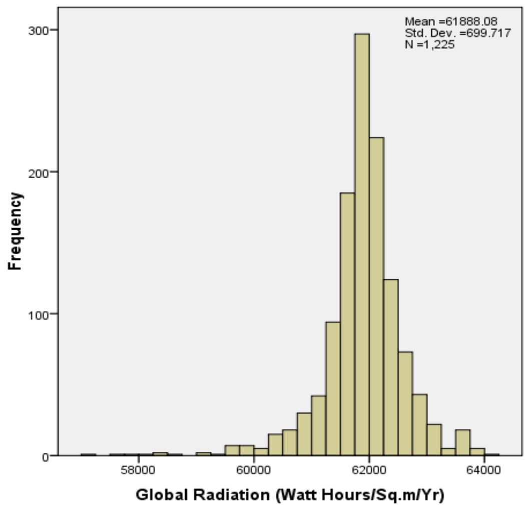


Figure H60. Histogram for Global Radiation #700 of *Solar Analyst*

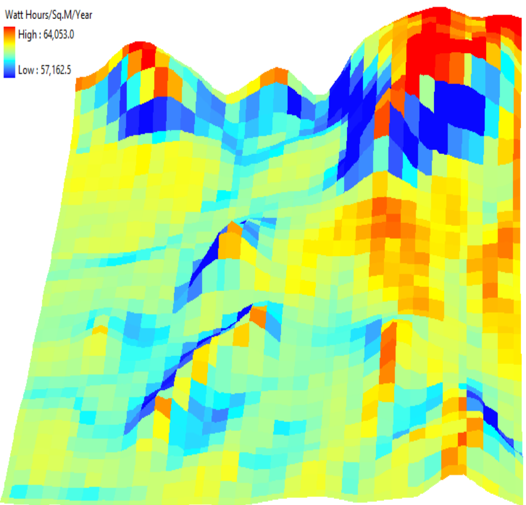


Figure H61. Global Radiation #800 of *Solar Analyst*

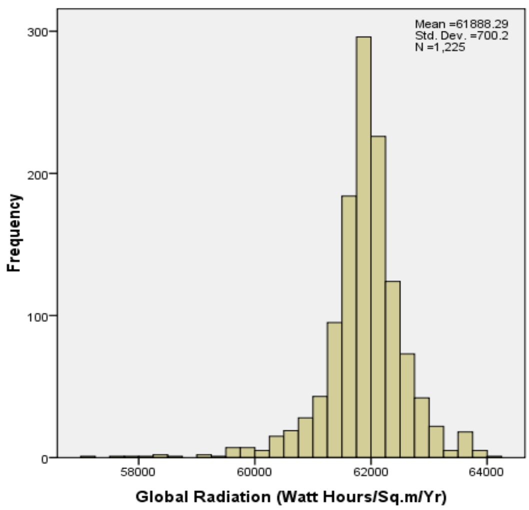


Figure H62. Histogram for Global Radiation #800 of *Solar Analyst*

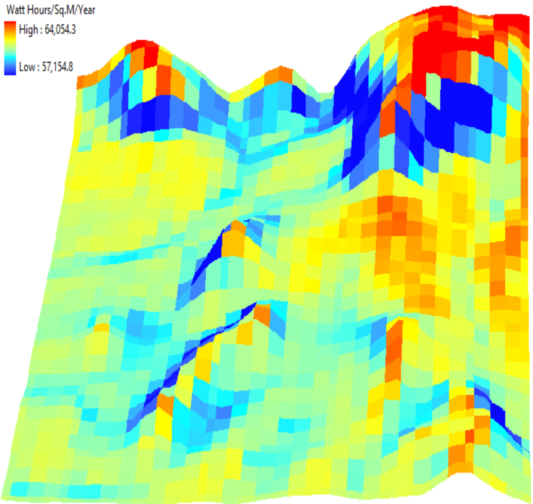


Figure H63. Global Radiation #900 of *Solar Analyst*

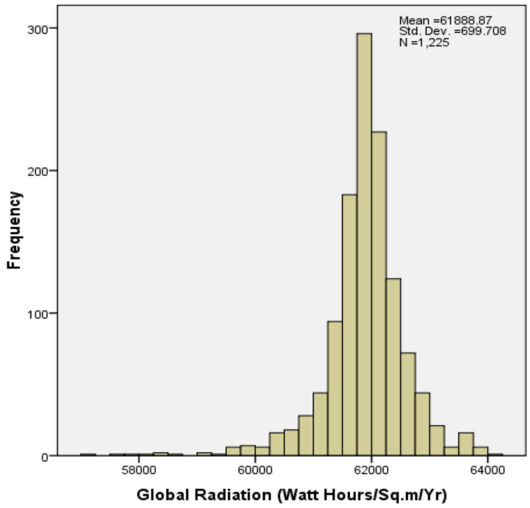


Figure H64. Histogram for Global Radiation #900 of *Solar Analyst*

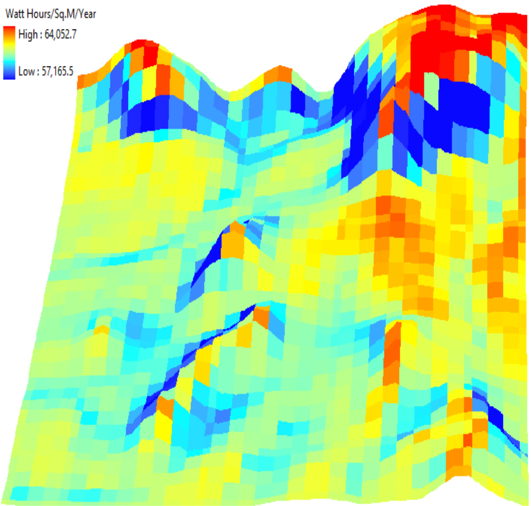


Figure H65. Global Radiation #1000 of *Solar Analyst*

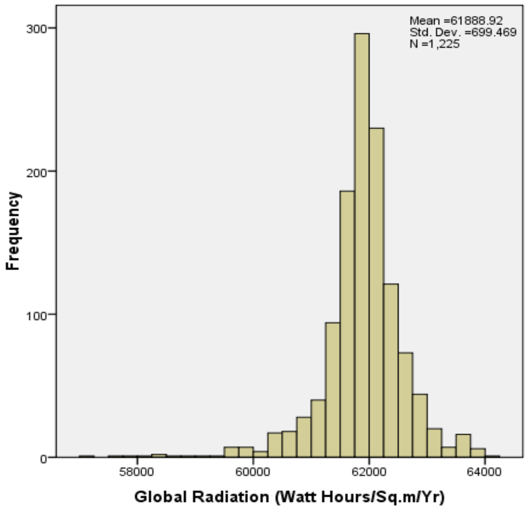


Figure H66. Histogram for Global Radiation #1000 of *Solar Analyst*

Direct Duration

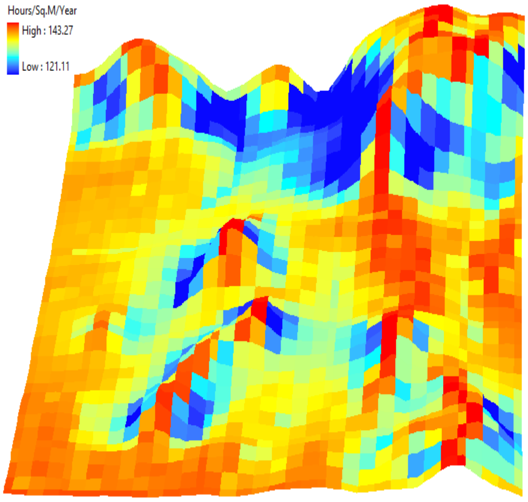


Figure H67. Direct Duration #1 of *Solar Analyst*

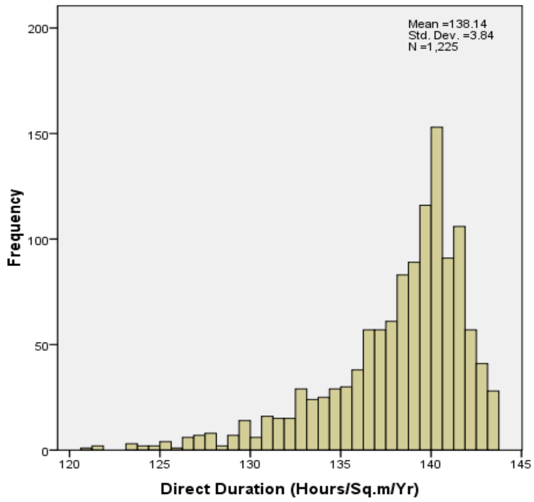


Figure H68. Histogram for Direct Duration #1 of *Solar Analyst*

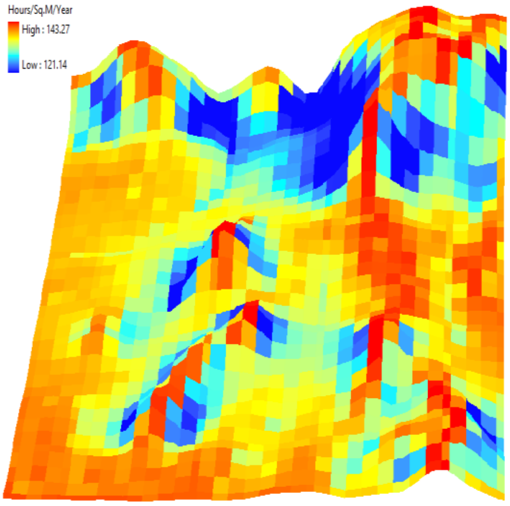


Figure H69. Direct Duration #100 of *Solar Analyst*

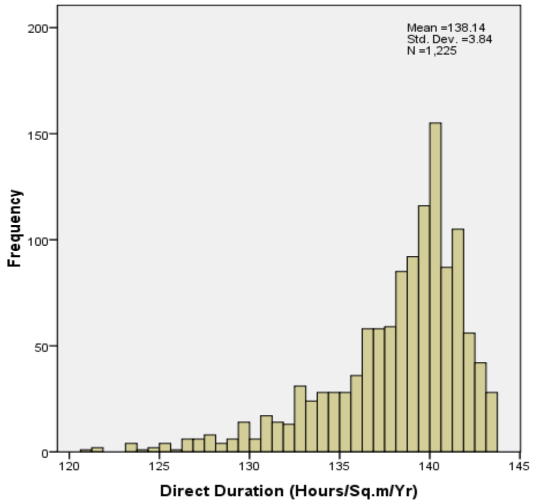


Figure H70. Histogram for Direct Duration #100 of *Solar Analyst*

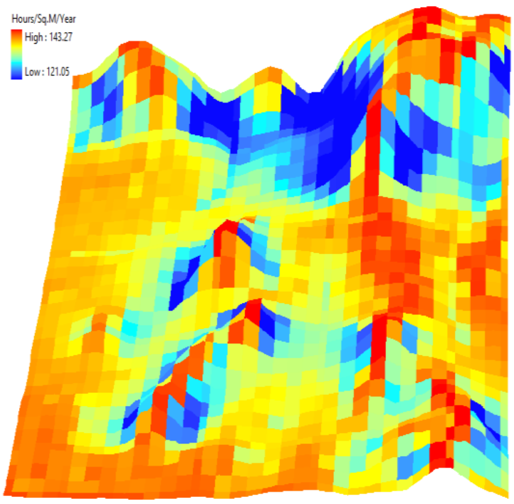


Figure H71. Direct Duration #200 of *Solar Analyst*

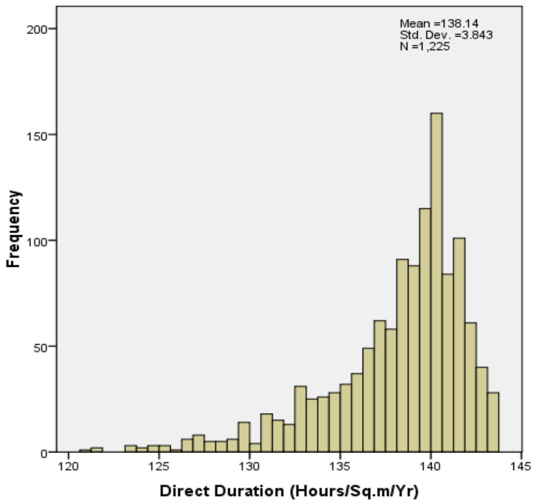


Figure H72. Histogram for Direct Duration #200 of *Solar Analyst*

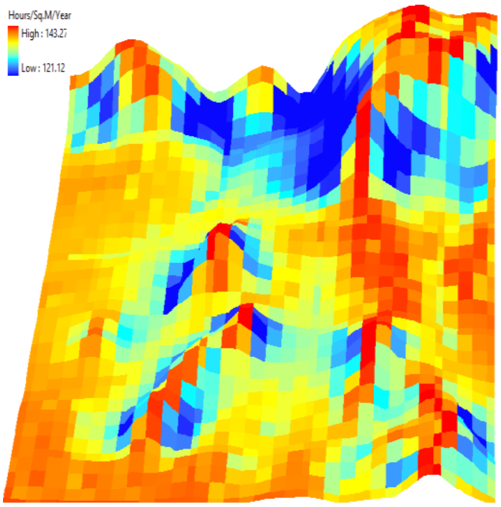


Figure H73. Direct Duration #300 of *Solar Analyst*

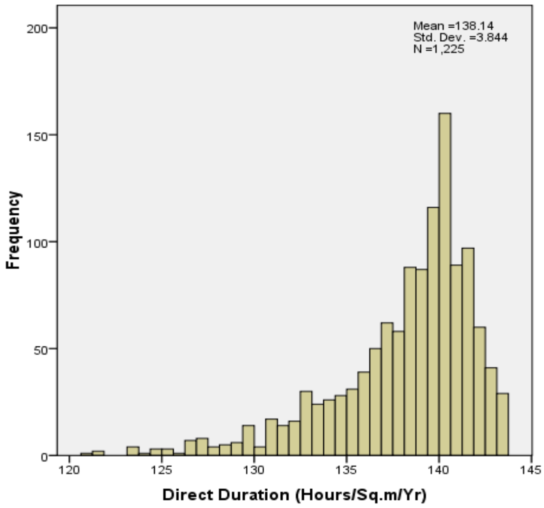


Figure H74. Histogram for Direct Duration #300 of *Solar Analyst*

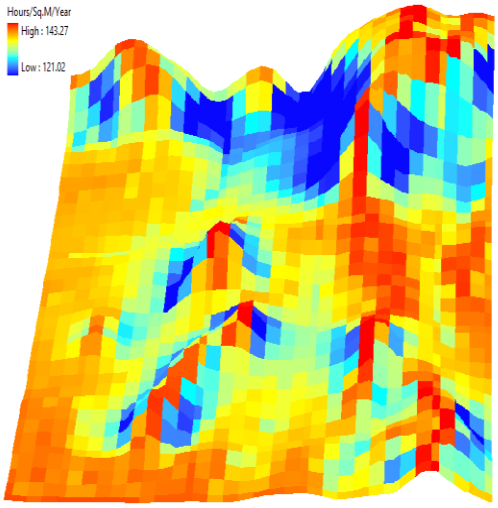


Figure H75. Direct Duration #400 of *Solar Analyst*

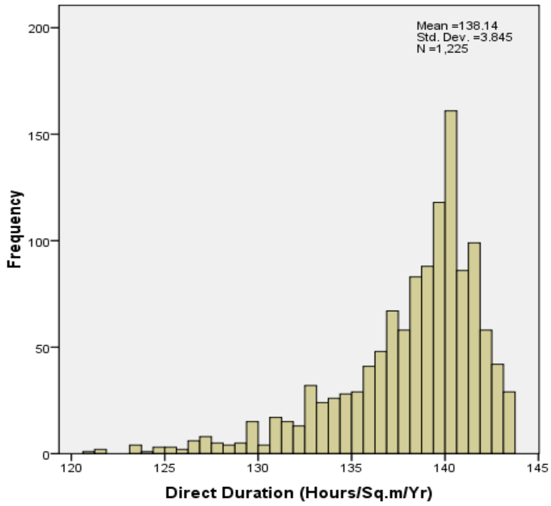


Figure H76. Histogram for Direct Duration #400 of *Solar Analyst*

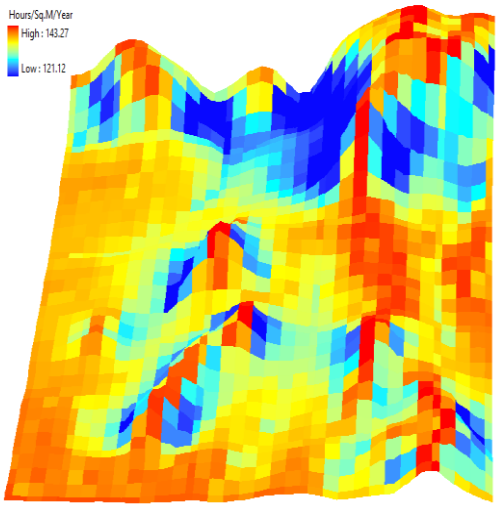


Figure H77. Direct Duration #500 of *Solar Analyst*

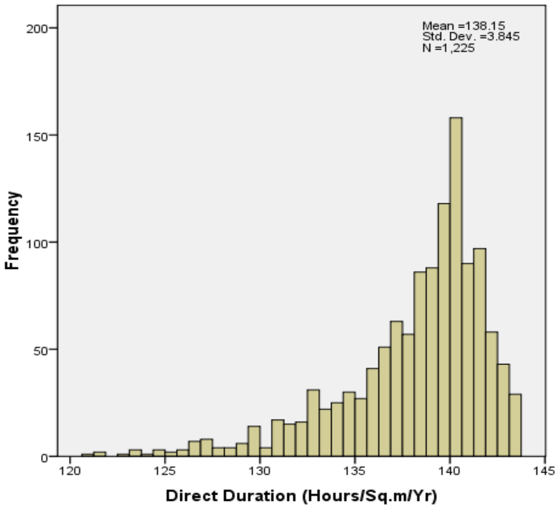


Figure H78. Histogram for Direct Duration #500 of *Solar Analyst*

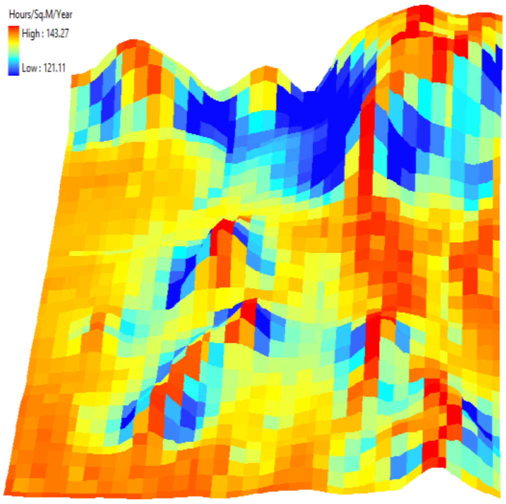


Figure H79. Direct Duration #600 of *Solar Analyst*

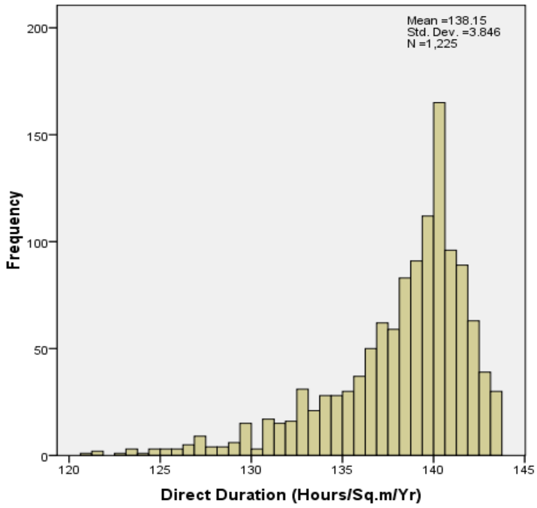


Figure H80. Histogram for Direct Duration #600 of *Solar Analyst*

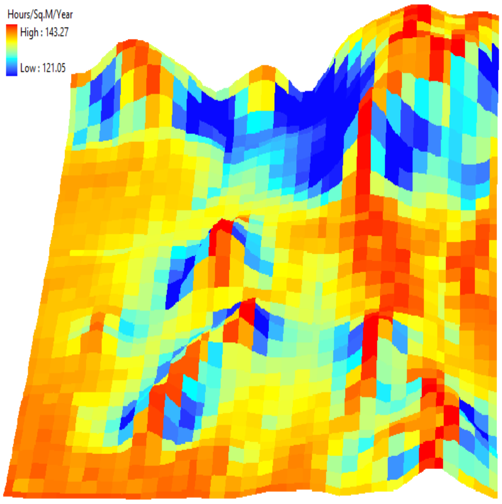


Figure H81. Direct Duration #700 of *Solar Analyst*

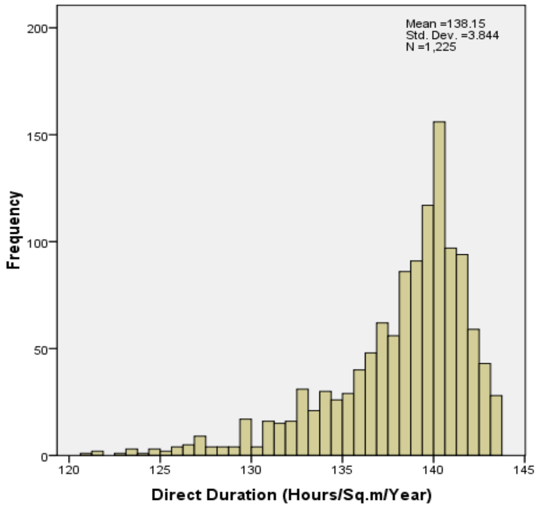


Figure H82. Histogram for Direct Duration #700 of *Solar Analyst*

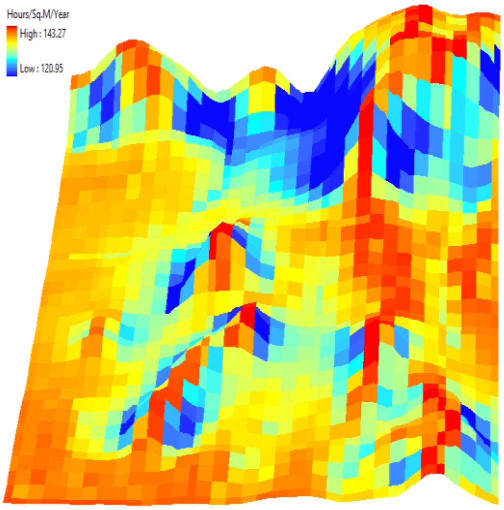


Figure H83. Direct Duration #800 of *Solar Analyst*

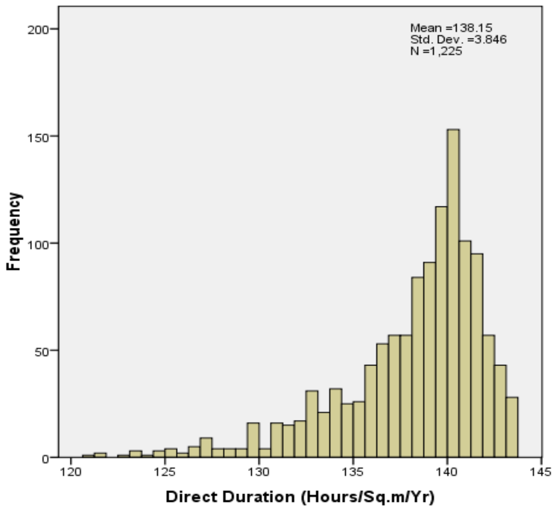


Figure H84. Histogram for Direct Duration #800 of *Solar Analyst*

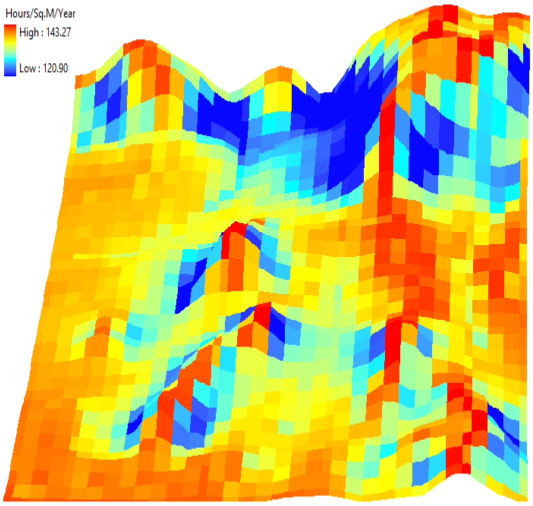


Figure H85. Direct Duration #900 of *Solar Analyst*

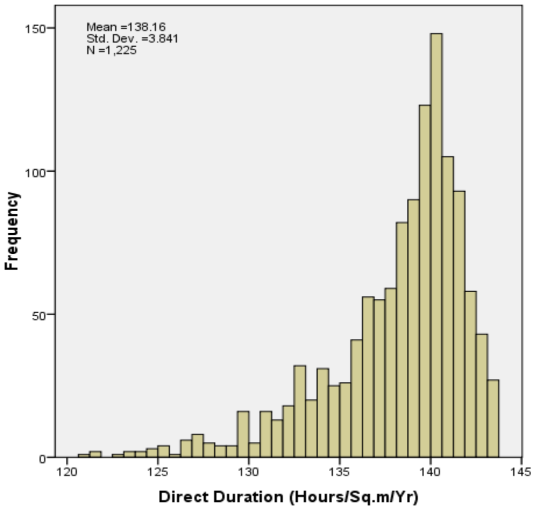


Figure H86. Histogram for Direct Duration #900 of *Solar Analyst*

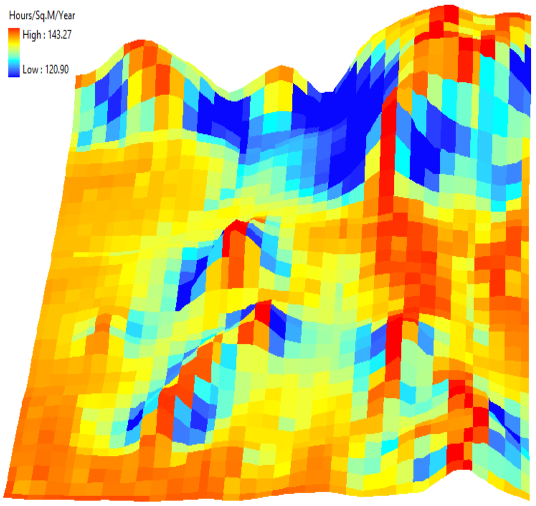


Figure H87. Direct Duration #1000 of *Solar Analyst*

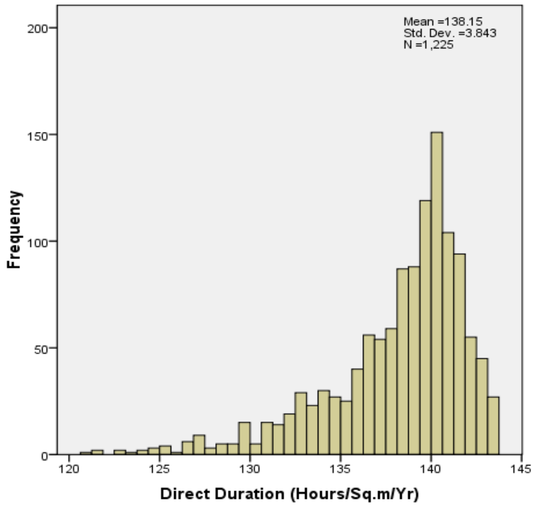


Figure H88. Histogram for Direct Duration #1000 of *Solar Analyst*

APPENDIX I

SELECTED OUTPUTS OF *R.SUN* (GRASS GIS) FROM THE REALIZED DEMS

Direct Radiation

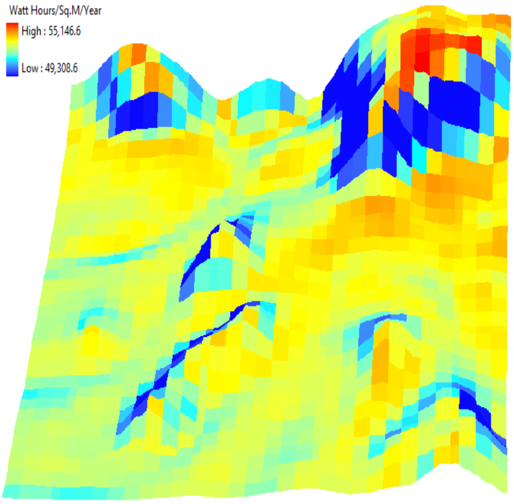


Figure I1. Direct Radiation #1 of *r.sun*

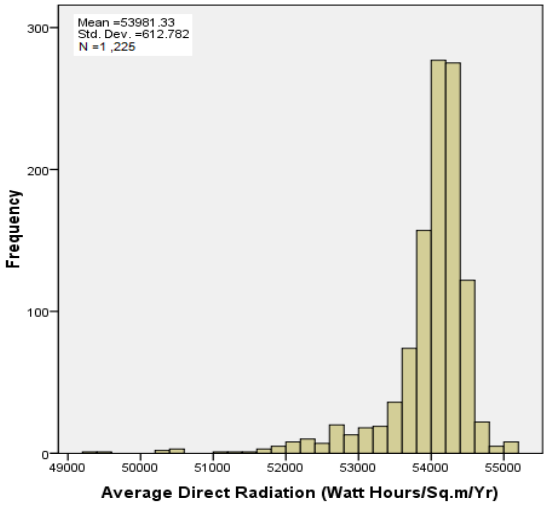


Figure I2. Histogram for Direct Radiation #1 of *r.sun*

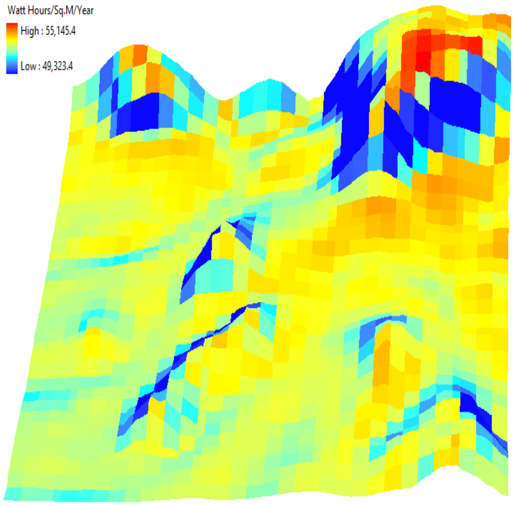


Figure I3. Direct Radiation #100 of *r.sun*

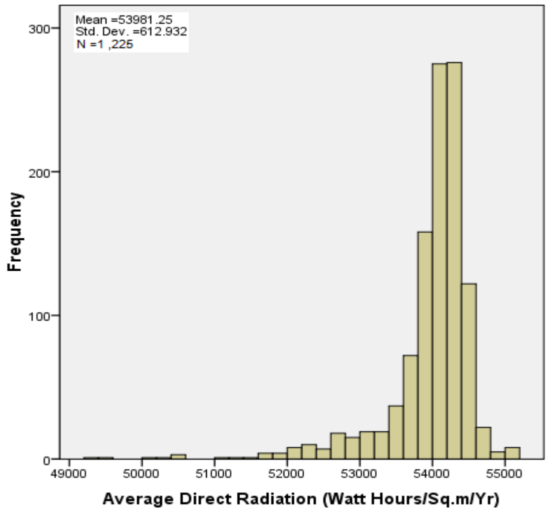


Figure I4. Histogram for Direct Radiation #100 of *r.sun*

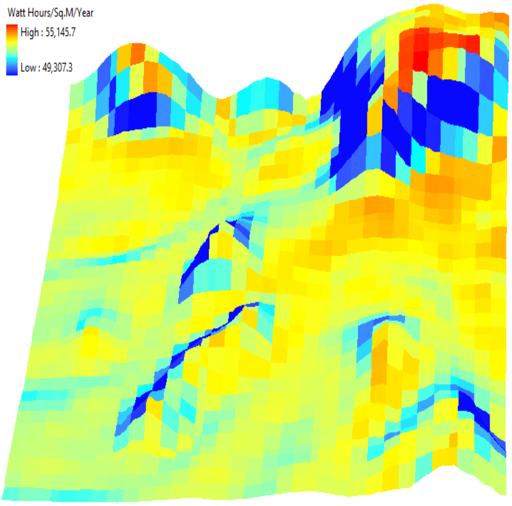


Figure I5. Direct Radiation #200 of *r.sun*

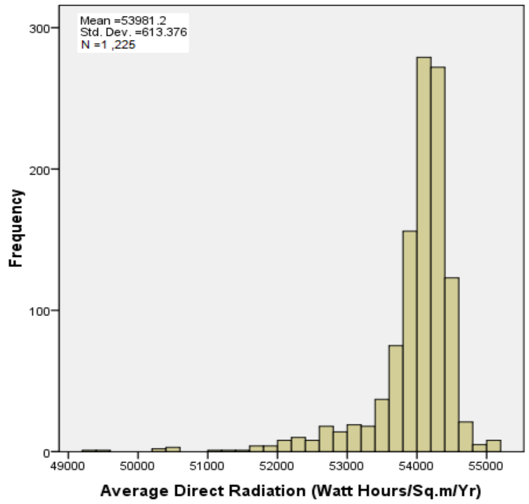


Figure I6. Histogram for Direct Radiation #200 of *r.sun*

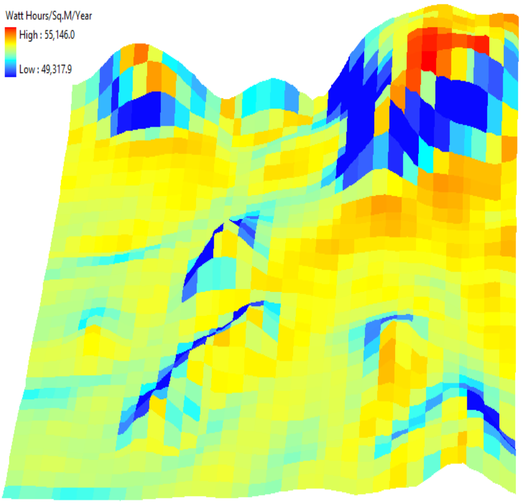


Figure I7. Direct Radiation #300 of *r.sun*

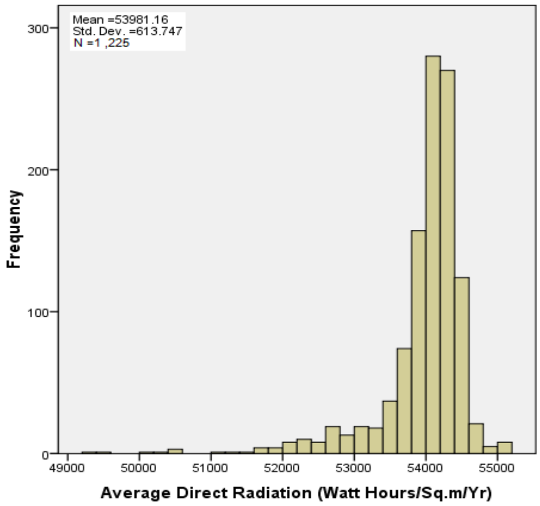


Figure I8. Histogram for Direct Radiation #300 of *r.sun*

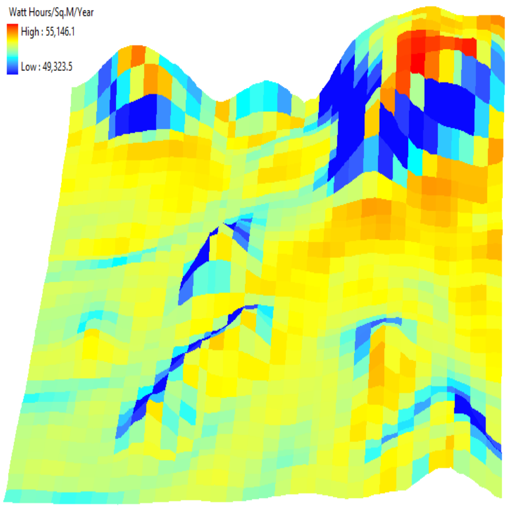


Figure I9. Direct Radiation #400 of *r.sun*

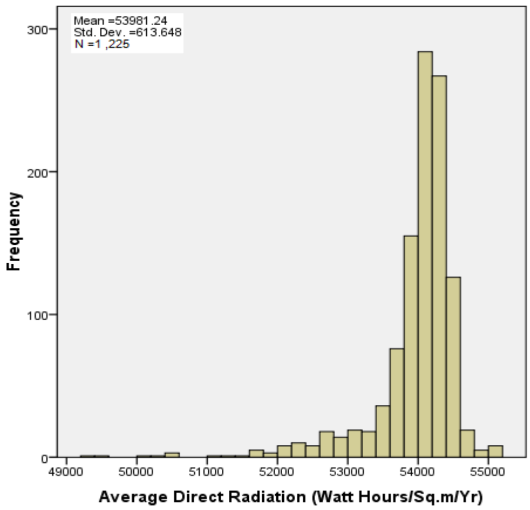


Figure I10. Histogram for Direct Radiation #400 of *r.sun*

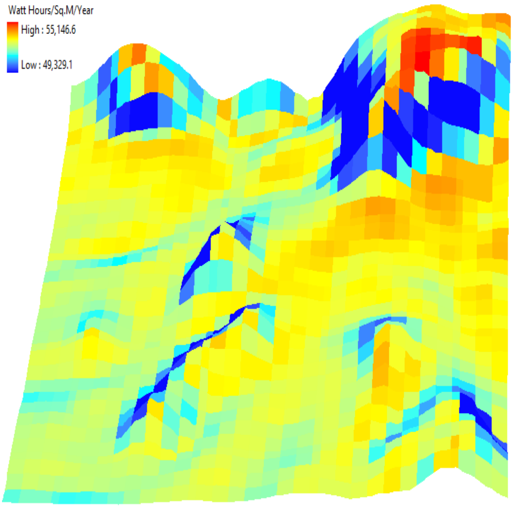


Figure I11. Direct Radiation #500 of *r.sun*

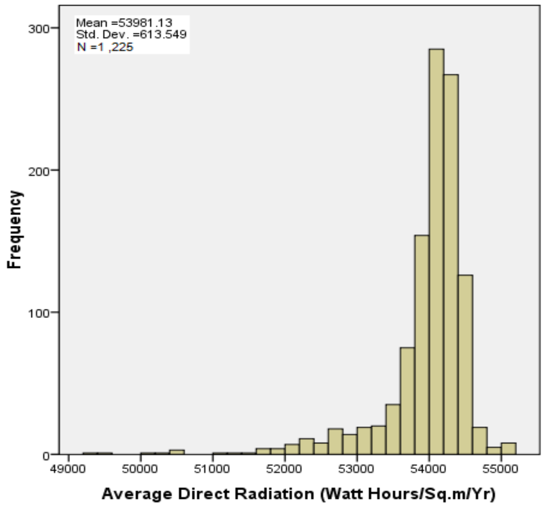


Figure I12. Histogram for Direct Radiation #500 of *r.sun*

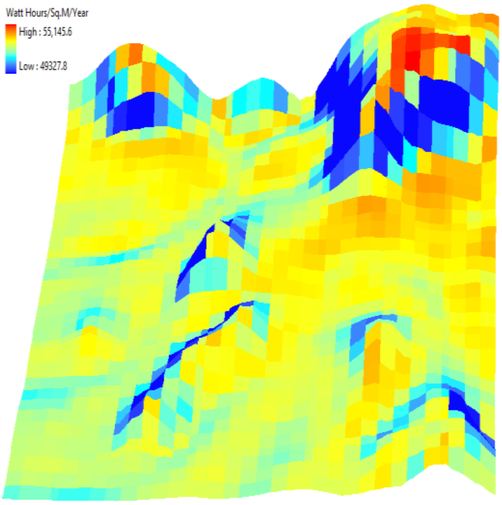


Figure I13. Direct Radiation #600 of *r.sun*

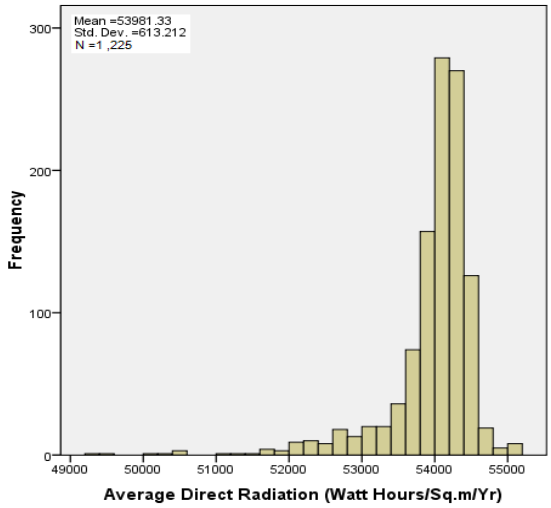


Figure I14. Histogram for Direct Radiation #600 of *r.sun*

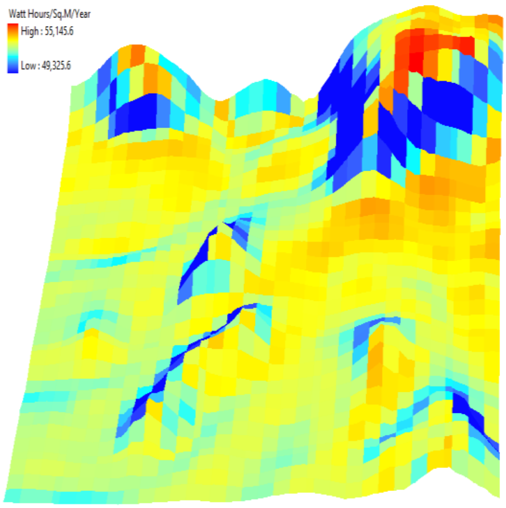


Figure I15. Direct Radiation #700 of *r.sun*

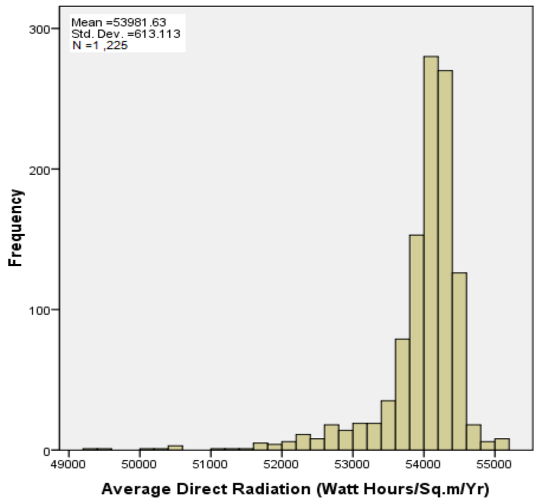


Figure I16. Histogram for Direct Radiation #700 of *r.sun*

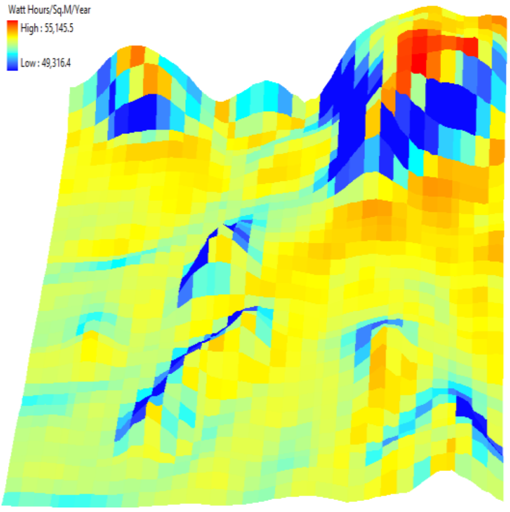


Figure I17. Direct Radiation #800 of *r.sun*

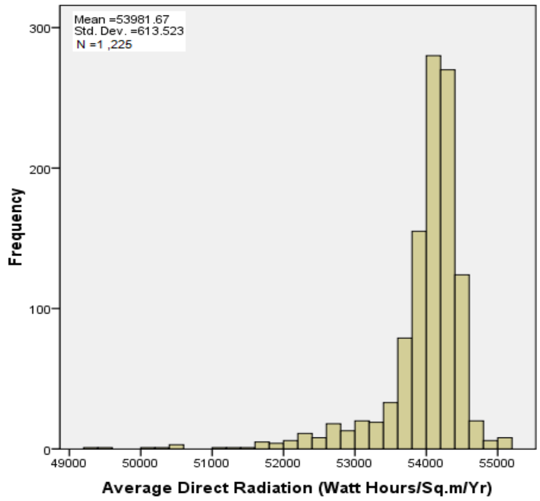


Figure I18. Histogram for Direct Radiation #800 of *r.sun*

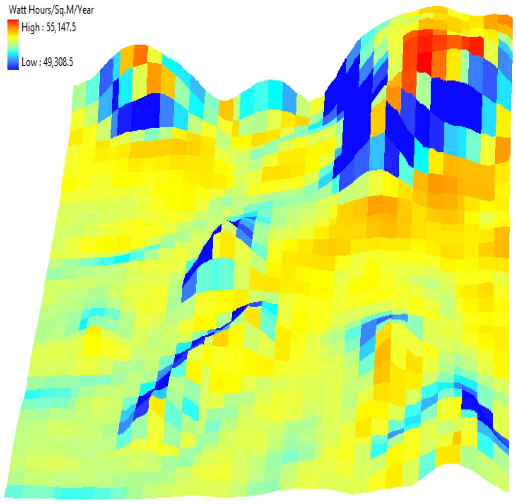


Figure I19. Direct Radiation #900 of *r.sun*

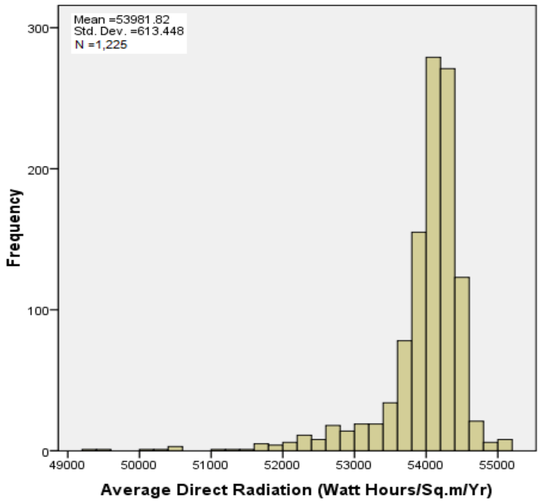


Figure I20. Histogram for Direct Radiation #900 of *r.sun*

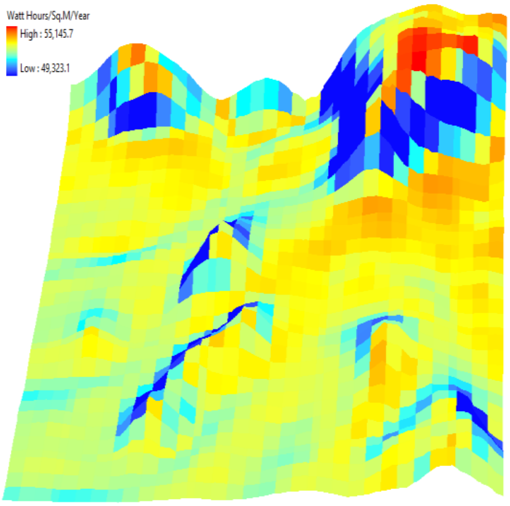


Figure I21. Direct Radiation #1000 of *r.sun*

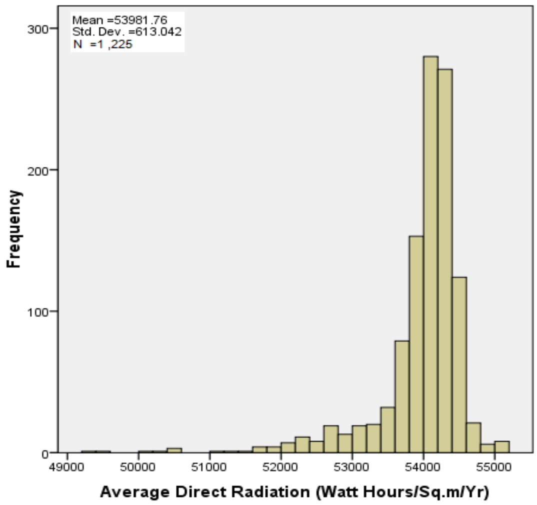


Figure I22. Histogram for Direct Radiation #1000 of *r.sun*

Diffuse Radiation

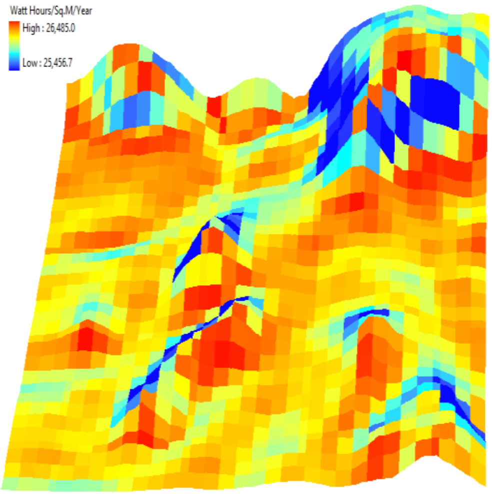


Figure I23. Diffuse Radiation #1 of *r.sun*

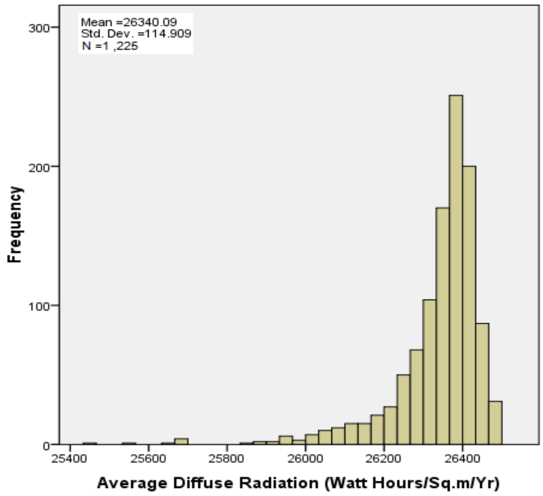


Figure I24. Histogram for Diffuse Radiation #1 of *r.sun*

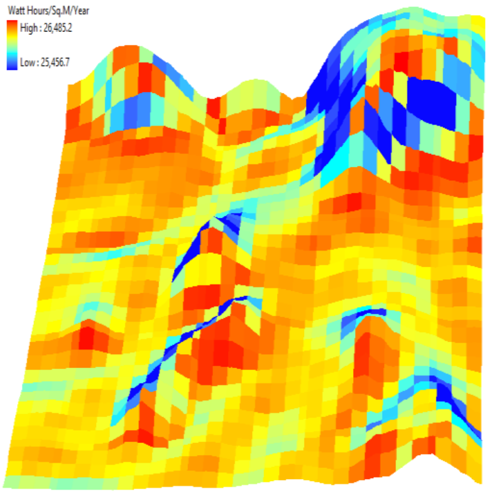


Figure I25. Diffuse Radiation #100 of *r.sun*

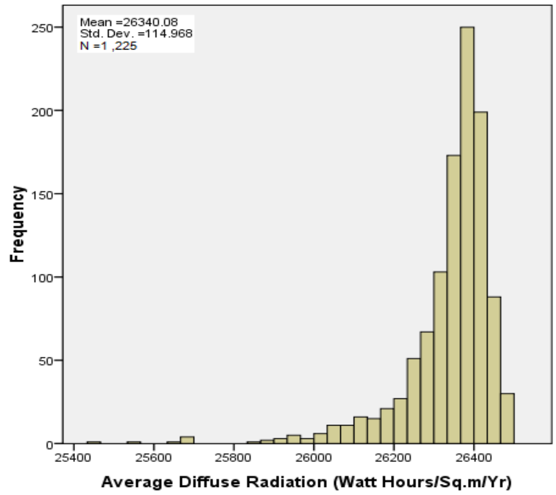


Figure I26. Histogram for Diffuse Radiation #100 of *r.sun*

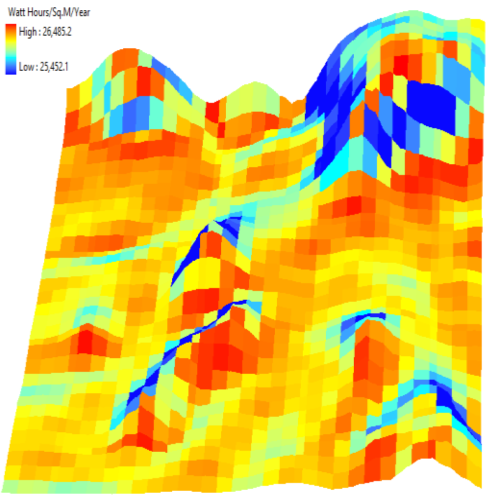


Figure I27. Diffuse Radiation #200 of *r.sun*

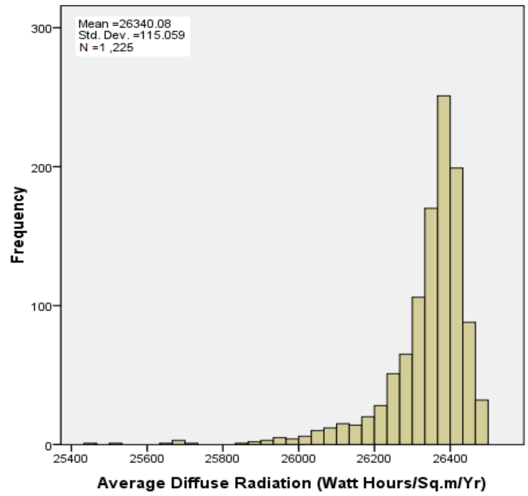


Figure I28. Histogram for Diffuse Radiation #200 of *r.sun*

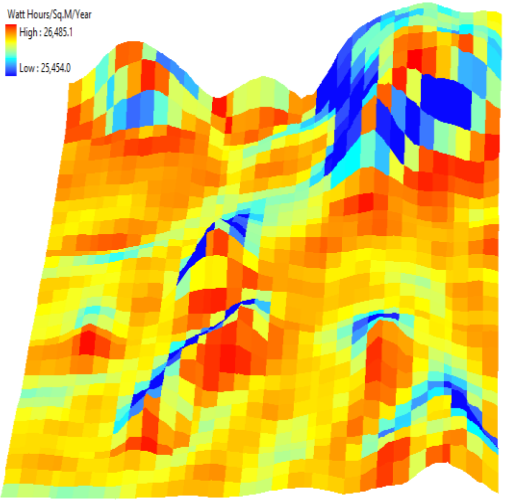


Figure I29. Diffuse Radiation #300 of *r.sun*

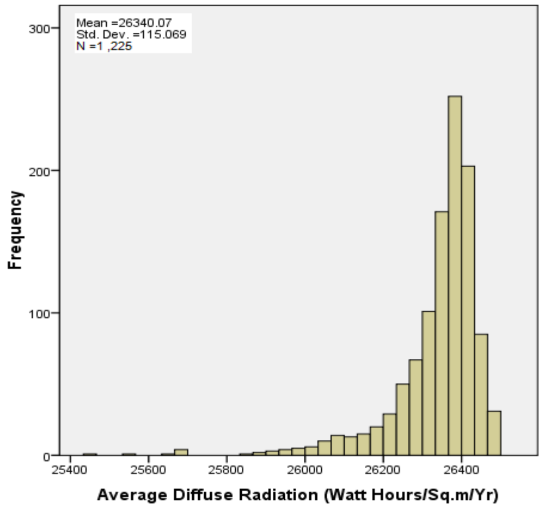


Figure I30. Histogram for Diffuse Radiation #300 of *r.sun*

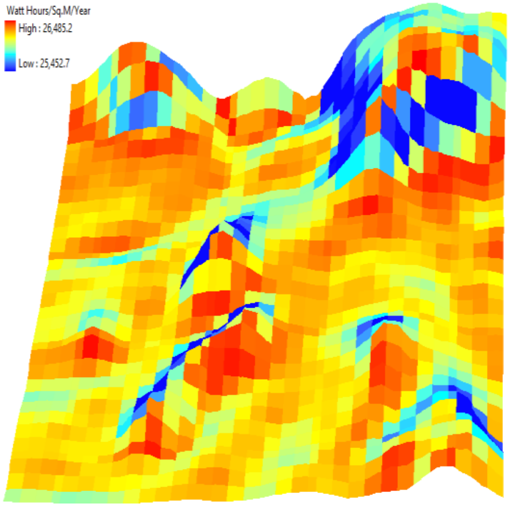


Figure I31. Diffuse Radiation #400 of *r.sun*

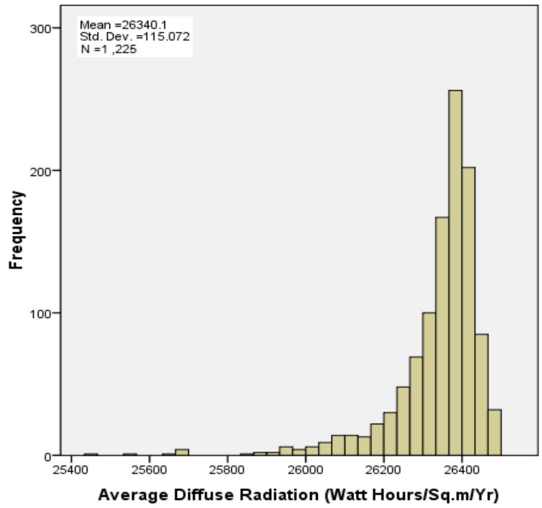


Figure I32. Histogram for Diffuse Radiation #400 of *r.sun*

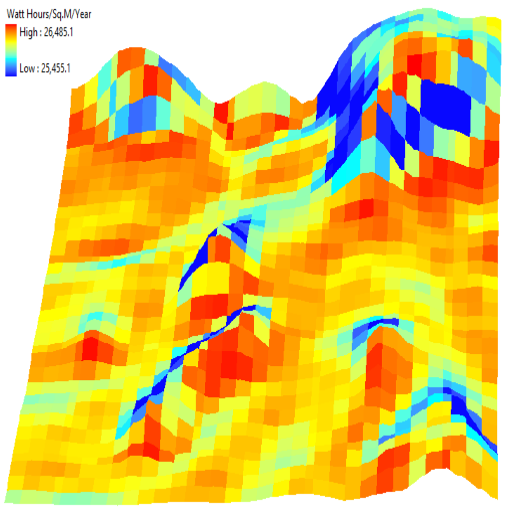


Figure I33. Diffuse Radiation #500 of *r.sun*

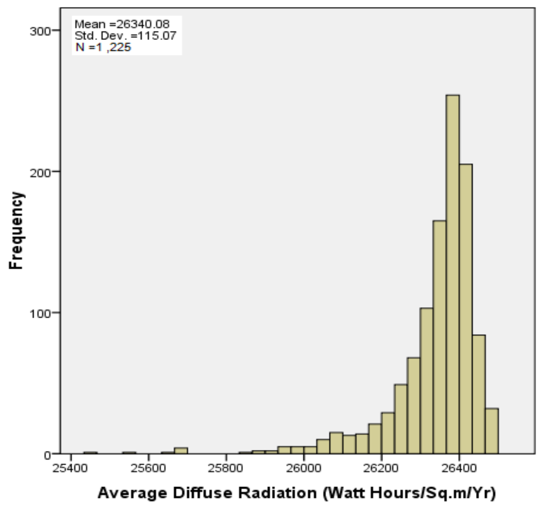


Figure I34. Histogram for Diffuse Radiation #500 of *r.sun*

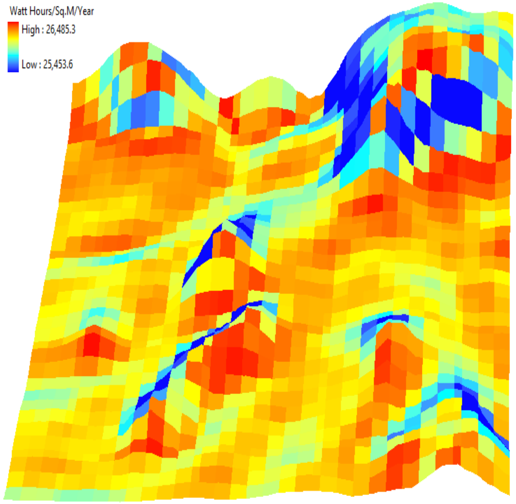


Figure I35. Diffuse Radiation #600 of *r.sun*

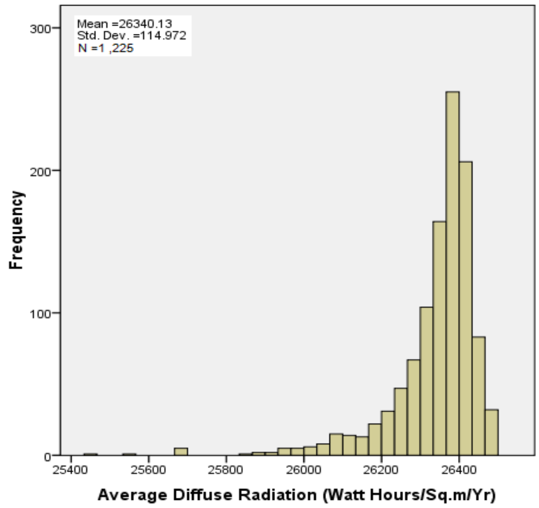


Figure I36. Histogram for Diffuse Radiation #600 of *r.sun*

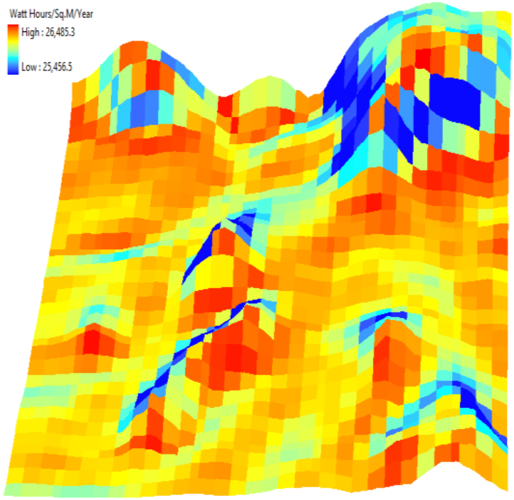


Figure I37. Diffuse Radiation #700 of *r.sun*

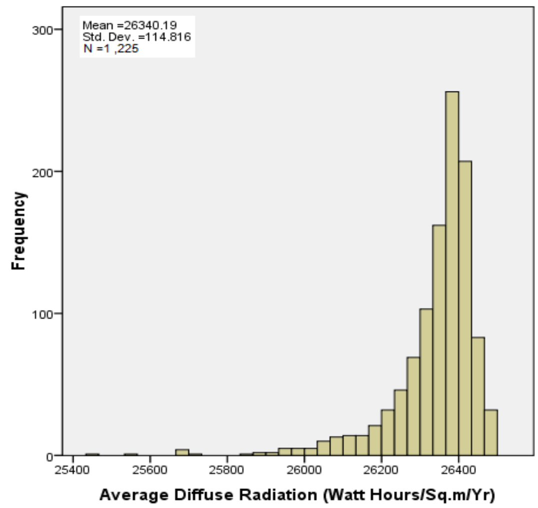


Figure I38. Histogram for Diffuse Radiation #700 of *r.sun*

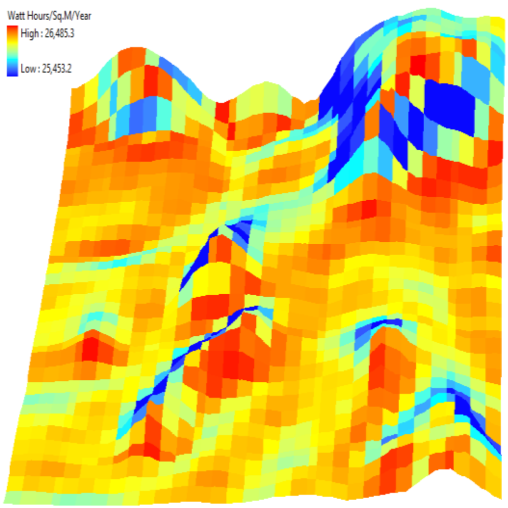


Figure I39. Diffuse Radiation #800 of *r.sun*

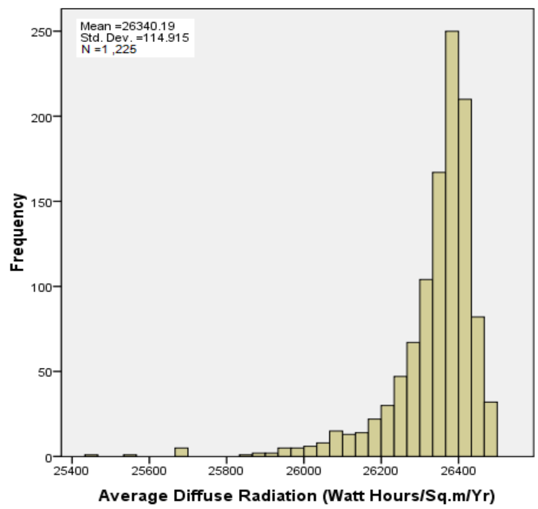


Figure I40. Histogram for Diffuse Radiation #800 of *r.sun*

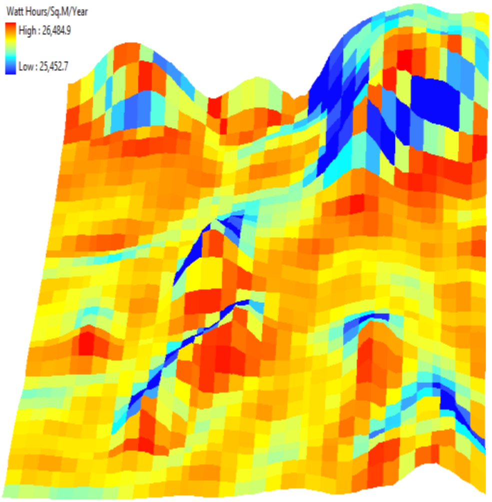


Figure I41. Diffuse Radiation #900 of *r.sun*

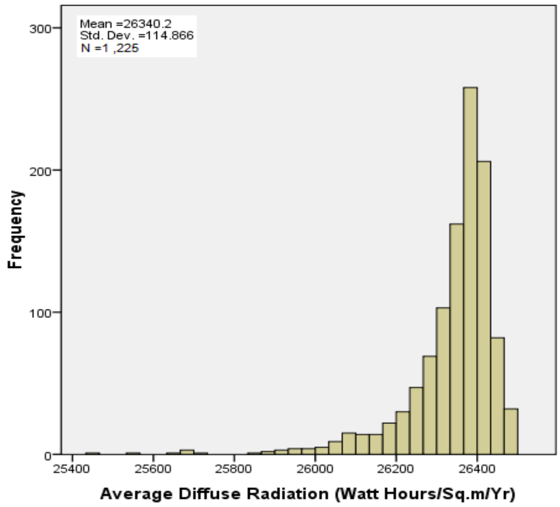


Figure I42. Histogram for Diffuse Radiation #900 of *r.sun*

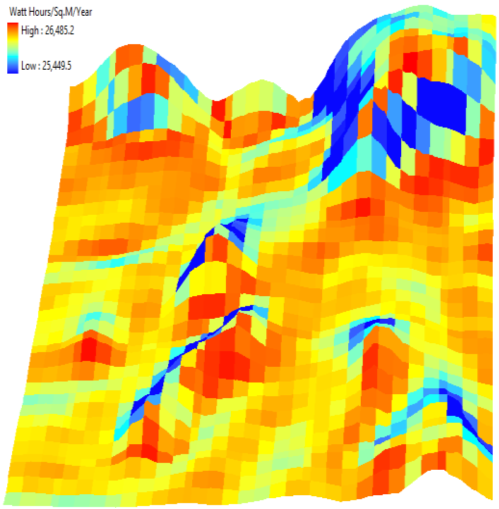


Figure I43. Diffuse Radiation #1000 of *r.sun*

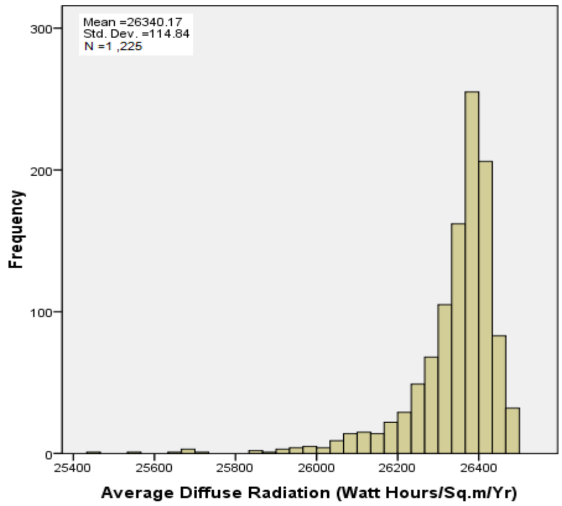


Figure I44. Histogram for Diffuse Radiation #1000 of *r.sun*

Global Radiation

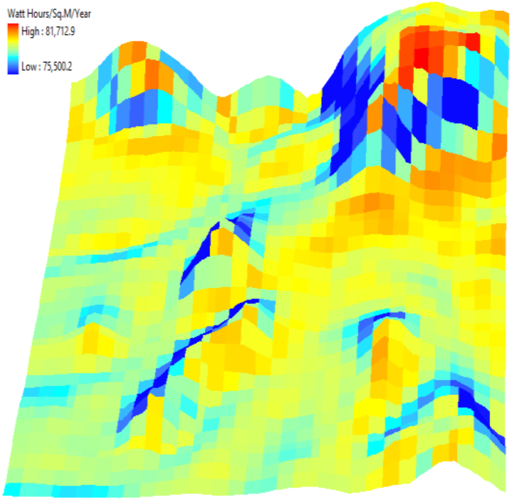


Figure I45. Global Radiation #1 of *r.sun*

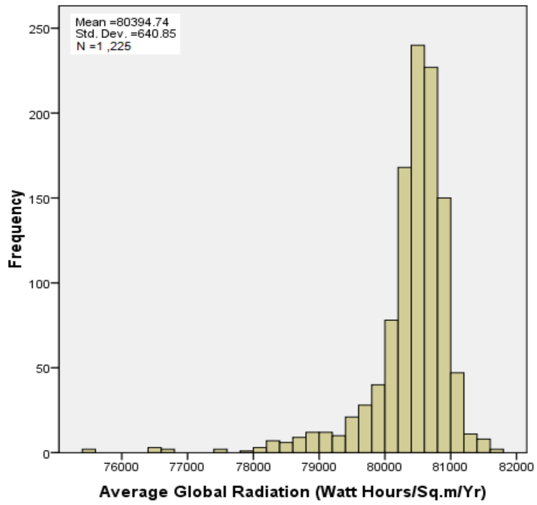


Figure I46. Histogram for Global Radiation #1 of *r.sun*

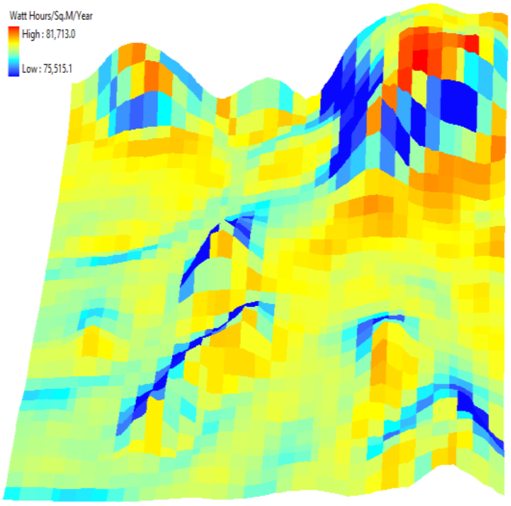


Figure I47. Global Radiation #100 of *r.sun*

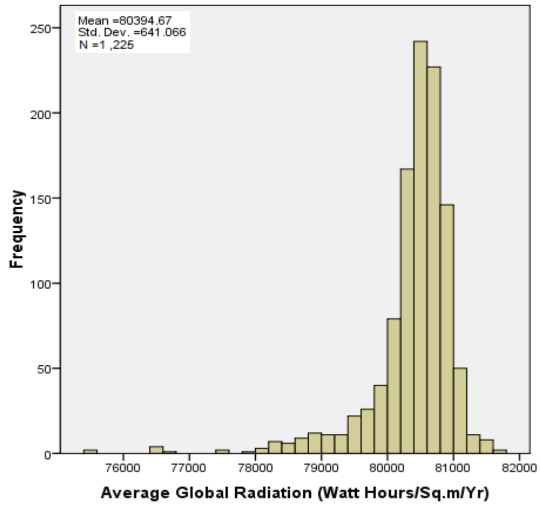


Figure I48. Histogram for Global Radiation #100 of *r.sun*

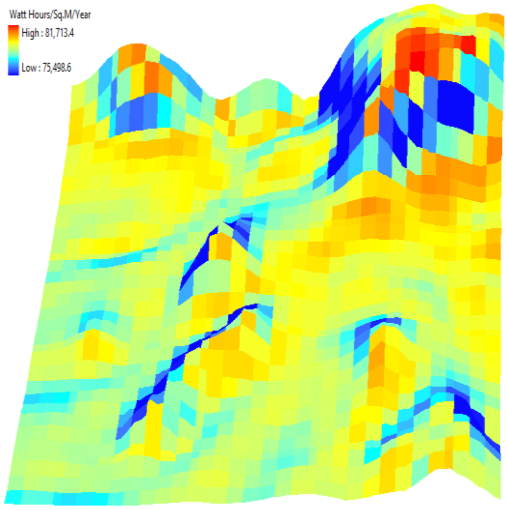


Figure I49. Global Radiation #200 of *r.sun*

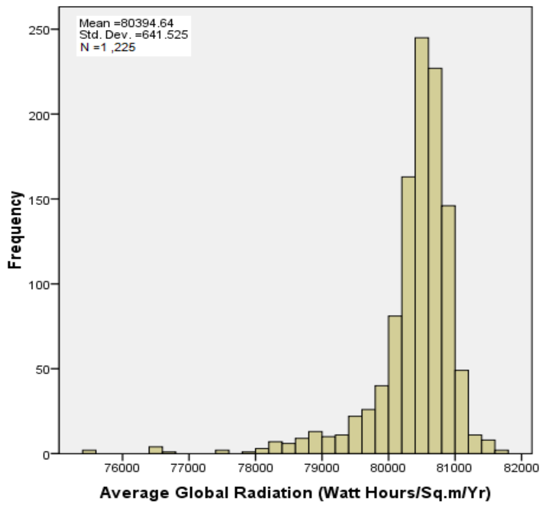


Figure I50. Histogram for Global Radiation #200 of *r.sun*

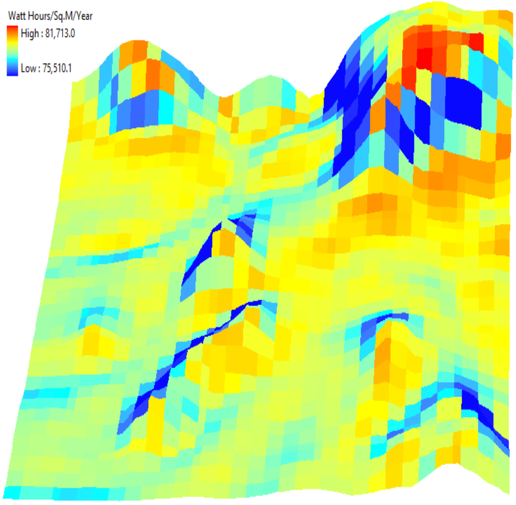


Figure I51. Global Radiation #300 of *r.sun*

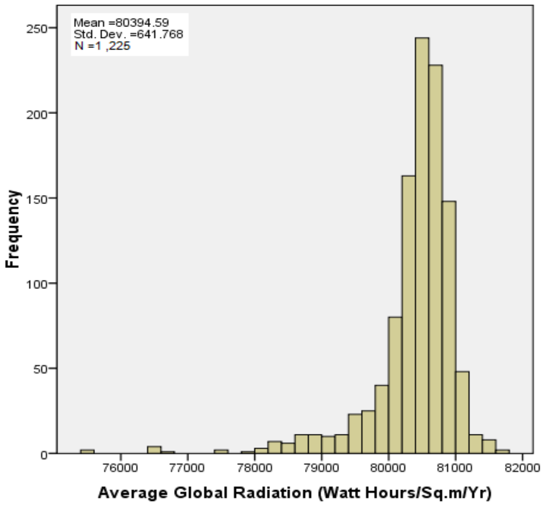


Figure I52. Histogram for Global Radiation #300 of *r.sun*

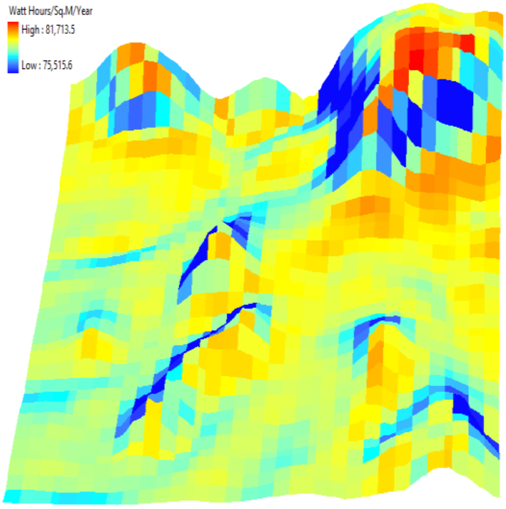


Figure I53. Global Radiation #400 of *r.sun*

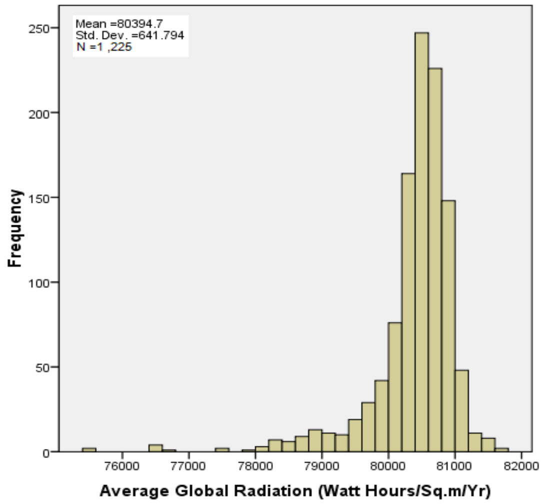


Figure I54. Histogram for Global Radiation #400 of *r.sun*

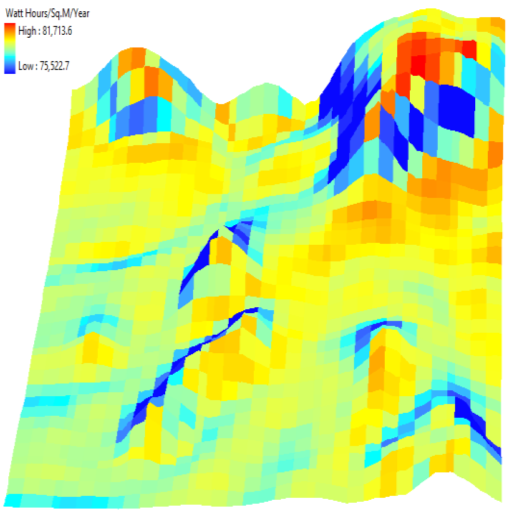


Figure I55. Global Radiation #500 of *r.sun*

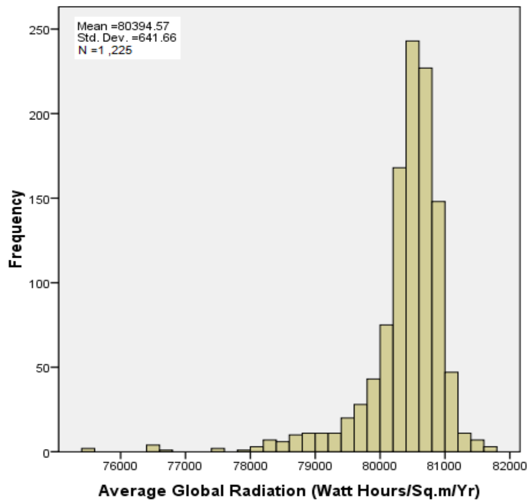


Figure I56. Histogram for Global Radiation #500 of *r.sun*

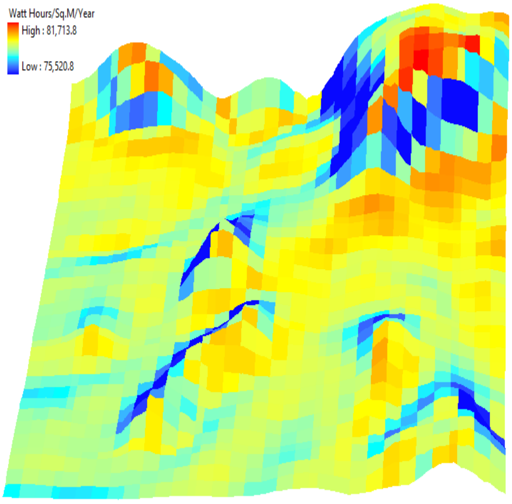


Figure I57. Global Radiation #600 of *r.sun*

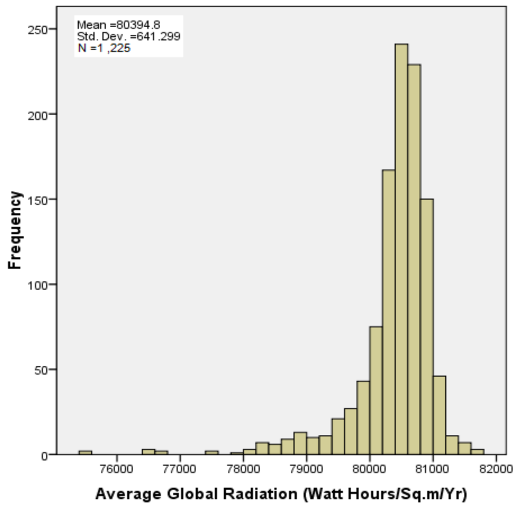


Figure I58. Histogram for Global Radiation #600 of *r.sun*

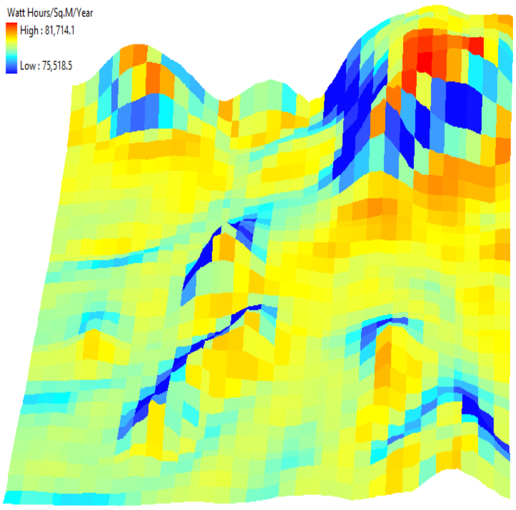


Figure I59. Global Radiation #700 of *r.sun*

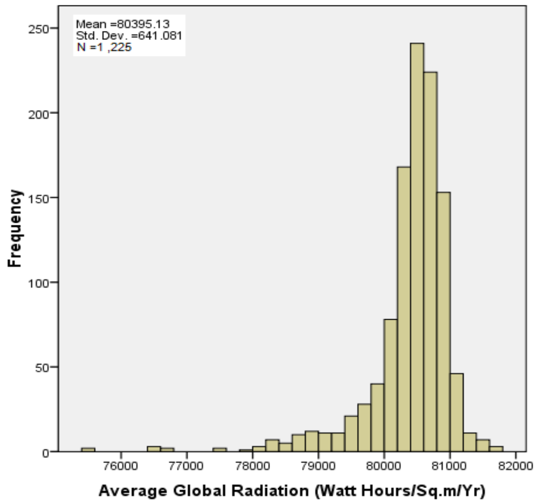


Figure I60. Histogram for Global Radiation #700 of *r.sun*

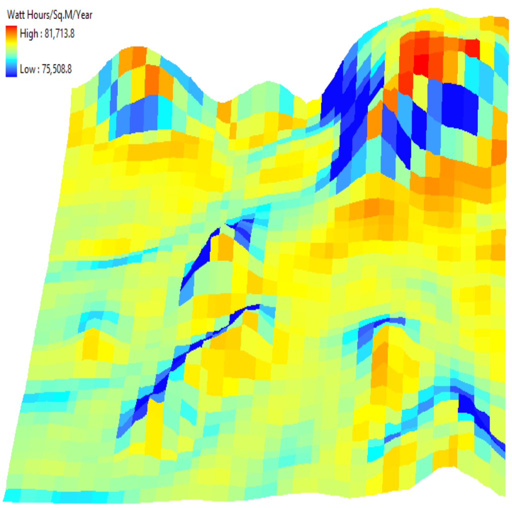


Figure I61. Global Radiation #800 of *r.sun*

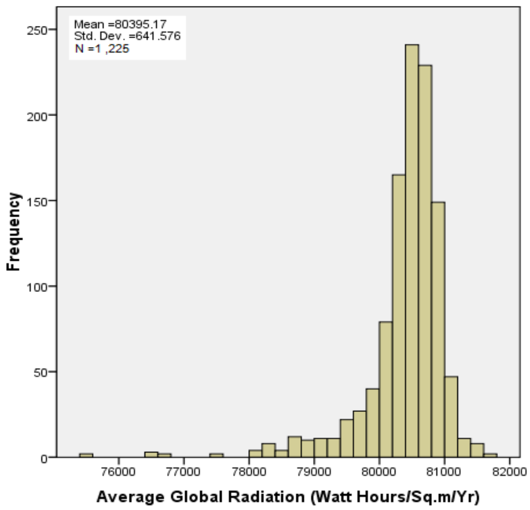


Figure I62. Histogram for Global Radiation #800 of *r.sun*

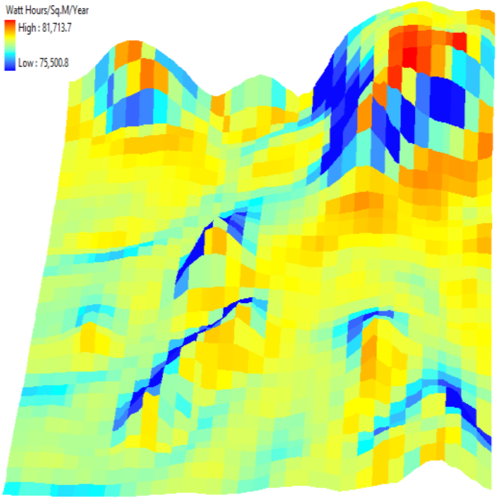


Figure I63. Global Radiation #900 of *r.sun*

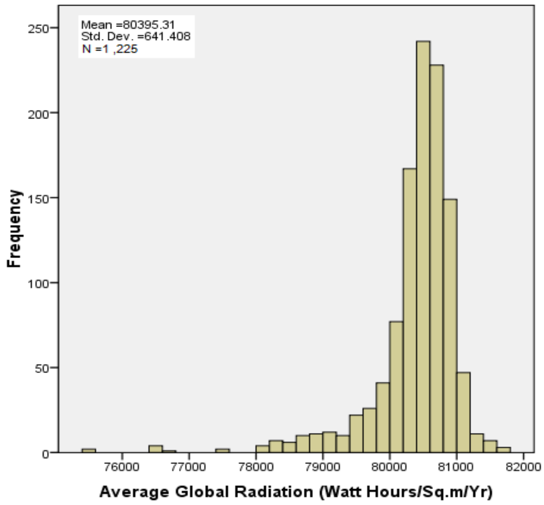


Figure I64. Histogram for Global Radiation #900 of *r.sun*

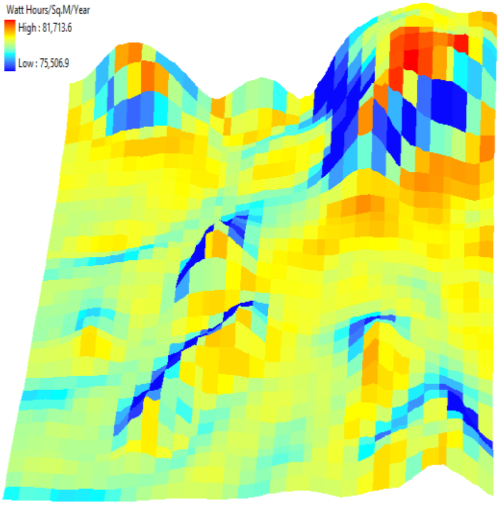


Figure I65. Global Radiation #1000 of *r.sun*

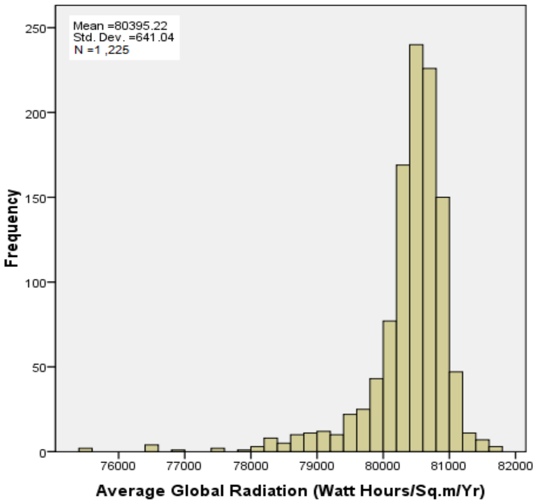


Figure I66. Histogram for Global Radiation #1000 of *r.sun*

Direct Duration

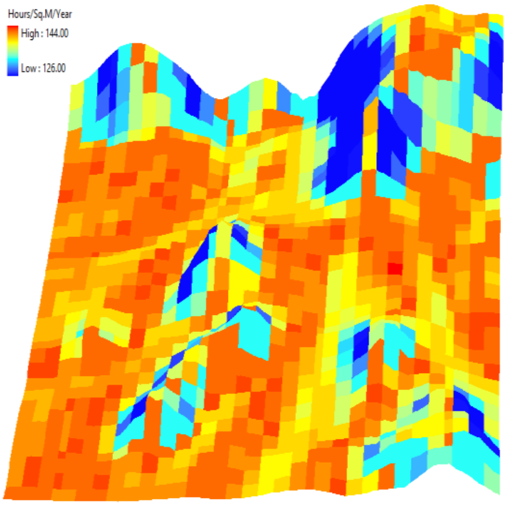


Figure I67. Direct Duration #1 of *r.sun*

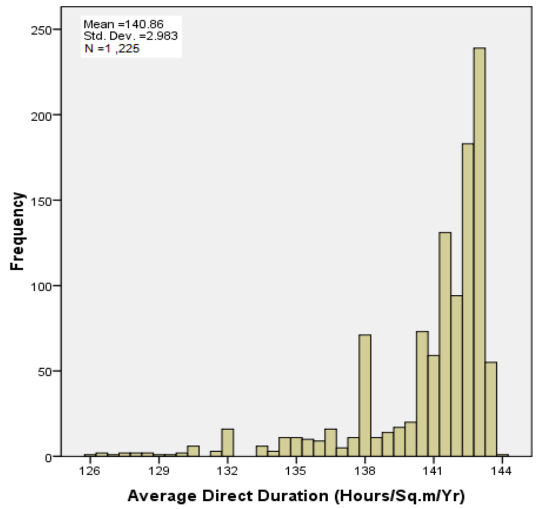


Figure I68. Histogram for Direct Duration #1 of *r.sun*

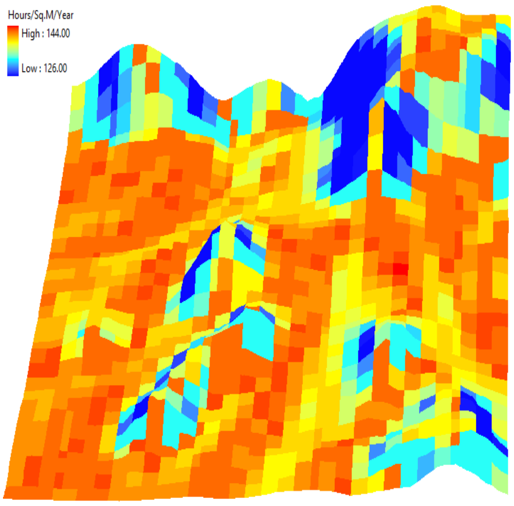


Figure I69. Direct Duration #100 of *r.sun*

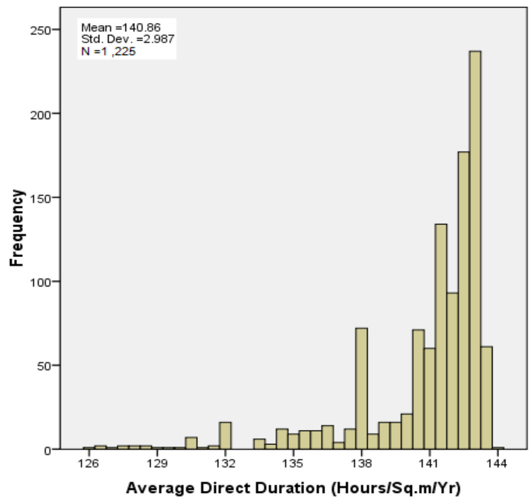


Figure I70. Histogram for Direct Duration #100 of *r.sun*

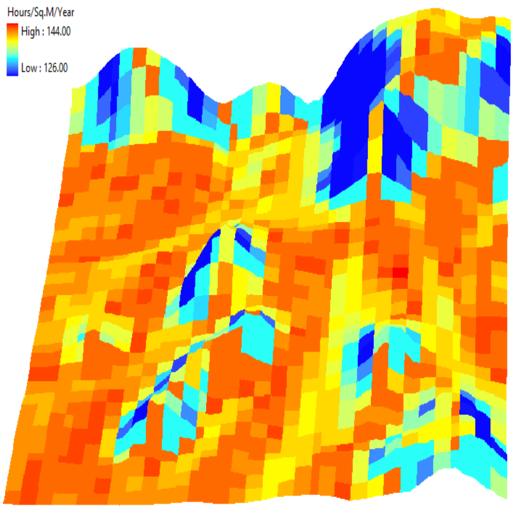


Figure I71. Direct Duration #200 of *r.sun*

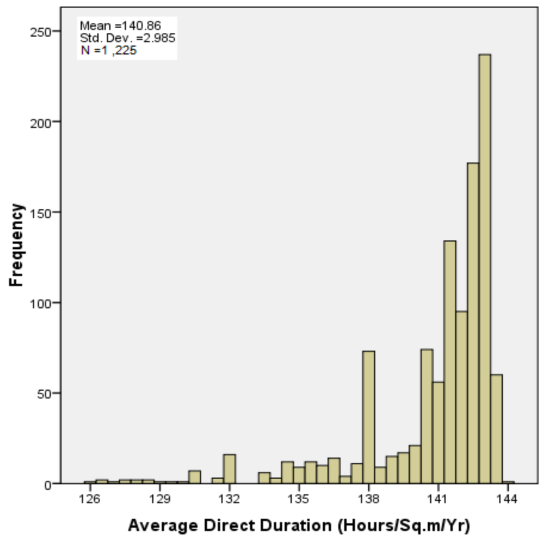


Figure I72. Histogram for Direct Duration #200 of *r.sun*

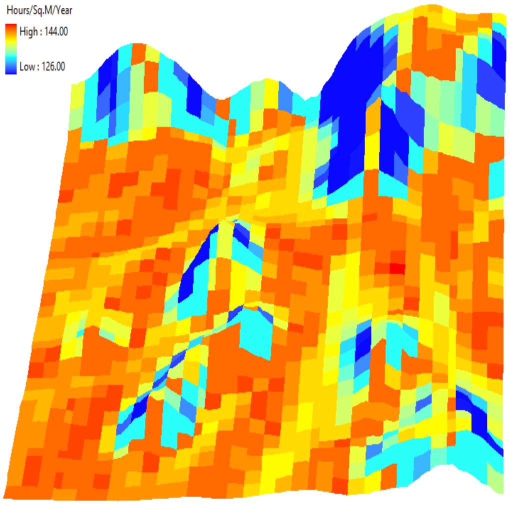


Figure I73. Direct Duration #300 of *r.sun*

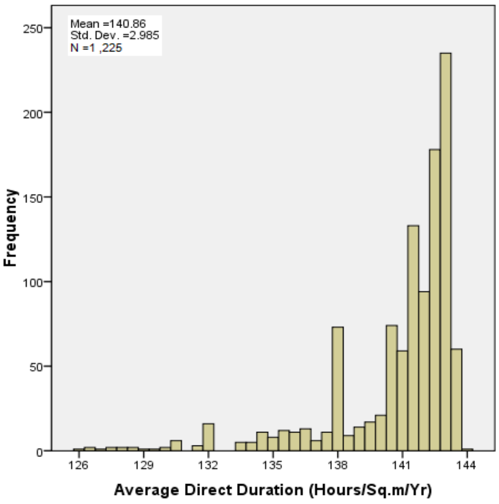


Figure I74. Histogram for Direct Duration #300 of *r.sun*

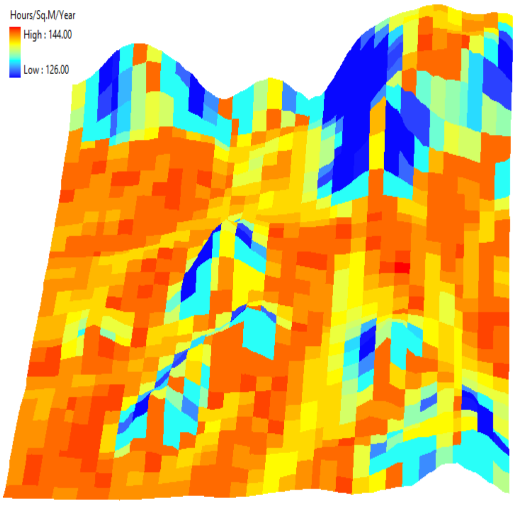


Figure I75. Direct Duration #400 of *r.sun*

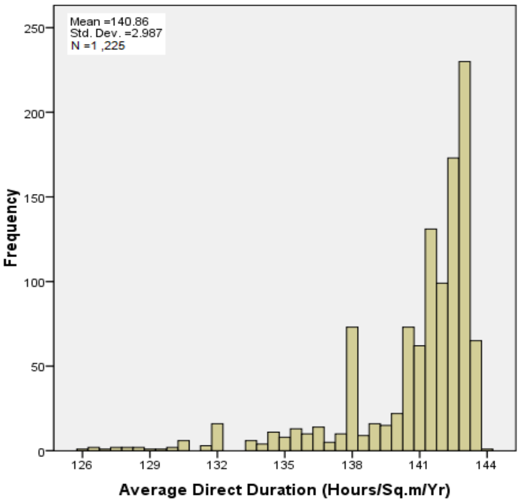


Figure I76. Histogram for Direct Duration #400 of *r.sun*

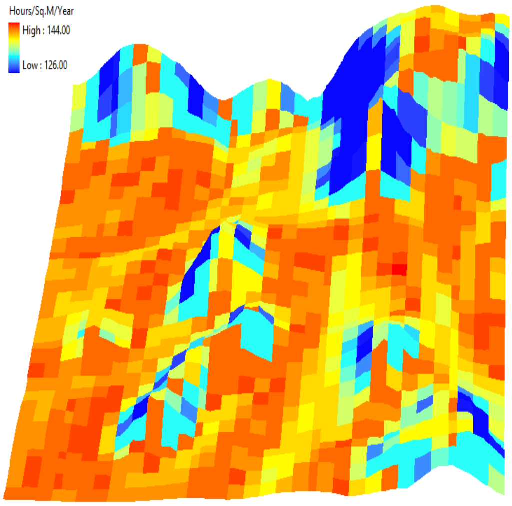


Figure I77. Direct Duration #500 of *r.sun*

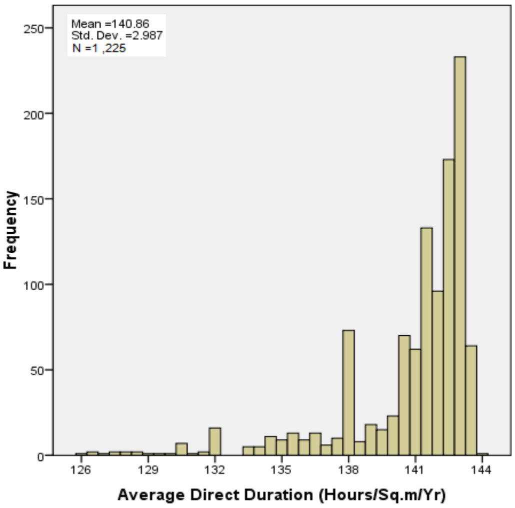


Figure I78. Histogram for Direct Duration #500 of *r.sun*

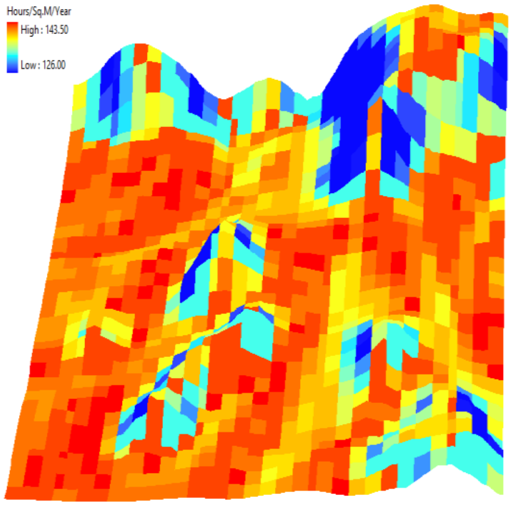


Figure I79. Direct Duration #600 of *r.sun*

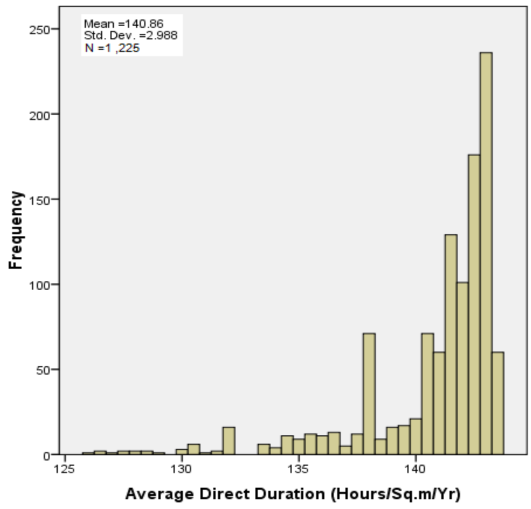


Figure I80. Histogram for Direct Duration #600 of *r.sun*

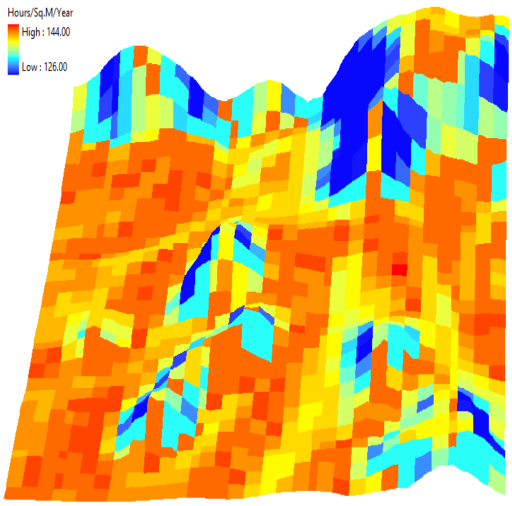


Figure I81. Direct Duration #700 of *r.sun*

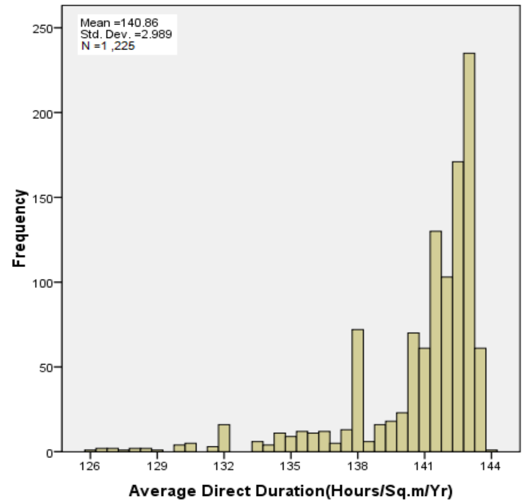


Figure I82. Histogram for Direct Duration #700 of *r.sun*

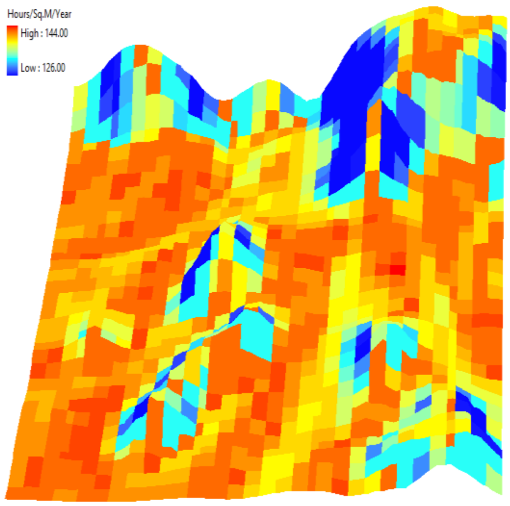


Figure I83. Direct Duration #800 of *r.sun*

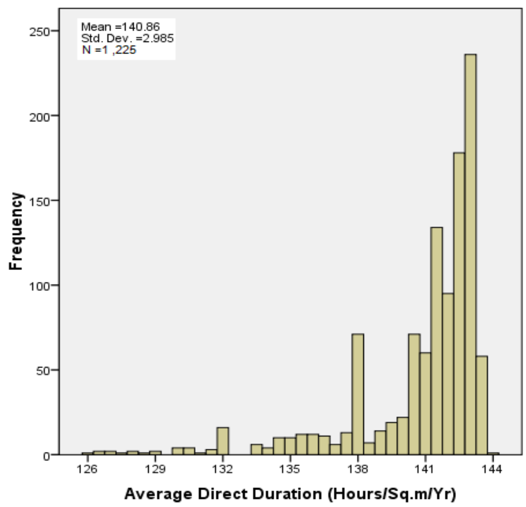


Figure I84. Histogram for Direct Duration #800 of *r.sun*

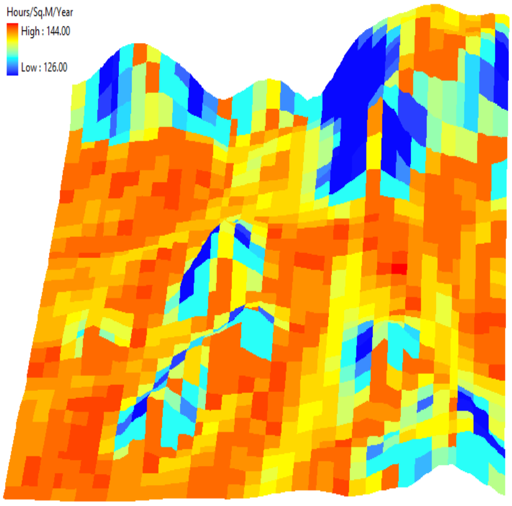


Figure I85. Direct Duration #900 of *r.sun*

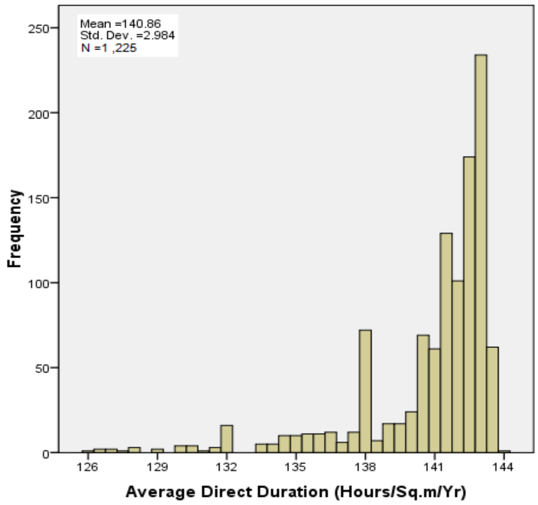


Figure I86. Histogram for Direct Duration #900 of *r.sun*

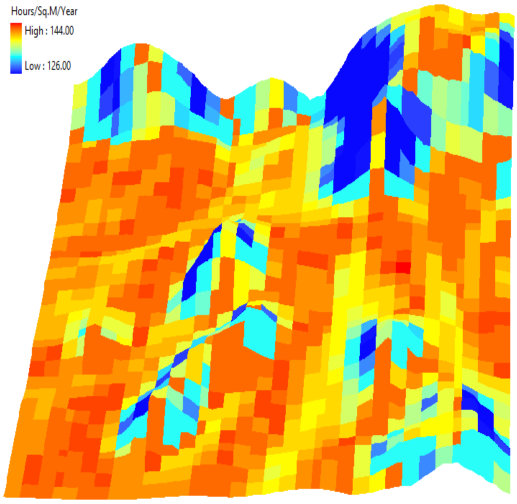


Figure I87. Direct Duration #1000 of *r.sun*

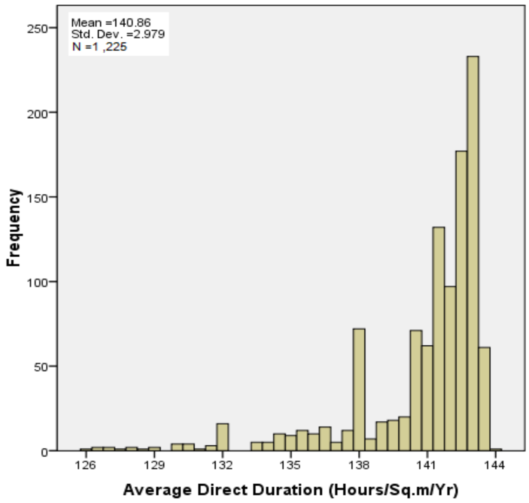


Figure I88. Histogram for Direct Duration #1000 of *r.sun*

APPENDIX J

SLOPE MAPS GENERATED BY *SPATIAL ANALYST* EXTENSION (ARCGIS) USING THE REALIZED DEMS

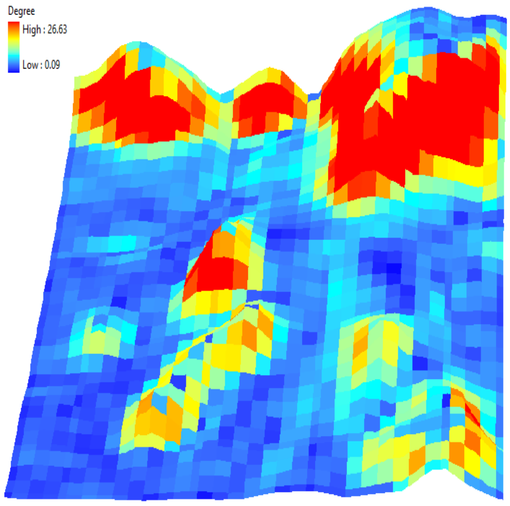


Figure J1. Slope #1 using *Spatial Analyst*

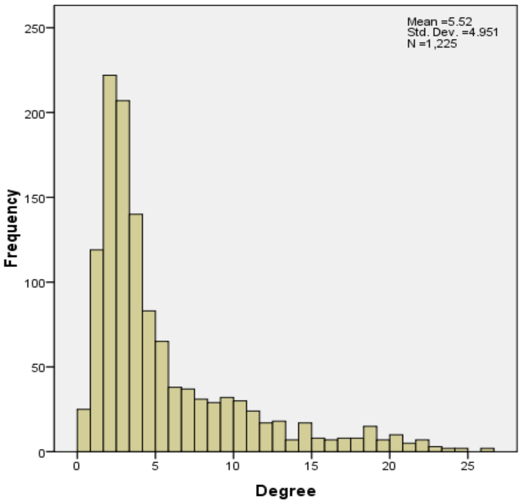


Figure J2. Histogram for Slope #1 using *Spatial Analyst*

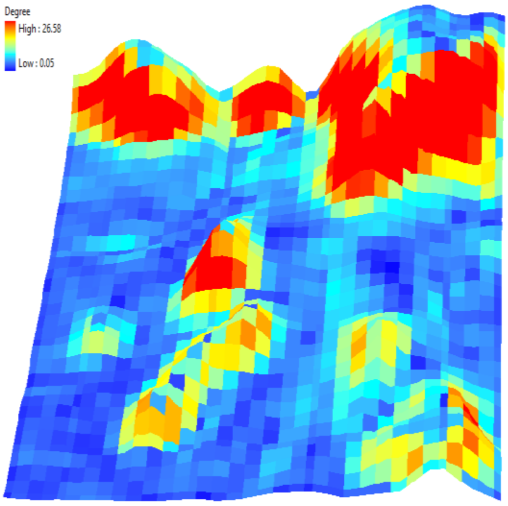


Figure J3. Slope #100 using *Spatial Analyst*

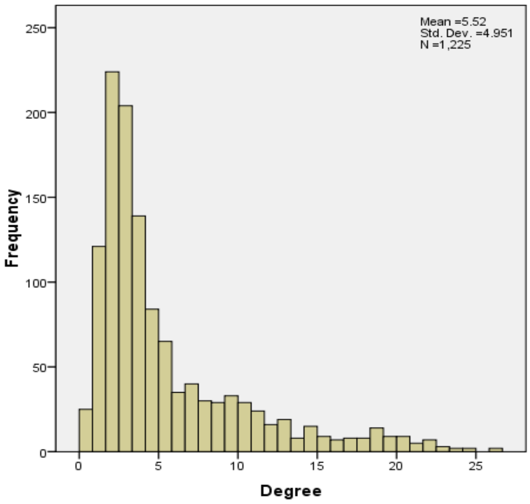


Figure J4. Histogram for Slope #100 using *Spatial Analyst*

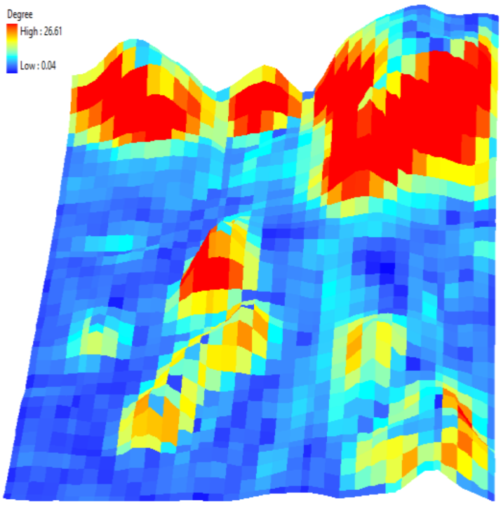


Figure J5. Slope #200 using *Spatial Analyst*

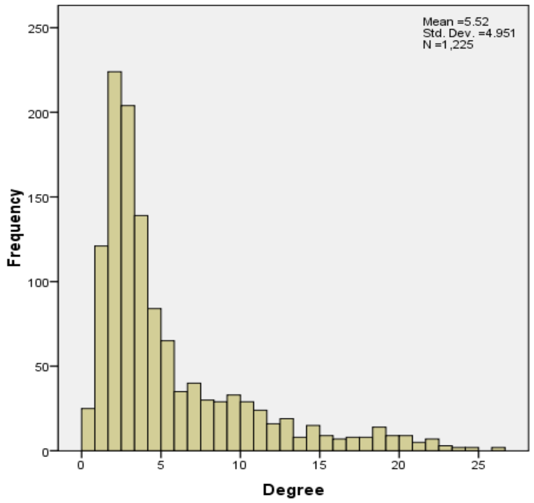


Figure J6. Histogram for Slope #200 using *Spatial Analyst*

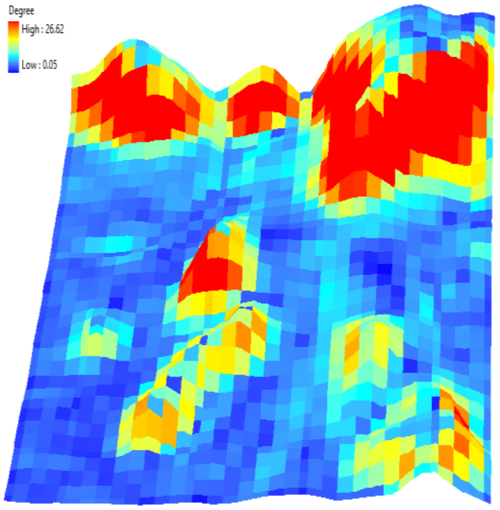


Figure J7. Slope #300 using *Spatial Analyst*

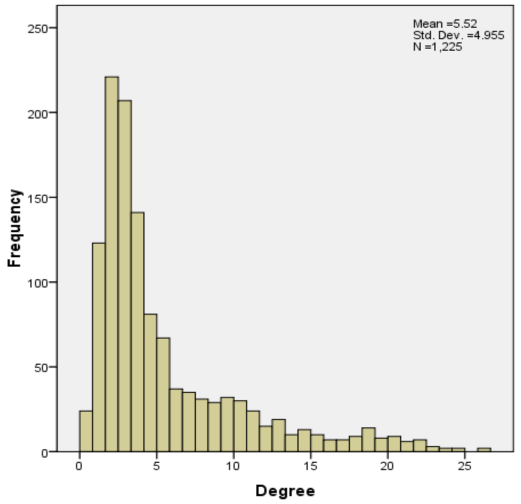


Figure J8. Histogram for Slope #300 using *Spatial Analyst*

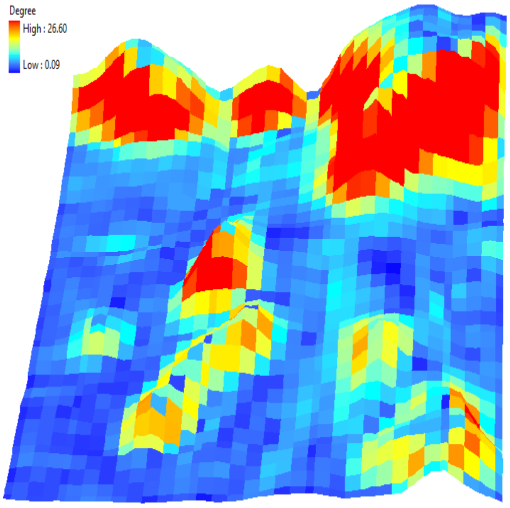


Figure J9. Slope #400 using *Spatial Analyst*

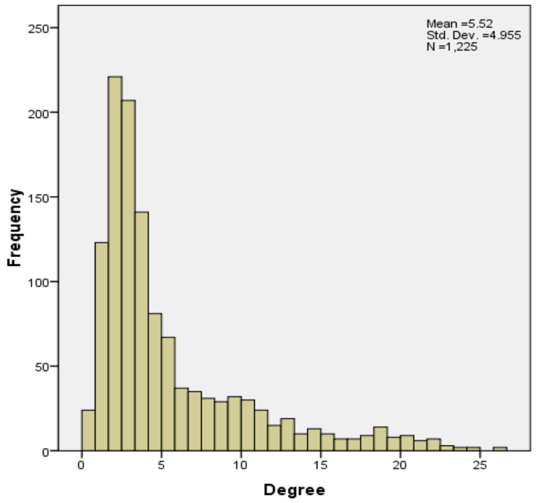


Figure J10. Histogram for Slope #400 using *Spatial Analyst*

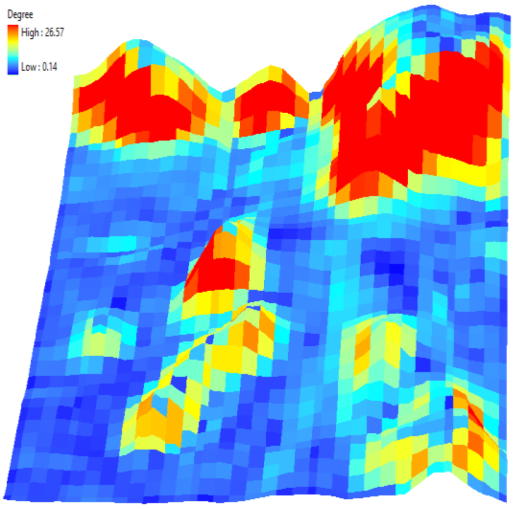


Figure J11. Slope #500 using *Spatial Analyst*

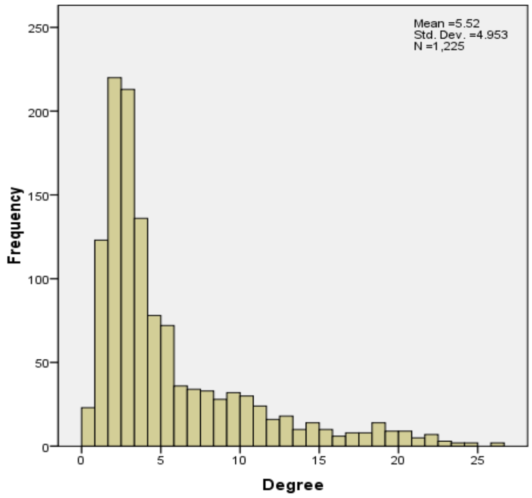


Figure J12. Histogram for Slope #500 using *Spatial Analyst*

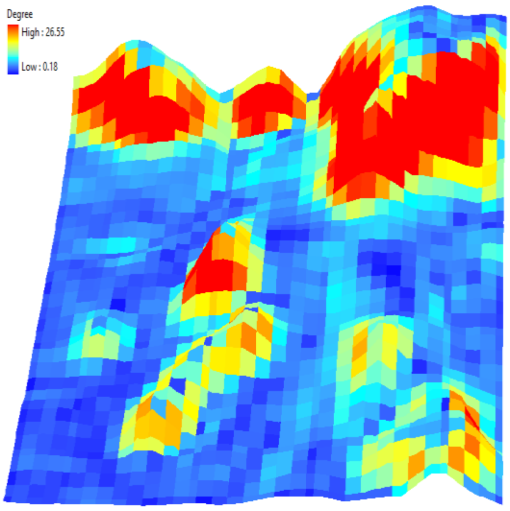


Figure J13. Slope #600 using *Spatial Analyst*

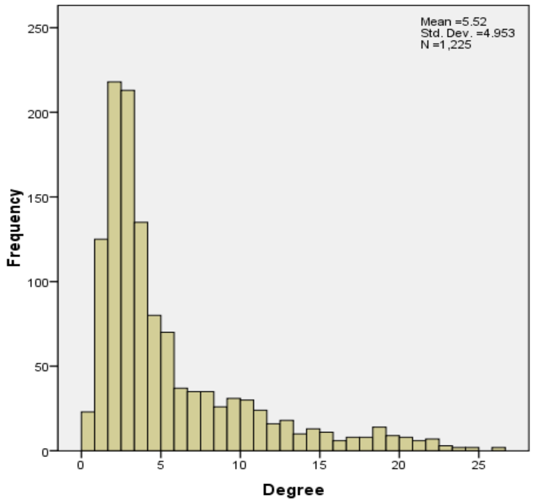


Figure J14. Histogram for Slope #600 using *Spatial Analyst*

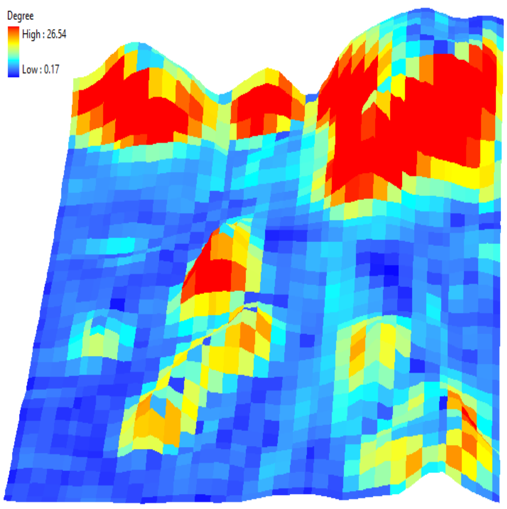


Figure J15. Slope #700 using *Spatial Analyst*

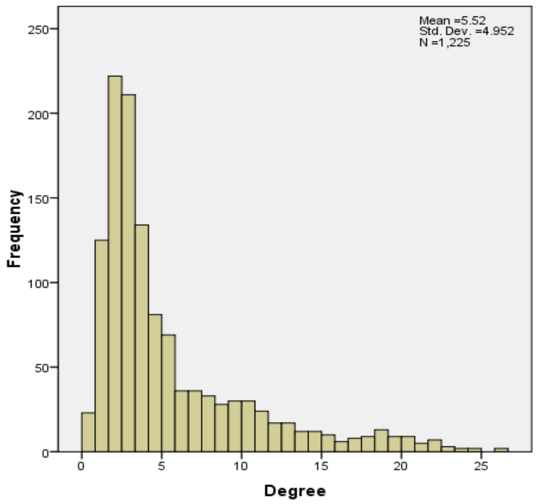


Figure J16. Histogram for Slope #700 using *Spatial Analyst*

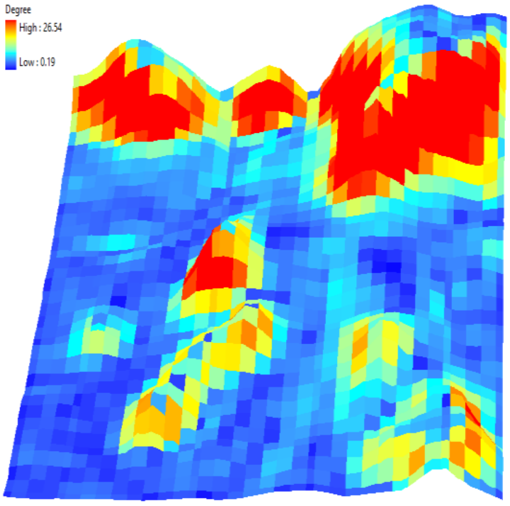


Figure J17. Slope #800 using *Spatial Analyst*

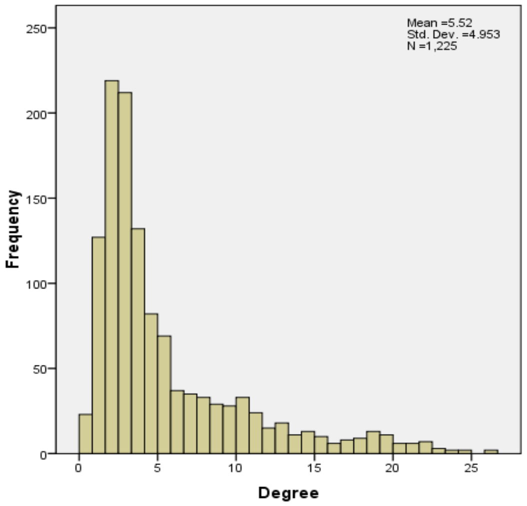


Figure J18. Histogram for Slope #800 using *Spatial Analyst*

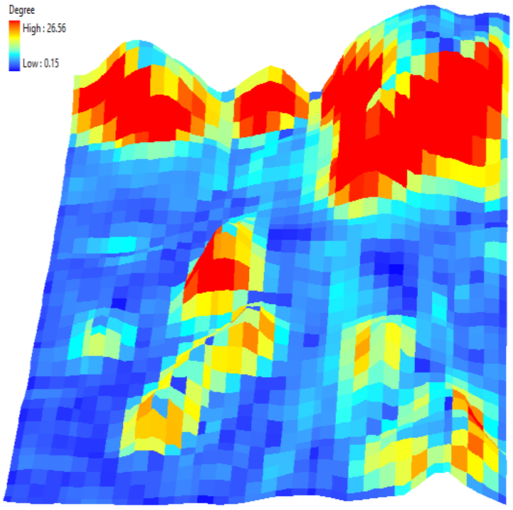


Figure J19. Slope #900 using *Spatial Analyst*

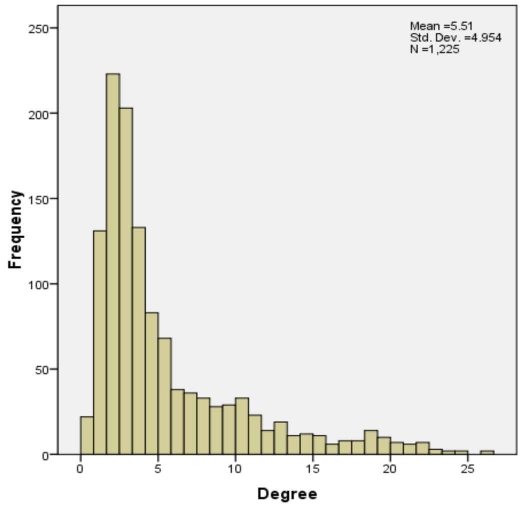


Figure J20. Histogram for Slope #900 using *Spatial Analyst*

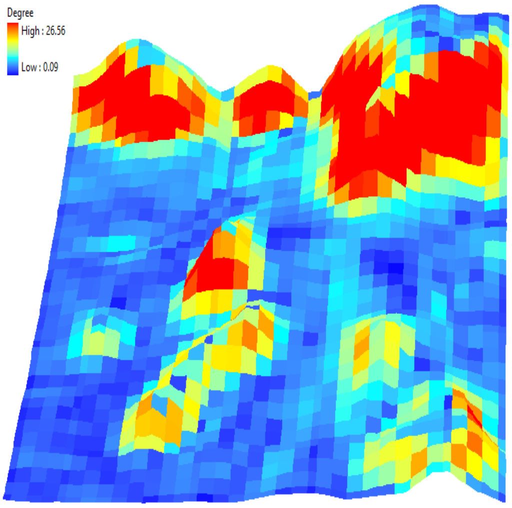


Figure J21. Slope #1000 using *Spatial Analyst*

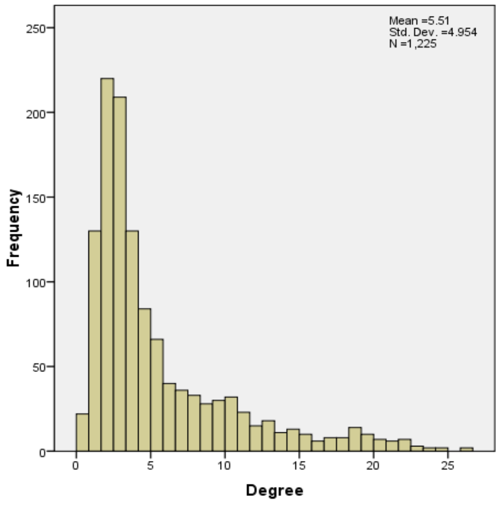


Figure J22. Histogram for Slope #1000 using *Spatial Analyst*

APPENDIX K

ASPECT MAPS GENERATED BY *SPATIAL ANALYST* EXTENSION (ARCGIS) USING THE REALIZED DEMS

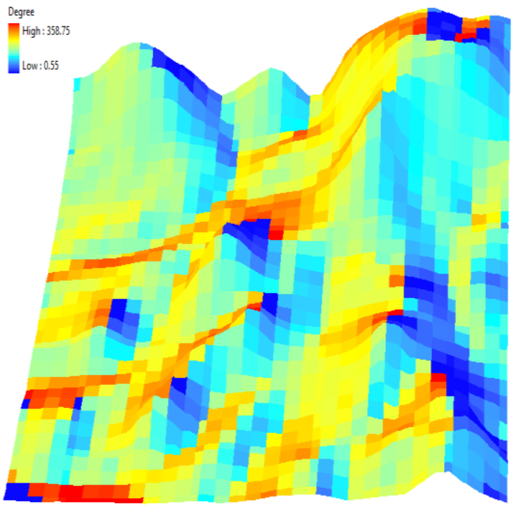


Figure K1. Aspect #1 using *Spatial Analyst*

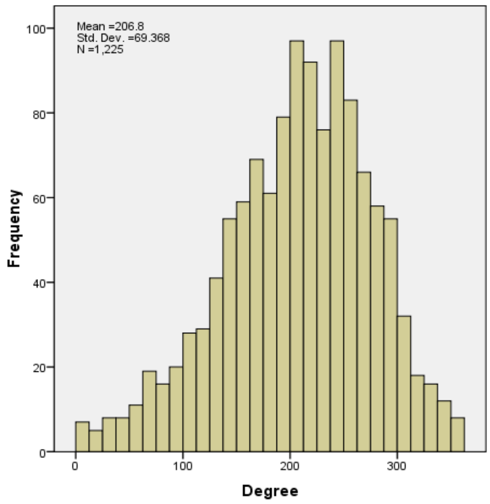


Figure K2. Histogram for Aspect #1 using *Spatial Analyst*

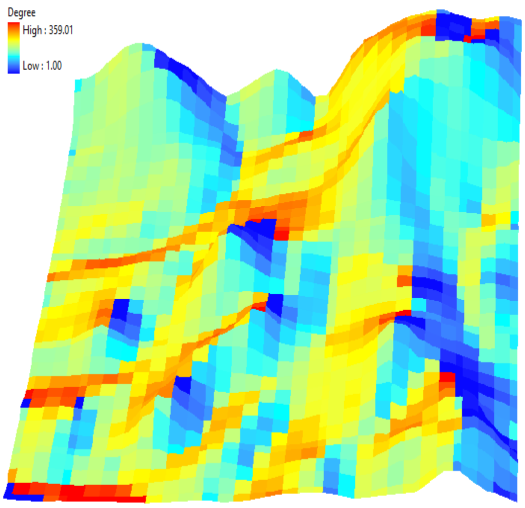


Figure K3. Aspect #100 using *Spatial Analyst*

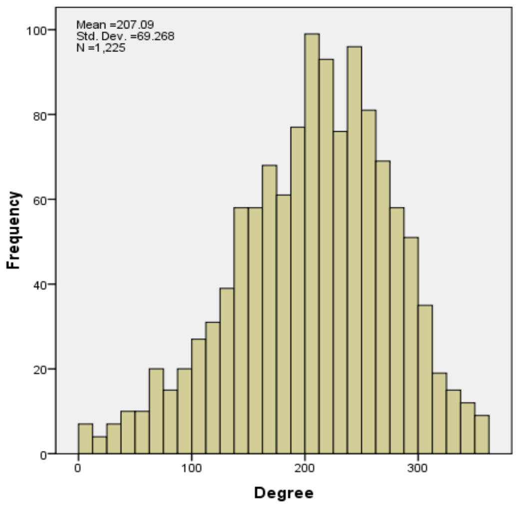


Figure K4. Histogram for Aspect #100 using *Spatial Analyst*

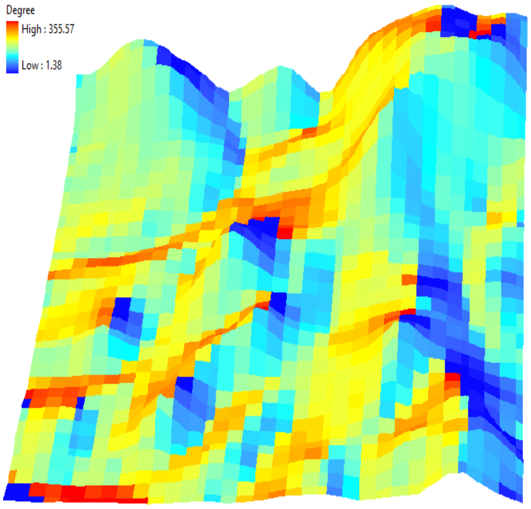


Figure K5. Aspect #200 using *Spatial Analyst*

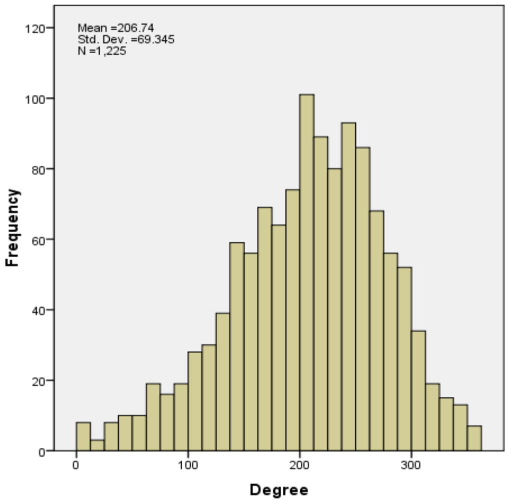


Figure K6. Histogram for Aspect #200 using *Spatial Analyst*

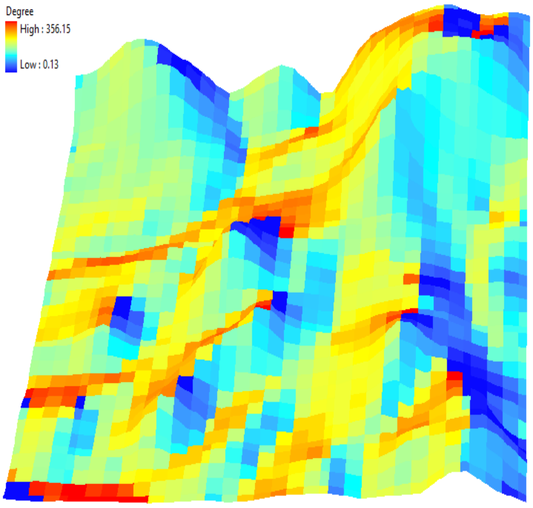


Figure K7. Aspect #300 using *Spatial Analyst*

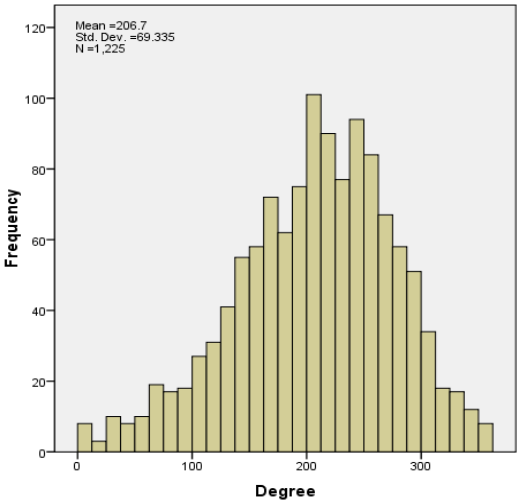


Figure K8. Histogram for Aspect #300 using *Spatial Analyst*

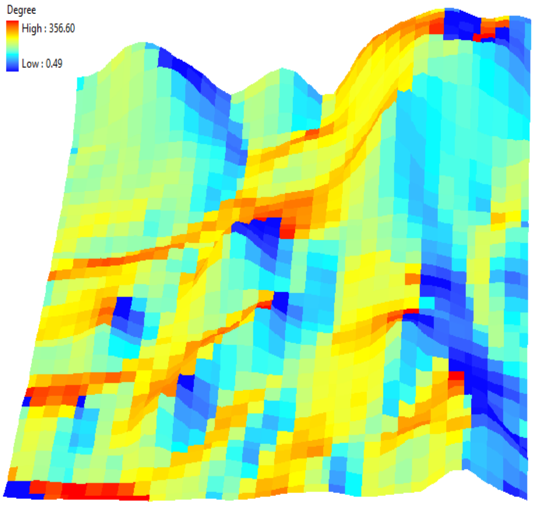


Figure K9. Aspect #400 using *Spatial Analyst*

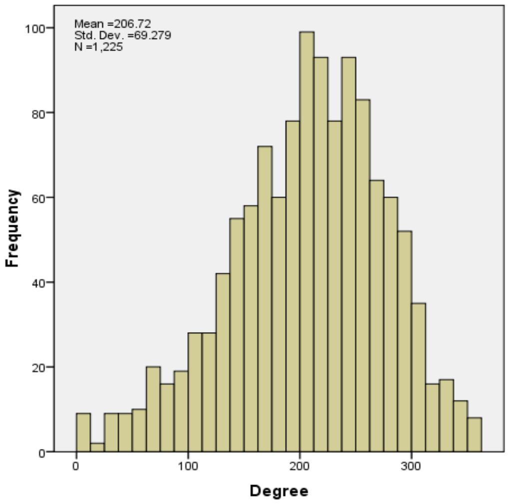


Figure K10. Histogram for Aspect #400 using *Spatial Analyst*

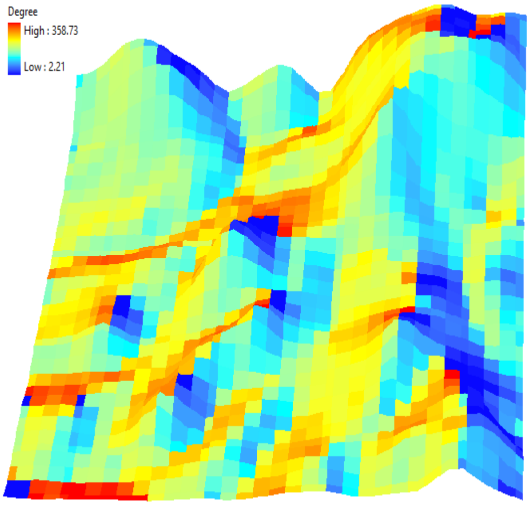


Figure K11. Aspect #500 using *Spatial Analyst*

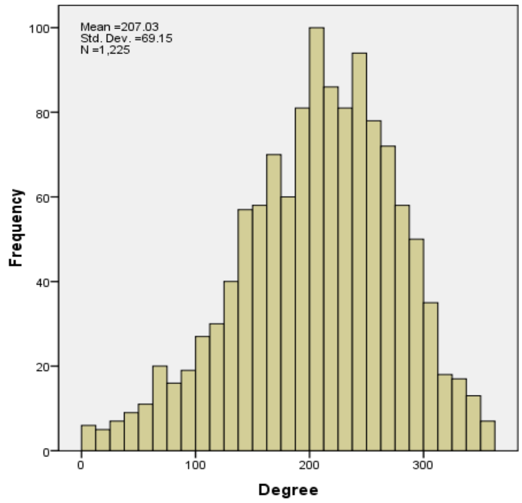


Figure K12. Histogram for Aspect #500 using *Spatial Analyst*

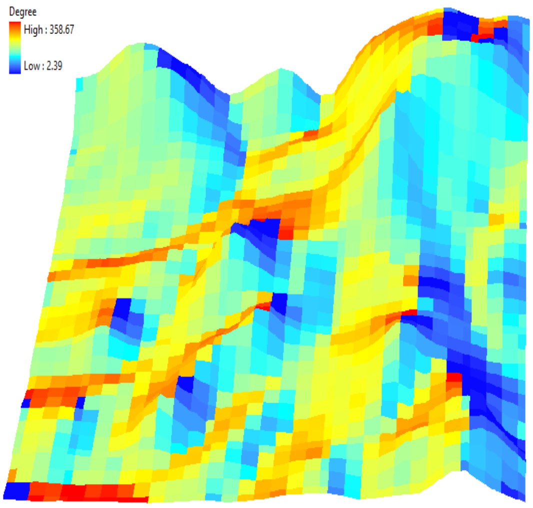


Figure K13. Aspect #600 using *Spatial Analyst*

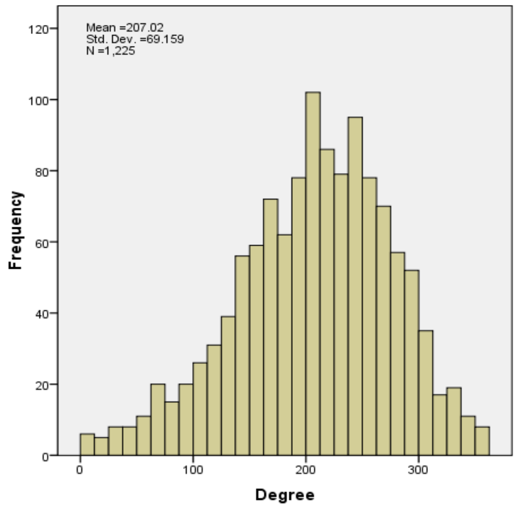


Figure K14. Histogram for Aspect #600 using *Spatial Analyst*

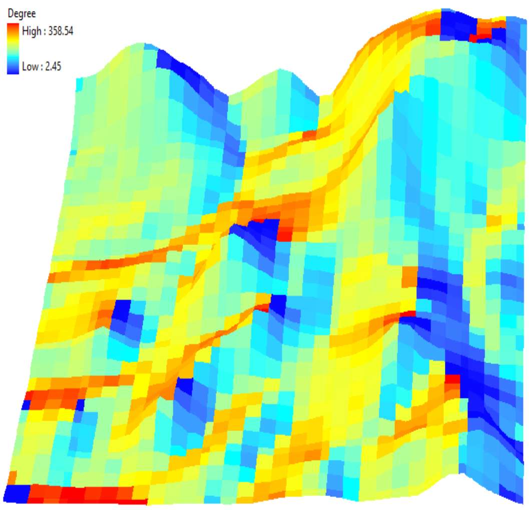


Figure K15. Aspect #700 using *Spatial Analyst*

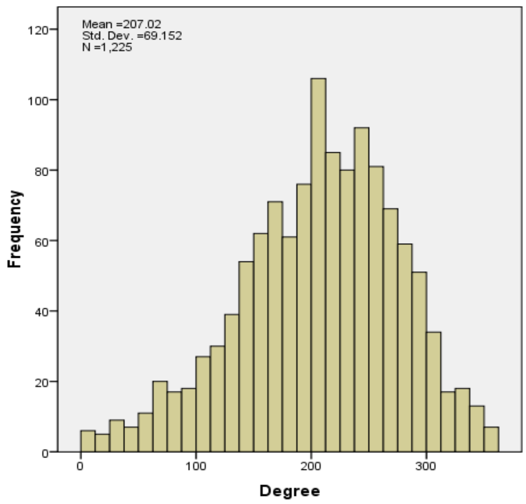
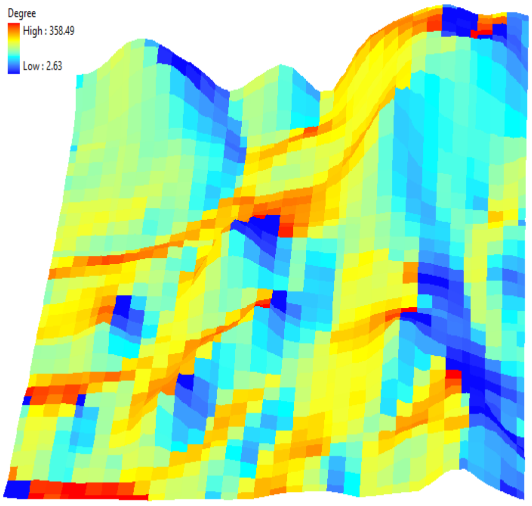


Figure K16. Histogram for Aspect #700 using *Spatial Analyst*



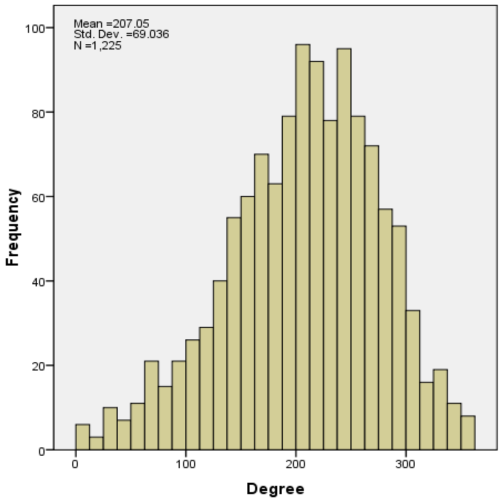
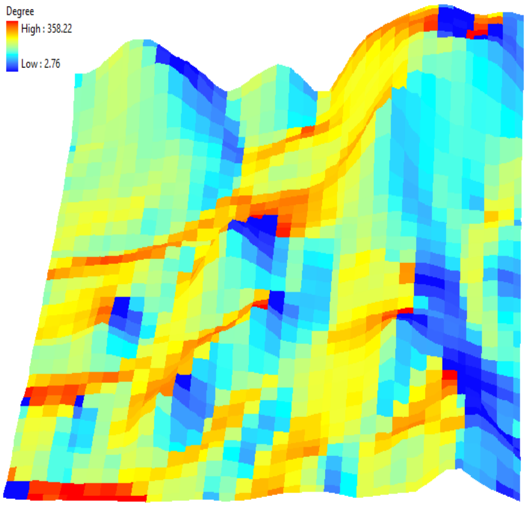


Figure K17. Aspect #800 using *Spatial Analyst*

Figure K18. Histogram for Aspect #800 using *Spatial Analyst*



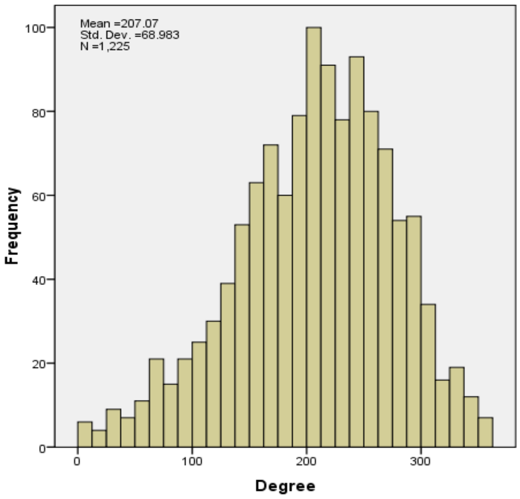
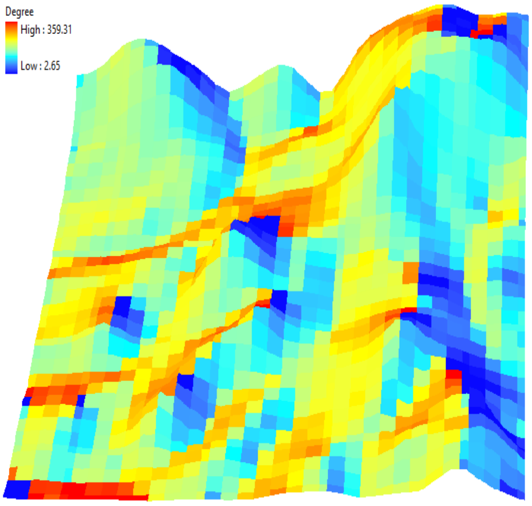


Figure K19. Aspect #900 using *Spatial Analyst*

Figure K20. Histogram for Aspect #900 using *Spatial Analyst*



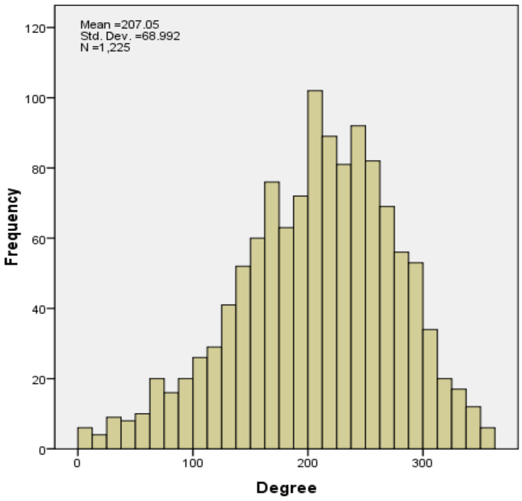


Figure K21. Aspect #1000 using *Spatial Analyst*

Figure K22. Histogram for Aspect #1000 using *Spatial Analyst*

APPENDIX L

TESTS OF NORMALITY FOR THE *SOLAR ANALYST* AND *R.SUN* OUTPUTS

I. Tests of Normality for the Solar Analyst outputs

A. Direct radiation

i. Mean

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
dir_mean_sa	.113	1225	.000	.893	1225	.000

a. Lilliefors Significance Correction

ii. Standard Deviation

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
dir_std_sa	.155	1225	.000	.789	1225	.000

a. Lilliefors Significance Correction

iii. Coefficient of Variaiton

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
dir_cov_sa	.152	1225	.000	.792	1225	.000

a. Lilliefors Significance Correction

B. Diffuse radiation

i. Mean

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
dif_mean_sa	.117	1225	.000	.909	1225	.000

a. Lilliefors Significance Correction

ii. Standard Deviation

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
dif_std_sa	.115	1225	.000	.826	1225	.000

a. Lilliefors Significance Correction

iii. Coefficient of Variaiton

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
dif_cov_sa	.117	1225	.000	.826	1225	.000

a. Lilliefors Significance Correction

C. Global radiation

i. Mean

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
glo_mean_sa	.111	1225	.000	.904	1225	.000

a. Lilliefors Significance Correction

ii. Standard Deviation

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
glo_std_sa	.148	1225	.000	.805	1225	.000

a. Lilliefors Significance Correction

iii. Coefficient of Variaiton

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
glo_cov_sa	.148	1225	.000	.808	1225	.000

a. Lilliefors Significance Correction

D. Direct duration

i. Mean

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
ddur_mean_sa	.128	1225	.000	.887	1225	.000

a. Lilliefors Significance Correction

ii. Standard Deviation

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
ddur_std_sa	.208	1225	.000	.704	1225	.000

a. Lilliefors Significance Correction

iii. <u>*Coefficient of Variaiton*</u>

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
ddur_cov_sa	.247	1225	.000	.693	1225	.000

a. Lilliefors Significance Correction

II. Tests of Normality for the*r.sun* outputs

A. Direct radiation

i. Mean

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
dir_mean_rsun	.193	1225	.000	.745	1225	.000

a. Lilliefors Significance Correction

ii. Standard Deviation

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
dir_std_rsun	.257	1225	.000	.704	1225	.000

a. Lilliefors Significance Correction

iii. Coefficient of Variaiton

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
dir_cov_rsun	.162	1225	.000	.863	1225	.000

a. Lilliefors Significance Correction

B. Diffuse radiation

i. Mean

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
dif_mean_rsun	.158	1225	.000	.775	1225	.000

a. Lilliefors Significance Correction

ii. Standard Deviation

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
dif_std_rsun	.268	1225	.000	.665	1225	.000

a. Lilliefors Significance Correction

iii. Coefficient of Variaiton

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
dif_cov_rsun	.186	1225	.000	.869	1225	.000

a. Lilliefors Significance Correction

C. Global radiation

i. Mean

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
glo_mean_rsun	.170	1225	.000	.772	1225	.000

a. Lilliefors Significance Correction

ii. Standard Deviation

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
glo_std_rsun	.185	1225	.000	.726	1225	.000

a. Lilliefors Significance Correction

iii. Coefficient of Variaiton

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
glo_cov_rsun	.281	1225	.000	.346	1225	.000

a. Lilliefors Significance Correction

D. Direct duration

i. Mean

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
ddur_mean_rsun	.218	1225	.000	.755	1225	.000

a. Lilliefors Significance Correction

ii. *Standard Deviation*

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
ddur_std_rsun	.175	1225	.000	.702	1225	.000

a. Lilliefors Significance Correction

iii. *Coefficient of Variaiton*

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
ddur_cov_rsun	.146	1225	.000	.705	1225	.000

a. Lilliefors Significance Correction

APPENDIX M

MANN-WHITNEY U TESTS FOR THE *SOLAR ANALYST* AND *R.SUN* OUTPUT

DIRECT RADIATION

a. Mean

Step 1: Hypotheses

H_a: The mean direct radiation of realized DEMs of *Solar Analyst* and *r.sun* models are statistically the same.

H_b: The mean direct radiation of realized DEMs of *Solar Analyst* and *r.sun* models are statistically different.

Step 2: Significance Level

$\alpha = 0.05$

Step 3: Rejection Region

Reject the null hypothesis if p-value ≤ 0.05

Step 4: Test Statistic

Test Statistics[a]

	dir_mean_sa_rsun
Mann-Whitney U	34.000
Wilcoxon W	750959.000
Z	-42.855
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: group

p-value = Asymp. Sig. (2-tailed) = 0.000

Step 5: Decision

Since p-value = 0.000 < 0.05 = α, we shall reject the null hypothesis.

Step 6: State conclusion in words

At the α = 0.05 level of significance, there is enough evidence to conclude that there is a difference in the mean direct radiation of realized DEMs of *Solar Analyst* and *r.sun* models.

b. Standard Deviation

Step 1: Hypotheses

H_c: The standard deviation for mean direct radiation of realized DEMs of *Solar Analyst* and *r.sun* models are statistically the same.

H_d: The standard deviation for mean direct radiation of realized DEMs of *Solar Analyst* and *r.sun* models are statistically different.

Step 2: Significance Level

$\alpha = 0.05$

Step 3: Rejection Region

Reject the null hypothesis if *p*-value ≤ 0.05

Step 4: Test Statistic

Test Statistics[a]

	dir_std_ sa_rsun
Mann-Whitney U	.000
Wilcoxon W	750925.0
Z	-42.857
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: group

p-value = Asymp. Sig. (2-tailed) = 0.000

Step 5: Decision

Since *p*-value = 0.000 < 0.05 = α, we shall reject the null hypothesis.

Step 6: State conclusion in words

At the α = 0.05 level of significance, there is enough evidence to conclude that there is a difference in the standard deviation for mean direct radiation of realized DEMs of *Solar Analyst* and *r.sun* models.

c. Coefficient of Variaiton

Step 1: Hypotheses

H_e: The coefficient of variation of mean direct radiation for realized DEMs of *Solar Analyst* and *r.sun* models are statistically the same.

H_f: The coefficient of variation of mean direct radiation for realized DEMs of *Solar Analyst* and *r.sun* models are statistically different.

Step 2: Significance Level

$\alpha = 0.05$

Step 3: Rejection Region

Reject the null hypothesis if *p*-value ≤ 0.05

Step 4: Test Statistic

Test Statistics[a]

	dir_cov_sa_rsun
Mann-Whitney U	.000
Wilcoxon W	750925.0
Z	-42.857
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: group

p-value = Asymp. Sig. (2-tailed) = 0.000

Step 5: Decision

Since *p*-value = 0.000 < 0.05 = α, we shall reject the null hypothesis.

Step 6: State conclusion in words

At the α = 0.05 level of significance, there is enough evidence to conclude that there is a difference in the coefficient of variation of mean direct radiation for realized DEMs of *Solar Analyst* and *r.sun* models.

DIFFUSE RADIATION

a. Mean

Step 1: Hypotheses

H_g: The mean diffuse radiation of realized DEMs of *Solar Analyst* and *r.sun* models are statistically the same.

H_h: The mean diffuse radiation of realized DEMs of *Solar Analyst* and *r.sun* models are statistically different.

Step 2: Significance Level

$\alpha = 0.05$

Step 3: Rejection Region

Reject the null hypothesis if *p*-value ≤ 0.05

Step 4: Test Statistic

Test Statistics[a]

	dif_mean_sa_rsun
Mann-Whitney U	.000
Wilcoxon W	750925.000
Z	-42.857
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: group

p-value = Asymp. Sig. (2-tailed) = 0.000

Step 5: Decision

Since p-value = 0.000 < 0.05 = α, we shall reject the null hypothesis.

Step 6: State conclusion in words

At the α = 0.05 level of significance, there is enough evidence to conclude that there is a difference in the mean diffuse radiation of realized DEMs of *Solar Analyst* and *r.sun* models.

b. Standard Deviation

Step 1: Hypotheses

H_i: The standard deviation for mean diffuse radiation of realized DEMs of *Solar Analyst* and *r.sun* models are statistically the same.

H_j: The standard deviation for mean diffuse radiation of realized DEMs of *Solar Analyst* and *r.sun* models are statistically different.

Step 2: Significance Level

$\alpha = 0.05$

Step 3: Rejection Region

Reject the null hypothesis if p-value ≤ 0.05

Step 4: Test Statistic

Test Statistics[a]

	dif_std_sa_rsun
Mann-Whitney U	.000
Wilcoxon W	750925.0
Z	-42.857
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: group

p-value = Asymp. Sig. (2-tailed) = 0.000

Step 5: Decision

Since *p*-value = 0.000 < 0.05 = α, we shall reject the null hypothesis.

Step 6: State conclusion in words

At the α = 0.05 level of significance, there is enough evidence to conclude that there is a difference in the standard deviation for mean diffuse radiation of realized DEMs of *Solar Analyst* and *r.sun* models.

c. Coefficient of Variaiton

Step 1: Hypotheses

H_k: The coefficient of variation of mean diffuse radiation for realized DEMs of *Solar Analyst* and *r.sun* models are statistically the same.

H_l: The coefficient of variation of mean diffuse radiation for realized DEMs of *Solar Analyst* and *r.sun* models are statistically different.

Step 2: Significance Level

$\alpha = 0.05$

Step 3: Rejection Region

Reject the null hypothesis if *p*-value ≤ 0.05

Step 4: Test Statistic

Test Statistics[a]

	dif_cov_sa_rsun
Mann-Whitney U	.000
Wilcoxon W	750925.0
Z	-42.861
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: group

p-value = Asymp. Sig. (2-tailed) = 0.000

Step 5: Decision

Since *p*-value = 0.000 < 0.05 = α, we shall reject the null hypothesis.

Step 6: State conclusion in words

At the α = 0.05 level of significance, there is enough evidence to conclude that there is a difference in the coefficient of variation of mean diffuse radiation for realized DEMs of *Solar Analyst* and *r.sun* models.

GLOBAL RADIATION

a. Mean

Step 1: Hypotheses

H_m: The mean global radiation of realized DEMs of *Solar Analyst* and *r.sun* models are statistically the same.

H_n: The mean global radiation of realized DEMs of *Solar Analyst* and *r.sun* models are statistically different.

Step 2: Significance Level

$\alpha = 0.05$

Step 3: Rejection Region

Reject the null hypothesis if *p*-value ≤ 0.05

Step 4: Test Statistic

Test Statistics[a]

	glo_mean_sa_rsun
Mann-Whitney U	.000
Wilcoxon W	750925.000
Z	-42.857
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: group

p-value = Asymp. Sig. (2-tailed) = 0.000

Step 5: Decision

Since *p*-value = 0.000 < 0.05 = α, we shall reject the null hypothesis.

Step 6: State conclusion in words

At the α = 0.05 level of significance, there is enough evidence to conclude that there is a difference in the mean global radiation of realized DEMs of *Solar Analyst* and *r.sun* models.

b. Standard Deviation

Step 1: Hypotheses

H_p: The standard deviation for mean global radiation of realized DEMs of *Solar Analyst* and *r.sun* models are statistically the same.

H_q: The standard deviation for mean global radiation of realized DEMs of *Solar Analyst* and *r.sun* models are statistically different.

Step 2: Significance Level

$\alpha = 0.05$

Step 3: Rejection Region

Reject the null hypothesis if p-value ≤ 0.05

Step 4: Test Statistic

Test Statistics[a]

	glo_std_sa_rsun
Mann-Whitney U	.000
Wilcoxon W	750925.0
Z	-42.857
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: group

p-value = Asymp. Sig. (2-tailed) = 0.000

Step 5: Decision

Since p-value = 0.000 < 0.05 = α, we shall reject the null hypothesis.

Step 6: State conclusion in words

At the α = 0.05 level of significance, there is enough evidence to conclude that there is a difference in the standard deviation for mean global radiation of realized DEMs of *Solar Analyst* and *r.sun* models.

c. Coefficient of Variaiton

Step 1: Hypotheses

H_r: The coefficient of variation of mean global radiation for realized DEMs of *Solar Analyst* and *r.sun* models are statistically the same.

H_s: The coefficient of variation of mean global radiation for realized DEMs of *Solar Analyst* and *r.sun* models are statistically different.

Step 2: Significance Level

$\alpha = 0.05$

Step 3: Rejection Region

Reject the null hypothesis if p-value ≤ 0.05

Step 4: Test Statistic

Test Statistics[a]

	glo_cov_ sa_rsun
Mann-Whitney U	.000
Wilcoxon W	750925.0
Z	-42.879
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: group

p-value = Asymp. Sig. (2-tailed) = 0.000

Step 5: Decision

Since p-value = 0.000 < 0.05 = α, we shall reject the null hypothesis.

Step 6: State conclusion in words

At the α = 0.05 level of significance, there is enough evidence to conclude that there is a difference in the coefficient of variation of mean global radiation for realized DEMs of *Solar Analyst* and *r.sun* models.

DIRECT DURATION

a. Mean

Step 1: Hypotheses

H_u: The mean direct duration of realized DEMs of *Solar Analyst* and *r.sun* models are statistically the same.

H_v: The mean direct duration of realized DEMs of *Solar Analyst* and *r.sun* models are statistically different.

Step 2: Significance Level

$\alpha = 0.05$

Step 3: Rejection Region

Reject the null hypothesis if p-value ≤ 0.05

Step 4: Test Statistic

Test Statistics[a]

	ddur_mean_ sa_rsun
Mann-Whitney U	341723.000
Wilcoxon W	1092648.000
Z	-23.350
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: group

p-value = Asymp. Sig. (2-tailed) = 0.000

Step 5: Decision

Since p-value = 0.000 < 0.05 = α, we shall reject the null hypothesis.

Step 6: State conclusion in words

At the α = 0.05 level of significance, there is enough evidence to conclude that there is a difference in the mean direct duration of realized DEMs of *Solar Analyst*and *r.sun* models.

b. Standard Deviation

Step 1: Hypotheses

H_x: The standard deviation for mean direct duration of realized DEMs of *Solar Analyst* and *r.sun* models are statistically the same.

H_y: The standard deviation for mean direct duration of realized DEMs of *Solar Analyst* and *r.sun* models are statistically different.

Step 2: Significance Level

$\alpha = 0.05$

Step 3: Rejection Region

Reject the null hypothesis if p-value ≤ 0.05

Step 4: Test Statistic

Test Statistics[a]

	ddur_std_sa_rsun
Mann-Whitney U	683640.000
Wilcoxon W	1434565.000
Z	-3.810
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: group

p-value = Asymp. Sig. (2-tailed) = 0.000

Step 5: Decision

Since p-value = 0.000 < 0.05 = α, we shall reject the null hypothesis.

Step 6: State conclusion in words

At the α = 0.05 level of significance, there is enough evidence to conclude that there is a difference in the standard deviation for mean direct duration of realized DEMs of *Solar Analyst* and *r.sun* models.

c. Coefficient of Variaiton

Step 1: Hypotheses

H_{z0}: The coefficient of variation of mean direct duration for realized DEMs of *Solar Analyst* and *r.sun* models are statistically the same.

H_{z1}: The coefficient of variation of mean direct duration for realized DEMs of *Solar Analyst* and *r.sun* models are statistically different.

Step 2: Significance Level

$\alpha = 0.05$

Step 3: Rejection Region

Reject the null hypothesis if *p*-value ≤ 0.05

Step 4: Test Statistic

Test Statistics[a]

	ddur_cov_sa_rsun
Mann-Whitney U	701316.000
Wilcoxon W	1452241.000
Z	-2.799
Asymp. Sig. (2-tailed)	.005

a. Grouping Variable: group

p-value = Asymp. Sig. (2-tailed) = 0.005

Step 5: Decision

Since *p*-value = 0.005 < 0.05 = α, we shall reject the null hypothesis.

Step 6: State conclusion in words

At the α = 0.05 level of significance, there is enough evidence to conclude that there is a difference in the coefficient of variation of mean direct duration for realized DEMs of *Solar Analyst*and *r.sun* models.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name:	Ismaila, Abdur-Rahman Belel
Email:	belelismaila@yahoo.com
Nationality:	Nigerian
Date of Birth:	12th March 1977
Marital Status:	Married
Address:	Department of Urban and Regional Planning, School of Environmental Sciences, Modibbo Adama University of Technology, P.M.B. 2076 Yola, Adamawa State, Nigeria

EDUCATION

2007 – 2013	**Ph.D. -** Geodetic and Geographic Information Technologies, Graduate School of Natural and Applied Sciences, Middle East Technical University, Ankara, Turkey
1999 – 2004	**B.Tech. (URP): 1st Class -** Department of Urban and Regional Planning, School of Environmental Studies, Modibbo Adama University of Technology, Yola, Nigeria
1997 – 1999	**National Diploma -** Department of Urban and Regional Planning, Colledge of Environmental Studies, Kaduna Polytechnic, Kaduna, Nigeria

PROFESSIONAL EXPERIENCE

2006 – Present	**Graduate Assistant -** Department of Urban and Regional Planning, School of Environmental Studies, Modibbo Adama University of Technology, Yola, Nigeria
2012 – 2013	**GIS Analyst -** Palye Mühendislik, Ankara - Turkey
2011	**GIS Analyst -** Toprak Su İnşaat, Ankara, Turkey
2006 – 2007	**Managing Partner** - Softnetics Nigeria Ltd., Yola, Nigeria
2003	**Industrial Trainee -** Federal Capital DevelopmentAuthority, Abuja, Nigeria
1998	**Industrial Trainee -** Ministry of Lands and CountryPlanning, Kaduna, Nigeria

LANGUAGES

English, Turkish, Fulfulde and Hausa (Native - Languages)

ACADEMIC AWARDS AND HONOURS

- Vice Chancellor's Prize for the Best Academic Performance (2002/2003 Academic Year).
- Dean's Prize for the Best Graduand in the School of Environmental Sciences (2002/2003 Academic Year).
- David A. Garnvwa's Prize for the Best Graduand in Urban and Regional Planning (2002/2003 Academic Year).
- Best Student (NDI) URP (1997/1998 Academic Year).
- Best Student (Pre-ND) URP (1996/1997 Academic Year).

PUBLICATIONS

Thesis

- **Ismaila, A.B.** (2013). Investigation of digital elevation model uncertainty in GIS-based solar radiation models using Markov Chain Monte Carlo simulation. Ph.D. thesis submitted to the Department of Geodetic and Geographic Information Technologies, Middle East Technical University, Ankara, Turkey.
- **Ismaila, A.B.** (2004). The application of GIS in data management of water supply facilities of Yola - Town, Adamawa State. B. Tech Dissertation submitted to the Department of Urban and Regional Planning, Modibbo Adama University of Technology (Formerly Federal University of Technology), Yola, Nigeria.
- **Ismaila, A.B.** (1999). A master plan for Down - Quarters (1999 - 2014), Kaduna. National Diploma Dissertation submitted to the Department of Urban and Regional Planning, Kaduna Polytechnic, Kaduna, Nigeria.

International

Chapter in a Book

- Yaylaci, E.D., **Ismaila, A.B.**, Uskay, O., and Düzgün, Ş. (2011). Spatial analyses of electricity supply and consumption in Turkey for effective energy management and policy-making. In: Schmidt M., Onyango V., Palekhov D. (eds): Implementing Environmental and Resource Management. Berlin, Germany. *Springer Verlag.* Pages 151-166.

Conference Papers/Conferences Attended

- Umar, U.A., Khamidi, M.F. and **Ismaila, A.B.** (2013). Transformations and impact of green building and sustainable development in 21st century. In: Benian, E., Kartal, S. (eds). "Farkindalik" Awareness, 8th International Symposium on Architect Sinan. *Trakya University, Faculty of Engineering & Architecture, Department of Architecture*. 25-26 April 2013, Edirne, Turkey. ISBN 978-975-374-159-0. Pages 487 - 494.

- **Ismaila, A.B.** and Usul, N. (2013). A GIS-based spatial analysis of health care facilities in Yola, Nigeria. In: Rückemann, C.P., Wilhelms, W. (eds). GEOProcessing 2013: The Fifth International Conference on Advanced Geographic

Information Systems, Applications, and Services. February 24 - March 1, 2013, Nice, France. ISBN 978-1-61208-251-6. Pages 46-53.

- **Ismaila, A.B.** (2013). Evaluating land use change in rapidly urbanizing Nigeria: Case study of Yola, Adamawa State. In: Ivan, I., Longley, P., Horak, J., Fritsch, D., Cheshire, J., Inspektor, T. (eds). GIS Ostrava 2013 - Geoinformatics for City Transformations. Ostrava: *VŠB - Technical University of Ostrava*, 21. - 23. 1. 2013. ISBN 978-80-248-2974-6. Page 93 - 107.

- **Ismaila, A.B.** and Usul, N. (2011). A GIS-based spatial analysis of health care facilities in Yola, Nigeria. 2011 Esri International User Conference, Website: http://proceedings.esri.com/library/userconf/ proc11/papers/3271_63.pdf.

- Kalayci, E., Aslan, M., Weber, G.-W. and **Ismaila, A.B.** (2009). Stochastic portfolio optimization in the Nigerian electricity sector. *"EURO XXIII 2009"*, Bonn, Germany. July 5-8. Page 6.

- Weber, G.-W., Alparslan-Gok, S.Z., Ozturk, B.A., Kalayci, E., Aslan, M.H., **Ismaila, A.B.** and Kropat, E. (2008). Eco-Finance Networks under Uncertainty, Seminar on Financial Mathematics and Seminar on Applied Mathematics in Life and Human Sciences and Economy in IAM, METU, and YUUP Biyotip Group "Bioinformatics", METU, October 31, 2008.

National

Journal

- Husain, M.A. and **Ismaila, A.B.** (2008). The use of GIS for decision taking by Adamawa State Water Board: Implication on planning - related agencies in Adamawa State. *Journal of Technology and Development*, Vol. 10.
- Husain, M.A. and **Ismaila, A.B.** (2006). The routines and requirements for the creation of a Geographical Information System (GIS): The case of GIS database for the Water Board Corporation, Yola. *Environ: Journal of Environmental Studies*, 2 *(6): 47-56.*