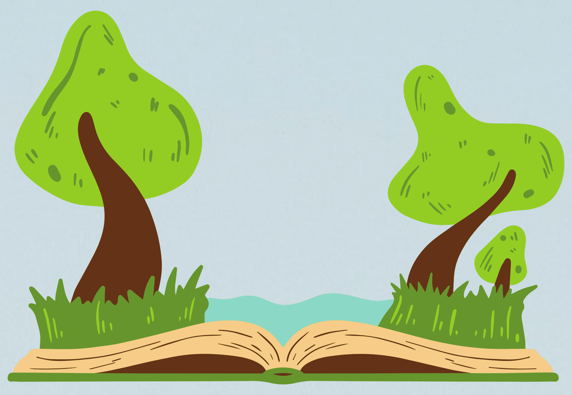
How can the vegetation growth be improved using weather data?

Group 11 - Tran Thi Ngoc Bich, Dong Hyeok Lee, Luke de Vos, Hidde Fokkelman, Boyd Smetsers

01

Introduction



The availability of fresh water is of great importance to our daily life. Therefore, a failure in water management could heavily affect people, farms, factories, and the entire ecosystem. To tackle the problem Waterschap Aa en Maas, the company responsible for managing water in streams in the area around the rivers in the north-east of Noord-Brabant, introduced RAM. RAM is a dashboard which shows the amount of water being held back by vegetation and sends notifications to users in order to mow the vegetation out of the stream.

02

Stakeholders

Regarding the dashboard, the stakeholders and their interests are defined as:

- **Employees of Water Aa en Maas:** minimized workload and money spent on mowing, prevention of flooding
- **Property owners in the area**: well-managed streams, no damaged properties
- **Area manager**: protecting endangered species, prevention of flooding

03

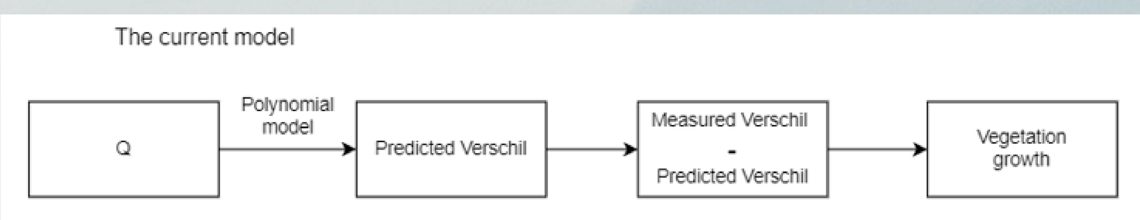
Methodology

Our project aims to answer the research question:
" **How can the vegetation growth estimation be improved using weather data?"**

To achieve such a goal, we will first go through the current model and data that Aa en Maas uses. Subsequently, problems of the current model will be discussed with the use of statistical tests and time-series plots. We will then conclude by presenting an evaluation of the problem, as well as giving suggestions for the stakeholders.

04

The current model of Aa en Maas



The basic idea of this model: there is more vegetation during summer -> the relationship between Verschil and Q during summer and winter are different -> more water being held back in a river.

In detail, the current model:

1. Take Q during summer and predicts Verschil based on the relationship found during winter.
2. Compare this to the actual measured Verschil during summer.
3. The difference between predicted and measured Verschil is a proxy of vegetation growth, assuming that this difference is caused by vegetation growth.

05

Data understanding & data preparation

- Negative values in features data (measured Q and Verschil in different weir compartments) are considered as measurement errors.
- Features data contain a considerable number of measurement errors, especially in Verschil measurements.

	VERSCHIL	Q
NEGATIVE VALUES	23.1%	0.01%
POSITIVE VALUES	76.9%	99.9%

- Before analyzing data, measurement errors were deleted.

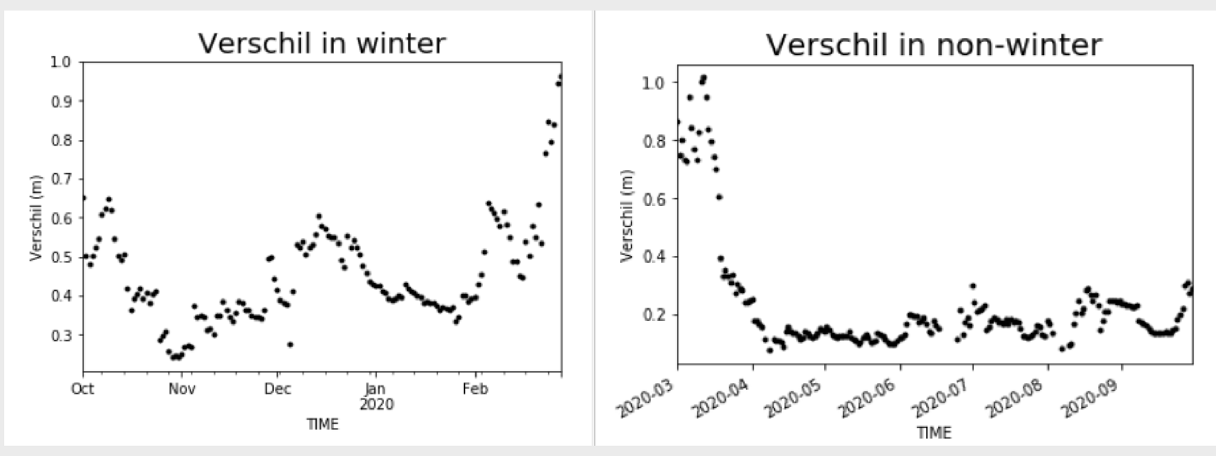
06

Problems of the current model

1. Actual vegetation growth can not be measured.
-> Not feasible to evaluate the model accuracy.

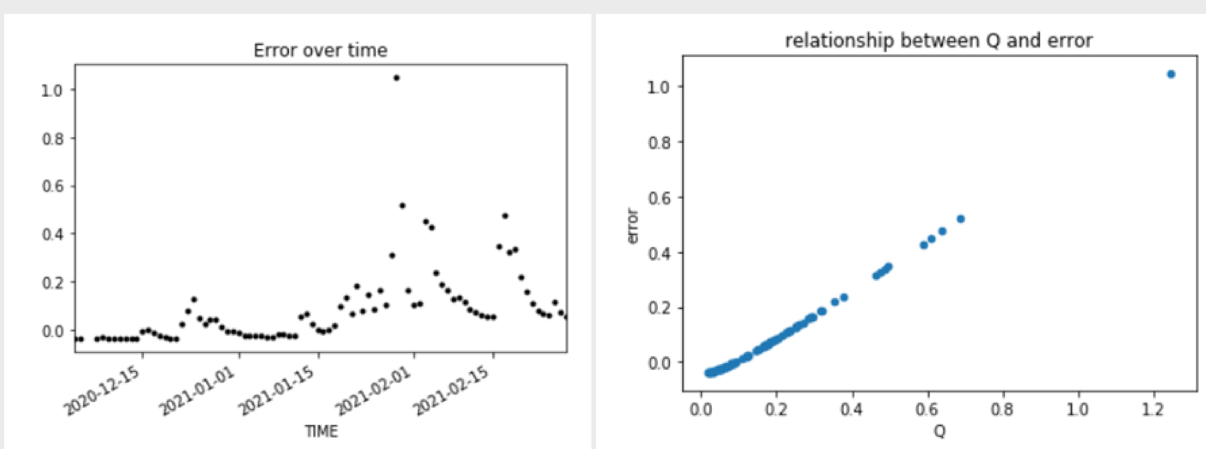
2. The significant variable is omitted in the current model

- Comparison of Verschil between winter and non-winter



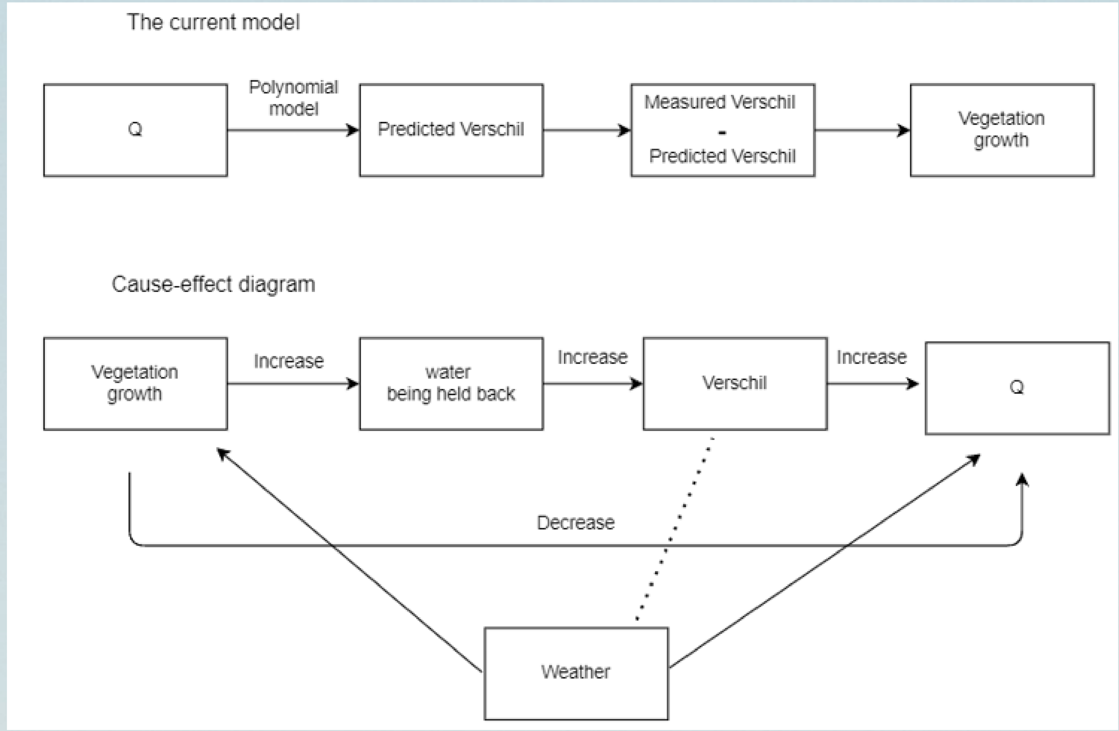
The current model assumes that Verschil is increased due to vegetation growth
-> Verschil is supposed to be largest during summer.
However, in the most weir compartments, it is found that Verschil reaches its peak in March and the variance in the winter is bigger than in the summer.
-> Verschil is affected more by another factor than vegetation growth.

- Heteroskedasticity and Endogeniety



The problem of heteroskedasticity and endogeniety, presented in the left and right plot respectively, indicates that the model omitted an important variable. They were found in a polynomial model of the current model that predicts Verschil given Q.
- The variation of the error term is not constant and varies a lot in January and February.
- Q has a strong correlation with the error term
-> a significantly important variable is omitted in the model.

3. Reversed causal relationship



The above diagram shows the causal relationship between variables in the current model and in reality. That is, the current model uses a responsible variable (Q) as an explanatory variable.
The independent and dependent variable are switched in the current model.
-> does not provide a true model
-> current model would always output biased results.

-Weather factors like wind or rainfall influence both vegetation growth (Peiyu Zhang, 2019)[1] and Q (Rodolfo Nóbrega, 2013)[2], but this is not considered in the current model. Yet, there is no existing research that shows the relationship between Verschil and weather.

07

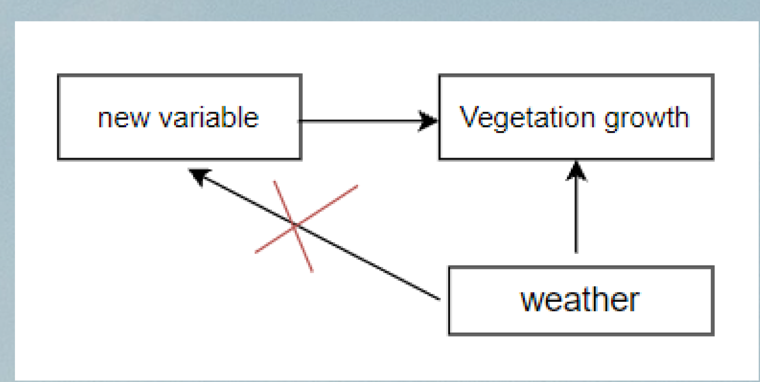
Limitation

No data of weather factors such as temperature, wind and rainfall in each river
-> It is not feasible to estimate how much weather affects vegetation growth and Q in Noord-Brabant.
-> An empirical proof to show that the current model omitted a significant factor (weather) can not be provided.

08

Suggestion

-Model improvement



It is necessary to know how much weather affects Vegetation growth and Q; however, it is too cost-consuming, as sensors detecting wind, temperature and rainfall would have to be installed in each river. Thus, it is recommended to **find a variable that is related to vegetation growth, but not to weather**. For example, finding animals or microorganisms that are related to vegetation in river in ecology system and counting the number of them in a specific area would be a good proxy to estimate vegetation amount in a river.

-Dashboard
There should be notification on the dashboard that the accuracy of the model is not known so that users are prevented from being too dependent on it.

References

1. Peiyu Zhang, Effects of Rising Temperature on the Growth, Stoichiometry, and Palatability of Aquatic Plants
2. Rodolfo Nóbrega, Understanding the relationship between rainfall and river discharge: trends in an Amazonian watershed