|  |
| --- |
| A Model to Identify Micro/Mini Hydropower Sites with High Potential Based on Freely Available Datasets |
| In Southern Mindanao, Philippines |
|  |
|  |



Msc. thesis report Geo-Information Science

A Model to Identify Micro Hydropower Sites with High Potential Based on Freely Available DatasetsIn Southern Mindanao, Philippines

B.P. Ottow Bsc.  
Student Registration Number: 890529633110  
For course: MSc Thesis Geo-information Science and Remote Sensing (GRS-80436)  
Supervisors: dr. ir. Gerard Heuvelink and dr. ir. Sytze de Bruin

Wageningen, 20 March 2015

Wageningen University

# Abstract

Clean, renewable energy is in high demand. One of the most cost-effective kinds is hydropower, which is also one of the most feasible in rural off grid villages. Methods to find spots that are high in potential of Watts are irreproducible, subjective, and require local knowledge. Several GIS model are starting to evolve to guide in this process. The objective of this study was to develop and test a spatial model to identify potential micro-hydropower sites based on global and freely available data and open source software for the south of Mindanao Island, Philippines. The scripting language R and the GIS package SAGA GIS have been used to process the free data into potential spots. ASTER GDEM, TRMM, and MOD16 provide the necessary input data. The resulting maps show potential spots on a fine resolution. Fluctuations in potential over the year are made clear with the use of time steps of a month. Also uncertainty about groundwater flows are shown by the use of three different scenarios. This model gives a quick and extensive overview for an early exploration of hydro power sites.

Keywords: hydro power, MHP, ASTER GDEM, TRMM, MOD16, R

Table of Contents

[Abstract iii](#_Toc414628384)

[1 Introduction 2](#_Toc414628385)

[1.1 Context and background 2](#_Toc414628386)

[1.2 Problem definition 3](#_Toc414628387)

[1.3 Research objective and research questions 3](#_Toc414628388)

[1.3.1 Main objective 3](#_Toc414628389)

[1.3.2 Research questions 4](#_Toc414628390)

[1.4 Study area 4](#_Toc414628391)

[2 Methodology 6](#_Toc414628392)

[2.1 Increased spatial and temporal resolution 7](#_Toc414628393)

[2.2 Increasing channel length by using costrasters for head calculation 7](#_Toc414628394)

[2.3 Including groundwater storage 8](#_Toc414628395)

[3 Results 9](#_Toc414628396)

[3.1 Increased spatial resolution 9](#_Toc414628397)

[3.2 Increased temporal resolution and making use of costrasters 10](#_Toc414628398)

[3.3 Including scenarios for different storage factors 17](#_Toc414628399)

[4 Discussion 18](#_Toc414628400)

[5 References 19](#_Toc414628401)

[A. Pseudo-code of the model 20](#_Toc414628402)

# Introduction

## Context and background

Hydropower on a small-scale is one of the most cost-effective energy technologies

to be considered for rural electrification in less developed countries [1]. The potential of small-scale hydropower is especially large in mountainous areas, which are often also areas where the electrical grid coverage is limited [2]. This makes micro-hydropower very useful for many parts in the world.

Hydropower is generally classified according to capacity. Micro Hydropower (MHP) has a capacity smaller than 500 kW; Mini Hydropower below 2 MW [3]. A MHP can provide enough electrical power for a farm up to a village [4].

MHPs exist in two forms. The first is to build a dam to make a water reservoir. A cheaper and more simple solution is the run-of-river scheme where water is diverted from the main river. In the latter form the MHP has to deal with a varying water flow but avoids the necessity for constructing a dam [3]. This research will focus on the run-of-river scheme variant. A schematic view of the construction is shown in Figure 1.

Before an exact location for an MHP is examined in detail, multiple potential spots have to be identified. Next, these spots should be visited to determine if the local conditions indeed meet the requirements, including detailed measurements and calculation on the potential. Once this is done a design for the installation can be made.

To identify potential sites, normally contour maps are visually interpreted or local information is gathered from people living in the area of interest [5]. However, this information is often incomplete, subjective, irreproducible and requires local knowledge or presence.

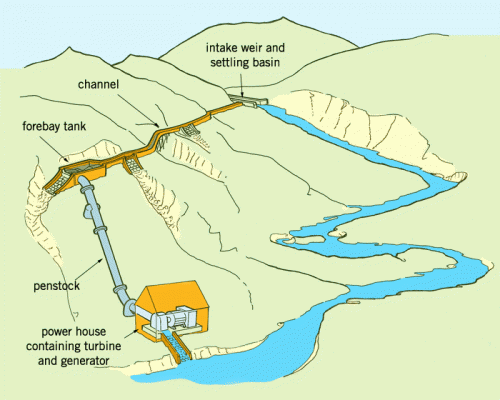


Figure An MHP following the run-of-river scheme. Water flows from the river via the intake weir into an almost horizontal side channel. After several hundreds of meters the water flows via a pipe (the penstock) downhill to the turbine in the power house. Source: practicalaction.org

## Problem definition

As concluded above, finding potential sites for MHPs in the usual way is laborious, subjective and prone to error. Recently, Feizizadeh (2012) developed an ArcGIS model to overcome these problems and to produce a map with potential MHP spots based on digital spatial data. The model makes use of yearly precipitation and evapotranspiration maps and a Digital Elevation Model (DEM) with 250m resolution.

There are also other models to locate potential hydropower sites [6-8], but these all require river runoff measurements. An overview of mainly commercial tools for hydropower potential calculation is given by [9].

The model of Feizizadeh (2012) makes use of an automated method to identify potential sites. It is relatively easy to implement but there are still some disadvantages which need more research. These disadvantages are: 1) the method is dependent on expensive proprietary software and data; 2) it uses an unrealistic runoff estimate; 3) it does not take the run-of-river scheme into account; 4) it uses too coarse temporal and spatial resolutions.

1. To be able to use this model easily on a new area for identification of potential sites, the model should only work with data that are freely available. Not only the data should be free but also the software that is used. When the model works with only freely available global data it is made sure that it is implementable everywhere in the world and also with a low budget.
2. There are some aspects in the model that are not realistic. For instance, the runoff part of the model is a poor representation of reality. The model assumes that runoff is precipitation subtracted by evapotranspiration. In this assumption groundwater flow is neglected. In steep mountainous areas the groundwater flow is on average relatively small (because of the steep slopes) [10] compared to the surface flow, but it still is a significant part.
3. The current model calculates the potential hydropower per grid cell, while in reality the MHP side channel is often longer than one cell, meaning that the cumulative hydraulic head of a chain of cells should be used instead.
4. The run-of-river scheme does not have a reservoir to store water. This makes that the year-round potential for the MHP cannot be based on the average annual runoff but should take the temporal variability in runoff into account. The current model does not regard seasonality of a river, e.g. the difference between the rainy season and dry season. It will be highly unwanted if the MHP has to be shut down for a longer period during the dry season.  
   Rivers in mountainous areas are mostly quite small and narrow, so to be able to have an accurate estimate of the elevation at the river bed the resolution (or support) of the DEM should be ideally smaller than the width of the river.
5. Using only freely available global data for this model brings in another problem. Especially free and global DEMs have large but unknown elevation errors. DEM errors will propagate through the model, which is not considered in the Feizizadeh model. It could be that the model will be that sensitive to errors in the DEM that the model result is too inaccurate to be useful for practical purposes. How to model the error in the DEM is not straightforward and includes some assumptions to be made. Various studies have looked to the vertical error in freely available DEMs [11-13]. Characterizing the spatial correlation of the errors goes one step further, multiple studies tried to find a best practice for that [14-17]. Specifically focusing on hydrological parameters derived from DEMs is done by [18, 19]. Furthermore, studies try to relate the error to specific properties of the landscape like morphology, elevation or land cover [13, 20].

If the problems mentioned above would be resolved, then this would yield a substantial improvement of the Feizizadeh model and would stimulate its practical application.

## Research objective and research questions

### Main objective

Develop and test a spatial model to identify potential micro-hydropower sites based on global and freely available data and open source software for the south of Mindanao Island, Philippines.

### Research questions

1. Can the Feizizadeh (2012) model be improved by:
   1. Increased spatial resolution and increased temporal resolution to regard seasonal fluctuations?
   2. Increasing channel length by using costrasters for head calculation?
   3. Including groundwater storage?
2. How can the error in the DEM be modelled?
3. How does DEM error propagate through the model, and how does it influence the location of potential sites?
4. What is the difference of the results of this model with the ‘unimproved’ model? Can the quality of the result be judged as improved?

## Study area

The test area for the model is the southern part (south of 7 degrees latitude) of Mindanao Island, Philippines. Figure 2 shows a map of the area. It is a mountainous island with peaks rising over 3000 meters. The island has a tropical climate with high precipitation values in the rainy season. There is a known planned location for an MHP at N6° 23’ 02”, E 125° 10’ 31”. For the performance of the model it is best to make the area as small as possible but to get accurate results it should at least contain a whole catchment. For the tests the catchment of the planned location is used.

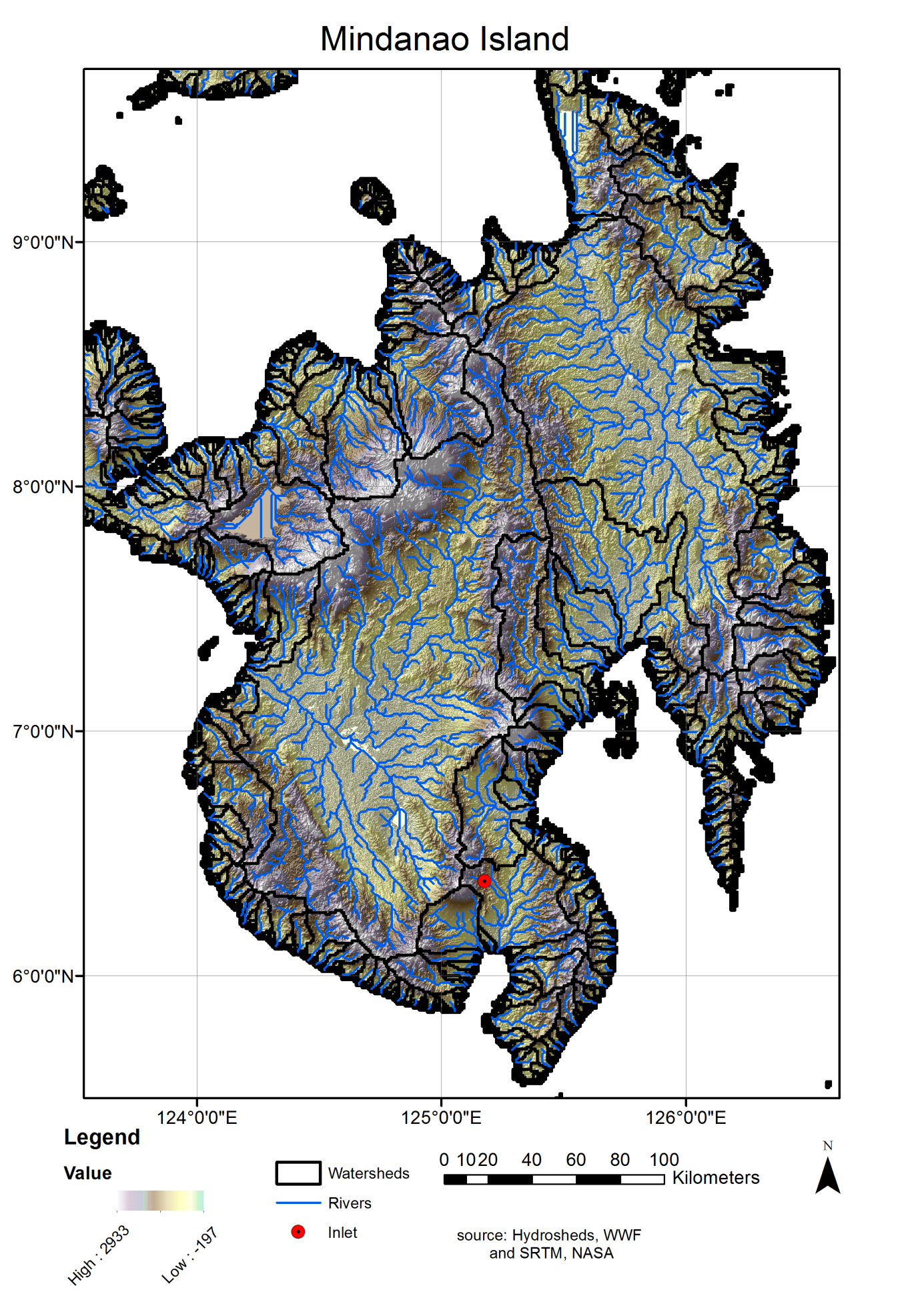


Figure Overview of Mindanao Island, Philippines. Including catchment boundaries (black), water streams (blue) and a planned location of an MHP. The study area is the catchment in with the red dot.

# Methodology

The model of Feizizadeh is the starting point for the model. The model a DEM and precipitation, and evapotranspiration maps as input. The precipitation is subtracted by the evapotranspiration per cell on a yearly basis. The DEM is filled to make it hydrologically correct. Next, the DEM is used to derive the flow direction and flow accumulation per cell. With the earlier calculated runoff and the flow accumulation, the runoff per cell can be calculated.

The DEM is also used to calculate the head of each cell. The head is calculated as the difference of the elevation of the cell and the elevation of the neighbouring cell with the lowest elevation. This head is multiplied by the runoff and the gravitation acceleration to result in the potential power generation of that cell.

The entire process is summarized in Figure 3. This research initially used the Feizizadeh model with only (global) open data as input. From now on, this model will be referred to as the default model.



Figure Scheme of the initial model [21]

## Improving the model

### Increased spatial and temporal resolution

To increase the spatial resolution, the ASTER GDEM Digital Elevation Model is used as input. The spatial resolution is 30m.

The run-of-river scheme for MHP does not buffer water in a reservoir so, to have it running year round, it is dependent on the lowest runoff in the river per year. Another possibility is to shut the system down in the dry season. In order to take this temporal restriction into account, the seasonal fluctuations in runoff in the river are estimated in the calculation of the potential.

To reach this, the temporal resolution for the input is increased to a monthly resolution. We decided to mark a cell as a potentially good site when the power potential is at least 10 months of the year above the chosen minimum-potential.

**Data**

Figure Process for head calculation.

The data that is used for the DEM is ASTER GDEM. Version two of this dataset was released in 2011. The coverage is 83 degrees north to 83 degrees south with a resolution of 1 arcsecond (approximately 30 m). The vertical accuracy varies in space. Studies in Japan and the US reported an accuracy of 20 m (95% confidence interval) [22], while a study in India reported an RMSE of 6.08m [11]. Both studies found statistical dependencies between accuracy and land use type as well as slope steepness.

Precipitation data is collected from TRMM (Tropical Rainfall Measuring Mission). TRMM provides data from 50 degrees north to 50 degrees south with a resolution of 4 km or 0.25 degree on a 3-hourly scale. The monitoring started in January 1998 and still goes on [23]. The dataset we used is an averaged per month dataset of the 2B31 product (a combined Precipitation Radar (PR) / TRMM Microwave Imager (TMI) rain-rate product). It has a 4 km horizontal resolution [24]

Evapotranspiration (ET) data come from the MOD16 product. MOD16 has global coverage with a 1 km resolution and 8-day, monthly and annual ET. The dataset covers the time period 2000 – 2010 [25].

The modelling environment is the R statistical software [26] with R Studio [27] as Integrated Development Environment. From R also SAGA GIS [28] is used with the RSAGA package[29]. This is all free and open source software.

### Increasing channel length by using costrasters for head calculation

The head is calculated from the DEM. The default model only used the direct neighbour of the cell to calculate the head. With a spatial resolution of 30 meters this would mean that the side-channel would (only) be 30 meters. The new model makes this value variable and for the user to set. A standard value is set to 500 meters. The side-channel cannot be built such that the water should flow up, so the channel should always stay near the river and the water should also exit in the same river again. This is the main assumption for the model in deciding which cells participate in the calculation of the head for a particular cell.

In Figure 4, the process of calculating the head for the entire area is visualized. It is only necessary to calculate the head for the cells which have enough runoff to make a hydropower feasible, that is why the calculation is only been done for the cells with a runoff above the minimum runoff.

The R function for this step tends to be very slow for large rasters. Because of that it is possible to call the function via a wrapper which splits up the big DEM into smaller bits, then calls the head function for each part and then sews the results back together.

### Including groundwater storage

Regarding groundwater flows the default model uses the most simple option; no groundwater flow. Over a whole year the estimate of runoff by subtracting Evapotranspiration from Precipitation works quite well because in a whole year the water storage in the area will not change much. On smaller intervals though, the storage changes significantly. In the rainy season the soil gets ‘filled’ with water, which flows in to the rivers again in the dry season. This is dependent on the soil depth to bedrock, land cover (vegetation), slope, soil type and other factors [10].

With a higher temporal resolution the default model is highly inaccurate but without much more local data about the soil and morphology of the area it is troublesome to make a better estimate. Furthermore it could cost the model a big deal of performance to include a good groundwater model. Following this arguments we chose to keep the model simple but include different scenario’s.

To start with, the model calculates the default runoff (precipitation minus evapotranspiration). Following, a factor (0, 0.3 and 0.6) of storage per time step (month) is calculated. The following formula is used:

with Q is the new runoff, q is the default runoff, and f is the factor. The resulting runoffs per scenario are multiplied by the head which gives per scenario and per month a map of power potential. This overall process is visualized in Figure 5.



Figure Overall process of the use of the three scenario's

## Modelling DEM error

ASTER GDEM is a Digital Elevation Model made of images from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). The instrument was built onboard NASA’s Terra spacecraft which was launched in 1999. The sensor makes along-track stereoscopic images in the infrared spectral band. Version 2 was released in 2011 and uses data acquired from 2000 to 2010. Infrared does not pierce clouds, so an image taken on a cloudy day cannot be used for the GDEM generation. This results in varying amounts of images per area and per pixel. The more images there are, the more elevation measurements there are, making the error in the values smaller. Some areas are covered in clouds (or their shadow) on all the images. To fill these gaps, data of SRTM is used. SRTM is a global DEM from a Space shuttle radar with a 3 arc seconds resolution (about 90 m). When downloading an ASTER GDEM tile, a quality file is also delivered. The quality file tells how many scenes are used to calculate the elevation. For the study-area these numbers are shown in Figure 6. The frequencies of the numbers in the area are shown in the histogram in Figure 7. 12% of the area was in all the scenes covered with clouds, these areas are mainly in the high mountains at the borders of the catchment.

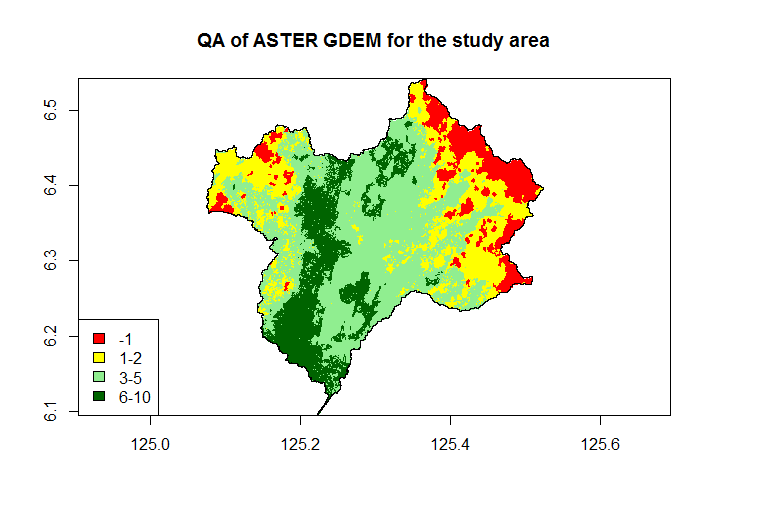


Figure Quality assessment of the ASTER GDEM for the study area. The numbers are the amount of scenes used in the processing, -1 means no scenes were present, there SRTM version 3 is used to fill the gaps.

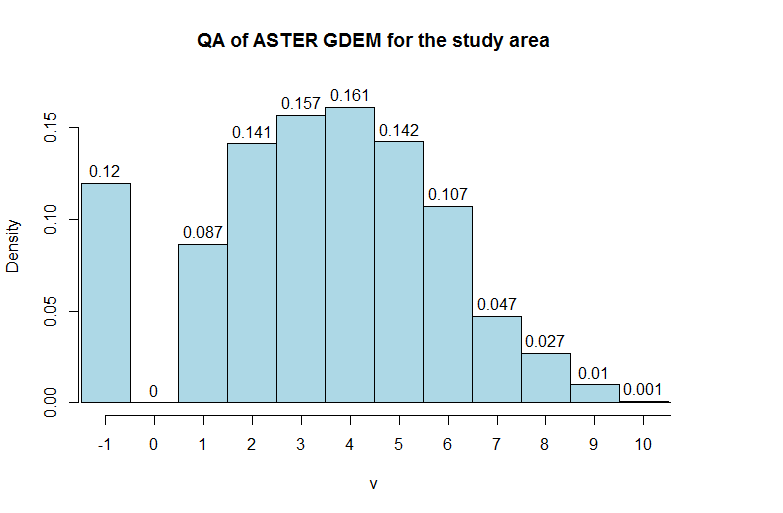


Figure Frequency of the numbers in the QA file of the ASTER GDEM for the study area. X-axis is amount of scenes used to calculate the elevation. -1 means no scenes were present and SRTM version is used to fill the gaps.

The method for modelling the DEM error followed Temme et al [14].

The true elevation can be represented as a statistical model:

where is the deterministic elevation map (DEM) and is a random variable which consists of an autocorrelated part and a pure noise part. The mean and standard deviation can be estimated from a set of control points.

DEM errors are usually spatially autocorrelated. It was assumed that spatial correlation depends on the separation distance vector between the locations, which can be characterized by the semivariogram. Isotropy, so independence to the direction of the vector, was also assumed.

The semivariogram can be estimated from observed DEM errors at control points, provided the number of observations is sufficiently large.

For the study area there were no control points available. The errors occurring are assumed to be similar to that of a mountainous area elsewhere in the world. A highly accurate (and fine resolution) Lidar dataset from the area south-west of Mt Rainier in the US (North: 46.8489°, South: 46.7167°, East: -121.7450°, West: -121.9394°) is used as control points dataset. With these points a semivariogram of the ASTER GDEM error of that area has been computed. This semivariogram is used as error model of the DEM in the study area.

More information about the dataset can be found at the OpenTopography website (<http://opentopo.sdsc.edu/gridsphere/gridsphere?gs_action=datasetMetadata&cid=geonlidarframeportlet&otCollectionID=OT.052013.26910.1>).

## DEM error propagation

To calculate how the DEM error propagates through the model, a Monte Carlo (MC) simulation was used. According to [Temme et al. (2009](#_ENREF_22)) the method comprises:

1. repeat N times:
   1. generate a set of realizations from the joint probability distribution of
   2. for this set of realizations , compute and store the output
2. compute and store sample statistics from the N outputs .

Where is an uncertain input, is the number of uncertain inputs which will be set to 1, in this case only the DEM. then is a realization of a DEM derived from the error model. is the output and is the model.

The error propagation only considered DEM error so each realization was just one realization of the DEM, while precipitation and evapotranspiration are not treated as uncertain.

The whole procedure is repeated 200 times.

In the model, the DEM is used for and has influence on two different parts. 1. To calculate the slope, a difference in the slope will result in a difference in the potential. 2. To derive the stream network, a difference here will result in a different location of the river, so also a different location of a high-potential spot, which can be a relatively big shift. Furthermore it is possible that the amount of cells in the upstream area of a cell in a river changes which yields a difference in runoff so in potential.

The Monte Carlo runs are run using 10.000 Watt as minimum potential which should be reached for a minimum of 10 months per year. If a cell meets these conditions it is marked as a potential site. Each run produces for each scenario (1.0, 0.7 and 0.4 factor storage) a map with ones (potential site) and zeroes (not a potential site). The 200 maps summed and divided by 2 gave a chance map with values per cell as a percentage of the runs where it is a potential site.

@Aggregated.

A lower limit for potential will be set to mark a spot as a high potential spot. The limit will be about 5 kW. For each simulation this will give a raster with zeroes and ones, zero for a not-high-potential spot and one for a high potential spot. Cumulated with the 500 results a chance map will be generated with the chance that a cell is a potential spot. A probable shift will be difficult to analyze in these maps, for this problem a different kind of visualization will be found like kernel density mapping. This method will show areas instead of points of high potential.

Difference of the results between improved and default models

The differences in the results of the default and improved models will be analyzed. A difference map will be generated and interpreted. Also the maps of the certainty of the output will be used in this analysis.

An area within the study-area is already investigated about the potential for hydropower and soon a project to realize a MHP will start. The analysis will include a visual inspection how the results compare to the locally investigated potential and with background maps and images. In December, possibly someone will visit the study area and can gather more data that can be used for comparison.

# Results

The resulting R script can be viewed and downloaded from GitHub (<https://github.com/boukepieter/thesis>). The pseudo-code for the full script can be found in Annex A.

Improving the default model

At first the default model has been built in an R script. The result is a single map of the potential. Figure 6 shows the result. The maximum potential in the catchment is 1.39 MW. To be able to compare the results the shown maps will be zoomed in to a part of the catchment in the north-west and the power potential colourscale will be kept constant.

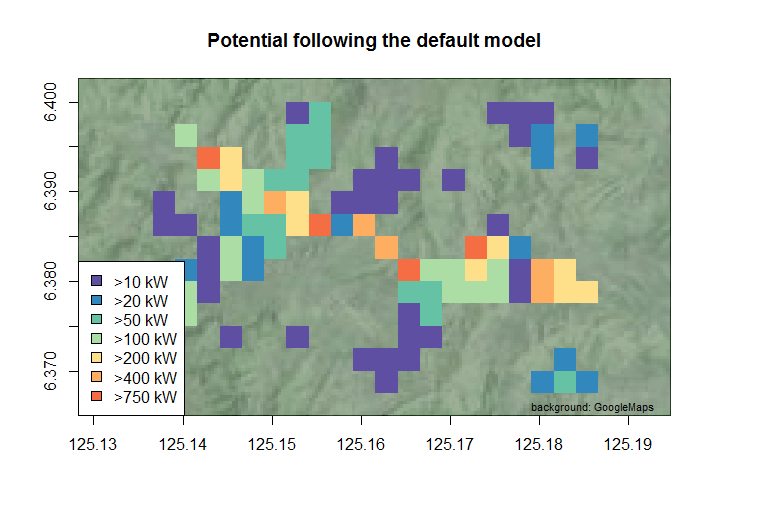


Figure Result of the default model

### Increased spatial resolution

Figure 7 is the map with the result. The potentials are far lower than in the default model, this is because the head between neighbouring cells is on average much lower. This is logical since the cell size is much smaller (from 270m to 30m), and thus the side-channel length is shorter. The following step solves this problem.

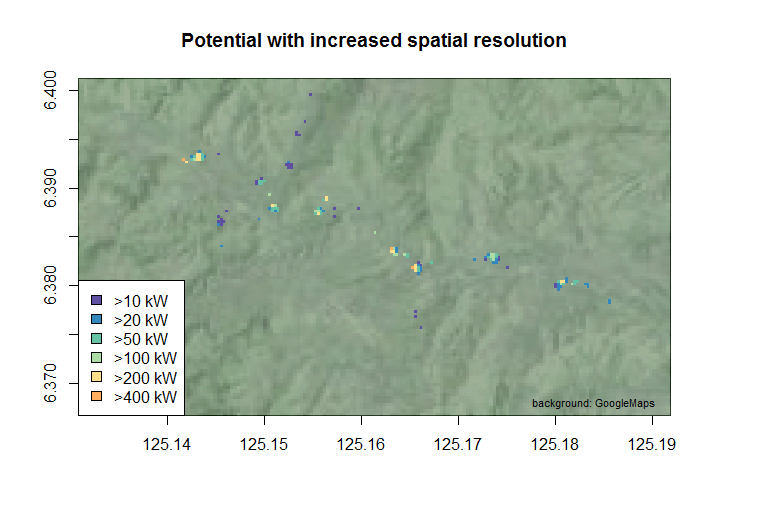


Figure The result of the model with increased spatial resolution

### Increased temporal resolution and making use of costrasters

Figures 8-13 show the result. There is a big difference per month due to the difference in rainfall. In February and December the evapotranspiration is higher than the precipitation causing a negative runoff and a negative or no potential.

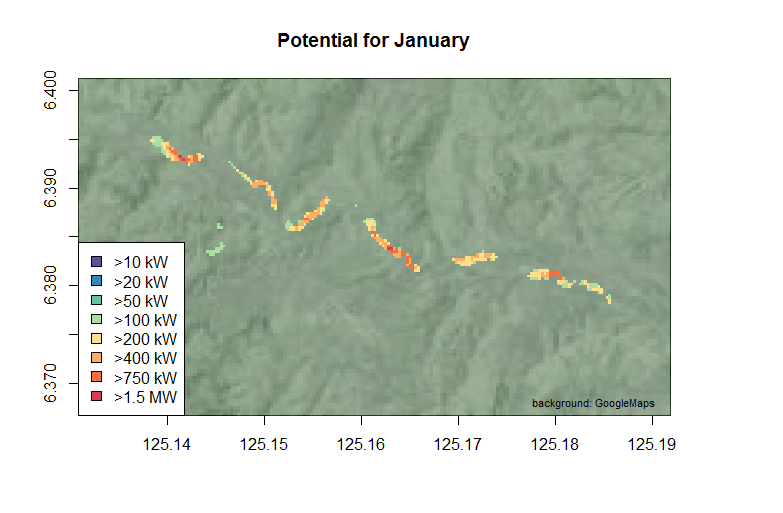
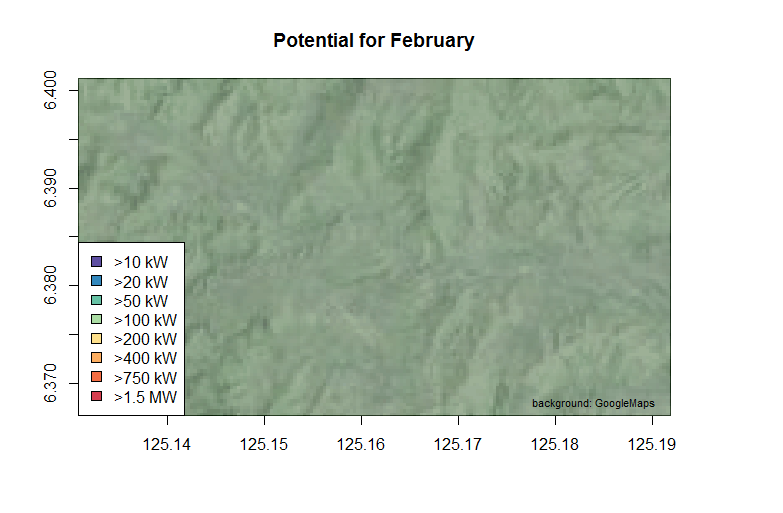
 

Figure Result of model with increased spatial and temporal resolution and with using costrasters for the head-calculation. For January and February

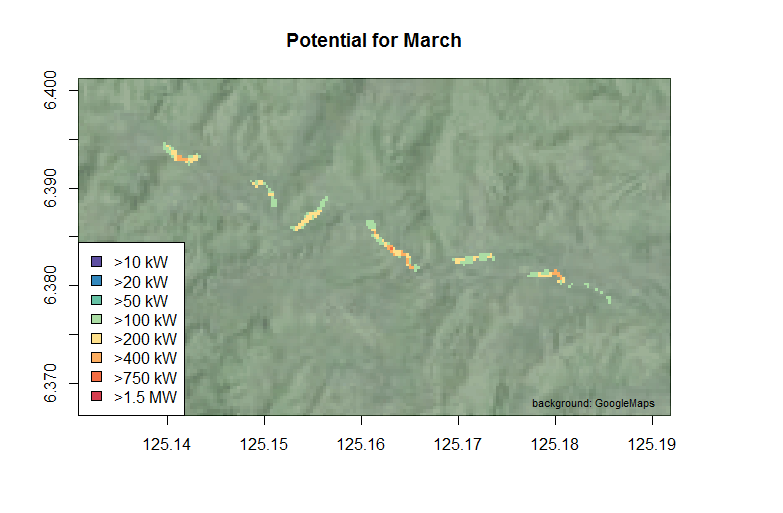
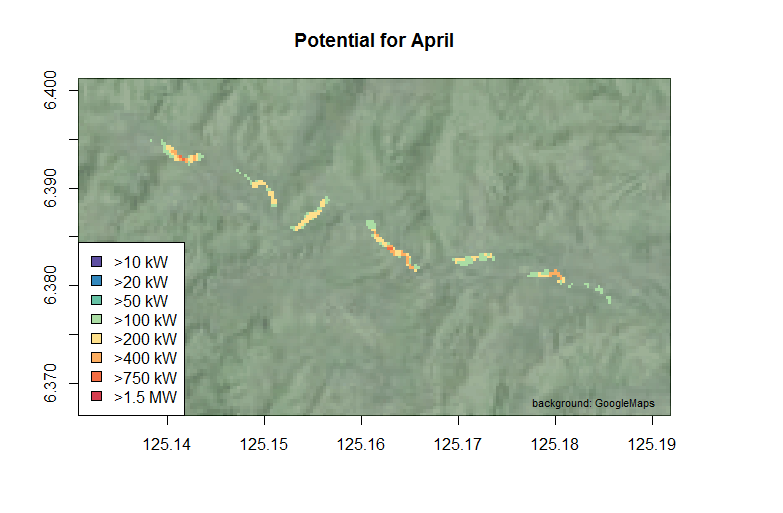
 

Figure Result of model with increased spatial and temporal resolution and with using costrasters for the head-calculation. For March and April.

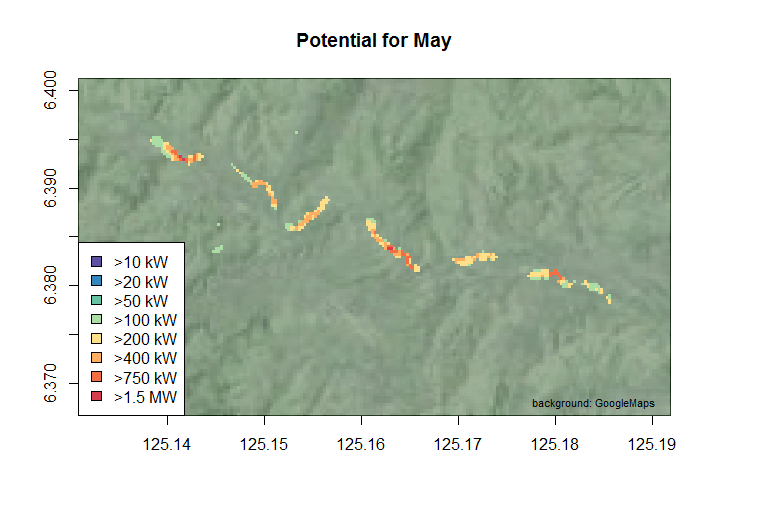
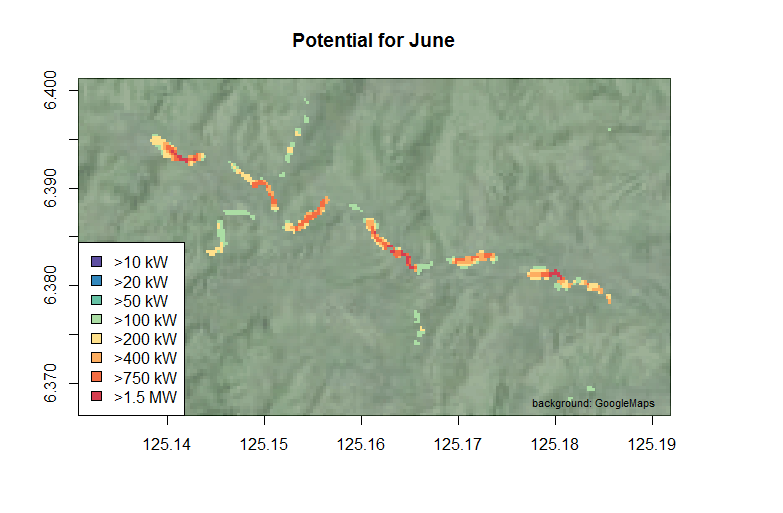
 

Figure Result of model with increased spatial and temporal resolution and with using costrasters for the head-calculation. For May and June.

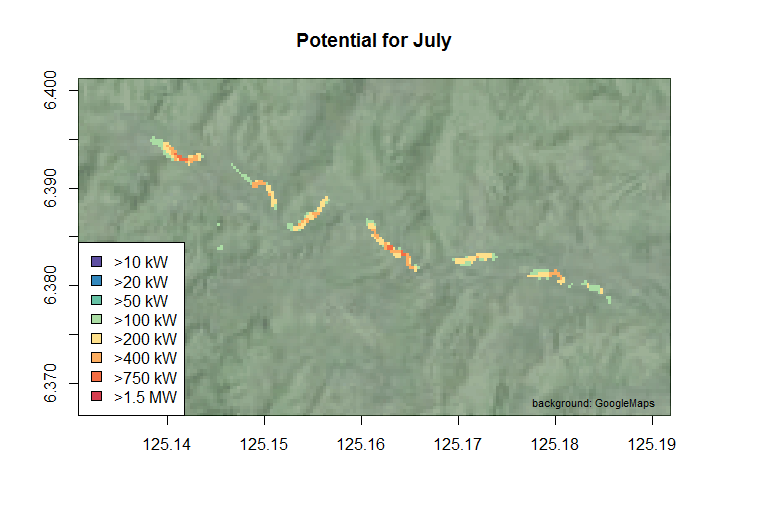
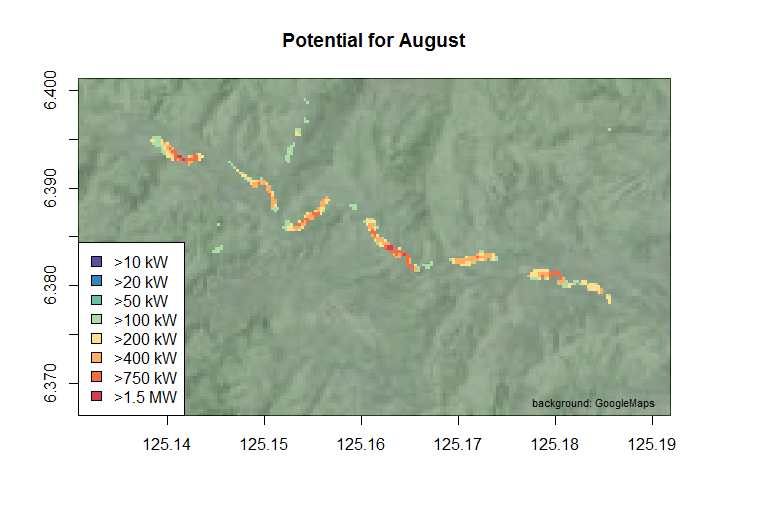
 

Figure Result of model with increased spatial and temporal resolution and with using costrasters for the head-calculation. For July and August.

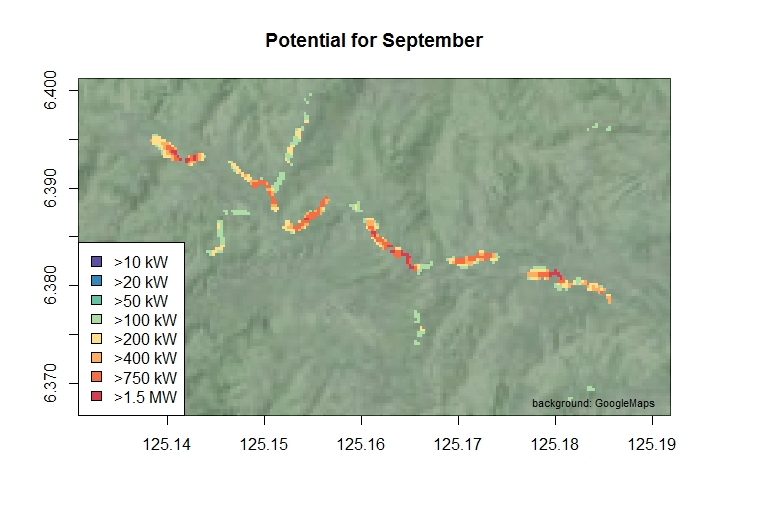
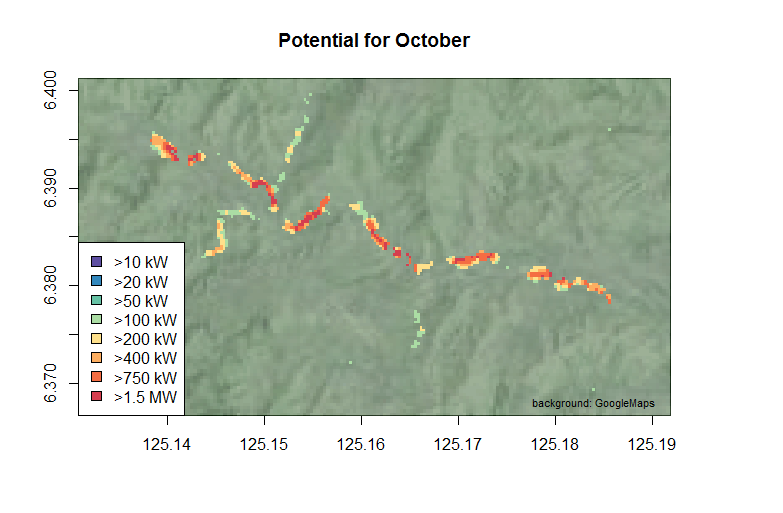
 

Figure Result of model with increased spatial and temporal resolution and with using costrasters for the head-calculation. For September and October.

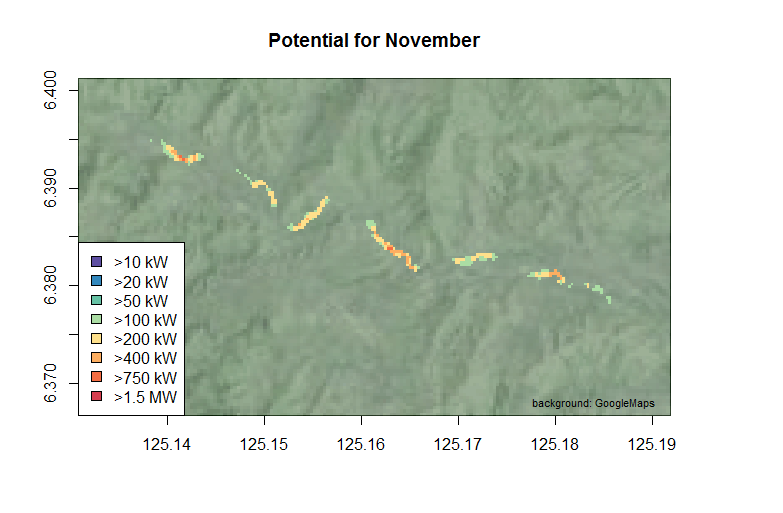
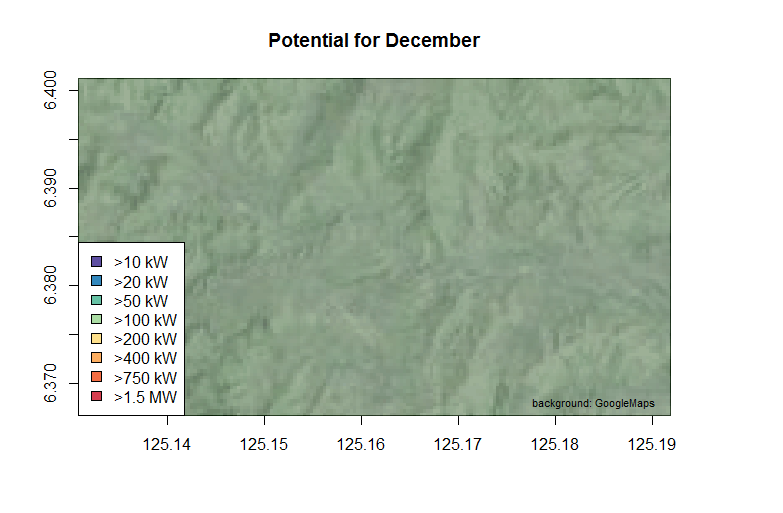
 

Figure Result of model with increased spatial and temporal resolution and with using costrasters for the head-calculation. For November and December.

### Including scenarios for different storage factors

The influence of the different scenarios for storage are shown in Figure 14, here is the runoff plotted for one specific point in the river for the different months and scenarios.

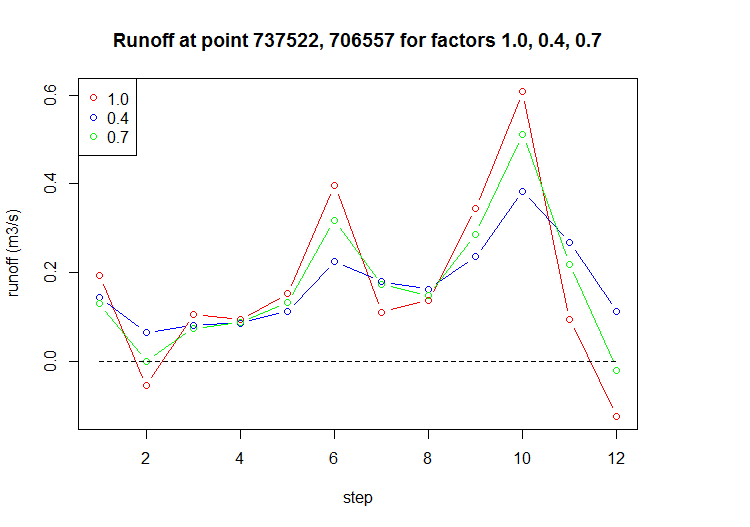


Figure Runoff per month for one point in the river. The factors are in runoff-factors per month.

## Modelling DEM error

The reference lidar dataset of the area near Mt Rainier first had to be pre-processed before it could be compared with the ASTER GDEM of that area. The dataset came in 132 subsets of 1 by 1 km containing totally about 1,000,000,000 points. First the points had to be filtered to get only the points on the ground (so for example not the points of a first return on top of a tree) and after that the number of points had to be reduced to a workable number of points for the comparison. To calculate which points are on the ground the ‘lasground’ tool of the LAStools package has been used. For each subset this tool is used followed by a selection of a random sample of 25 points from the result. The code to perform this in R is shown in Figure 17.

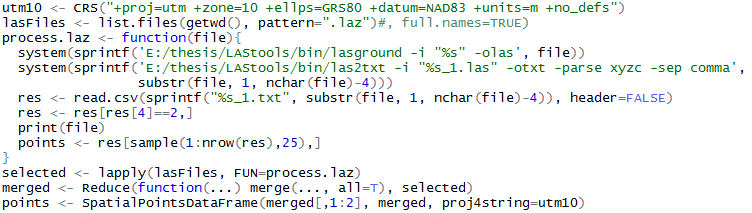


Figure R code to pre-process the lidar dataset. For each subset the ground points are selected, the las file is translated to a txt and 25 points are sampled from the result.

The resulting points are shown in Figure 18

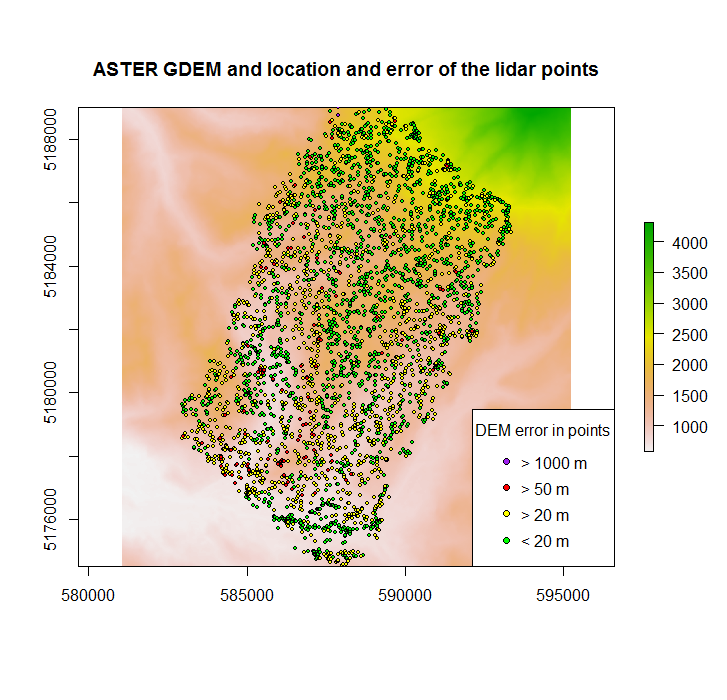


Figure The selected points on the ASTER GDEM background. The colour of the points corresponds to the error.

@visualisatie van punten

@absolute verschillen, verwijderen van de uitschieters

@histogram

@variogram

## DEM error propagation

# Discussion

The acquired results show an improved insight of the possibility for Micro Hydro-Power in the area. The finer spatial resolution give more confidence about the calculated head for the spots. Furthermore the usability of the model is improved by using a model-parameter for the side channel-length. Seasonality in runoff and thus in potential is also made clear in using time steps.

Unfortunately the certainty in the model is still not clear. There is no validation been done and also no uncertainty propagation. A first step has been made by introducing three scenarios in groundwater storage which give insight in the uncertainty in the runoff. For the uncertainty in the DEM a follow up study would be required.

# References

1. Paish, O., *Small hydro power: technology and current status.* Renewable & Sustainable Energy Reviews, 2002. **6**(6): p. 537-556.

2. Razan, J.I., et al., *A Comprehensive Study of Micro-Hydropower Plant and Its Potential in Bangladesh.* ISRN Renewable Energy, 2012. **2012**: p. 10.

3. Awan, A.B. and Z.A. Khan, *Recent progress in renewable energy - Remedy of energy crisis in Pakistan.* Renewable & Sustainable Energy Reviews, 2014. **33**: p. 236-253.

4. Adhau, S.P., R.M. Moharil, and P.G. Adhau, *Mini-hydro power generation on existing irrigation projects: Case study of Indian sites.* Renewable & Sustainable Energy Reviews, 2012. **16**(7): p. 4785-4795.

5. Unknown, *Manuals and Guidelines for Micro-hydropower Development in Rural Electrification Volume I*. 2009, Department of Energy Energy Utilization Management Bureau: Manila Philippines.

6. Kusre, B.C., et al., *Assessment of hydropower potential using GIS and hydrological modeling technique in Kopili River basin in Assam (India).* Applied Energy, 2010. **87**(1): p. 298-309.

7. Larentis, D.G., et al., *Gis-based procedures for hydropower potential spotting.* Energy, 2010. **35**(10): p. 4237-4243.

8. Yi, C.S., J.H. Lee, and M.P. Shim, *Site location analysis for small hydropower using geo-spatial information system.* Renewable Energy, 2010. **35**(4): p. 852-861.

9. Punys, P., et al., *Tools for Small Hydropower Plant Resource Planning and Development: A Review of Technology and Applications.* Energies, 2011. **4**(9): p. 1258-1277.

10. Critchley, W.S., Klaus, *A Manual for the Design and Construction of Water Harvesting Schemes for Plant Production*. 1991, Rome: Food and Agriculture Organization of the United Nations.

11. Mukherjee, S., et al., *Evaluation of vertical accuracy of open source Digital Elevation Model (DEM).* International Journal of Applied Earth Observation and Geoinformation, 2012. **21**(1): p. 205-217.

12. Gonga-Saholiariliva, N., et al., *Techniques for quantifying the accuracy of gridded elevation models and for mapping uncertainty in digital terrain analysis.* Progress in Physical Geography, 2011. **35**(6): p. 739-764.

13. Tetsushi Tachikawa, M.K., Akira Iwasaki,, et al., *ASTER Global Digital Elevation Model Version 2 - Summary of Validation Results.* 2011.

14. Temme, A.J.A.M., et al., *Geostatistical simulation and error propagation in geomorphometry*, in *Developments in Soil Science*. 2009. p. 121-140.

15. Heuvelink, G.B.M., *Error propagation in environmental modelling with GIS*. Error propagation in environmental modelling with GIS. 1998.

16. Fisher, P., *Improved modeling of elevation error with Geostatistics.* GeoInformatica, 1998. **2**(3): p. 215-233.

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

18. Wechsler, S.P., *Uncertainties associated with digital elevation models for hydrologic applications: A review.* Hydrology and Earth System Sciences, 2007. **11**(4): p. 1481-1500.

19. Hengl, T., G.B.M. Heuvelink, and E.E. van Loon, *On the uncertainty of stream networks derived from elevation data: the error propagation approach.* Hydrology and Earth System Sciences, 2010. **14**(7): p. 1153-1165.

20. Erdoǧan, S., *Modelling the spatial distribution of DEM error with geographically weighted regression: An experimental study.* Computers and Geosciences, 2010. **36**(1): p. 34-43.

21. Feizizadeh, B.H., Eva Maria, *GIS-based Procedures of Hydropower Potential for Tabriz Basin, Iran.* 2012.

22. Tachikawa, T., et al. *Characteristics of ASTER GDEM version 2*. in *International Geoscience and Remote Sensing Symposium (IGARSS)*. 2011.

23. TRMM, *TRMM Senior Review Proposal.* 2011.

24. Bookhagen, B., *High resolution spatiotemporal distribution of rainfall seasonality and extreme events based on a 12-year TRMM time series, in review.* (in review).

25. Mu, Q.Z., Maosheng; Running, Steven W. , *Brief Introduction to MODIS Evapotranspiration Data Set (MOD16).* 2011.

26. project, R. [cited 2015 March 20th]; Available from: <http://www.r-project.org/>.

27. Studio, R. [cited 2015 March 20th]; Available from: <http://www.rstudio.com/>.

28. Analyses, S.f.A.G. [cited 2015 March 20th]; Available from: <http://www.saga-gis.org/>.

29. Brenning, A. [cited 2015 March 20th]; Available from: <http://cran.r-project.org/web/packages/RSAGA/index.html>.

30. Beekhuizen, J., et al., *Effect of DEM uncertainty on the positional accuracy of airborne imagery.* IEEE Transactions on Geoscience and Remote Sensing, 2011. **49**(5): p. 1567-1577.

31. Shi, W., B. Wang, and Y. Tian, *Accuracy Analysis of Digital Elevation Model Relating to Spatial Resolution and Terrain Slope by Bilinear Interpolation.* 2014. **46**(4): p. 445-481.

# A. Pseudo-code of the model

Installing packages

Preprocessing

Setting parameters

Load functions and copy scripts

Start model

Model:

Load packages

Setup saga environment

Create output and step folders

Load in ET and P datasets

Fill and clip DEM (watershed.R):

DEM to saga grid

Fill DEM

Calculate streams

Calculate strahler order

Select strahler order > 7

Set water (from unfilled DEM) in strahler dataset to NA

Set water (from unfilled DEM) in filled DEM to NA

Calculate basin and subbasins from strahler dataset

Project point of interest to UTM

Extract basin number from point

Select basin and vectorize

Read in shape

Crop filled DEM on shape

Return cropped filled DEM

Resample P to ET

Calculate runoff: P - ET

Resample runoff to DEM

Write resampled runoff to step folder

For each timestep (month):

Runoff to saga grid

Copy saga DEM to new file

Resample new file to the right grid (runoff grid)

Calculate cell accumulation

Remove intermediate file

Read in results as rasterstack

Calculate runoff in m3/s (cell acc \* cell\_size ^2 / 24 / 3600 / 1000)

Define storage scenario factors

Calculate runoffs per factor (storage.fun):

Define matrix of runoff factors

Multiply matrix with runoff values

Calculate head from elevations (HeadOnRiver.large):

Calculate neighbourhood (channel-length / cellsize)

Calculate in how many squares to split up (row and columns / splitSize)

For each square:

Calculate begin and end cell with buffer

Change the cellnumber for border squares

Make a new raster with the values of the original raster for this square

Make a new raster with the values of the original runoff for this square

Make multicore cluster

Start log file

For each square:

Write square number to log file

Calculate head (HeadOnRiver):

Define cells which should be calculated (runoff > minimum runoff)

Store number of rows of the square

Store number of columns of the square

Store resolution of the cells

Store number of cells for the neighbourhood

Copy DEM to head raster

Make dataframe with cellnumbers which should be visited

For each cell:

Make raster of cell and neighboorhood with runoff values (1 or NA)

Make a transition raster (gdistance package)

Geo-correct the raster

Define origin cell

Calculate costraster

Calculate max elevation difference for origin cell to the cells within channelLength range

Write square number and resulting raster metadata to log file

Stop multicore cluster

Sew all the squares (without the buffer) together to one head raster

Write the head raster to the step folder

Crop the runoffs to the head raster

If head < minimum head; set head to NA

Calculate potentials (runoff \* head \* 9.81 \* 1000)

Calculate high potentials (if potential > minimumPotential)

Calculate number of cells with high potential

Print number of points with high potential (per month per scenario)

Write potentials to tiff files

Write high potentials to tiff files