https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

The dynamics of peak head responses at Dutch canal dikes and the impact of climate change

Bart Strijker[1,2], Matthijs Kok[1,2]

[1]Hydraulic Engineering, Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands
[2]Risk and disaster management Unit, HKV Consultants, Lelystad, the Netherlands

Correspondence to: Bart Strijker (b.strijker@tudelft.nl)

Abstract. The management of water and flood risk levels in low-lying polder regions depends on the performance of canal dikes. Heavy rainfall can lead to peak hydraulic heads within the dikes affecting their stability, which can induce dike breaches. Variations in head responses and head statistics are both relevant for regional flood risk analysis of canal dike systems. This study examined the dynamics of peak heads in canal dikes on a national scale using time series models calibrated on observed heads. Various model structures are evaluated and a nonlinear model performed the best. These models were used to simulate long-term head time series. Subsequently, dike clusters were identified based on the coincidence of peak heads, allowing for the identification of dikes where peaks are caused by (dis)similar types of rainfall events. The differences and similarities in peak head response were related to physical dike characteristics. While no single significant relationship emerged, the soil type combined with the width of the dike appears to be important factors influencing the variation in head responses. However, the presence of the same soil type and dike widths in multiple clusters indicates that these characteristics do not yield a definitive outcome for the head response. Moreover, peak head statistics across various dikes were derived and indicated that extreme and yearly load conditions are relatively close to each other, with a median decimation height of only 15 centimeters. The head statistics are affected by climate change, characterized by increasing winter precipitation and summer evaporation. By 2100, extreme peak heads are expected to occur between 3 times less and 8 times more frequently, depending on the climate scenario and the type of canal dike.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

1. Introduction

Catastrophic dike failures have occurred throughout history due to various causes, such as storm surges, ice drifts, and extreme weather conditions like heavy rainfall or drought, and have involved several failure mechanisms, including overflow and overtopping, external erosion, piping, and inner slope instability (Özer et al.,2019b; Van Baars & Van Kempen, 2009). For many dikes along rivers and coasts, inner slope instability occurs due to the infiltration of water into the dike body and its foundation, leading to higher head levels and pore-water pressures, decreasing effective stresses and decreasing shear strength of the soil (Sharp et al., 2013; Frank et al., 2004; Ridley et al, 2004). The infiltration of water can be caused by high water levels against the dike as well as heavy rainfall (Rikkert, 2022; Van Baars & Van Kempen, 2009). For dikes with controlled water levels and limited fluctuations, such as canal dikes, the infiltration of water caused by heavy rainfall can be more significant and is considered a primary mechanism of dike failure. Canal dikes are among others present in polders, which can be found in coastal and alluvial lowlands all over the world, like Netherlands, Bangladesh, Vietnam, England and China (Martín-Antón et al.,2016; Morton & Olson, 2018; Lendering et al., 2018; Tran & Weger, 2018; Triet et al., 2018; Manh et al., 2013; Warner et al., 2018). The water levels in these reclaimed areas are artificially regulated by an internal drainage system with canals, where water levels can reach several meters above the surrounding terrain (see Fig. 1). This makes the surrounding low-lying areas vulnerable to floods in the event of a canal dike breach. These canal dikes are found not only in polders but also along internal waterways and irrigation canals worldwide, where major dike failures have occurred throughout history (Gildeh et al., 2019).

To manage flood risks in an embanked area, the performance of dikes plays a crucial role. The failure probabilities of the individual elements or dike stretches contribute to the reliability of the canal dike system and the flood risk level in an area (Vanmarcke, 1977; Kanning, 2012; Jongejan et al., 2020). The Netherlands has an extensive system of canal dikes critical for managing its low-lying geography and protecting against flooding with a total length of more than 10,000 km (Pleijster and van der Veeken, 2015). In general, the failure probability of a dike system increases with the total length of the dike system, due to partial correlation or independency between different individual elements (Kanning, 2012; Vrijling & van Gelder, 2002). This phenomenon is known as the length effect, which is caused by both dependencies between resistance and loading conditions in space, and differs for each failure mechanism. The inner slope instability is a failure mechanism of the Dutch canal dikes that contributes significantly to the calculated failure probability and flood risks in the polders, where the load primarily consists of high hydraulic head peaks (Lendering et al., 2018; Rikkert et al., 2022; Van Baars and Van Kempen, 2009). The variations of the canal water levels in the drainage systems of Dutch polders are small (up to tens of centimeters), while the observed hydraulic head fluctuations are an order of magnitude larger than water level fluctuations. Therefore, the fluctuations of the hydraulic heads are primarily driven by rainfall and evaporation. Whether the head levels at two nearby canal dikes both result in extreme load conditions after a heavy rainfall event, depends on the head response of

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

the dikes. Variations in the head response can cause extreme loads to occur after different weather events and influence the system's reliability.

Figure 1 Two examples of canal dikes in the Netherlands: Duifpolder polder is shown at the left and the Drooggemaakte Geer- en Kleine Blankaardpolder is depicted at the right. In both images, the canal has permanent high water levels close to the crest level (on the left side) and the low-lying polder is mainly used for agriculture (on the right side). The head difference between the canal water level and water level in the polder are 2.7m (Duifpolder) and 4.3m (Drooggemaakte Geer- en Kleine Blankaardpolder). The red tubes protect the measuring equipment and the piezometers used to measure the hydraulic head. Photos by EURECO/Cyril Liebrand, 2022.

To calculate the failure probability of individual canal dikes and dike systems, information about peak head responses is essential. Currently, there is limited understanding regarding the variability in head responses and head statistics in canal dikes, which is partly due to the lack of measurements and the extensive computation time required for groundwater models. Multiple studies have modelled the effects of rainfall and evaporation on the phreatic surface in dikes using different approaches (Rikkert, 2022; Jamalinia et al., 2019; van Esch, 2012). Multi-year measurements of hydraulic head levels are often lacking in dikes, which makes modelling exercises difficult to validate. Furthermore, the validation of models is hindered by the heterogeneity of dikes and the unknown field hydraulic conductivities, potentially influenced by burrowing animals, plant roots or cracks, and resulting flow paths. This makes that there is also little known about the effects of climate change on head statistics and failure probabilities of canal dikes. Future climate projections indicate increasing temperatures with summers becoming hotter and drier, and winters warmer and wetter and are expected to affect the stability of slopes. Although the impact has been studied for both natural slopes (e.g. Moore et al. 2010) and earthworks (e.g. Huang et al., 2024; Rouainia et al. 2020), for canal dikes, with different boundary conditions and head dynamics, the effects are studied to a limited extent. This study aims to assess the dynamics of peak hydraulic heads in canal dikes on a national level, caused by heavy rainfall events, by analysing the variation in head responses and head statistics. Furthermore, the impact of climate change on the head statistics is quantified, which is an indication of how flood risks in Dutch polders are changing in the future.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

The approach developed for assessing the dynamics of hydraulic heads is shown in Fig. 2. The basis consists of head observations and time series models with impulse response functions (IRF). After collecting head observations in canal dikes, time series models are set up to simulate hydraulic heads in canal dikes, using precipitation and evaporation as the explanatory time series. Several model structures are assessed and the model structure that has the overall best performance is used. Only models that meet the specified reliability criteria (minimum goodness of fit and sufficient time series length compared to model parameters) are selected, resulting in a set of models, that explain the fluctuations of the observed head levels in different dikes with a variety of head responses (step 1). These models are forced with 30 years of precipitation and evaporation time series, corresponding to different climate scenarios (current and future conditions). This is done to obtain extended head time series that encompass more extreme events (step 2), facilitating the analysis of both the dynamics and statistics associated with extreme occurrences. The variation in head responses is quantified by analysing the coincidence of the head peaks across canal dikes, selected using the Peaks-Over-Threshold method, and classifying different clusters of dikes with similar head responses. The variation and similarity of head responses are related to several physical dike characteristics, like soil type and dike profile to search for explanations of the differences found (step 3). Finally, a generalized Pareto distribution (GPD) is fitted to the head peaks that describe the probability of occurrence of peak values. The variations and properties of the head statistics are analysed, as well as the impact of climate change (step 4).

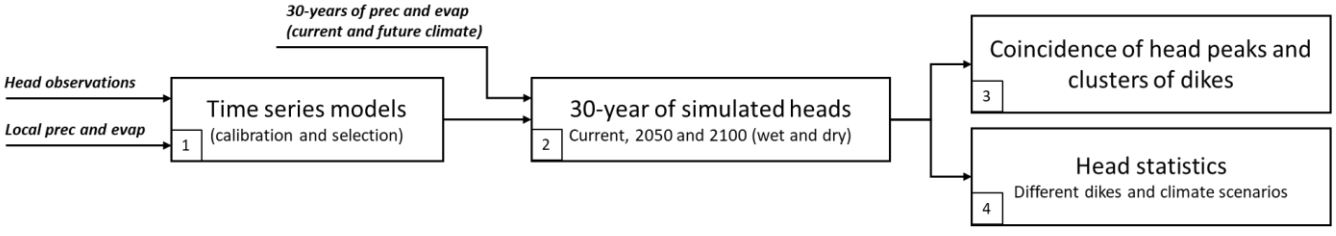

Figure 2 An overview of the approach and its accompanying steps.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

2. Study area and data

2.1. Dutch canal dike system

The Netherlands has lived with the threat of flooding for centuries and this vulnerability is controlled by a regulated system of flood defences, where distinction can be made between primary and regional flood defences. The Primary flood defences are located along major bodies of water, like the sea, the major rivers and the large lakes, while the regional defences are located along drainage canals, man-made lakes, compartmentation dikes and smaller rivers. In general, a breach in regional defences will have a smaller impact than a breach in the primary defences, though it can still have considerable consequences. This study focusses on a subset of the regional flood defences, namely the canal dikes. The canal dikes are primarily located in the Western and Northern part of the Netherlands, where the polders are located (see Fig. 3). These cultivated lowland areas serve as agricultural land as well as for human settlement. Many cities, villages and small communities are situated throughout the polders. The water inside the polder is separated from the outside water by the primary flood defences, and the polder drainage systems manage the water inside the primary flood defences. The water is managed by discharging or pumping the polder water into canals (also called the *boezem* in Dutch), after which the canals release the water into the outside water, either naturally or using pumps (Steenbergen et al., 2009). The water levels in the canals are higher than the polder water levels and dikes are located along these canals, resulting in a flood vulnerability within the polders (see Fig. 3). The subsurface of canal dikes is characterized by low-permeable soils that mainly consist of clay and peat and in the past many canal dikes breached (Van Baars & van Kempen, 2009).

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

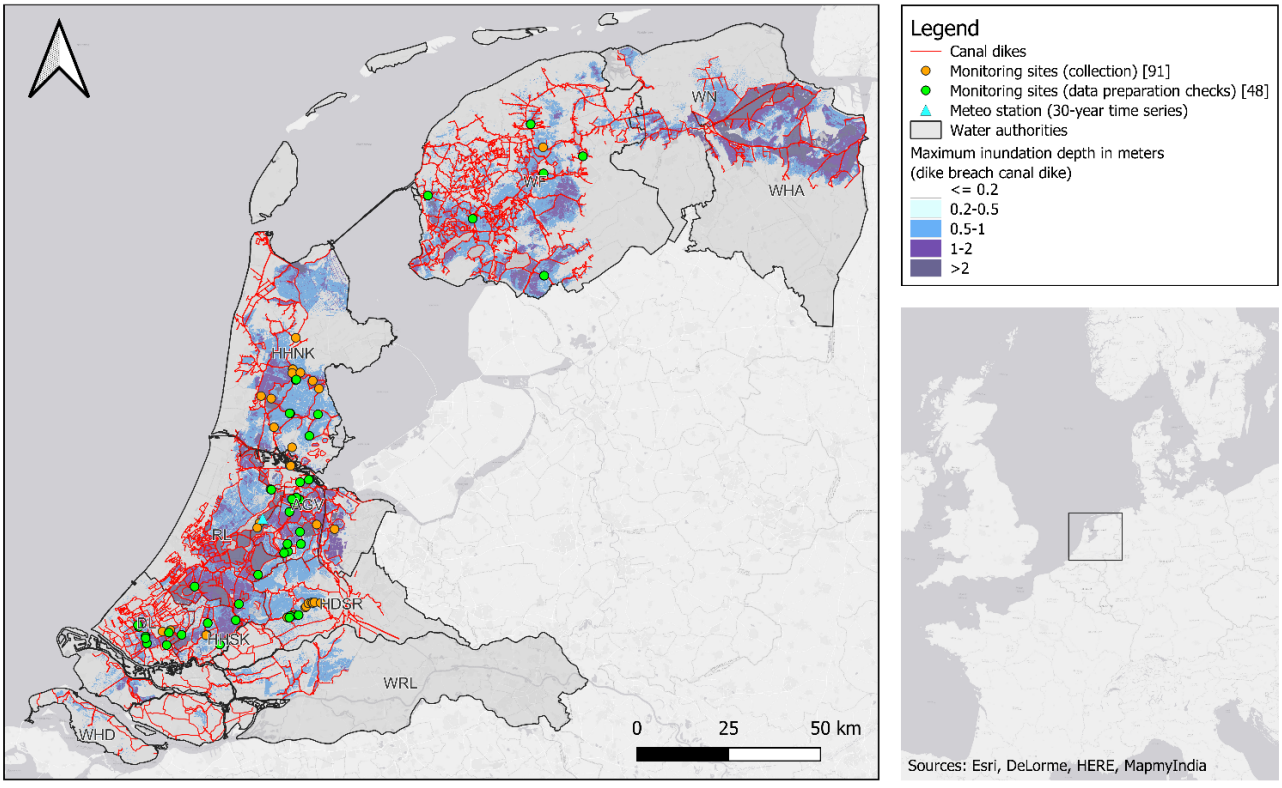

Figure 3 Overview of the study area with the canal dike system (subset within the regional flood defence system that are located in the provinces South-Holland, North-Holland, Friesland, Groningen and Utrecht). Monitoring sites are indicated by circle markers, with orange and green denoting collected data and the date checked and utilized for further analysis, respectively. Maximum inundation depths are depicted to illustrate potential flood risk in the polders.

2.2. Data collection

2.2.1. Head observations and preparation checks

Head observations in canal dikes were collected and received from seven Dutch regional water authorities, namely Hoogheemraadschap Schieland & de Krimpenerwaard (HHSK), Hoogheemraadschap Delfland (DL), Hoogheemraadschap Rijnland (RL), Hoogheemraadschap Hollands Noorderkwartier (HHNK), Weterksip Fryslân (WF), Waterschap Amstel, Gooi and Vecht (AGV) and Hoogheemraadschap De Stichtse Rijnlanden (HDSR). The crest levels of the dikes with head observations vary from around NAP+1,5 m to NAP-2,5 m (all elevations are relative to the Dutch reference level called NAP, which is approximately mean sea-level) and polder water levels ranging from NAP-2 m to NAP-6.5 m, highlighting that many canal dikes lie below sea level.

In total, 258 head observations at 91 monitoring sites were collected in this study (see Fig. 3). Multiple head observations can be located at one monitoring site, where piezometers are aligned within a dike cross-section, for example measurements

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

in the crest, inner slope and toe of the dike. Monitoring sites located near each other, for example 20m, may exhibit varying responses due to differences in the underlying soil conditions. Consequently, no filters were applied based on the proximity of monitoring sites. The heads were measured with automatic pressure loggers, with hourly measurement intervals in the period between 2006 and 2023 and were resampled to daily mean values for the analysis. For further analysis, only time series that are longer than 2.5 years were selected. Additionally, any time series exhibiting visual anomalies or odd behaviour such as pronounced drift, absence of fluctuations, or inexplicable jumps were removed from the dataset. The monitoring sites were also checked, whether the head observations are located in a dike, since sometimes the dike is more like a quay without a slope. The head dynamics in a dike is complex and also location-dependent within the dike profile, since the head in the outer crest can respond differently than head levels in the inner slope. Only head observations that are located in the talud zone or mid-slope of the dike were used (see Fig. 4). This is the area between the top of the dike at the polder side (inner crest) and the toe of the dike and where most variations in groundwater levels can be expected. This is because it is farthest from the regulated water levels in the canal and polder, which are maintained at the target level.

The resulting dataset consists of 108 head time series at 48 monitoring sites, consisting of phreatic head levels measured in the dike body. The average length of the time series at the monitoring sites is 5.0 years and varies between approximately 3 and 9 years.

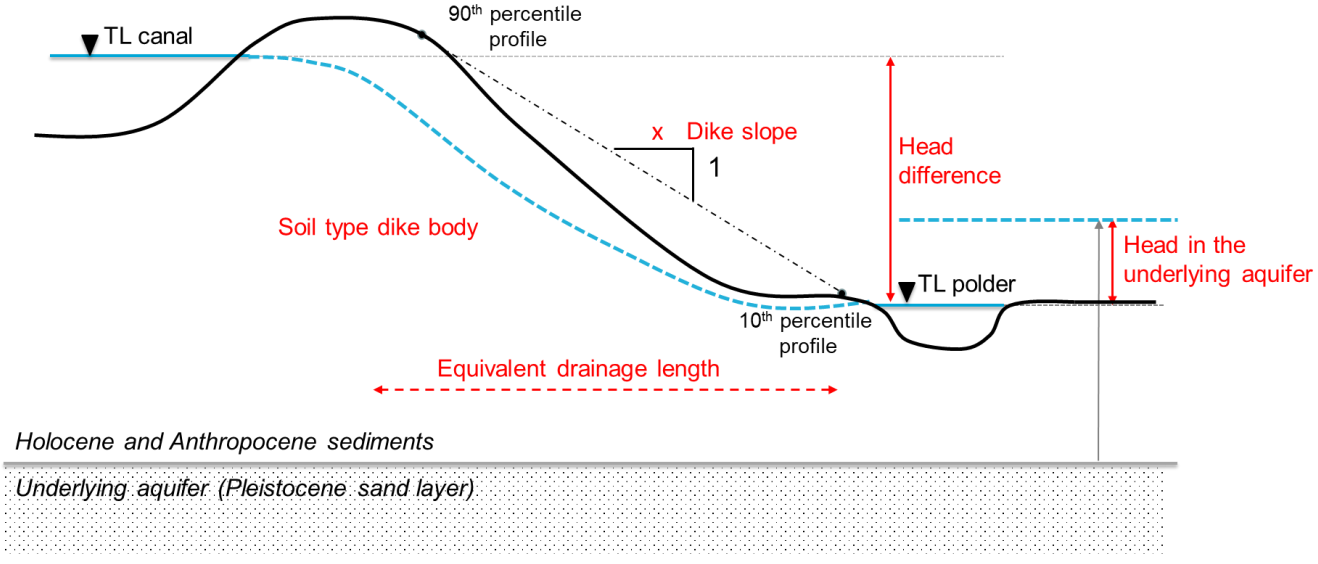

Figure 4 The canal dike with regulated water levels (TL=target level) on both sides (canal and polder) and fluctuations of the phreatic surface, caused by precipitation and evaporation. In this study, the groundwater levels in the talud zone of the dike are analysed.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

2.2.2. Precipitation and evaporation

The Dutch meteorological institute KNMI provides several data products about weather and climate. In this study, two data products for rainfall and evaporation (on daily basis) were used, serving the purpose of 1) getting the best estimate of the historic local weather conditions at the head observation sites (for the calibration of time series models) and 2) getting long term time series of the weather representing the current and future climate situation in the Netherlands (for extending the head time series). First, the local historic weather conditions were derived from rainfall and evaporation maps of the Netherlands, namely the RADAR-derived precipitation amounts and inverse distance weighting (IDW) interpolated evaporation amounts based on the KNMI ground stations. The RADAR-derived precipitation amounts were derived by combining and correcting RADAR-images from the Dutch precipitation RADAR stations and calibrating them with ground observations of precipitation (Wolters et al. 2023). The evaporation maps give estimates of the daily Makkink reference evapotranspiration derived from ground observations of the global radiation and the average daily temperature (De Bruin, 1987). Secondly, to derive extended head time series encompassing more extreme events, there was made use of 30-year of precipitation and evaporation time series corresponding to different (current and future) climate scenarios (Van Dorland et al., 2023). These time series are a representation of the climate and not an estimate of the actual weather. Since this study focuses on variations in peak heads caused by different head responses and not caused by the spatial variations in the nature of the load (dimensions of weather events), the 30-year time series at one location was used. The station Aalsmeer, close to Amsterdam, was used which lies rather central in the Western Netherlands (see Fig. 3). In total nine 30-year time series were used, corresponding to different time horizons (2005, 2050 and 2100), emission scenarios (high SSP5-8.5 and low SSP1-2.6) and regional response (wet and dry-trending).

2.3. Physical dike characteristics

The relationship between the head response of a dike and several physical dike characteristics is examined. The characteristics include the soil type of the dike body, the head difference (water level difference between the canal and polder), the dike slope, the equivalent drainage length and the head in the underlying Pleistocene aquifer (see Fig. 5). Dike slopes can be seen as a measure of the hydraulic gradient within the dike body, which influences the horizontal groundwater flow and the head response. In addition, steep dike slopes can reduce recharge, as it can increase surface runoff that limits the possibility for water to infiltrate into the dike. The slope of a canal dike was not always obvious, since the dikes don't always have a typical geometry, where the profile is irregular and crest, berm and toe of the dike are not clear. The dike slope was obtained from an elevation map of the Netherlands (AHN4 with a horizontal resolution of 0.5m) and was calculated by referencing two percentile points on the elevation profile (10th and 90th percentile), as an approximation of the slope between crest and toe of the dike. The target levels (TL) of the canal and polder were obtained from the local water authorities, which were used to calculate the head difference. The equivalent drainage length was determined by dividing the

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

head difference by the estimated dike slope, providing an estimate of the width of the dike. The head levels in the underlying aquifer were based on the Netherlands Hydrological Instrument (de Lange, 2014), where the average head in the period between 2011 and 2018 is calculated. This head level was referenced to the target level of the polder to get an indicator of upward or downward seepage in the dike body. The subsurface of the dike was based on borehole descriptions or CPTs at the monitoring sites. In the absence of soil investigation, a detailed three-dimensional model (GeoTOP) of the upper 30 meters of the subsurface of the Netherlands were used (Stafleu et al., 2012).`

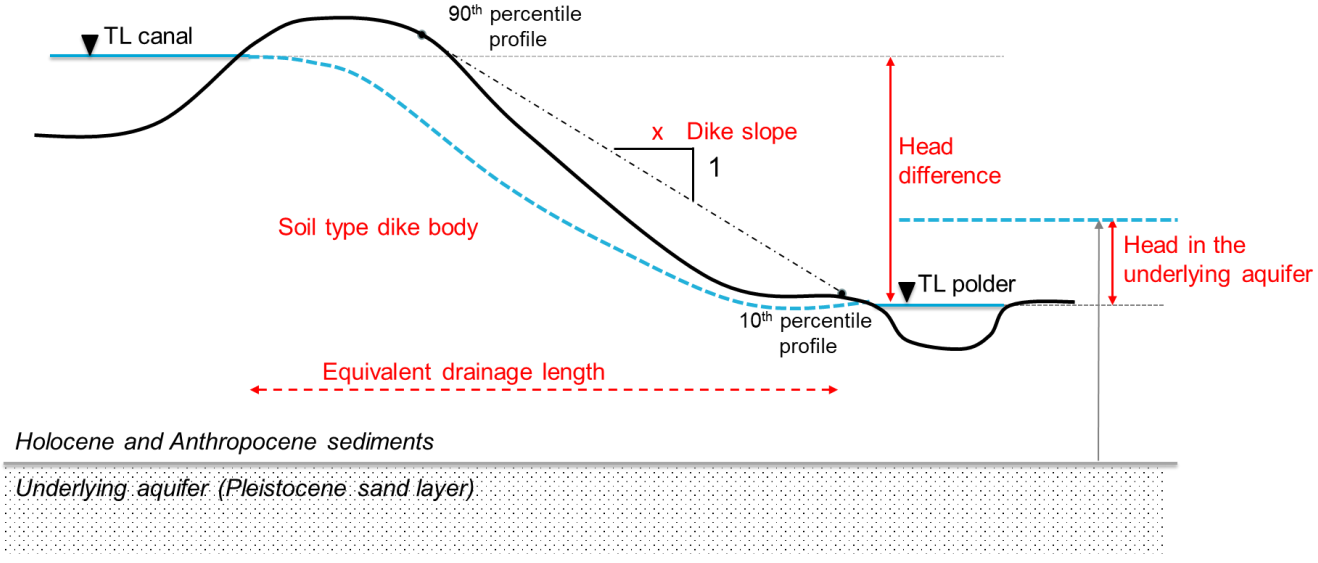

Figure 5 A simplified cross-sectional profile of a canal dike showing the five dike characteristics (in red) considered. The vertical grey arrow indicate the head level in the underlying aquifer

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

3. Method

3.1. Groundwater modelling in dikes

Several modelling approaches can be used to model the hydraulic head in canal dikes. Commonly used approaches comprise numerical groundwater models, like Hydrus-2D, PlaxFlow and MODFLOW, that are solutions to (systems of) differential equations that describe the flow of groundwater (Šimůnek et al., 1999; McDonald & Harbaugh, 2003). These approaches need detailed information on material behaviour for both unsaturated and saturated soils. In the case of canal dikes, Van Esch (2012) showed that it remains difficult to reproduce observed hydraulic heads in dikes, because of the uncertain conceptualization of the subsurface, spatial heterogeneity, and applied boundary conditions. Time series modelling is a simplified and abstract representation of head fluctuations at one point resulting from the complex 3D movement of water in the dike (Bakker & Schaars, 2019). It is a data-driven approach that can estimate the contribution of independent stresses (rainfall, evaporation, water levels, etc.) on the observed head levels derived exclusively from observed data. This approach is used in this study.

3.1.1. Time series models

The basic principles of time series analysis comes from the statistical sciences (Box & Jenkins, 1970). Transfer function noise (TFN) modelling is a subfield within time series analysis that aims to convert one or more input series into an output series using a statistical model. Von Asmuth et al. (2002) presented a novel form of TFN models that relies on the concepts of convolution and predefined impulse response functions (IRF) and is used for many applications within groundwater science. Predefined response functions are used to estimate the effect of a unit pulse of a driver, like precipitation, on the head response. The head response is simulated through the convolution of various drivers with their response functions. The basic model structure of a TFN model to simulate heads may be written as:

$$h(t) = \sum_{m=1}^{M} h_m(t) + d + r(t)$$

where $h(t)$ is the observed heads, $h_m(t)$ is the contribution of stress m to the head, d is the base elevation of the model, and $r(t)$ are the residuals. The number of stresses M in each model varies based on the selected model structure. The contribution of stress m to the head is computed through convolution:

$$h_m(t) = \int_{-\infty}^{t} S_m(t)\theta_m(t-\tau)d\tau$$

Where S_m is a time series of stress m, and θ_m is the associated impulse response function. A variety of impulse response function can be used to simulate the effects of certain stresses, where commonly used impulse response function are the scaled Gamma and the exponential response function. The exponential response function is the simplest response function with only two parameters and may be used for stresses that have an immediate effect on the head, like the shallow head levels in the canal dikes (up to a few meters below ground level). Together with the small geohydrological dimensions (the

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

dimensions of the dike are typically only a few tens of meters), a relatively rapid response is expected, which is confirmed by measurements where head observations respond quickly to rainfall, despite the presence of low-permeable soils.

3.1.2. Various model structures

The head fluctuations are explained by the contribution of various hydrological stresses that are convoluted with a response function. Various model structures can be chosen, incorporating different stresses with different response functions. A stress can also be the combination of multiple stresses. The net groundwater recharge *R(t)* is frequently used as stress that is the derived from rainfall *P(t)* and Makkink reference evaporation $E_p(t)$ as inputs (e.g., von Asmuth et al. 2008):

$$R(T) = P(t) - fE_p(t)$$

Where the parameter f is the so-called evaporation factor used to scale the reference evaporation. This model is referred to as a linear recharge model, after which the net recharge *R(t)* is substituted by S_m in and then convoluted with a response function to determine the impact of recharge on the head.

Next to this linear recharge model, there are also non-linear recharge models that model recharge using soil-water balance concepts, like FlexModel or Berendrecht, where the model becomes more complex with additional model parameters (Collenteur et al., 2021; Berendrecht et al., 2006). These models can account for the nonlinear response of the head to precipitation and evaporation, where the recharge can be modelled as connecting reservoirs, like an interception and root zone reservoirs, including short-term water retention in the soil in the calculations. Another non-linear model structure is the Threshold autoregressive self-exciting open-loop (TARSO) model (Knotters and Gooijer, 1999). This structure consists of two regimes (upper and lower) which are separated by a threshold. Each regime has its own exponential response function with corresponding drainage levels, but only when the head reaches the upper drainage level, the upper response function becomes active. Therefore, this model can be useful when the head response is different above a certain head level.

The head time series were modelled using time series models as implemented in the Python open-source package Pastas (version 1.5.0) (Collenteur et al., 2019). Time series models were set up, assuming that the head dynamics in canal dikes are primarily influenced by rainfall and evaporation, and only these two stresses are included in the model.
This assumption is supported by the observation that canal water levels fluctuated minimally (on the order of tens of centimeters) during the measurement period, while the observed averaged head range was more than one meter. Additionally, the models demonstrated an overall good fit. In total, four different model structures were employed (see Table 1). The model structure with the highest averaged goodness-of-fit across all models was used for further analysis. Selecting a single model structure ensures consistent comparison across different locations and simplifies the interpretation of results.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

Table 1 Model structures employed in the study and their characteristics, like type of recharge model, impulse response function (IRF) and number of fitted parameters.

	Recharge	IRF	Number of fitted parameters
Lineair – Exp	Lineair	Exponential	4
Lineair – Gamma	Lineair	Gamma	5
Flex-model	Non-lineair	Exponential	7
TARSO-model	Non-lineair	Exponential	7

To illustrate the performance of different model structures, Fig. 6 gives an example for the monitoring site at Molenlaan (site DL4). The linear models (exponential and gamma response functions) are not able to model the head response for the full range of head levels, as can be seen in the scatterplot. For this location, it appeared that the head response was nonlinear. The TARSO-model provides a better fit across the entire range of head levels.

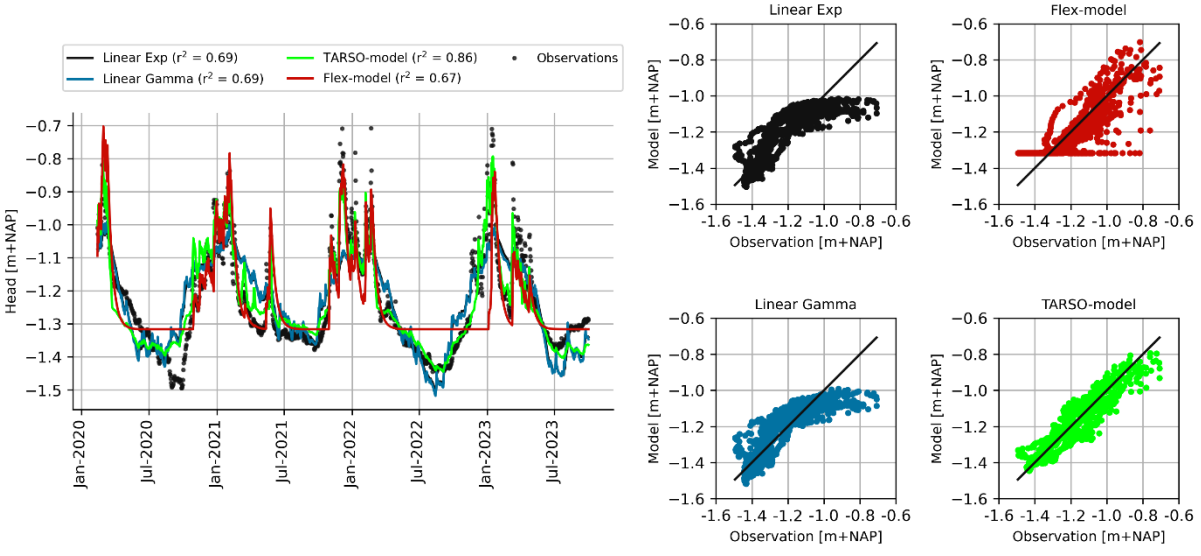

Figure 6 The performance of different model structures at Molenlaan (DL4). The left graph shows the simulated and observed head levels and on the right four scatterplots of the model structures showing the observed and simulated daily head levels (the black line indicates the 1:1 line).

3.1.3. Model calibration and selection

The time series models are used to characterize and simulate the geohydrological behaviour of canal dikes with a single deterministic parameter set. For every head time series, time series models were set up, where the full series were used for calibration to maximize data utilization. This was done because the length of the available time series was limited and to avoid the missing information of the head response when splitting up the data for model calibration and model validation.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

Although, validating models indicate that the models are performing well and are adequate to achieve good quality model predictions in post-validation model application, several studies showed that the most robust models are achieved when all data are used for calibration (Shen et al., 2022; Arsenault et al., 2018), which is in line with the goal of this study. Overfitting is mitigated by employing time series models with up to 7 parameters, and using head calibration data with more than 1000 data points. The model parameters were estimated using the least squares method, employing a warmup period of 10 years and without incorporating a noise model to represent the residuals. After choosing the best model structure, the calibrated models of that structure were evaluated using two criteria. These two criteria were used to determine whether a model is reliable for further analysis:

- Goodness-of-fit: the model's goodness of fit, measured by the R-squared (R^2), must be equal to or greater than 0.7 in the calibration period, indicating a minimum acceptable level of fit. This is also known as the coefficient of determination which is a measure of how well observed outcomes are explained by the model.
- Response time: the 95% response time, the time it takes for 95% of the influence of an impulse (groundwater recharge) to dissipate, must not exceed the length of the measurement series. Time series should be long enough to cover the head response in order to estimate parameters accurately (Knotters & van Walsum, 1997). This criterion eliminates models for which the time series data isn't long enough considering the model parameters.

When there are multiple reliable models of head time series available at one monitoring location, the model that provided the best fit was selected as the representative model for that location.

3.2. Peak selection and extreme value analysis

Extreme value theory concerns the behaviour of the extremes of a series of observations or simulations, like hydraulic heads. Accurate estimates of exceedance probabilities of hydraulic head levels are crucial for the stability and safety of dikes and risk management. In order to do so, the occurrence of peak values within a time series are selected and their magnitude and frequencies are analysed. For these analysis, the peak events need to be filtered out of the time series such that the peaks of every are mutually independent from each other in time. In the case of hydraulic heads, peaks are seen to occur in groups: an extremely high hydraulic head is likely to be followed by another since the groundwater system of dikes contains autocorrelation or memory. The peaks-over-threshold (POT) method was used to select independent peaks by setting a threshold value above which events are considered extreme and a time window for identifying clusters separated by a given period, after which only the highest peak within each cluster was selected. The threshold was set at a low level, namely the 90th percentile of the analysed head series. Therefore, many peaks can be selected, but afterwards, only the highest number of peaks were selected equal to the length of the time series, making that on average one peak in every year was selected. This procedure was followed because the primary focus is on extreme values. A time window of 30 days was chosen to guarantee the selection of independent peaks.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

Next, a generalized Pareto distribution (GPD) was fitted to the peaks that describes the probability of occurrence of peak values. The GPD has three main forms, the Type I, Type II, and Type III distributions, which differ in the number of parameters and the flexibility of the tail behaviour. The cumulative distribution functions of the GPD are defined by:

$$f(x) = \begin{cases} 1 - \left(1 + \left(\xi \frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\xi}}, & \xi \neq 0 \\ 1 - \exp\left(-\frac{x-\mu}{\sigma}\right) \quad , & \xi = 0 \end{cases}$$

where x is the hydraulic head and the three parameters of the GPD are called the scale (σ), shape (ξ) and location (μ) parameters. When ξ=0 the GPD is equivalent to the exponential distribution. In the case of peak hydraulic heads, the exponential distribution is preferred due to its relative simplicity and its suitability for the process, as suggested by visual inspections of the tail of the distribution and the head peaks.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

4. Results

4.1. Modelled head responses in canal dikes

4.1.1. Reliable time series models

For each of the 108 head time series across 48 monitoring sites, time series models with various model structures were created and calibrated. Of the four different model structures, the TARSO model has the highest average goodness-of-fit (an average r^2 of 0.70) and performed as the best structure for 79% of the models. The second best model is the linear recharge model with the Gamma response function (an average r^2 of 0.66), while the Flex model structure performed the worst with an average goodness-of-fit of 0.32. So, the head levels in Dutch canal dikes can be best modelled with a non-linear recharge, outperforming linear models. This non-linear behaviour can be the result of various soil layers in the dike body, each with distinct hydraulic properties, and changes in infiltration rates or nonconstant storage capacities of the unsaturated zone during the dry season (Knotters & de Gooijer, 1999). Next, the calibrated TARSO models are evaluated using the reliability criteria (goodness-of-fit and response time) and for every monitoring site, the model that meets the reliability criteria and has the highest goodness-of-fit score is selected. Out of 48 monitoring sites at 38 sites (79%) models were developed with r^2 of 0.7 or larger, which is considered as one reliability criterion. The other reliability criterion, stating that the 95% response time should not exceed the length of the measurement series, dropped the number of sites with reliable models to 35 (73%). Plots of the measured and simulated heads during the measuring period for all selected monitoring sites are shown in Fig. A1 in the Appendices. This set of models is used for further analysis.

The model results of one of the selected time series models are shown in Fig. 7, to illustrate the model outcome and performance. The simulated heads show a good fit with the observations with a R-squared value of 0.84. The model overestimates the heads in the first summer, which was a particularly dry summer. These differences may be due to inaccurate precipitation data used in the model (as summer precipitation can be highly localized) or disturbanes during installation affecting the head levels. However, these factors were not investigated as the focus lies on high head levels. The other dry summer in 2022 is modelled more properly. As can be seen in the scatterplot, the simulated and observed heads are close to the one-to-one relationship between the two variables. The (block) response functions for both the upper and lower regime, determined by the time series model, are shown in the lower right graph in Fig. 7. The peak of the block response represents the increase in head level that would occur if the recharge were 1 mm in one day. Therefore, a block response peak of almost 4 means that the head level would rise by 4 mm with 1 mm of rainfall. The 95% response time, further on referred to as the response time, is a measure of the memory of the groundwater system and, in this case, represents the time it takes for 95% of the influence of an impulse (groundwater recharge) to dissipate.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

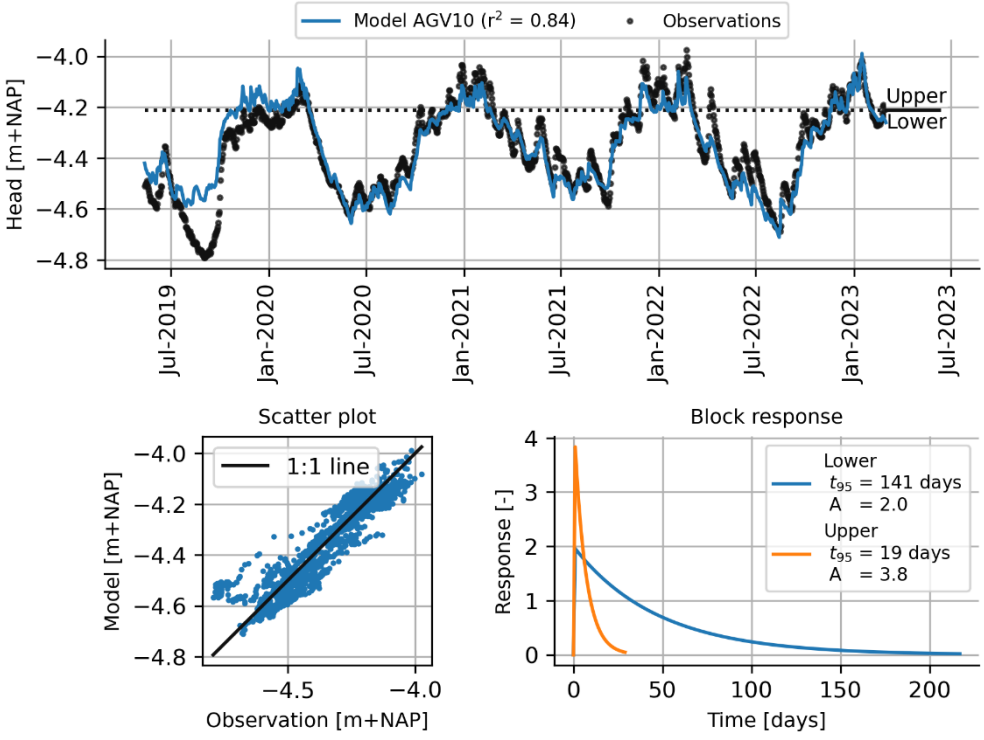

Figure 7 Illustration of the simulated head levels with the time series model in comparison with the observations for dike AGV10. The graph on top shows the observations and model series where also the upper and lower regime are indicated with the black dotted line and the bottom left shows the scatter plot of observed and simulated heads. The lower right graph shows the impulse response functions of the upper and lower regime.

4.1.2. Characteristics of impulse response functions

The head dynamics of various dikes are quantified by examining the characteristics of impulse response functions, by means of the peak of the block responses and the response times. Figure 8 shows these characteristics for both the upper and lower regime of the TARSO-model, where the colors indicate in which water authority region the dike is located. The response times of the upper regimes of the dikes are generally short, mostly ranging from about 2 to 50 days, with two exceptions where response times exceed 250 days. These two exceptions have a very small peak of the block response compared to other dikes, which have a large variation in the peak response of the upper regime, with values reaching up to nearly 15. This variation may result from different soil storage capacities and the redistribution of infiltrated water within the dike, which can accumulate in the talud zone causing large head responses. The response times of the lower regime have more variation than the upper regime, with most dikes ranging between 100 and 600 days. Meanwhile, the peak block for the majority is below 5. Initially, the variation of head responses across regions suggests no specific pattern, likely due to the heterogeneous subsoil conditions within the canal dike system, as shown by the random color distribution in Fig. 8.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

Two things stand out related to the non-linearity that apply to almost all locations when analysing the response functions, namely that the response time for the lower regime is longer than the upper regime (32 out of 33 sites) and the peak of the block response for the upper regime is higher than the lower regime (30 out of 33 sites). The difference in response time can be caused by head gradients (as a driver of groundwater flow) that depend on the head level itself, the presence of various soil layers with different permeabilities, or the fact that groundwater comes close to the surface, alters the storage coefficient, and/or horizontal redistribution of water occurs. The lower peak of the block response for the lower regime can be caused by non-linear processes in the unsaturated zone, where more water can be stored after dry periods with low head levels (Berendrecht, 2006).

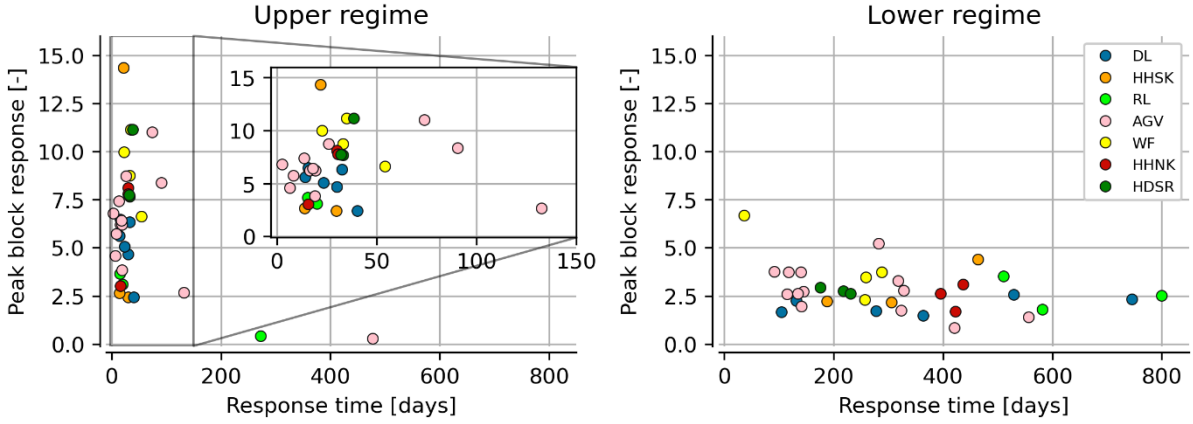

Figure 8 The characteristics of the impulse response functions at 33 canal dikes (peak of the block response and 95% response time) for the upper and lower regime, where the colors indicate the various water authority regions where the dikes are located. The inset at the upper regime shows a zoomed-in view.

4.2. Variation in peak head responses

4.2.1. Coincidence of head peaks and clusters

For dike stability and flood safety, the extreme peak head levels in dikes are of main interest. Among others, differences in peak behaviour between various canal dikes are relevant to estimate regional risk levels and understand the loading conditions of dikes, like what type of weather events cause peak head levels in dikes (e.g. different durations of precipitation events), and water authorities can also act accordingly. As seen in Fig. 8 the model parameters and resulting head response differ across dikes. Whether varying model parameters lead to different peak head behaviours was examined by analysing the coincidence of simulated head peaks of various dikes, all subjected to the same 30 years of rainfall and evaporation. Two peaks at different dikes are assumed to coincide, and are driven by similar weather patterns, when the peak moments occur on the same day. Based on the percentage of coinciding peak heads among the dikes, dike clusters are identified which show similar peak behaviour. The clusters are estimated using the k-means clustering algorithm (Hartigan & Wong, 1979), where the number of clusters (k) has to be given beforehand and is based on the mean Silhouette score of all samples (Rousseeuw, 1987) and the "elbow method", as implemented by Yellowbrick (Bengfort & Bilbro, 2019). The elbow method suggests that

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

the optimal number of clusters is k=4, while the Silhouette score indicates that either k=2 or k=4 could be optimal, with scores of 0.442 and 0.439 respectively. The selected number of clusters k=4 will be used for the remainder of the study. These four classified clusters of dikes (called cluster A, B, C and D) have distinctive peak behaviour as shown in the coincidence matrix (see left graph in Fig. 9). While the three clusters A, B and C are less distinct from each other with still moderately high percentages of coincident peaks, sometimes exceeding 50%, cluster D has a very distinct peak behaviour. This cluster consists of only two dikes of which the time series models have deviant impulse response functions; longer 95% response times and the smaller amplitudes of the block response in the upper regime compared to the other clusters (upper right graph in Fig. 9). Furthermore, distinctive peak behaviour between clusters is strongly influenced by the 95% response time of the upper regime (as depicted in the right graph in Fig. 9). In clusters A, B and C, the average 95% response time of the upper regime is 16 days, 32 days and 77 days, respectively. This twofold increase in response times for each cluster results in distinct rainfall events leading to peak heads.

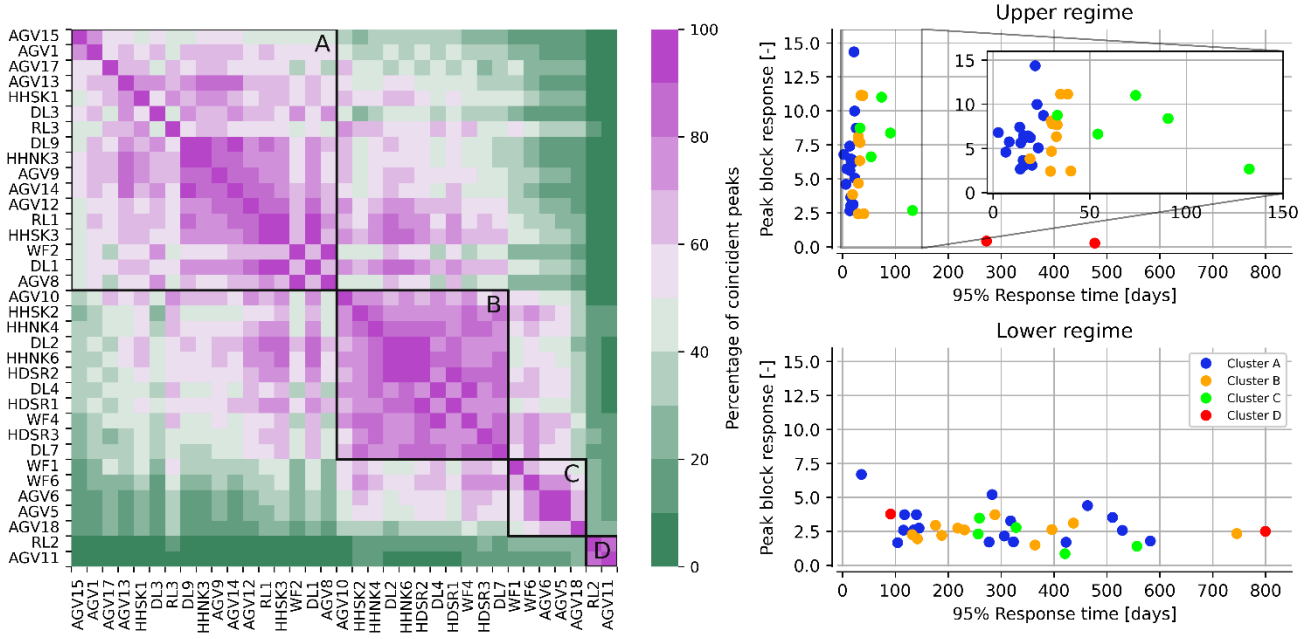

Figure 9 The coincidence matrix of the head peaks at the canal dikes, including four identified clusters (left graph). Within every cluster, the locations are ranked based on the 95% response time. The characteristics of the impulse response functions of the dikes within every cluster are shown in the right graph.

The differences in peak behaviour were also analysed by examining the average timing and distribution of peaks throughout the year (see Appendix A2). For all canal dikes, peaks typically occur in the winter half-year, shifting further into winter from cluster A to cluster D. This shift is due to different head responses, with later winter peaks linked to longer rainfall events and earlier winter peaks linked to shorter rainfall events.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

4.2.2. Clusters in relation to physical dike characteristics

The relationships between physical dike characteristics and the 95% response time, as well as clusters of dikes, were examined. Various statistical tests were employed to assess these relationships by calculating the p-value for different types of variables, both categorical and continuous; Wald Test for comparing two continuous variables, the Wald Chi-Squared Test is used for two categorical variables, and the Kruskal-Wallis Test is employed for one continuous and one categorical variable. The analyses of relationships with soil types were limited to clayey and peaty dikes, as the dataset includes only one sand dike and Dutch canal dikes are generally composed of these materials. Table 2 shows the p-values for the relationships between physical dike characteristics and both the clusters and the characteristics of the impulse response functions (upper and lower regime), showing that no significant (p-values greater than 0.05) relationship was found. However, the soil type appears to play an important role in the identification of the clusters. The statistical tests used to calculate the p-values can be inappropriate due to violations of test assumptions, small sample sizes, multiple comparisons and data dependency, all of which can lead to misleading statistical conclusions (Greenland et al., 2016). Therefore, the physical dike characteristics within every cluster are also visually analysed and, as expected from the statistical tests, there is a large variation of dike characteristics within every cluster, see Fig. 10. The clayey and peaty dikes appear in all dike clusters, with most clayey dikes in cluster B (see also Fig. A3 in the Appendix). Furthermore, small dikes, with drainage lengths less than 20 meters and often associated with steep slopes, are found only in Clusters A and B with the smallest 95% response times. Furthermore, the median equivalent drainage length of the clusters increases from Cluster A to Cluster D, associated with increasing 95% response times. This indicates the importance of dike geometry, where shorter distances from a drain (the canal or ditch) lead to faster dike drainage and smaller response times. This can be explained by the fact that the hydraulic gradient, which drives water towards the drain, increases with shorter distances. Yet, determining the cluster to which dikes with certain dike characteristics belong isn't straightforward, since all clusters include dikes with various characteristics.

Table 2 P-values of the relationship between the considered physical dike characteristics (rows) and the clusters of dikes and the 95% response time and peak block response for both the upper and lower regime (columns).

	Clusters	Upper regime		Lower regime	
		95% response time	Peak block response	95% response time	Peak block response
Soil type	0.05	0.23	0.23	0.23	0.23
Head underlying aquifer	0.14	0.52	0.23	0.11	0.67
Equivalent drainage length	0.19	0.19	0.31	0.22	0.39
Dike slope	0.35	0.31	0.90	0.96	1.0
Head difference	0.42	0.13	0.10	0.07	0.06

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

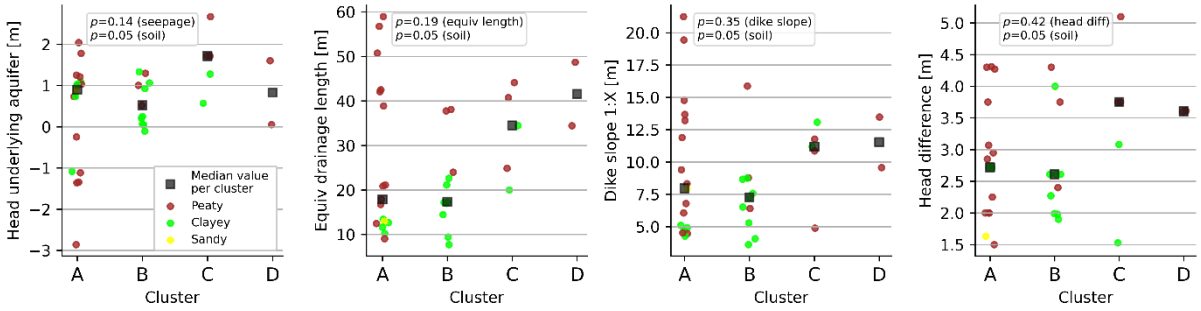

Figure 10 The occurrence of dike characteristics across clusters (A, B, C and D), where data points color-coded by soil type and median values per cluster are marked.

4.3. Statistics of head peaks

4.3.1. Head statistics and decimate height

The selected hydraulic head peaks in the 30-years simulated head time series are used to estimate the extreme value distribution of peak head levels by fitting an exponential distribution. The statistics of hydraulic head levels at a dike along the Beemster polder (HHNK3) are used as illustration, of which the resulting frequency line is shown in the left graph in Fig. 11. The decimate height, which is defined as the increase in head level that occurs when the return period of the head level experiences a tenfold increase, is a relevant characteristic of the load and dike safety (Wojciechowska, 2015; Schweckendiek, 2014). Low decimate heights indicate a small difference between significant heads and normally or yearly occurring heads. Measurements in dikes with small decimate heights makes more reliable deductions about the head levels during significant loads that improve the dike safety assessment than dikes with high decimate height (Wojciechowska, 2015). The decimate height at HHNK3 is approximately 7 cm, while across various dikes, the values range from nearly 0 to 50 centimeters with a median decimate height of 15 centimeters (as seen in the right graph in Fig. 11). Higher decimate heights are found at dikes with higher peak block responses in the upper regime, in combination with longer response times. For Cluster C, these decimate heights are smaller than 5 centimeters, mainly because of the very small peak block responses. For these dikes, the head levels that occur on average once every 1000 years are only expected to be 15 centimeters higher than the yearly occurring head levels. For dikes that are marginally stable, this increase in head level can still be a trigger for dike failure, but for the estimation of failure probabilities, this provides useful information for survived load events.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

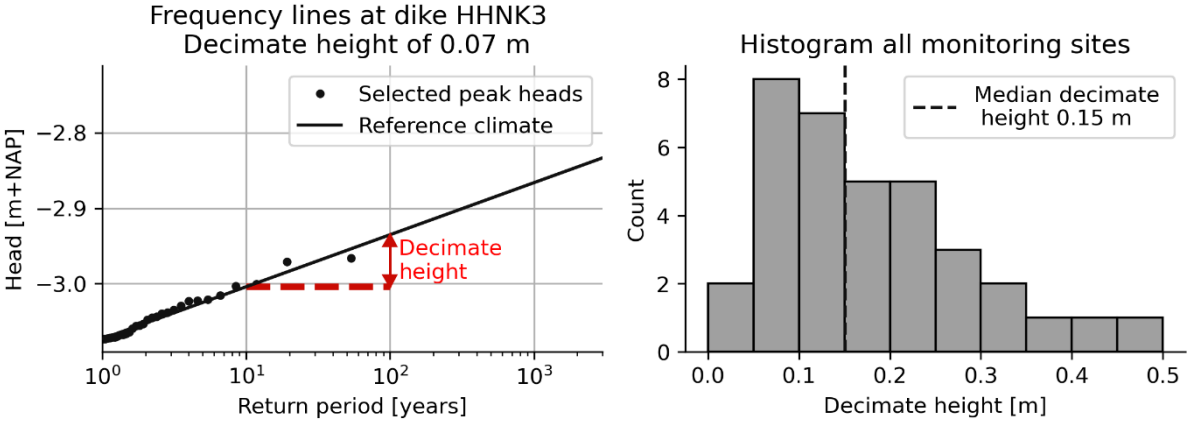

Figure 11 Left: Frequency lines of hydraulic head levels at a dike along the Beemster polder (HHNK3) based on precipitation and evaporation corresponding to the current climate. Right: a histogram of the decimation heights of all dikes considered of which the median decimation height is 15 centimeters.

4.3.2. Impact of climate change

The impact of climate change, which in this study includes changes in precipitation and evaporation, on the head statistics is quantified for different time horizons and climate scenarios (emission scenario and regional response). Precipitation and evaporation series from these different time horizons and climate scenarios are used to simulate head levels the time series models. Subsequently, peaks are selected and an extreme value distribution is fitted. For one location, the resulting frequency lines of various climate scenarios in the year 2100 are shown in Fig. 12 (left graph). For this location, all climate scenario's result in more extreme head levels or head levels to occur more frequently. In the Hn scenario, which has the largest increase in precipitation according to Van Dorland et al.(2023), the head level at a return period of 100 years only increases with 10 centimeters. Due to the small decimate height, the original head level, occurring once every hundred years, is expected to happen once every 15 years, indicating a sixfold increase in frequency. This factor is called the probability factor and expresses the impact of climate change on the frequency of occurrence of the head level with a return period of 100 years. The probability factor across all dikes ranges from about 3 times less frequent to 7 times more frequent across the four climate scenarios in 2100, as can be seen in the right graph in Fig. 12. In the year 2050, the probability factor varies between 4 times less and 2 times more frequently with an average value of 1, indicating the frequency of extreme head levels is not changing. The variation could not be linked to the clusters of dikes. However, the head statistics of dikes with longer response time of the lower regime appear to be less impacted by climate change, which can be explained by the fact that these dikes can dry out more during the dryer summer. The water storage of a more dried-out dike body is larger, causing heavy rainfall events after a dry summer to be stored effectively in the dike, causing head peaks to occur less frequently. This can be seen in the climate scenario with drying trends (Hd and Ld) in 2050 and 2100, where the precipitation increases, but the frequency of extreme head level occurrences remains largely unchanged. For climate scenario's in 2100 with a wetting trend, the frequency of occurrence of extreme load levels is increasing.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

485

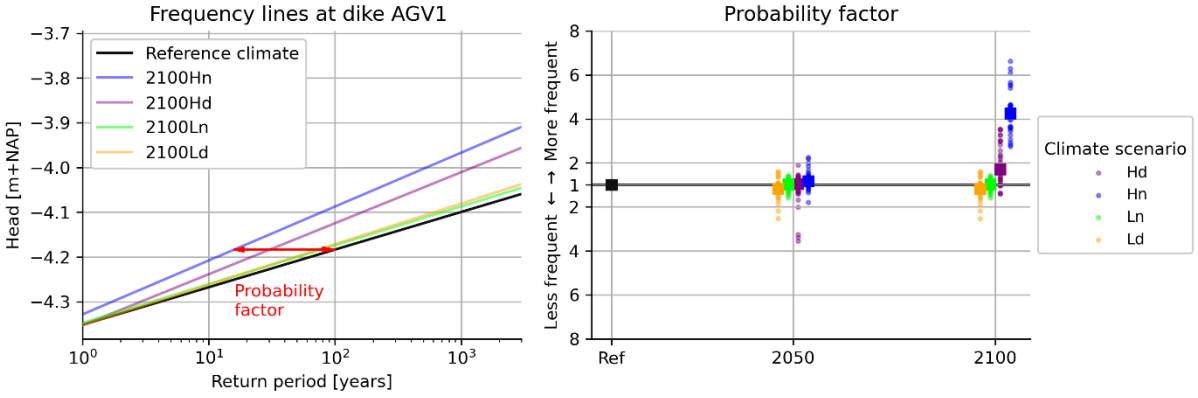

Figure 12 Left: The frequency lines of head levels for various climate scenarios at the dike in a polder south of Amsterdam. Right: the probability factor of all dikes (●) and the median value (■) for different time horizons and climate scenarios.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

5. Discussion

5.1. Limitations

The groundwater data used in this study are the result of contacting many water authorities. However, given the extensive canal dike system with thousands of kilometers of canal dikes and the heterogeneity of dike bodies, the number of available observations in terms of the number of monitoring locations and length of the measurement series is limited. Nevertheless, this paper tried to show the usefulness of observations and hopefully more measurement campaigns, lasting multiple years, will be organised. The head responses that can be adequately modelled using time series models were selected, applying reliability criteria to filter out potentially unreliable measurements. At the same time, reliable measurements can be filtered out, that do not fit our chosen models. The head response of canal dikes may be influenced by processes that are not accounted for in the selected models. Furthermore, as with any model, estimating hydraulic heads outside the measured range of head variations are uncertain, depending on the behaviour and changes in head responses in more extreme situations. Deviating head response behaviour can be expected due to, for example, additional non-linear processes, like surface run-off by excess rainfall amounts, which may make head levels to be less extreme than modelled.

In this study we tried to assess variation in head responses between different dikes, however, there can also be variation in head responses within the talud zone. At fourteen monitoring sites, there are multiple reliable models within the talud zone. For ten out of fourteen monitoring sites, the time responses of the upper regime lie only 10 days apart from each other, indicating very similar peak behaviour. Time series models are calibrated using head observations with an average length of 4.5 years. Knotters & Walsum (1997) found that the minimum length of observations needed to calibrate a TFN model adequately for shallow water-table depths is 4 years. This minimum number also depends on the model structure chosen and can differ for the TARSO-model with 7 fitted parameters. In this study, the model uncertainty associated with these time series models is not explicitly considered.

Different clustering methods exist, which can result in different clusters (Everitt et al., 2011). In this study, only the k-means clustering algorithm is considered. Additionally, the data on which a cluster analysis is applied, of course, determines the outcome. In our case, the percentages of coincidence of head peaks across various dikes are used, since this is, in our view, the closest parameter to our goal to evaluate the dynamics of peak heads across various canal dikes. For different purposes, different parameters and data can be chosen to cluster dikes. Clustering dikes is a practical way to manage variability, but in practice, there is no clear distinction or a clear-cut between different dikes in terms of head responses and peak behaviour; instead, it is a gradual shift across a spectrum of head responses. Furthermore, the variation in head responses and the found clusters were linked to dike characteristics, which were difficult to clearly determine, leading to uncertainty. For instance, the profile of the canal dike can be very irregular and in many cases there is no singular, defined slope. Additionally, canal dikes are typically made up of multiple soil types, making it difficult to classify a peaty or clayey dike. The head level in the

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

underlying aquifer is also uncertain and estimated based on models. The only straightforward characteristic is the head difference, as it is controlled by a human-made system. With all these uncertainties, it's challenging to define clear characteristics for dikes and identify patterns, if not impossible. Furthermore, we examined the relationship between the head response and various dike characteristics, looking at each characteristic separately. It is possible that considering the combination of various characteristics might reveal a clearer pattern. However, it is important to keep in mind that even if there's a relationship between head responses in canal dikes and dike characteristics, it is challenging to determine the local dike characteristics of each individual dike stretch, because of the heterogeneous nature of the dike system. This heterogeneity exists in both longitudinal and cross-sectional directions. As a result, it can be expected that the head responses of the canal dikes have large spatial variations and can even differ for dikes that are close to each other.

5.2. Implications

This study quantified variations in peak head responses of canal dikes, which are relevant for estimating regional or national flood risk levels in polders. Consider a canal dike ring along a polder where each individual dike section is assumed, for simplicity, to have a failure probability of 1/100 per year, with fully correlated load and strength characteristics. Under these conditions, the probability of flooding in the polder equals the failure probability of the individual dike sections. In contrast, if there were four different types of dike with statistically independent load conditions, but still with fully correlated strength characteristics, the flood probability in the polder increases to approximately 1/25 per year; a factor four higher than with fully correlated load characteristics. In general, this can be calculated by $P_{f,sys} = 1 - \left(1 - P_{f,sec}\right)^n$, where $P_{f,sys}$ is the failure probability of the dike ring (or flood probability of the polder), $P_{f,sect}$ is the failure probability of the individual dike sections and n is the number of statistically independent dike sections. The correlations in loads and strength are crucial for accurately estimating flood risk levels in regions, and this study provides valuable insights into the variation of loadings in dikes by considering the different head responses that affect these estimates.

Gariano and Guzzetti (2016) reviewed the literature about the impact of climate change on landslides, both natural and engineered slope, and concluded that the risk of shallow landslides can increase (triggered by short and intense rainfall events), while the risk of deep-seated landslide may decrease or show no significant change (related to long rainfall periods). This is primarily due to changing meteorological conditions that lead to higher head levels, reducing the shear strength, soil suction and cohesion, and increasing the weight (wet density) of slope materials, all of which contribute to increasing slope instability. Deep-seated landslides appear to decrease or show no significant change, because these types of landslides depend on monthly and/or seasonal rainfall amounts. These longer duration rainfall events are expected to decrease in regions, like the Alps (Rianna et al., 2014; Gariano and Guzzetti, 2016). Canal dikes show similar behaviour to shallow landslides, with, on average, an increased risk of instability in the future. These insights are relevant for dike safety

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

assessments and design, particularly when the groundwater levels do not raise to the slope surface, representing the maximum or worst-case loading condition.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

6. Conclusions and recommendations

This study assessed the dynamics of peak heads in Dutch canal dikes on a national level by investigating the variation in head responses, statistical properties of peak heads and the impact of climate change. This is done using time series models with impulse response functions that explain the head responses, using precipitation and evaporation as the explanatory time series. Head time series at 48 monitoring sites with 108 head time series were modelled using various model structures. The TARSO-model, which is a non-linear model consisting of two regimes (upper and lower) with their own response functions, outperformed the other (non-)linear models, indicating the presence of various soil layers with distinct hydraulic properties, and/or fluctuations in infiltration rates or nonconstant storage capacities in the unsaturated zone during the dry season. For all head time series a TARSO-model structure was calibrated and evaluated, using two reliability criteria (goodness-of-fit and response time), after which at 35 dikes (73%) at least one time series model met the criteria and the best model at every location was selected for further analysis.

First, four clusters of dikes were identified, each consisting of dikes where peaks were caused by similar weather events and were distinct from dikes in other clusters. The differing peak behaviours and relevance of various weather events in these clusters can be explained by the 95% response times of the upper regime of these dikes. While three clusters were less distinct from each other, with still moderately high percentages of coincident peaks, sometimes exceeding 50%, one cluster has a very distinct peak behaviour. This cluster consists of only two dikes of which the 95% response times of the time series model were the longest and the amplitudes of the block response were the smallest. Next, the clusters of dikes and variations in head responses were related to physical dike characteristics, namely the dike slope, head difference, the head in the underlying aquifer and soil type. While no single significant relationship was found, the soil type combined with the width of the dike appear to be important factors influencing the variation in head responses. However, the presence of the same soil type and dike widths in multiple clusters indicates that these characteristics do not yield a definitive outcome for the head response

Secondly, the statistics of peak head levels were derived across various canal dikes and it was found that the median decimate height is only 15 centimeters and ranges from nearly 0 to 50 centimeters. This indicates that yearly occurring head levels are on average relatively close to those during extreme events. These decimate heights were mainly linked to the peak of the block response in the time series models, which cannot be directly related to the dike clusters, since the response time was the main differentiating factor. One cluster was characterized by very small peak block responses in the upper regime, where decimate heights smaller than 5 centimeters were found. For these dikes, the head levels that on average occur once every 1000 years are expected to be only 15 centimeters higher than the head levels that occur annually.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

Finally, the impact of climate change, by means of changing precipitation and evaporation rates in future climates, on the head statistics is quantified for a range of emissions scenarios and various time horizons. It was found that head levels with a return period of 100 years are expected to occur about 3 times less frequently to 7 times more frequently in the year 2100, depending on the climate scenario and the type of canal dike. Dryer summers can cause extreme loads to occur less frequently since more water can be stored in a more dried-out dike body, however, most climate scenarios indicate an increase in the frequency of occurrence of extreme load levels in the year 2100, caused by a wetting trend in winter. The variation in the impact of climate change can be attributed to differing head responses in dikes, with various drying and wetting behaviours, such as different durations and timing of heavy rainfall, leading to distinct peak head levels.

The head statistics at canal dikes are analysed and the impact of climate change is assessed by means of changing weather conditions. However, this is not the only way climate change can influence the head statistics, since the wetting and drying cycle can also result in soil deterioration with changing soil parameters occurring both seasonally and gradually over time (Stirling et al., 2021). Dixon et al. (2019) observed large variability in hydraulic conductivity in the uppermost 1,0 m of which it is unknown how this impacts head levels deeper in the dike. Quantification of all the effects of future climate scenarios is challenging, where both the hydraulic and mechanical behaviour of soils are impacted. The translation towards slope instability is even more complex. Vardon (2015) identified the major climatic variations expected to influence geotechnical infrastructure: increasing temperature (causing soil drying), increasing mean rainfall (causing reduction in soil suctions), increasing drought events (leading to drying and soil desiccation) and increasing intense precipitation (causing soil erosion and hydromechanical failure). To comprehend and quantify changes in head responses within dikes, an initial step involves continuous long-term monitoring, preferably exceeding 10 years. These measurement series are scarcely available and underscore once again the importance of long-term head observations in dikes.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

Appendices

A1. Time series models

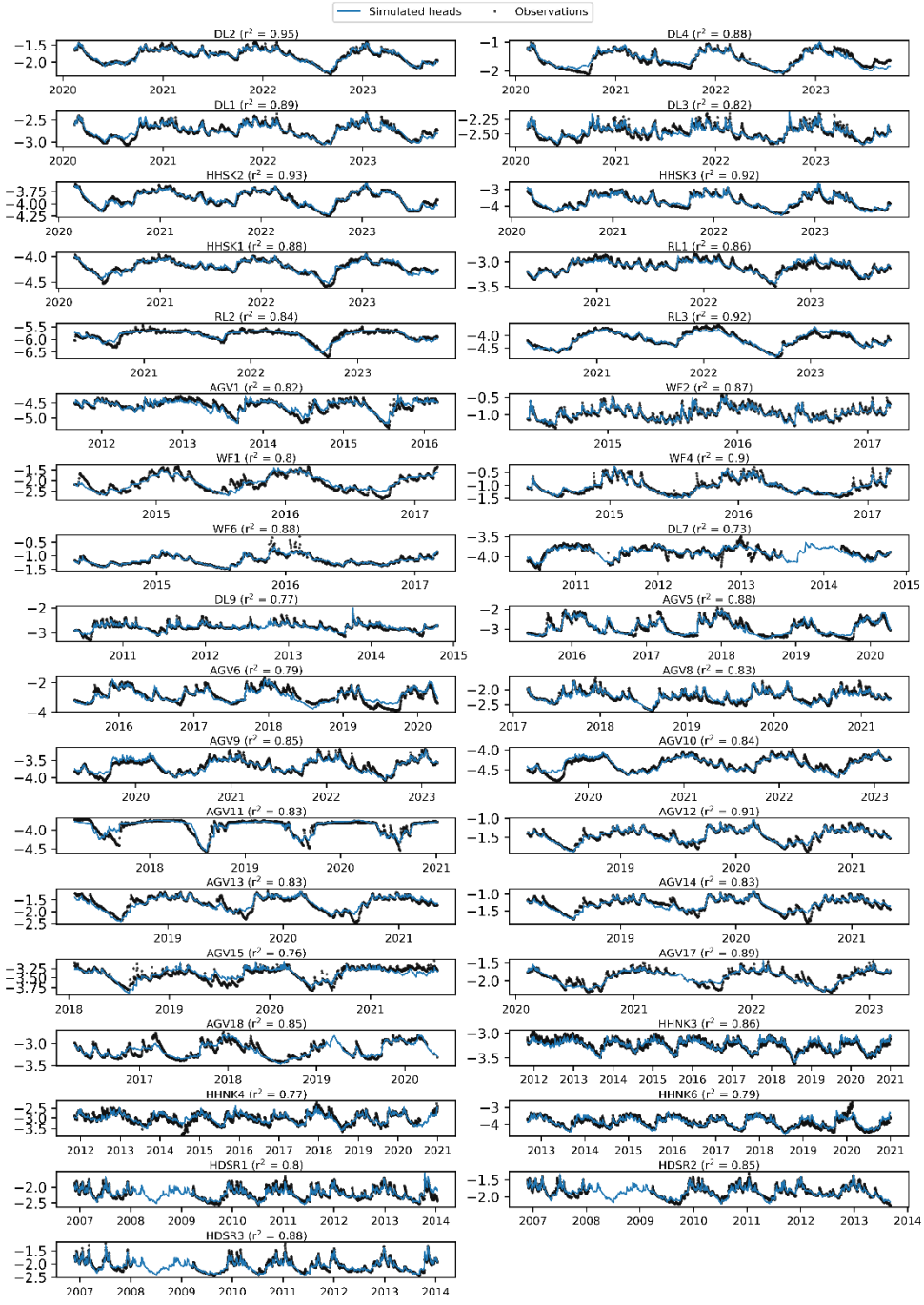

Figure A1. The observed and simulated heads for all monitoring wells.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

A2. Seasonality

The dynamics of the peak heads were analysed by quantifying the seasonality of the dikes, which is measured by using the average timing and temporal concentration of the selected head peaks. The method to determine the average timing of head peaks involves circular statistics, and it is extensively described in Hall & Blosch (2018). The average timing of the head peaks is the average date on which peaks have occurred during the time series. The average timing of head peaks can be the result of a wide range of peak dates during the year and therefore the temporal concentration of peaks occurrence within the year is considered using the concentration index. The concentration index of peak dates around the average timing serves as a measure of how well the seasonality is defined for a specific dike, with 0 indicating evenly distributed peaks during the year and 1 indicating that all peaks occur on the same date.

Seasonality varies across the dikes, but the average timing of these peaks occurs in the winter half-year, as shown in the right graph in Figure A2. The average timing shifts further into the winter from cluster A to Cluster D, with increasing temporal concentrations. This behaviour is also illustrated by the probability density distributions of the peaks of four dikes in the left graph in Figure A2. The vertical dashed line indicates the average timing, which moves further into the winter, while also the density functions become narrower.

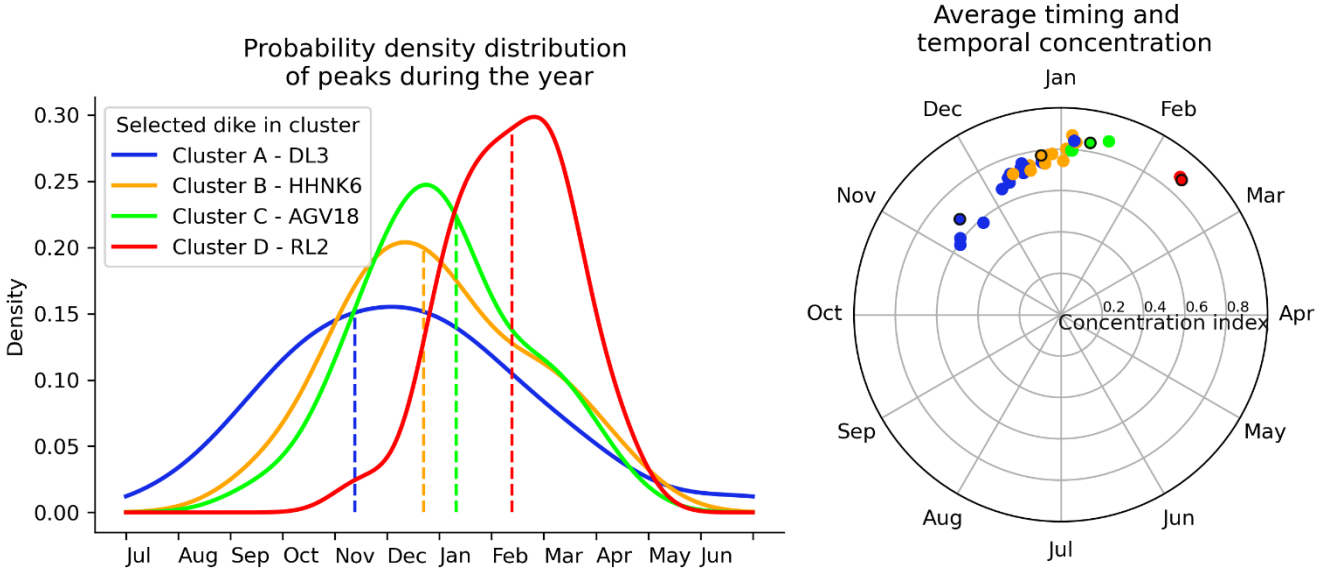

Figure A2. Right: The average timing and temporal concentrations for the considered dikes, with colors representing the dike clusters (refer to the legend in the left graph). Left: The probability density distribution of the peaks throughout the year for four selected dikes from different clusters, with the vertical dashed line indicating the average timing. These four dikes are highlighted with a black edge around the circle in the right graph.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

A2.Physical dike characteristics and their relation to head responses

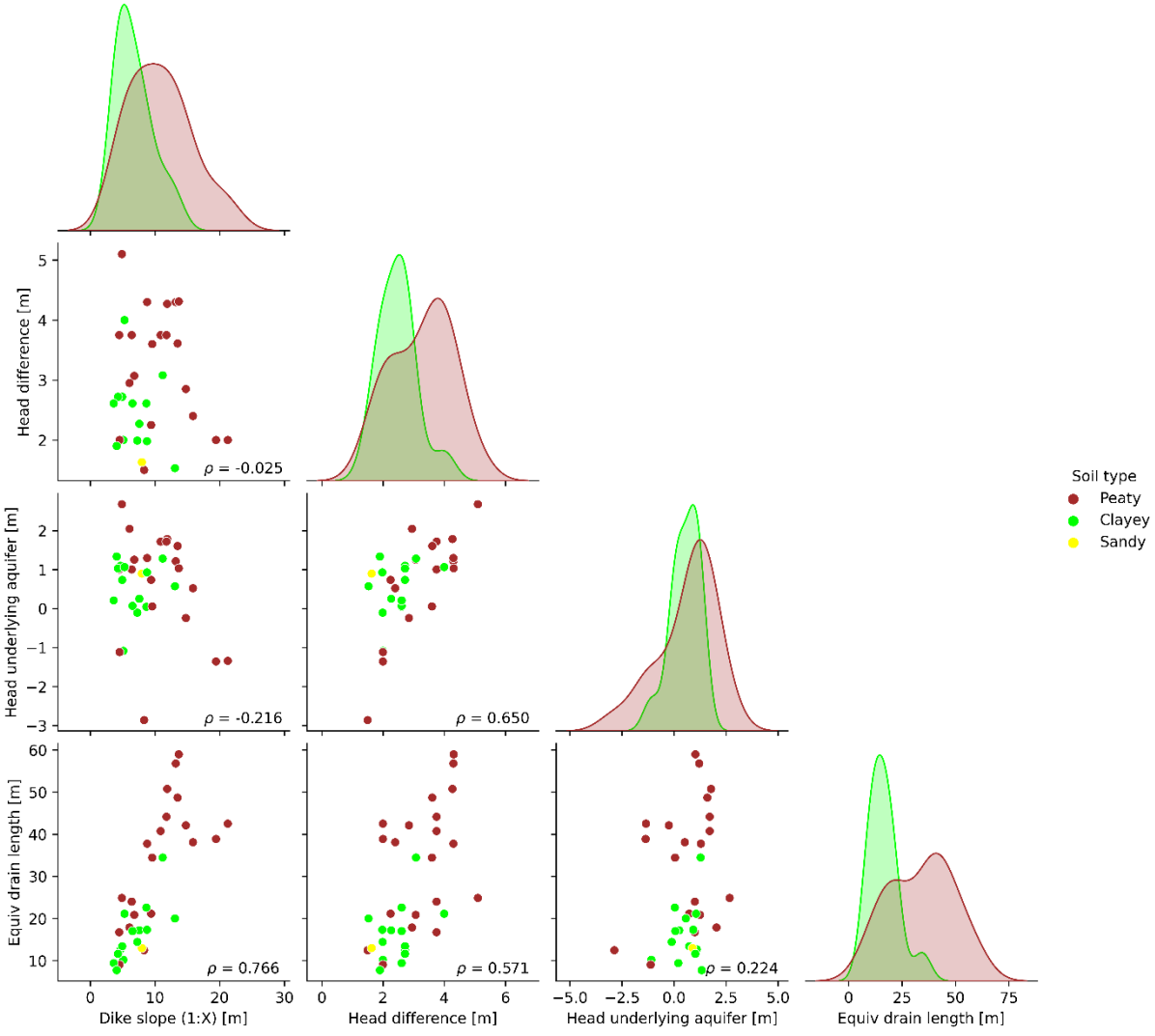

Figure A3. The relationships between the considered physical dike characteristics are illustrated by scatterplots, with the soil type of the dike indicated by color. The Pearson correlation coefficients between the variables are displayed in the bottom right corner. The diagonal plots show the univariate distributions, highlighting the marginal distribution of each variable, with distinctions made based on soil type.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

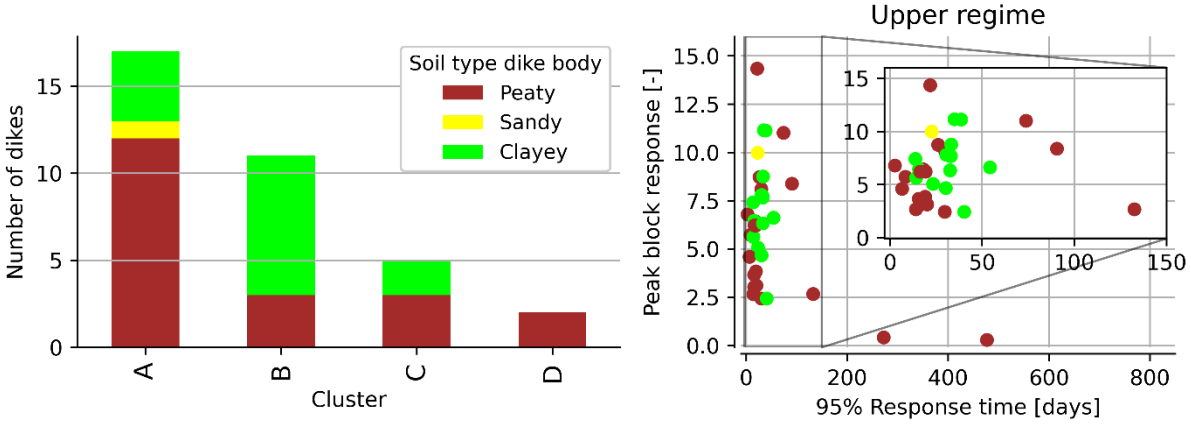

Figure A4. Stacked bars of the soil type of the dike body for the three clusters of dikes. Right: the characteristics of the impulse response functions (95% response time and peak block response) where the colors indicate the soil type of the dike body.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

Data availability statement

The measurement data used to establish the model set, consisting of 108 head time series across 48 monitoring sites and the local historic rainfall and potential evaporation, are available at 4TU.ResearchData (Strijker, 2024). This dataset contains:

- An overview of the monitoring sites and piezometers, along with their geographic locations (CSV-file and shapefile)
- Hydraulic head time series from the piezometers (CSV-file)
- Time series of local precipitation and potential evaporation (CSV-file)

A readme file is added that describes the data files. The data source has a CC0 license, which entails the waiver of all copyright and related rights, enabling unrestricted use of the data for any purpose. Authors appreciate being informed when using the data by contacting the corresponding author.

Acknowledgements

This research was supported by the STOWA and Rijkswaterstaat. Furthermore, the authors express their gratitude to all the water authorities who provided the necessary data: Hoogheemraadschap Schieland & de Krimpenerwaard (HHSK), Hoogheemraadschap Delfland (DL), Hoogheemraadschap Rijnland (RL), Hoogheemraadschap Hollands Noorderkwartier (HHNK), Weterksip Fryslân (WF), Waterschap Amstel, Gooi and Vecht (AGV) and Hoogheemraadschap De Stichtse Rijnlanden (HDSR).

Competing interests

The contact author has declared that none of the authors has any competing interests

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

References

Bakker, M., & Schaars, F. (2019). Solving groundwater flow problems with time series analysis: you may not even need another model. Groundwater, 57(6), 826-833.

Bengfort, B., & Bilbro, R. (2019). Yellowbrick: Visualizing the scikit-learn model selection process. Journal of Open Source Software, 4(35), 1075.

Berendrecht, W. L., Heemink, A. W., Van Geer, F. C., & Gehrels, J. C. (2006). A non-linear state space approach to model groundwater fluctuations. Advances in water resources, 29(7), 959-973.

Box, G. E. P., & Jenkins, G. M. (1970). Time Series Analysis: Forecasting and Control.

Collenteur, R. A., Bakker, M., Caljé, R., Klop, S. A., & Schaars, F. (2019). Pastas: open source software for the analysis of groundwater time series. Groundwater, 57(6), 877-885.

Collenteur, R. A., Bakker, M., Klammler, G., & Birk, S. (2021). Estimation of groundwater recharge from groundwater levels using nonlinear transfer function noise models and comparison to lysimeter data. Hydrology and Earth System Sciences, 25(5), 2931-2949.

De Bruin, H. A. R. (1987). From penman to makkink. In Evaporation and Weather: Technical Meeting 44, Ede, The Netherlands 25 March 1987. The Hague, Netherlands. 1987. p 5-31. 1 fig, 4 tab, 34 ref..

De Lange, W. J., Prinsen, G. F., Hoogewoud, J. C., Veldhuizen, A. A., Verkaik, J., Essink, G. H. O., ... & Kroon, T. (2014). An operational, multi-scale, multi-model system for consensus-based, integrated water management and policy analysis: The Netherlands Hydrological Instrument. Environmental Modelling & Software, 59, 98-108.

Dixon, N., Crosby, C. J., Stirling, R., Hughes, P. N., Smethurst, J., Briggs, K., ... & Hudson, A. (2019). In situ measurements of near-surface hydraulic conductivity in engineered clay slopes. Quarterly Journal of Engineering Geology and Hydrogeology, 52(1), 123-135.

Everitt, B. S., Landau, S., Leese, M., and Stahl, D.: Cluster Analysis, Wiley Online Library, 2011.

Frank, R., Bauduin, C., & Driscoll, R. M. (2004). Designers' Guide to Eurocode 7: Geotechnical Design: Designers' Guide to EN 1997-1. Eurocode 7: Geotechnical Design-General Rules. Thomas Telford.

Gariano, S. L., & Guzzetti, F. (2016). Landslides in a changing climate. Earth-Science Reviews, 162, 227-252.

Gildeh, H. K., Hosseini, P., Zhang, H., Riaz, M., & Acharya, M. (2019). Canal embankment failure mechanism, breach parameters and outflow predictions. In Sustainable and Safe Dams Around the World/Un monde de barrages durables et sécuritaires (pp. 61-74). CRC Press.

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. European journal of epidemiology, 31(4), 337-350.

Hall, J., & Blöschl, G. (2018). Spatial patterns and characteristics of flood seasonality in Europe. Hydrology and Earth System Sciences, 22(7), 3883-3901.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. Journal of the royal statistical society. series c (applied statistics), 28(1), 100-108.

Huang, W., Loveridge, F. A., Briggs, K. M., Smethurst, J. A., Saffari, N., & Thomson, F. (2024). Forecast climate change impact on porewater pressure regimes for the design and assessment of clay earthworks. Quarterly Journal of Engineering Geology and Hydrogeology, 57(1), qjegh2023-015.

Jamalinia, E., Vardon, P. J., & Steele-Dunne, S. C. (2020). The impact of evaporation induced cracks and precipitation on temporal slope stability. Computers and Geotechnics, 122, 103506.

Jongejan, R. B., Diermanse, F., Kanning, W., & Bottema, M. (2020). Reliability-based partial factors for flood defenses. Reliability Engineering & System Safety, 193, 106589.

Kanning, W. (2012). The Weakest Link - Length Effects for Piping. Phd thesis, Delft University of Technology, Delft, The Netherlands.

Knotters, M., & De Gooijer, J. G. (1999). TARSO modeling of water table depths. Water Resources Research, 35(3), 695-705.

Knotters, M., & Van Walsum, P. E. V. (1997). Estimating fluctuation quantities from time series of water-table depths using models with a stochastic component. Journal of Hydrology, 197(1-4), 25-46.

Lendering, K., Schweckendiek, T., & Kok, M. (2018). Quantifying the failure probability of a canal levee. Georisk: assessment and management of risk for engineered systems and geohazards, 12(3), 203-217.

Manh, N. V., Merz, B., & Apel, H. (2013). Sedimentation monitoring including uncertainty analysis in complex floodplains: a case study in the Mekong Delta. Hydrology and Earth System Sciences, 17(8), 3039-3057.

Martín-Antón, M., Negro, V., del Campo, J. M., López-Gutiérrez, J. S., & Esteban, M. D. (2016). Review of coastal land reclamation situation in the world. Journal of Coastal Research, (75 (10075)), 667-671.

McDonald, M. G., & Harbaugh, A. W. (2003). The history of MODFLOW. Ground water, 41(2), 280.

Moore, R., Carey, J. M., & McInnes, R. G. (2010, November). Landslide behaviour and climate change: predictable consequences for the Ventnor Undercliff, Isle of Wight. Geological Society of London.

Morton, L. W., & Olson, K. R. (2018). The pulses of the Mekong River Basin: Rivers and the livelihoods of farmers and fishers. Journal of Environmental Protection, 9(04), 431.

Özer, I. E., van Damme, M., & Jonkman, S. N. (2019). Towards an international levee performance database (ILPD) and its use for macro-scale analysis of levee breaches and failures. Water, 12(1), 119.

Pleijster, E. J., van der Veeken, C., Jongerius, R., & Luiten, E. (2015). Dijken van Nederland. Nai010 uitgevers.

Rianna, G., Zollo, A., Tommasi, P., Paciucci, M., Comegna, L., & Mercogliano, P. (2014). Evaluation of the effects of climate changes on landslide activity of Orvieto clayey slope. Procedia Earth and Planetary Science, 9, 54-63.

Ridley, A., McGinnity, B., & Vaughan, P. (2004). Role of pore water pressures in embankment stability. Proceedings of the Institution of Civil Engineers-Geotechnical Engineering, 157(4), 193-198.

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

Rikkert, S. J. H. (2022). A system perspective on flood risk in polder drainage canal systems (Doctoral dissertation, Delft University of Technology).

Rikkert, S. J. H., Kok, M., Lendering, K., & Jongejan, R. (2022). A pragmatic, performance-based approach to levee safety assessments. Journal of Flood Risk Management, 15(4), e12836.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 20, 53-65.

Rouainia, M., Helm, P., Davies, O., & Glendinning, S. (2020). Deterioration of an infrastructure cutting subjected to climate change. Acta Geotechnica, 15(10), 2997-3016.

Schweckendiek, T. (2014). On reducing piping uncertainties: A Bayesian decision approach.

Sharp, M., Wallis, M., Deniaud, F., Hersch-Burdick, R., Tourment, R., Matheu, E., ... & Simm, J. (2013). The international levee handbook.

Shen, H., Tolson, B. A., & Mai, J. (2022). Time to update the split-sample approach in hydrological model calibration. Water Resources Research, 58(3), e2021WR031523.

Šimůnek, J., Šejna, M., & Van Genuchten, M. T. (1999). The HYDRUS-2D software package for simulating the two-dimensional movement of water, heat, and multiple solutes in variably-saturated media: Version 2.0. US Salinity Laboratory, Agricultural Research Service, US Department of Agriculture.

Stafleu, J., Maljers, D., Busschers, F. S., Gunnink, J. L., Schokker, J., Dambrink, R. M., ... & Schijf, M. L. (2012). GeoTop modellering. TNO report, 10991.

Steenbergen, C. M., Reh, W., Nijhuis, S., & Pouderoijen, M. T. (2009). The Polder Atlas of The Netherlands: Pantheon of the Lowlands. Uitgeverij Thoth.

Stirling, R. A., Toll, D. G., Glendinning, S., Helm, P. R., Yildiz, A., Hughes, P. N., & Asquith, J. D. (2021). Weather-driven deterioration processes affecting the performance of embankment slopes. Géotechnique, 71(11), 957-969.

Strijker, B. (2014): Dataset in support of the dynamics of peak head responses at Dutch canal dikes and the impact of climate change. 4TU.ResearchData. [Dataset]. https://doi.org/10.4121/4004f445-b71b-4996-bd7d-10b1fafbc86b

Tran, D. D., & Weger, J. (2018). Barriers to implementing irrigation and drainage policies in An Giang Province, Mekong Delta, Vietnam. Irrigation and drainage, 67, 81-95.

Triet, N. V. K., Dung, N. V., Fujii, H., Kummu, M., Merz, B., & Apel, H. (2017). Has dyke development in the Vietnamese Mekong Delta shifted flood hazard downstream?. Hydrology and Earth System Sciences, 21(8), 3991-4010.

Van Baars, S., & Van Kempen, I. M. (2009). The causes and mechanisms of historical dike failures in the Netherlands. E-Water Official Publication of the European Water Association, 1-14.

Van Dorland, R., Beersma, J., Bessembinder, J., Bloemendaal, N., Van Den Brink, H., Brotons Blanes, H., ... & Van Der Wiel, K. (2023). Knmi national climate scenarios 2023 for The Netherlands. Scientific report WR-23-02, Royal Netherlands Meteorological Institute (KNMI).

https://doi.org/10.5194/egusphere-2024-1495
Preprint. Discussion started: 20 September 2024

Van Esch, J. M. (2012, June). Modeling groundwater flow through embankments for climate change impact assessment. In Proceedings, 19th International Conference on Computational Methods in Water Resources (CMWR 2012) University of Illinois at Urbana–Champaign, USA.

Vanmarcke, E. H. (1977). Reliability of earth slopes. Journal of the Geotechnical Engineering Division, 103(11), 1247-1265.

Vardon, P. J. (2015). Climatic influence on geotechnical infrastructure: a review. Environmental Geotechnics, 2(3), 166-174.

Von Asmuth, J. R., Bierkens, M. F., & Maas, K. (2002). Transfer function-noise modeling in continuous time using predefined impulse response functions. Water Resources Research, 38(12), 23-1.

Von Asmuth, J. R., Maas, K., Bakker, M., & Petersen, J. (2008). Modeling time series of ground water head fluctuations subjected to multiple stresses. Groundwater, 46(1), 30-40.

Vrijling, J. K., & Van Gelder, P. H. A. J. M. (2002). Probabilistic design in hydraulic engineering. Lecture notes CT5310.

Warner, J. F., van Staveren, M. F., & van Tatenhove, J. (2018). Cutting dikes, cutting ties? Reintroducing flood dynamics in coastal polders in Bangladesh and the netherlands. International journal of disaster risk reduction, 32, 106-112.

Wojciechowska, KA. (2015). Advances in operational flood risk management in the Netherlands. [Dissertation (TU Delft), Delft University of Technology]. K.A. Wojciechowska. https://doi.org/10.4233/uuid:5d719beb-bbbf-4fde-8e10-526785315fd1

Wolters, E.; Hakvoort, H.; Bosch, S.; Versteeg, R.; Bakker, M.; Heijkers, J.; Talsme, M.; Peerdeman, K. Meteobase: Online neerslag-en referentiegewasver-dampingsdatabase voor het Nederlandse waterbeheer. Meteorologica 2013, 1, 15–18.