

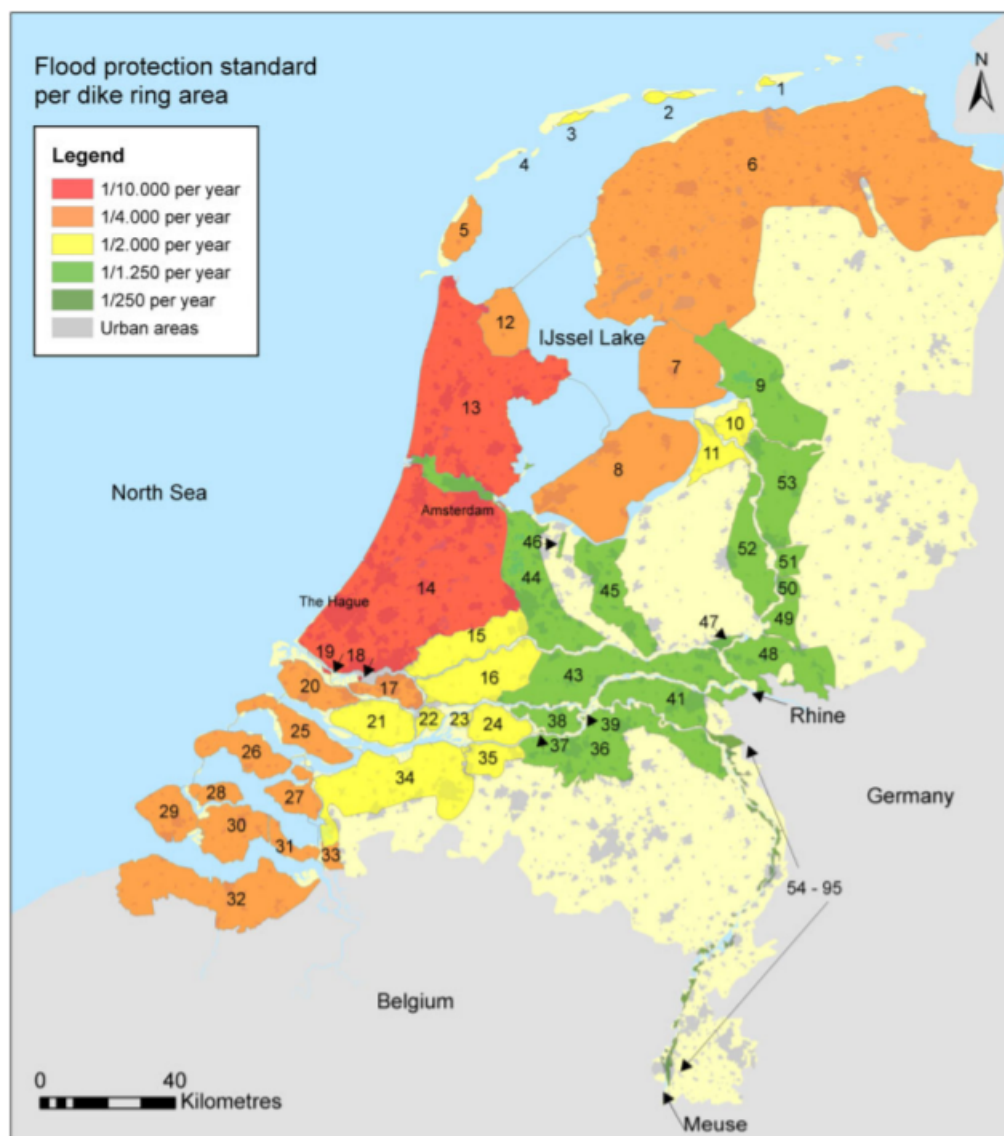
## **Economically Efficient Flood Protection Standards for The Netherlands**

ISyE 6501 Introduction to Analytics Modeling

Course Project

By Frank Kornet

Date 18 April 2019



Source: [https://www.deltares.nl/app/uploads/2014/12/kind2014\\_JFRM1.pdf](https://www.deltares.nl/app/uploads/2014/12/kind2014_JFRM1.pdf) (page 104, Fig 1)

## **1. Introduction**

For the ISyE 6501 Introduction to Analytics Modeling course project we were to select one success story from three different web sites that contain lists of success stories. The detailed instructions for the project were as follows:

In this course project, your job is to think carefully about what analytics models and data might have been required.

- (1) Browse the short overviews of the projects. Read a bunch of them – they’re really interesting. But don’t try to read them all unless you have a lot of spare time; there are lots!
- (2) Pick a project for which you think at least three different Analytics models might have been combined to create the solution.
- (3) Think carefully and critically about what models might be used to create the solution, how they would be combined, what specific data might be needed to use the models, how it might be collected, and how often it might need to be refreshed and the models re-run. DO NOT find a description online (or elsewhere) of what the company or organization actually did. I want this project to be about your ideas, not about reading what someone else did.
- (4) Write a short report describing your answers to (3).

## **2. Selected Project**

It did not take me long to find a success story that interested me, and that I believed to have at least three different Analytics models underlying the success story solution.

I selected the “Economically Efficient Flood Protection” success story at <https://www.informs.org/Impact/O.R.-Analytics-Success-Stories>.

The article interested me for several reasons. The first reason was that I was born and raised in The Netherlands. So, I am quite familiar with to the idea of living and working under sea-level. The second reason was that during my career, I traveled extensively, and, during these trips, I was asked many questions about dikes, and The Netherlands being under water. The third reason is that, living in Houston now, I am still no stranger to flooding.



### **Economically Efficient Flood Protection**

Protection against flooding is a vital  
issue in the Netherlands

**Read More**

The project was carried out between 2006 and 2011. The Dutch government endorsed it in 2012. The solution itself is based on large Operations Research optimization model. The model is mostly a once off effort and resulted in several articles being published. The model determines the most efficient solution for The Netherlands until 2050.

My focus, however, is on turning this into an evolving model instead of a once off model. The reason for this focus is that I believe that climate change and its impact is poorly understood and, as a result, needs to be closely monitored to ensure that the models and solutions developed remain valid. A major challenge is that there is a large time lag between a decision to increase the dike height and the dikes being heightened. Delaying a decision too long to increase the dikes could result in flooding and significant damages (in the order of billions of Euros). On the other hand, increasing the dikes' height too soon could lead to significant extra investments that are not needed (again, in the order of billions of Euros).

### **3. Background**

Before diving into the three analytics models that underly the success story's solution, I want to provide a quick background to put things into perspective and help the reader to understand the issue and solution better.



The Netherlands is basically a delta where two of Europe's main rivers (Rhine and Meuse) go out to the sea. The Dutch have been building dikes for many centuries to protect against the sea **and** the rivers. A dike is basically a human made construction to keep water from flowing inland. There are many types of dikes, but their purpose is the same: keep water out.

The picture on the left shows how The Netherlands would look like without dikes. If you compare that to the picture on the front page, you can see that a large chunk of The Netherlands would be under water. Source: [https://en.wikipedia.org/wiki/Flood\\_control\\_in\\_the\\_Netherlands](https://en.wikipedia.org/wiki/Flood_control_in_the_Netherlands)

Most people only consider the sea to be a problem, but the rivers can equally pose serious flooding problems if they break the

dikes that fence them off.

The picture on the front page shows the fifty-four primary dike rings that protect The Netherlands from flooding. Each dike ring is managed by a separate Water Control Board (WCB). The WCB members are selected through local elections, and the WCB has the right to levy taxes. The first WCB was established in the twelfth century. So, they are an established part of Dutch society and culture.

About 55 percent of all economic activity takes place in dike ring 13 and 14. These are closest to the North Sea. The area is known as Randstad, and consist of major cities like Rotterdam, The Hague, Utrecht, Haarlem, and Amsterdam that are growing slowly together into a mega metropolis.

Over the centuries the Dutch have pumped most of the lakes dry. The land that was freed up was then used for farming or creating villages and towns for people to live. For example, Haarlemmer Meer is now a large part of Haarlem, but used to be a lake. The same is true for IJssel Meer, which consist of large parts that have been reclaimed land from the sea. The lake is closed off from the North Sea using the Afsluitdijk (completed 25<sup>th</sup> September 1933). On that day, the South Sea became the IJssel Meer.

When reviewing the document, I want the reader to keep in mind that The Netherlands is a relatively small country (~ 16,412 square miles). As a result, the solutions are unique to The Netherlands and its geography as well as its size. Just for fun, consider the state where I live today, Texas is much bigger than The Netherlands (~ 268,597 square miles, so 16 times bigger). The solutions that work in The Netherlands would have to be tailored for Texas to be successful.

### **4. Dike Ring**

A dike ring closes off an area with dikes. The dikes protect against the sea and/or the river. The area within the dike ring can have additional dikes to protect critical areas inside the dike ring from flooding in case the primary dike rings are flooded.



Since The Netherlands is basically a sea delta, there is a lot of peat, and the Dutch used peat for many centuries. The picture on the left shows what peat looks like. More information about peat (including the picture) see <https://en.wikipedia.org/wiki/Peat>.

The Dutch dug out the peat and let it dry in the sun. They used it for building huts, roofs, agriculture, and heating. This increased the gap between the sea level and the land level in many areas. A further challenge is that the peat soil is weak and slowly sinking (estimated to be about 1 cm per annum). So, even without rising sea and rivers levels the gap is widening.

A significant number of inland dikes are made from peat. These dikes work well if the soil is wet. With the summers getting warmer, however, they cause the dikes to dry up and weaken. As a result, the Dutch have recently seen an increasing number of dikes break and had to make emergency fixes to avoid more dikes breaking.

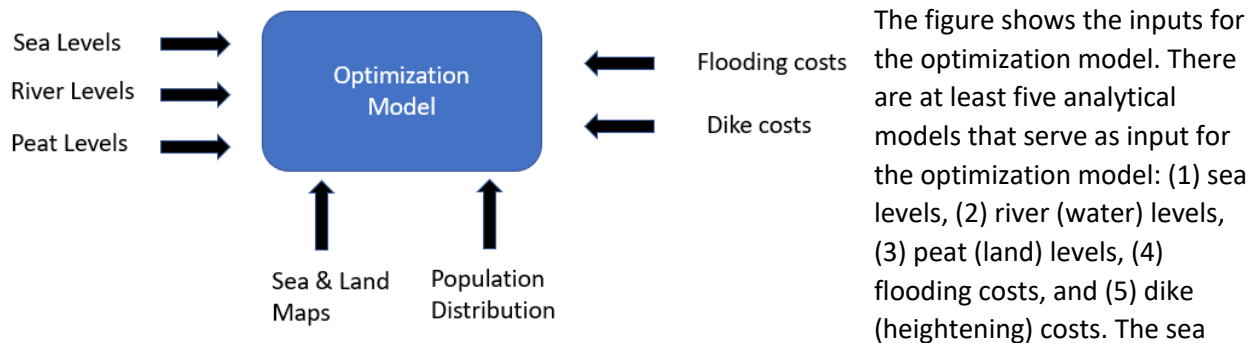
The protection level per dike ring is legislated in law. It is expressed in terms of in how many years a single flooding event can be expected. For example, 1/100 means that you can expect a flooding to take place once in a hundred years. Another way of thinking about this is to see it as a probability, i.e. the probability of flooding is 1% or 0.01.

The latter is probably the best way of looking at the risk of flooding, as it shows that a probability of 1% does not mean a flooding will happen only once every 100 years. Flooding could happen multiple times during a one-hundred-year period due to random effects. It, basically, is a binomial distribution, with

one hundred trials and  $x$  possible successes (i.e. flooding events), where  $x$  is between zero and one hundred.

Depending on economic assets and assets the protection levels for the dike rings vary from 1/10,000 to 1/250 in The Netherlands. In the US the protection level is 1/100. The protection levels for dike rings 13 and 14 is 1/10,000 (a factor 100 higher than what is used in US coast line).

### **5. Overall Model**



and land maps are static as well as the population distribution. Over time as the model gets better, these might be dynamic, and directly derived from other governmental databases.

The optimization model is targeted to make a twenty-year forecast of the optimal model. Instead of running it once, the model will be run annually to ensure that the selected and planned model is still optimal. The time horizon is chosen to allow a long enough time horizon to be able to heighten dikes over long stretches of land.

The length of the dike rings is around 3,500 km long (the coast line itself is about 800 km long). Increasing the dike rings by 10 cm takes many years to complete. The actual time needed depends on which dike rings need to be increased (type of dike ring, material used for dike ring, the state of the dike ring, and whether there is space to increase the dike ring). A project can do about 30 - 50 km per annum and the associated cost is measured in hundreds of millions of Euros per annum.

The original model was developed in a response to a new rule in 2008 that all the dike rings risk protection had to be increased ten-fold. The costs associated with this new rule were estimated to be 11.5 billion Euros. The optimization model was developed to see whether this cost could be brought down while still offering comparable protection against rising sea and river water levels. Note that the ten-fold increase costs around three million Euro per km.

With the background material covered now, we will now go into details of each of the analytical models. First, we will focus on sea levels, river levels, and peat levels. Then, we cover flooding costs and dike costs. With these analytical models in place, we then turn our attention to the optimization model.

### **6. Sea Levels**

Sea levels are measured using gauges spread along the North Sea coast line. The sea levels are influenced by the moon. With full moon the sea levels are higher than with new moon. There are moon tables that summarize the impact. These can be used to normalize measured sea levels for moon cycles.

The next thing that influences sea levels is the number of storms that take place any given month. There are more storms in the period from September to March. Especially, the combination of storms and full moon result in something called a spring tide and the sea level seen at such an event is higher than what occurs without a spring tide.

The last time, a spring tide broke the dikes in The Netherlands was 31 January 1953 when a significant part of the southern west part of The Netherlands was flooded. The damage costs were estimated to be about 15 per cent of gross domestic product (GDP). Around 1,800 people died as a result of the flooding. Other countries affected by the storm were UK, Belgium, Denmark and France. This led to a large investment in extra flood defense systems, called the Delta Werken (the Delta Works). The Delta Werken were completed in stages and the last piece was completed in 1998. The total cost at the time was estimated to be around 1 billion Euro (1959). Today's equivalent amount would be around 4 – 5 billion Euros.

Although the underlying data is strictly a time series model, I will not use a time series model like exponential smoothing or ARIMA, because they work best for short-term forecasting when data is stable without many peaks and valleys. We need long term forecasting, and the data has peaks and valleys. So, these models are not suited for our forecasting needs.

Instead, I will use a linear regression model that uses the following features: year, month, day, moon cycle correction, and storms correction. To train the model we will use data from the last twenty years.

Moreover, annually, we will test whether the measured sea levels values fall within the expected range of values predicted by the linear regression. If needed, the coefficients will be updated to reflect the latest annual data.

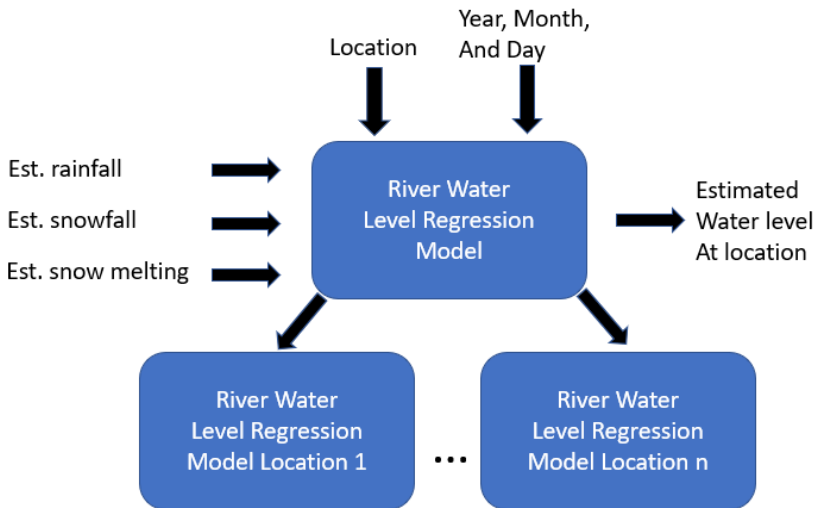
Now that we have covered determining the sea level, the next analytical model is determining the river water level at different locations.

### **7. River Water Levels.**

The main rivers' (Rhine's and Meuse's) water comes from the Alps, and the water levels are heavily influenced by the amount of snow that melts in spring, the amount of rain, and the locks and levees along the way that control the flow and path the rivers take. To keep things simple, I will assume that no new major investments will be made in the rivers' paths. As a result, I will only focus on the rain fall and amount of snow that melts as predictors.

Along the rivers there are gauges that measure the water levels of the river. Unlike the sea the rivers are non-tidal, and there is no ebb and flow. The position of the moon has little influence on the water levels as the amount of water is too small compared to the amount of water in the seas. There is a seasonal pattern, but it is much weaker than for sea levels. So, I will not try to correct for that by normalizing the data.

Using the river gauges' data for the last twenty-years, I will develop a linear regression model. Before building the river water level regression model, I would augment the data with the estimated amount of monthly snowfall in the Alps, the monthly estimated rainfall, and the estimated monthly amount of snow melted. With these predictors, I would build a regression model.



The linear regression model to predict river water levels would include the following predictors: year, month, day, estimated monthly rain, estimated monthly snow fallen, and estimated monthly snow melted. I will build a separate model for each main river. The year, month, and day, furthermore, will be combined into number of days since a start date.

As the model is used to predict water levels in the future the amount of rainfall, snow fallen and the snow melting rate is unknown. To estimate these in the future, I will develop three extra models to predict these. These will be tailor made (left hand side in figure above).

Rainfall is seasonal and needs to be normalized to determine the underlying trend. Then to predict a monthly rainfall amount in the future, we use a linear regression model for the underlying trend over time, and then correct for the season. To determine the regression model, we use data from the past twenty years.

Snowfall is similarly seasonal, and typically happens from November – April. Within this period there is again a seasonal element, and we need to normalize for that to determine the underlying trend. For the underlying trend we build a linear regression model using data from the past twenty years, and then correct for seasonal effect.

The snow melting rate is similarly estimated using data from the past twenty years. The melting happens mostly in spring, summer, and Fall. The snow starts melting when the temperature in the mountains rises above freezing point, and the amount that melts depends on how high the temperature in the Alps is. So, the regression model is based upon an estimated monthly temperature. The assumption is that within a year it is not possible to melt all the snow in the mountains (i.e. you cannot runout of snow in a given year). The temperature in the Alps is normalized for seasonal effects to determine the underlying trend (if any). We then build a linear regression model for the trend that we will use to predict the amount of water melting in a given month.

Note: the reader may think that the above model for estimating the river water levels is too complicated, but it turns out that the dike rings are most vulnerable to flooding from the rivers. There have been several near dike breaks in the last decade or so. And this trend is expected to increase. So, it is important to build a more holistic model that can track this trend and allow the WCBs to increase the dike heights if needed.



We covered sea and river levels. The last level that we need to cover is peat (land) level, which is next. After that we will turn our focus on flooding costs and dike costs.

### **8. Peat Land Levels**

The peat land levels sink by an average of 1cm a year. So, a simple estimation model will be used. All depths below sea level in The Netherlands are measured against “Normaal Amsterdam Peil” (Normalized Amsterdam Level) or NAP. The norm is based on the level of the IJ river in 1891 at which the dike would break. Most people in The Netherlands think of it as roughly the North Sea level. The depth is expressed in terms of NAP as follows  $\text{NAP} - 2\text{m}$  (meaning 2 meters below NAP).

For a limited set of locations, we know the depth level in NAP. For a given point we take the closest four NAP measurements and average the depth across these four points. With this estimate, we then correct for a 1 cm reduction per annum. Note that this approach is very similar to what the k-nearest neighbor (KNN) model for classification does. However, here we calculate an average and then adjust for sinking over time. KNN is used for classification, and that is not what we need here.

With this, we have covered all the different levels (sea, river, and peat). The next block of analytical models we will cover is flooding costs and dike costs. We start with the flooding costs.

### **9. Flooding Costs**

The flooding costs are modeled using econometric methods. These are basically multiple regression equations that depend on each other. This was not covered in our course, as it requires a more advanced level of mathematical modeling skills. The flooding costs are calculated using the below simplistic econometric model:

$$\text{flood damage} = f(\text{area flooded, damage})$$

$$\text{damage} = f(\text{building damage, infra damage, housing damage, lost production})$$

$$\text{building damage} = f(\text{agriculture, manufacturing, services})$$

$$\text{infra damage} = f(\text{transportation})$$

$$\text{housing damage} = f(\text{people})$$

$$\text{lost production} = f(\text{agriculture, transportation, manufacturing, services})$$

Exogenous variables: area flooded, agriculture, manufacturing, services, transportation, and people

Each function  $f$  is a regression model (could strictly speaking be linear, exponential, and so on). The exogenous variables are stored per dike ring and used to calculate the overall flood damage. Note that except for area flooded, all exogenous variables are Euro amounts.

The flooding costs are calculated using the above model. The area flooded is calculated in the optimization model and that is then used together with the exogenous variables stored for each dike ring.



The building damage takes agriculture, manufacturing, and services as input parameters and then multiplies each economic sector by an appropriate factor. The infra damage is calculated in a similar manner and multiplies the transportation by a factor.

The housing damage depends on the number of people living in a dike ring and assumes that as more people live in a dike ring the cost per person goes up (land more expensive, more expensive to build, and so on).

Lost production is based upon the different economic sectors, and assumes that each sector has a different time line to recover. For example, agriculture production is fully restored in one to two years, transportation is fully restored in five years, and so on.

Note that the above econometric model assumes no interaction with the dike rings. That understates the impact of flooding and needs to be added over time to better reflect reality. The model also does not cater for economic growth and as a result the estimated flood damage is underestimated.

Next we look at dike costs, and after that will discuss the optimization model that uses the above analytical models as input.

### **10. Dike Costs**

The dike costs are calculated using a linear regression model for each dike ring. So, the regression model is a tree of regression models. The dike cost is calculated using four input variables: amount to raise, length of dike to raise, type of dike, and which dike ring. The regression model uses different unit costs per type of dike. The dike unit cost is then multiplied by amount to raise and length of dike to determine an estimate of dike raising costs.

The data for the regression models needs to be provided by the WCBs. Studies have been carried out in the past that provide an analysis of the relationship between these variables and found the relationships between the variables to be linear.

Periodically (every five years), the unit costs need to be reviewed and updated.

With this we have completed all five analytical models that serve as input for the optimization model. The optimization model is next.

### **11. Optimization Model**

Each dike ring has an assigned risk. The risk is used to determine what the expected water level is to breach the dike. There is a branch of statistics that analyzes extreme values with the assumption that the underlying process does not change (see [https://en.wikipedia.org/wiki/Extreme\\_value\\_theory](https://en.wikipedia.org/wiki/Extreme_value_theory)). There are two different methods for calculating extreme values: (1) Annual Maxima Series (AMS) and (2) Peak Over Threshold (POT). We will use the AMS method and generate annual maxima for the duration of 30 years.

The optimization model assumes that the assigned risk is translated into a height needed to cause flooding. So, we can calculate the height needed to breach a dike ring given a specified risk. The risk is the average number of events, e.g. 1/10,000 means that the dike ring will be breached every 10,000 years. The assumption is that the breach events are Poisson distributed, which means that the time between dike ring breaches is modeled using the exponential distribution.

Below is a first attempt at structuring the overall optimization model using Python pseudo code. It is incomplete as the original project took almost six years to complete and involved a large number of professional experts. But, it does give a sense of what is involved to make the overall optimization work. The first part is running a simulation for one year:

```
def simulate_one_year(dr, year, max_water_levels):  
  
    flooding_area = 0  
    for loc in all_locations_around_dike_ring(dr):  
        if max_water_levels[loc][year] > dike_height[(dr, loc)]:  
            flooding_area = flooding_area +  
                flood(loc, dike_height[(dr, loc)], max_water[loc])  
  
    if flooding_area > 0:  
        return flooding_cost(area, sectors[dr], nr_of_people[dr])  
    else:  
        return(0)
```

The main simulation is given below and leverages the `simulate_one_year()` routine outlined above:

```
for dr in dike_rings:  
    best_costs[dr] = Inf  
    best_height_incr[dr] = 0  
  
    for increase in 0:50 (step=5):  
  
        for loc in all_locations_in_dike_ring(dr):  
            max_water_levels[loc]=calculate_AMS_series(  
                risk, mean_level[(dr, loc)], 30)  
  
        total_damage = 0  
        total_dike_cost = dike_cost(increase, type_of_dike[dr])  
  
        for run in nr_runs:  
            for year in 1:30:  
                total_damage = total_damage + simulate_one_year(  
                    dr, year, max_water_levels)  
  
        avg_damage = total_damage / nr_runs  
        total_cost = total_dike_cost + avg_damage  
  
        if total_cost < best_cost[dr]:  
            best_cost[dr] = total_cost  
            best_height[dr] = increase  
        else  
            break out of increase for loop  
  
print(dr, best_cost[dr], best_height[dr])
```

The above optimization model assumes that it only makes sense to calculate the total cost per dike ring for a new height if the total cost of the previous height increase is lower than any of the previous total costs calculated for the dike ring. If this is not the case, it stops the simulation for the dike ring and assumes that it has found the optimal cost and height increase. This is a heuristic to minimize the time needed to calculate the optimal solution for each dike ring.

### **References**

- 1) S. Madrikas, S.C. Wheelwright, and V.E. McGee. *Forecasting, Methods and Applications*. Wiley: 1983, 2<sup>nd</sup> ed. Print.
- 2) H.M. Wagner. *Principles of Operations Research, with Applications to Managerial Decisions*. Prentice Hall International: 1975, 2<sup>nd</sup> ed.
- 3) J.M. Kind. *Economically efficient flood protection standards for The Netherlands*. Journal of Flood Risk Management 7 (2014), 103 - 117. Electronic copy can be found at [https://www.deltares.nl/app/uploads/2014/12/kind2014\\_JFRM1.pdf](https://www.deltares.nl/app/uploads/2014/12/kind2014_JFRM1.pdf)
- 4) Author unknown. *Extreme value theory*. Wikipedia.com, year unknown. [https://en.wikipedia.org/wiki/Extreme\\_value\\_theory](https://en.wikipedia.org/wiki/Extreme_value_theory). Last accessed April 16, 2019.
- 5) Author unknown. *Amsterdam Ordnance Datum*. Wikipedia.com, year unknown. [https://en.wikipedia.org/wiki/Amsterdam\\_Ordnance\\_Datum](https://en.wikipedia.org/wiki/Amsterdam_Ordnance_Datum). Last accessed April 16, 2019.
- 6) I. Haigh. *Sea Level Return Periods: What are they and How Do We Use them in SurgeWatch?* Surgewatch.com, 2016. <https://www.surgewatch.org/sea-level-return-periods-what-are-they-and-how-do-we-use-them-in-surgewatch/>. Last accessed April 16, 2019.
- 7) Hoogheemraadschap van Rijnland. *Flood Control in The Netherlands, A Strategy for Dike Reinforcement And Climate Change*. Rijnland Water Control Board, 2009. <https://www.rijnland.net/downloads/floodcontrolrijnland-1-1.pdf>. Last accessed April 16, 2019.
- 8) Author unknown. *Flood control in The Netherlands*. Wikipedia.com, year unknown. Last accessed April 17, 2019.
- 9) Author unknown. *Peat*. Wikipedia.com, year unknown. <https://en.wikipedia.org/wiki/Peat>. Last accessed, April 17, 2019.