Delft University of Technology

Testing for Trends in Dutch Climate Data

Are there Climate Trends and Extreme Weather Shifts?

Annika van Adrichem

TU Delft

Testing for Trends in Dutch Climate Data

Are there Climate Trends and Extreme Weather Shifts?

by

Annika van Adrichem

to obtain the degree of Bachelor of Science
at the Delft University of Technology,
to be defended publicly on Monday June 30, 2025 at 2:30 PM.

Student number:	5411866	
Project duration:	April 22, 2025 – June 30, 2025	
Thesis committee:	Prof. dr. ir. G. Jongbloed,	TU Delft, supervisor
	Prof. dr. H.M. Schuttelaars,	TU Delft

This thesis is confidential and cannot be made public until June 30, 2025.

TU Delft

Preface

This thesis was written as the final project for the Bachelor's degree in Mathematics at Delft University of Technology. It focuses on identifying trends in Dutch climate data using statistical methods, including regression techniques for temperature, precipitation, and extreme weather indicators.

I chose this topic because it allowed me to apply mathematical tools to a subject I care about. In a time where climate change is a highly urgent and widely discussed issue, working with climate data offers a meaningful way to contribute through quantitative analysis. The project has been both challenging and rewarding, and has deepened my appreciation for the role of statistics in environmental research.

I would like to thank my supervisor, G. Jongbloed, for their guidance and feedback. I also thank the other member of my examination committee, H.M. Schuttelaars, for their time and evaluation. Finally, I am grateful for the helpful discussions with fellow students.

Annika van Adrichem
Delft, June 2025

Contents

1
Introduction

Climate change is an increasingly urgent topic in environmental science, with strong evidence that human related greenhouse gas emissions are driving significant shifts in the Earth's climate system. Some of the most noticeable effects include the increasing frequency of hot days, a decrease in cold days, and heightened variability in precipitation patterns, including more frequent droughts and extreme wet events.

The Intergovernmental Panel on Climate Change (IPCC) concludes with high confidence in its Sixth Assessment Report that extreme weather events, including temperature and precipitation extremes, are becoming more frequent and intense as global temperatures rise [12]. The report shows that even small increases in global warming lead to statistically significant changes in climate extremes on both global and regional scales.

At the same time, research based on NASA's GRACE and GRACE-FO satellite data demonstrates that periods of significant dryness and excessive precipitation have become more common and severe from 2002 to 2021 [8]. This shows that warming is causing more intense movement of water worldwide.

Global climate trends are well known, but it is still important to understand how these changes affect specific regions and local areas. This thesis investigates climate developments in the Netherlands by analysing historical weather observations from the De Bilt station, maintained by the Royal Netherlands Meteorological Institute (KNMI). It focuses on detecting statistically significant trends in temperature and precipitation, including the frequency of extreme events such as tropical and ice days, as well as dry, wet, and extremely wet days. The primary research question it seeks to answer is: *To what extent do historical weather observations from De Bilt reveal statistically significant trends in temperature, precipitation, and extreme weather events over the past century?*

To answer this question, two types of climate time series are examined: continuous variables such as annual mean temperature and total precipitation, and count-based indicators such as the yearly number of tropical days, ice days, and extreme precipitation events. For continuous variables, both linear regression and isotonic regression are used to detect possible monotonic trends. For the count data involving extreme weather events, Poisson and Negative Binomial regression models are used to account for the discrete nature of the outcomes. Isotonic regression is also applied as a flexible, non-parametric alternative to identify monotonic trends in these event counts.

Chapter 2 introduces the dataset, detailing how daily observations are aggregated and how different types of extremes are defined. Chapter 3 presents the regression models used, outlines key assumptions, and includes visual summaries of the data. Chapter 4 describes the methods used to test for significant trends in the regression models. Chapter 5 reports the results, interpreting the model outcomes and evaluating the statistical evidence for trends in both continuous and count-based climate variables.

In sum, this thesis investigates whether local observational data support the presence of long-term changes in temperature and precipitation patterns, as well as in the frequency of extreme weather events, contributing to the broader understanding of climate change impacts at a regional scale.

2 Dataset Description and Preprocessing

All data used in this project comes from the Royal Netherlands Meteorological Institute (KNMI) [9]. The data is publicly accessible via their web-address. This study uses data exclusively from the De Bilt station, located centrally in The Netherlands. Due to its geographical position and long-standing measurement history, De Bilt is widely considered representative of national climate patterns. Additionally, De Bilt is also recognized as a long-term observing station by the World Meteorological Organization [13].

This research aims to examine whether climate change is detectable in Dutch historical weather data, with a focus on two key indicators: temperature and precipitation. Both long-term trends and the frequency of threshold-based events are considered. Analyses are done over different time periods, like yearly, monthly, and by season. This to capture different patterns of variability. For seasonal assessments, meteorological definitions of winter (December–February) and summer (June–August) are applied. Meteorological seasons are preferred over astronomical ones in this case, due to their simplicity and better alignment with the climatological characteristics of the Northern Hemisphere [11]. Below, the specific datasets and assumptions used for each analysis type are detailed. All data analysis, statistical evaluation, and visualization were performed using Python [6]. The code for the data preprocessing can be found in Appendix A.1.

2.1. Temperature Data

To study long-term temperature trends, this research uses the homogenized monthly average temperature series provided by KNMI. These data have been adjusted to correct for inhomogeneities caused by station relocations, changes in instrumentation, or observation methods. The homogenized dataset is considered suitable for time series analysis of climate change. It covers the period from 1901 to 2024. A visualization of this dataset is shown in Figure 2.1. Although it is difficult to draw definitive conclusions at a glance, there appears to be a decrease in the number of months with average temperatures at or below 0°C, and an increase in months averaging 20°C or more. More generally, if one considers a threshold around 10°C, the number of months exceeding this value seems to have increased over time.

When looking at trends, it could also be useful to consider annual and seasonal averages. An overview of those can be found in Figure 2.2. Looking at this figure, the trend appears to be positive for both winter and summer temperatures, as well as for the yearly average.

For event-based analysis, specifically the counting of

- **Tropical days**: days with a maximum temperature of 30 °C or higher,
- **Ice days**: days with a maximum temperature below 0 °C,

the non-homogenized daily temperature dataset is used. This choice is made because daily resolution is necessary to count these specific events. Non-homogenized in this context means that no corrections have been applied to account for factors such as weather station relocations or changes in instrumentation.

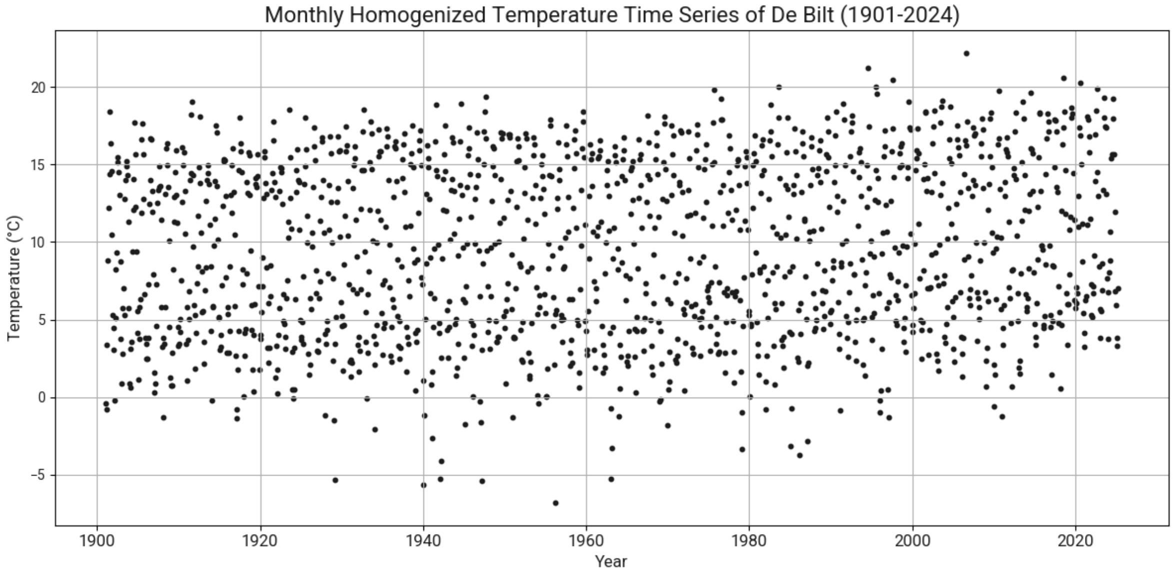

Figure 2.1: Monthly average temperatures at De Bilt, corrected for inhomogeneities in the historical record.

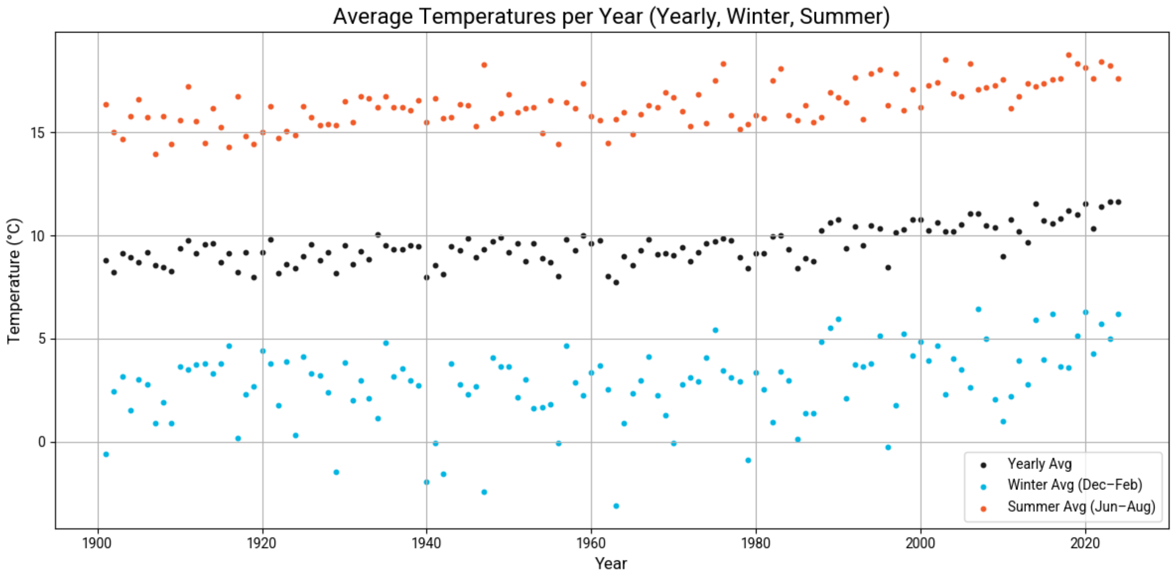

Figure 2.2: Average temperatures in each year, summer and winter measured in De Bilt.

As a result, the KNMI advises against using this dataset for trend analysis. However, since the data is only used here for counting occurrences above or below fixed thresholds, small inconsistencies are considered acceptable. The dataset may contain some inhomogeneities, but these are less impactful in this type of threshold-based counting analysis.

The results of the event count are presented in Figure 2.3. A quick inspection suggests a decreasing trend in the number of ice days and an increasing trend in the number of tropical days.

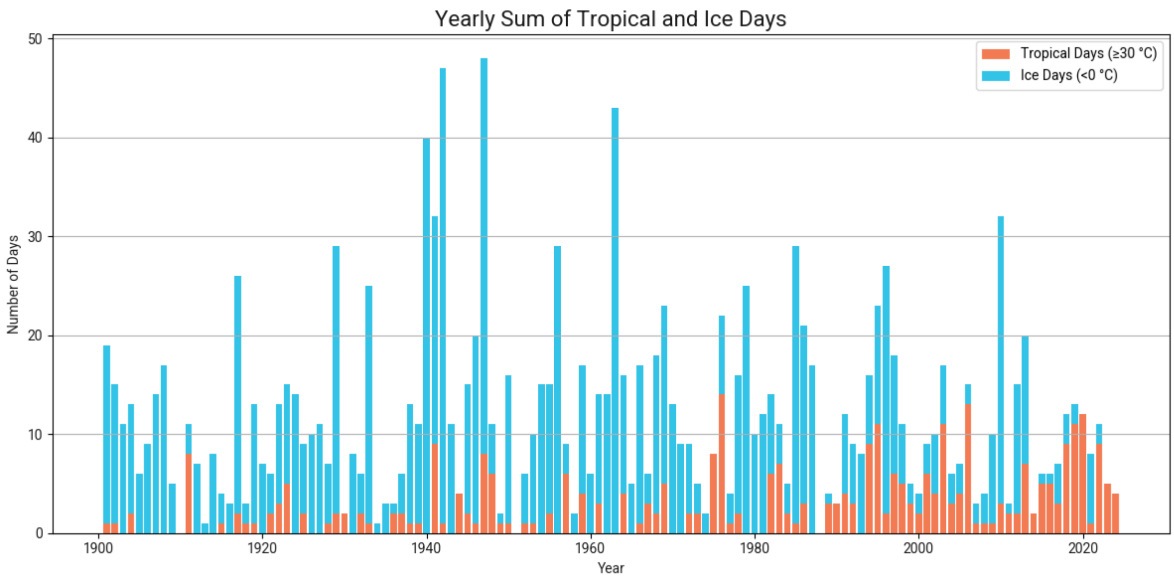

Figure 2.3: Annual counts of tropical days (maximum temperature ≥ 30°C) and ice days (maximum temperature < 0 °C) in De Bilt.

Mathematical Description of the Temperature Data. Let $t \in \{1901, 1902, \ldots, 2024\}$ denote the year and $m \in \{1, 2, \ldots, 12\}$ the month. The homogenized monthly average temperature is represented as

$$T_m(t) \in \mathbb{R},$$

where $T_m(t)$ is the average temperature in month m of year t (in degrees Celsius). The KNMI calculates this by first averaging the 24 hourly temperatures for each day to get a daily average. Then, it takes the average of all the daily averages in that month. Annual and seasonal aggregates are defined as

$$T^{\text{year}}(t) = \frac{1}{12} \sum_{m=1}^{12} T_m(t),$$
$$T^{\text{sum}}(t) = \frac{1}{3} \sum_{m \in \{6,7,8\}} T_m(t),$$
$$T^{\text{win}}(t) = \frac{1}{3} \left(T_{12}(t-1) + T_1(t) + T_2(t)\right)$$

For threshold-based event counting, we use non-homogenized daily data. Let t denote the year, and let $D(t)$ be the set of all days in year t. For each day $d \in D(t)$, let $T_{\max}(d, t)$ denote the maximum temperature on that day. Define the indicator functions

$$I_{\text{trop}}(d, t) = \mathbb{1}_{\{T_{\max}(d,t) \geq 30\}}, \quad I_{\text{ice}}(d, t) = \mathbb{1}_{\{T_{\max}(d,t) < 0\}}.$$

Then, the number of tropical and ice days in year t are given by

$$N^{\text{trop}}(t) = \sum_{d \in D(t)} I_{\text{trop}}(d, t), \quad N^{\text{ice}}(t) = \sum_{d \in D(t)} I_{\text{ice}}(d, t).$$

2.2. Precipitation Data

To study long-term precipitation trends over time, the homogenized monthly precipitation dataset from KNMI is used. This dataset is corrected for inhomogeneities such as changes in measurement instruments or station relocations, and is therefore suitable for time series analysis. It allows for the investigation of trends in average monthly and seasonal precipitation from 1906 onward. A first overview can be found in Figure 2.4. Here it seems that in the last few years, dryer years occur more often.

Like with the temperature data, it could be useful to have the same look at seasonal and annual averages to see if any trends can be discovered. When looking at Figure 2.5, where these events are visualised, it is hard to see any kind of relation or trend when it comes to precipitation.

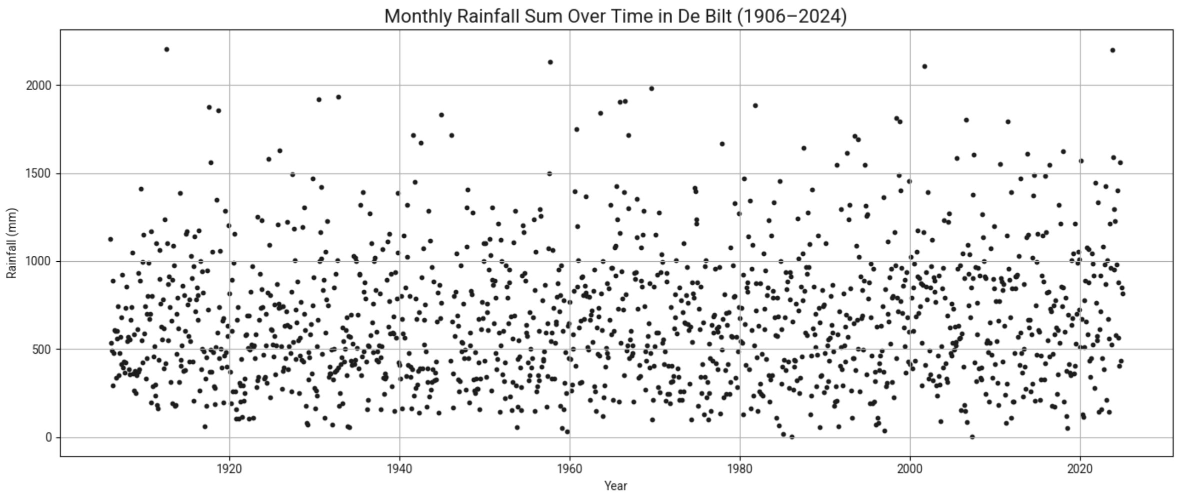

Figure 2.4: Homogenized monthly precipitation measured in De Bilt.

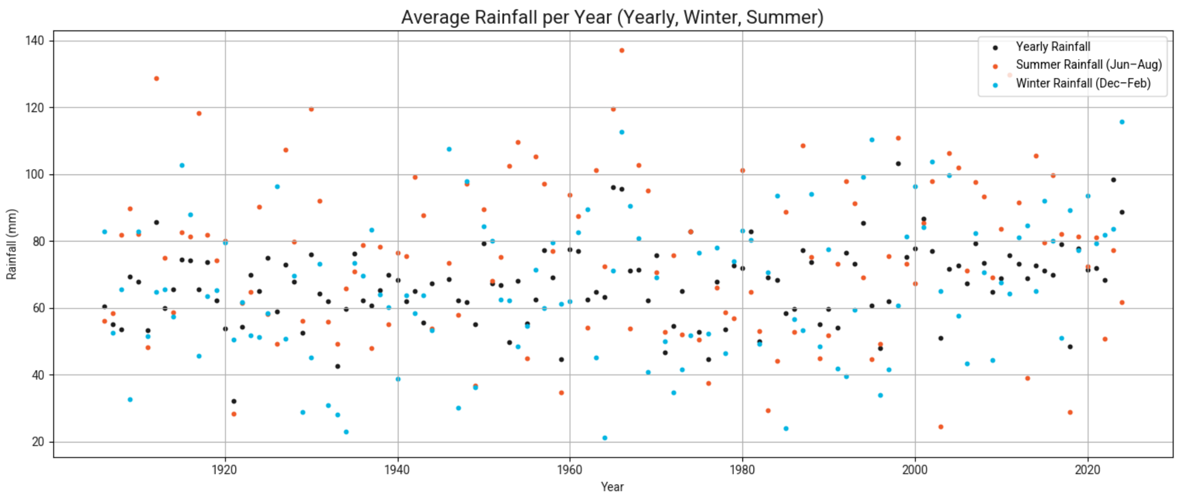

Figure 2.5: Average precipitation in each year, summer and winter measured in De Bilt.

In addition, this study evaluates the frequency of specific precipitation events using non-homogenized daily data. Specifically, the number of days per year that qualify as

- **Dry days**: days with less than 0.1 mm of precipitation[1],
- **Wet days**: days with precipitation between 10 mm and 20 mm[2],
- **Extreme wet days**: days with precipitation of 20 mm or more[3],

The thresholds of 10 mm and 20 mm align with internationally recognized climate indices such as R10mm and R20mm, developed by the Expert Team on Climate Change Detection and Indices (ETCCDI). These indices quantify the frequency of moderate and heavy precipitation events, and are widely applied in climate monitoring and impact assessments. R10mm represents the number of days per year with at least 10 mm of precipitation, R20mm likewise refers to days with 20 mm of precipitation. It is good to note that exact classifications of rainfall intensity are not standardized across Europe, these thresholds are broadly used in both global and regional analyses due to their relevance for hydrology, agriculture, and infrastructure planning. In the Dutch context, such classifications are not formally codified, but similar thresholds are used in research

[1] As defined in KNMI Technical Report TR-349 [2]
[2] Corresponds to the widely used R10mm index for moderate rainfall events [14]
[3] Consistent with the R20mm index, often used to represent intense rainfall [14]

and water management applications.

An overview of the annual counts of dry, wet, and extreme wet days are presented in Figure 2.6. Notably, the number of dry days appears to have increased over the past 60 years, particularly in more recent decades.

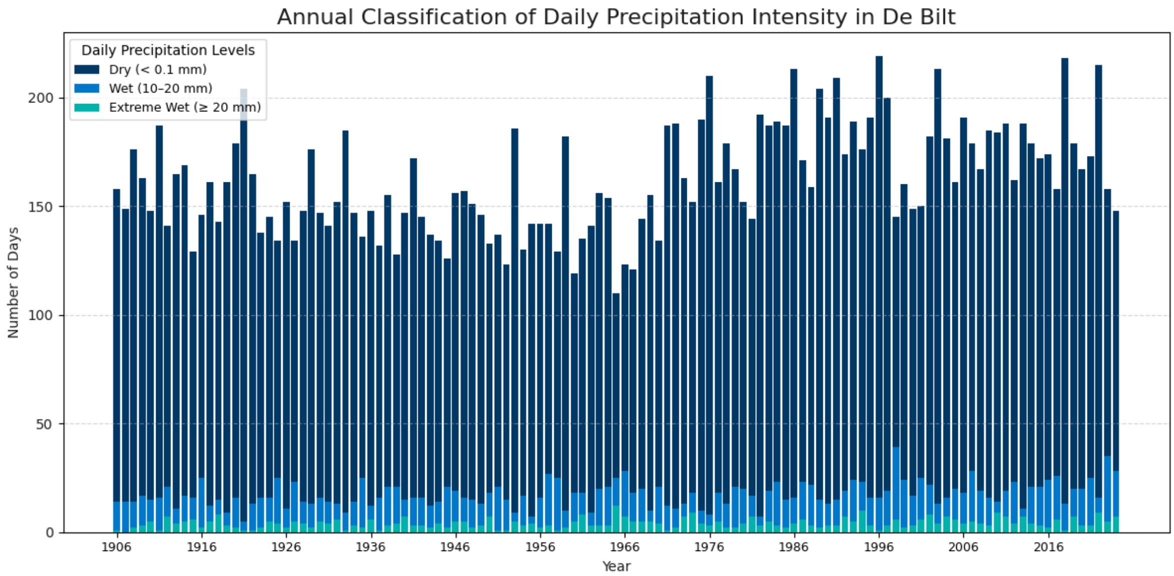

Figure 2.6: Yearly totals of dry (< 0.1 mm), wet (10–20 mm), and extreme wet (≥ 20 mm) days based on daily precipitation measurements in De Bilt.

Mathematical Description of the Precipitation Data. Let $P_m(t) \in \mathbb{R}_{\geq 0}$ denote the average monthly precipitation (in mm) in month m of year t. Aggregated values are

$$\begin{aligned} P^{\text{year}}(t) &= \sum_{m=1}^{12} P_m(t), \\ P^{\text{summer}}(t) &= \sum_{m \in \{6,7,8\}} P_m(t), \\ P^{\text{winter}}(t) &= P_{12}(t-1) + P_1(t) + P_2(t). \end{aligned}$$

Using daily data, let $R(d,t)$ denote the precipitation on day d of year t. We define the indicator functions

$$I_{\text{dry}}(d,t) = \mathbb{1}_{\{R(d,t)<0.1\}}, \quad I_{\text{wet}}(d,t) = \mathbb{1}_{\{10 \leq R(d,t)<25\}}, \quad I_{\text{ext}}(d,t) = \mathbb{1}_{\{R(d,t)\geq 25\}}.$$

Annual event totals are computed as

$$N^{\text{dry}}(t) = \sum_{d \in D(t)} I_{\text{dry}}(d,t), \quad N^{\text{wet}}(t) = \sum_{d \in D(t)} I_{\text{wet}}(d,t), \quad N^{\text{ext}}(t) = \sum_{d \in D(t)} I_{\text{ext}}(d,t).$$

3

Trend Analysis Methods

This chapter examines several statistical modelling approaches, namely Linear Regression, two types of Isotonic Regression, Poisson Regression, and Negative Binomial Regression. These models are used to identify trends in Dutch climate data described in Chapter 2. All models are applied to aspects of the same underlying datasets, which include annual or seasonal averages and counts of temperature and precipitation related variables.

Let i index the years in the dataset, and let t_i denote the actual calendar year (e.g., 1901, 1902, ..., 2024). Recall that we previously defined the following variable notation in Chapter 2, but now we express each variable as a function of the year index i:

- T_i: the average temperature in year i,
- T_i^{win}, T_i^{sum}: average temperatures in winter and summer of year i,
- N_i^{trop}: the number of tropical days (≥ 30 °C) in year i,
- N_i^{ice}: the number of ice days (< 0°C) in year i,
- P_i: the total annual precipitation in year i,
- P_i^{win}, P_i^{sum}: total precipitation in winter and summer of year i,
- $N_i^{\text{dry}}, N_i^{\text{wet}}, N_i^{\text{ext}}$: the number of dry (< 0.1 mm), wet (10–20 mm), and extreme wet (≥ 20 mm) days in year i.

The choice of model depends on the nature of the data so whether we have continuous or count data and the type of trend or relationship that is expected. This setup helps ensure that results from different models can be compared fairly and clearly. The Python code for all models can be found in Appendix A.2.

3.1. Linear Regression

Linear regression is a statistical method used to model the linear relationship between a dependent variable and one or more independent variables. It estimates the coefficients of a linear equation to predict the outcome variable based on input features, under the assumption that this relationship is approximately linear in the parameters.

In a simple linear regression model for a bivariate dataset $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, we assume the values $x_1, x_2, \ldots, x_n$ are fixed, while the corresponding y_i are realizations of random variables Y_i satisfying the model

$$Y_i = a + bx_i + \varepsilon_i.$$

Here, x_i is the observed value of the independent (predictor) variable, Y_i is the true (unobservable) dependent variable, and y_i is its observed realization. The parameter a is the intercept, representing the expected value of Y when $x = 0$, and b is the slope, representing the change in Y per unit increase in x. The term ε_i is a random error, assumed to follow a normal distribution with mean zero and constant variance σ^2, i.e., $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$, and the errors are independent.

This implies that the expected value of Y_i given x_i is

$$\mathbb{E}[Y_i] = a + bx_i.$$

To estimate the parameters a and b, the method of least squares is used. This involves minimizing the sum of squared differences between the observed values y_i and the predicted values $\hat{y}_i = a + bx_i$. The objective function is:

$$\sum_{i=1}^{n} (y_i - a - bx_i)^2.$$

Solving this minimization yields the least squares estimators:

$$b = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}, \quad a = \bar{y} - b\bar{x},$$

where $\bar{x}$ and $\bar{y}$ are the sample means of the x_i and y_i, respectively. More details can be found in *An Introduction to Mathematical Statistics* [3].

In this thesis, linear regression was implemented using the Ordinary Least Squares (OLS) method provided by the `statsmodels` package in Python. This computational approach was applied separately to yearly, summer, and winter datasets, allowing for an empirical estimation of trends.

3.1.1. Modelling Temperature and Precipitation Trends

We assume that the relationship between temperature or precipitation and time can be approximated linearly. The linear regression models are specified as

$$T_t = a_T + b_T t + \varepsilon_T, \quad \text{and} \quad P_t = a_P + b_P t + \varepsilon_P.$$

In these models

- a_T, a_P are intercept terms,
- b_T, b_P are slope coefficients representing the rate of change per year,
- $\varepsilon_T, \varepsilon_P \sim \mathcal{N}(0, \sigma^2)$ are normally distributed error terms.

The slope b_T provides an estimate of the average annual change in temperature, while b_P reflects the average change in precipitation per year.

3.1.2. Seasonal Trends

To capture seasonal effects, the same regression model is applied to temperature and precipitation data aggregated over meteorological seasons. For each year t, we extract seasonal averages such as

$$T_t^{\text{winter}},\ T_t^{\text{summer}}, \quad \text{and} \quad P_t^{\text{winter}},\ P_t^{\text{summer}},$$

and fit separate linear regression models of the form

$$T_t^{\text{season}} = a_{\text{season}} + b_{\text{season}} t + \varepsilon^{\text{season}} \text{and} \quad P_t^{\text{season}} = \alpha_{\text{season}} + \beta_{\text{season}} t + \eta^{\text{season}}.$$

Results of these linear regressions on seasonal temperature and precipitation data are shown in Figure 3.1 and Figure 3.2, respectively.

In the linear regression plot for temperature (Figure 3.1), we observe a positive trend in Winter, Summer, and Yearly averages. This indicates a gradual increase in temperature over the years for each period. The linear regression lines for all seasons show a steady increase, reflecting a potential warming trend in the climate over the study period.

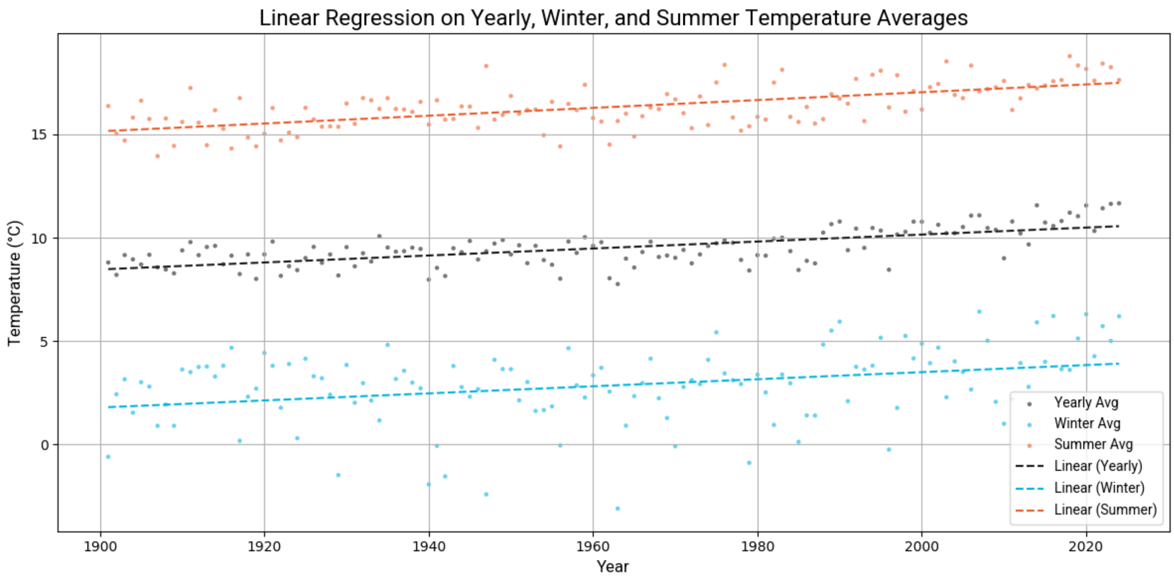

Figure 3.1: Linear regression trends of average yearly, summer (June–August), and winter (December–February) temperatures in De Bilt from 1901 to 2024. The method captures linear changes in seasonal and annual temperature averages.

The precipitation data is separated into different plots for Winter, Summer, and Yearly averages due to the overlap of the data points as we could see in Figure 2.5. This separation helps to better visualize the trends for each period and avoid confusion in the overlapping data. Figure 3.2 shows the linear regression plot for the precipitation data. Here, the regression lines for winter and yearly precipitation also show an increasing trend, confirming the potential gradual rise in precipitation. However, the summer rainfall data remains relatively flat, showing little to no change over the years.

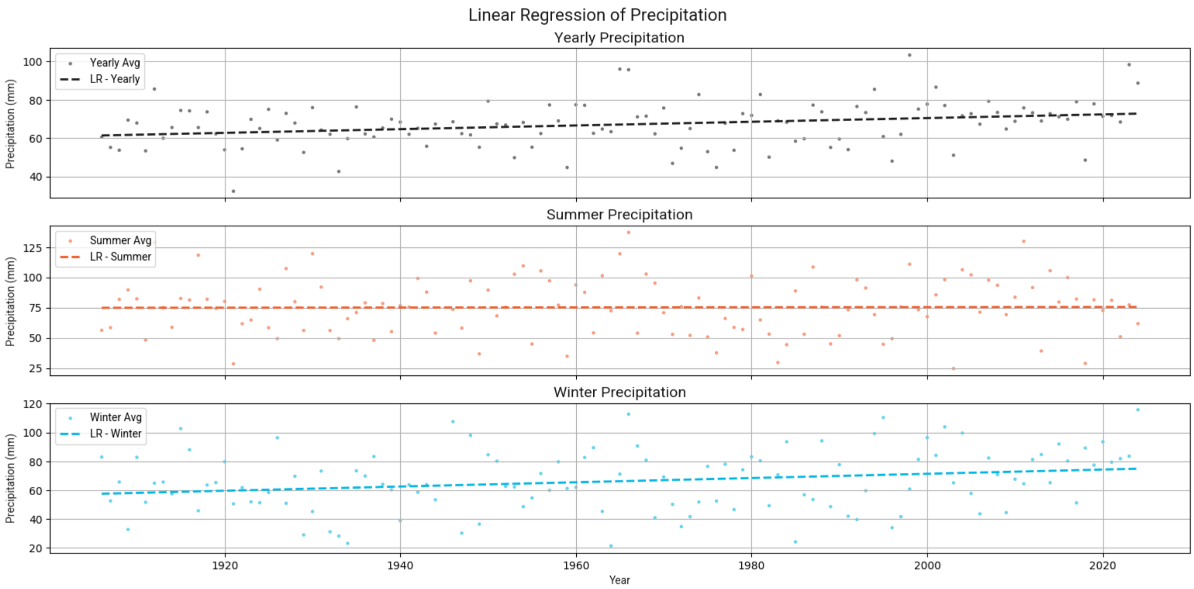

Figure 3.2: Linear regression applied to monthly precipitation averages for Winter, Summer, and Yearly data. The plot presents the original data points and the corresponding linear regression lines for each of the seasons (Winter, Summer) and yearly averages.

Investigating trends allows one to assess whether warming or cooling, wetting or drying trends are more pronounced in specific seasons. Statistical significance of the slopes can later be tested using hypothesis tests. These are discussed in Chapter 4.

3.2. Isotonic Regression for Continuous Data

Isotonic regression is a non-parametric method used to fit a monotonic function to data. Unlike linear regression, see Section 3.1, it does not assume a specific functional form, like linearity, but only that the underlying function is either *non-decreasing* or *non-increasing*.

Given observed data points

$$(x_1, y_1),\ (x_2, y_2),\ \ldots,\ (x_n, y_n) \quad \text{with} \quad x_1 < x_2 < \cdots < x_n,$$

isotonic regression seeks a function $f : \mathbb{R} \to \mathbb{R}$ that minimizes the sum of squared errors

$$\hat{f} = \arg \min_{f(x_1) \le f(x_2) \le \cdots \le f(x_n)} \sum_{i=1}^{n} \left(y_i - f(x_i)\right)^2,$$

subject to the monotonicity constraint $f(x_1) \le f(x_2) \le \cdots \le f(x_n)$.

The resulting estimator $\hat{f}$ is piecewise constant and provides the best fit while maintaining monotonicity. This method is particularly useful when the response is expected to structurally increase or decrease with the predictor, as in dose-response relationships or reliability analysis.

In this study, the isotonic regression model was fitted using the `IsotonicRegression` class from the `sklearn.isotonic` module in Python. This implementation uses an efficient version of the *pool-adjacent-violators algorithm* (PAVA), which works by iteratively averaging adjacent values that violate the monotonicity constraint until a monotonic sequence is achieved. This approach is computationally efficient and well-suited for large-scale applications.

A decreasing version of isotonic regression can be used by reversing the inequality constraint.

3.2.1. Application to Temperature and Precipitation

To detect consistent trends in climate indicators over time, isotonic regression is applied to temperature and precipitation time series measured at De Bilt. For continuous-valued indicators such as average temperature or seasonal precipitation, we assume the model

$$Y_t = f(t) + \varepsilon,$$

where t denotes the year, $Y_t \in \{T_t,\ T_t^{\text{win}},\ T_t^{\text{sum}},\ P_t,\ P_t^{\text{win}},\ P_t^{\text{sum}}\}$, f is an unknown monotonic function representing the long-term trend, and $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ is Gaussian noise where the variance may differ depending on the indicator.

Isotonic regression is well-suited for this setting because it allows us to estimate monotonic trends without assuming a specific functional form or model structure. This flexibility makes it appropriate for a variety of indicators, which may show increasing, decreasing, or flat patterns. The method is also robust to short-term variability, helping to highlight longer-term directional changes.

Seasonal trends are modelled by applying isotonic regression separately to T_t^{win}, T_t^{sum}, P_t^{win}, and P_t^{sum}. An example of isotonic regression applied to seasonal temperature data can be found in Figure 3.3.

When looking at the plot, the isotonic regression lines at the edges of the plot show more significant fluctuations due to the limited number of data points available for modelling, causing the regression to react more sensitively to the boundary values. This characteristic contrasts with the smoother behaviour observed in the central portion of the data, where more points provide a more stable fit. There is however a quite big jump for the annual and winter averages especially between 1980 and 2000. This jump could indicate a sudden shift in the temperature trend. This may be associated with a significant climate event, such as a change in regional climate patterns, or influences from global warming or other external factors. Overall, this positive trend aligns closely with the results from the linear regression in Figure 3.1.

For the isotonic regression on this precipitation data, results can be found in Figure 3.4. The plot shows how the data for each period is adjusted to smooth out fluctuations while retaining the overall trend. For the winter

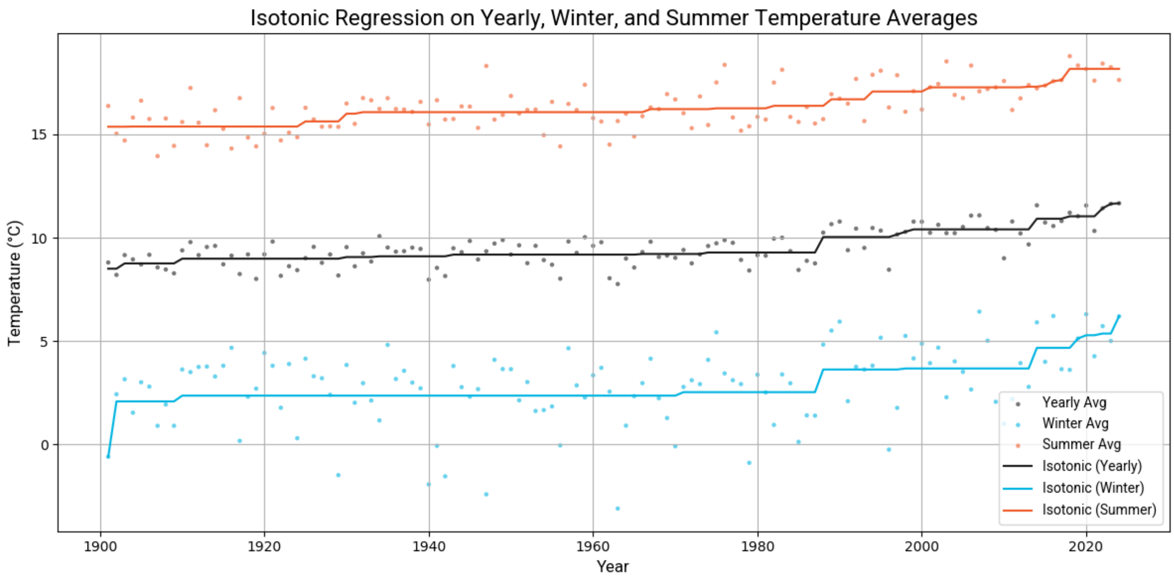

Figure 3.3: Isotonic regression trends of average yearly, summer (June–August), and winter (December–February) temperatures in De Bilt from 1901 to 2024. The method captures non-linear changes in seasonal and annual temperature averages.

and yearly rainfall, the isotonic regression curve shows a slight upward trend, indicating a general increase in precipitation over the years. This trend could reflect potential changes in climate or regional weather patterns. On the other hand, the summer rainfall data presents almost a straight line. This suggests a stable trend in summer precipitation over time, with less variation compared to the winter and yearly averages. These observations align with the result from the linear regression in Figure 3.2.

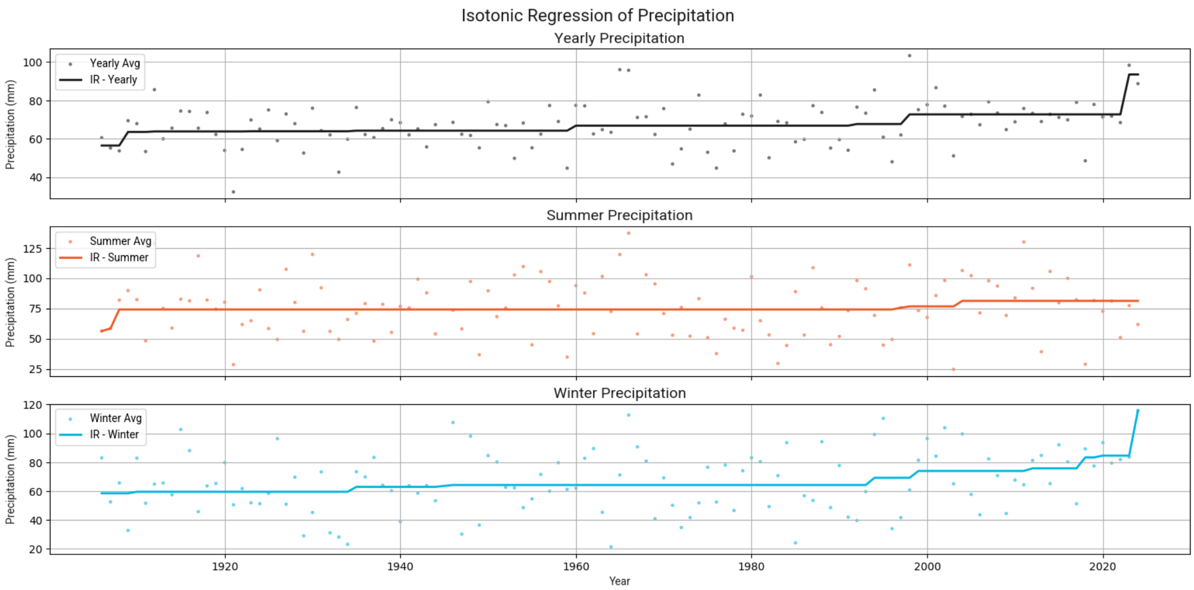

Figure 3.4: Isotonic regression applied to monthly rainfall averages for Winter, Summer, and Yearly data. The plot shows the original data points along with the smoothed curves obtained from isotonic regression for each of the seasons (Winter, Summer) and yearly averages.

3.3. Poisson Regression

Poisson regression is a fundamental method for modelling count data, where the response variable represents the number of times an event occurs within a fixed period or spatial domain. In this framework, the count variable Y_i for observation i is assumed to follow a Poisson distribution with mean λ_i, such that

$$Y_i \sim \text{Poisson}(\lambda_i), \quad \lambda_i > 0.$$

The expected count $\lambda_i = \mathbb{E}[Y_i]$ is linked to a predictor variable x_i through a log-linear relationship

$$\log(\lambda_i) = \alpha + \beta x_i,$$

which implies

$$\lambda_i = \exp(\alpha + \beta x_i).$$

Here, α is the intercept and β the slope or trend parameter. This log-link function guarantees that the predicted mean λ_i remains strictly positive.

The model coefficients α and β are typically estimated using maximum likelihood estimation (MLE). Under the Poisson model, the likelihood function for a set of independent observations $\{(x_i, y_i)\}_{i=1}^n$, where x_i denotes the predictor and y_i the corresponding count response, is given by

$$L(\alpha, \beta) = \prod_{i=1}^{n} \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!}, \quad \text{where } \lambda_i = \exp(\alpha + \beta x_i).$$

Taking the logarithm of the likelihood yields the *log-likelihood* function

$$\ell(\alpha, \beta) = \sum_{i=1}^{n} \left[y_i(\alpha + \beta x_i) - \exp(\alpha + \beta x_i) - \log(y_i!) \right].$$

The maximum likelihood estimates $\hat{\alpha}$ and $\hat{\beta}$ are obtained by solving for the values of α and β that maximize this log-likelihood. Since these equations do not have closed-form solutions and are typically solved using numerical methods such as Newton-Raphson or Iteratively Reweighted Least Squares (IRLS) [1]. I applied Poisson regression models using the Generalized Linear Model (GLM) framework in Python, implemented via the *statsmodels* library. This library estimates the model coefficients using the IRLS algorithm.

A distinctive feature of the Poisson distribution is that the mean and variance are equal, i.e.,

$$\mathbb{E}[Y_i] = \text{Var}(Y_i) = \lambda_i.$$

This property, known as *equidispersion*, often does not hold in practice. In many real-world datasets, particularly those involving environmental phenomena, we encounter *overdispersion*, where the variance exceeds the mean. Such excess variability may stem from unobserved heterogeneity, temporal correlation, or omitted covariates. If unaccounted for, overdispersion can lead to underestimated standard errors and misleading significance tests.

A common diagnostic for overdispersion involves the dispersion statistic, computed as the ratio of the Pearson chi-square statistic to the residual degrees of freedom

$$\hat{\phi} = \frac{\sum_i (y_i - \hat{\lambda}_i)^2 / \hat{\lambda}_i}{n - p},$$

where n is the number of observations and p the number of model parameters. Values of $\hat{\phi}$ significantly greater than 1 indicate overdispersion. In such cases, alternative models such as Isotonic regression or Negative Binomial regression are recommended. These Regression models are further discussed in Section 3.4 and 3.5 respectively.

In addition to overdispersion, the presence of a substantial number of zero counts can also impact model performance and the suitability of the Poisson distribution. As noted by Cameron and Trivedi [4], the Poisson model assumes that the variance equals the mean and may underpredict the frequency of zeros when the data contain more zeros than expected under this assumption. In the current dataset, zero counts occur relatively frequently for some variables: 37 years (29.8%) had no tropical days and 11 years (8.9%) had no ice days. In contrast, no years with zero counts were observed for wet, dry, or extreme wet day totals. Hilbe [7] suggests that when the proportion of zeros exceeds approximately 20–30%, standard Poisson regression may be deficient, motivating the use of alternative approaches such as Isotonic or Negative Binomial regression. These models can better accommodate excess zeros and overdispersion, as discussed earlier on.

3.3.1. Application to Tropical and Ice Days

For the tropical and ice day count data, Poisson regression models were fitted using the year t_i as the predictor variable. The general form of the models is

$$\log\left(\mathbb{E}\left[N_t^{\text{trop}}\right]\right) = \alpha_{\text{trop}} + \beta_{\text{trop}}\, t,$$
$$\log\left(\mathbb{E}\left[N_t^{\text{ice}}\right]\right) = \alpha_{\text{ice}} + \beta_{\text{ice}}\, t,$$

where α represents the intercept and β the slope coefficient indicating the trend over time for each count variable.

In the left plot of Figure 3.5, the increase in the number of tropical days is clearer than in Figure 2.3. It is especially noticeable that in the past 30 years, there has not been a single year without at least a few tropical days. The right plot shows the opposite trend: the number of years with many ice days, which were common in the mid-1900s, has decreased significantly. In the last 20 years, such years have become rare.

Figure 3.5 also illustrates the fitted Poisson regression models for the number of tropical and ice days over time. A clear upward trend is visible in the tropical days, indicating an increasing frequency of hot days as the years progress. Conversely, the ice days exhibit a subtle downward trend, suggesting a gradual decline in the number of cold days.

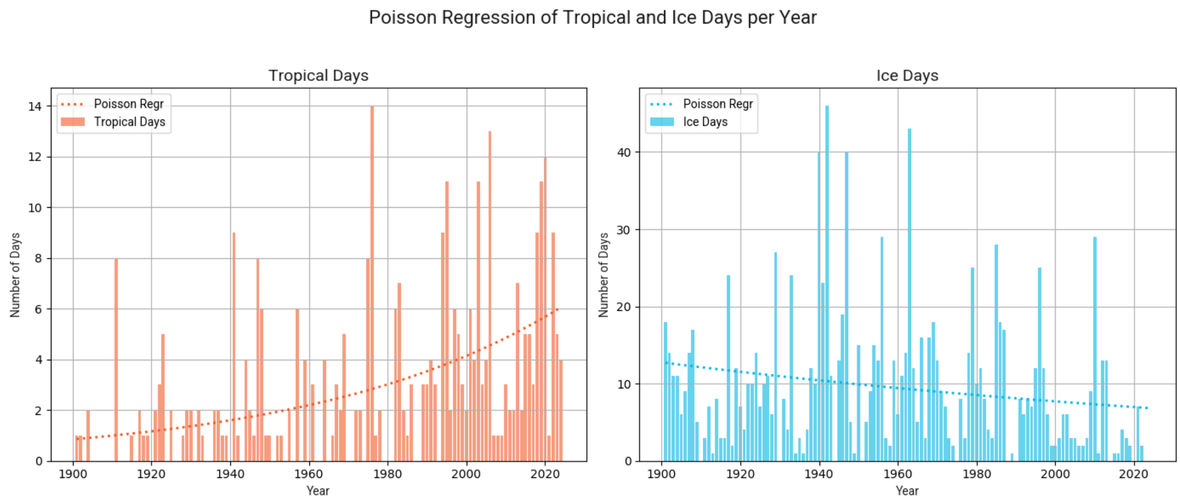

Figure 3.5: Poisson regression of extreme weather events per year. The left plot shows the number of Tropical Days (maximum temperature ≥ 30°C) per year with the Poisson regression fit, and the right plot shows the number of Ice Days (maximum temperature < 0°C) per year with the Poisson regression fit.

3.3.2. Application to Precipitation Event Counts

To model the temporal trends in the annual counts of dry, wet, and extreme wet days, we again apply Poisson regression. Specifically, for year t, the model expresses the logarithm of the expected count N_t as a linear function of time

$$\log\left(\mathbb{E}\left[N_t^{\text{dry}}\right]\right) = \alpha_{\text{dry}} + \beta_{\text{dry}}\, t,$$
$$\log\left(\mathbb{E}\left[N_t^{\text{wet}}\right]\right) = \alpha_{\text{wet}} + \beta_{\text{wet}}\, t,$$
$$\log\left(\mathbb{E}\left[N_t^{\text{ext}}\right]\right) = \alpha_{\text{ext}} + \beta_{\text{ext}}\, t,$$

where α represents the baseline log-count and β measures the trend over time for each precipitation category. Positive values of β indicate an increasing trend in the number of days, while negative values indicate a decrease. This framework allows us to quantify and test for significant changes in precipitation patterns over the observation period.

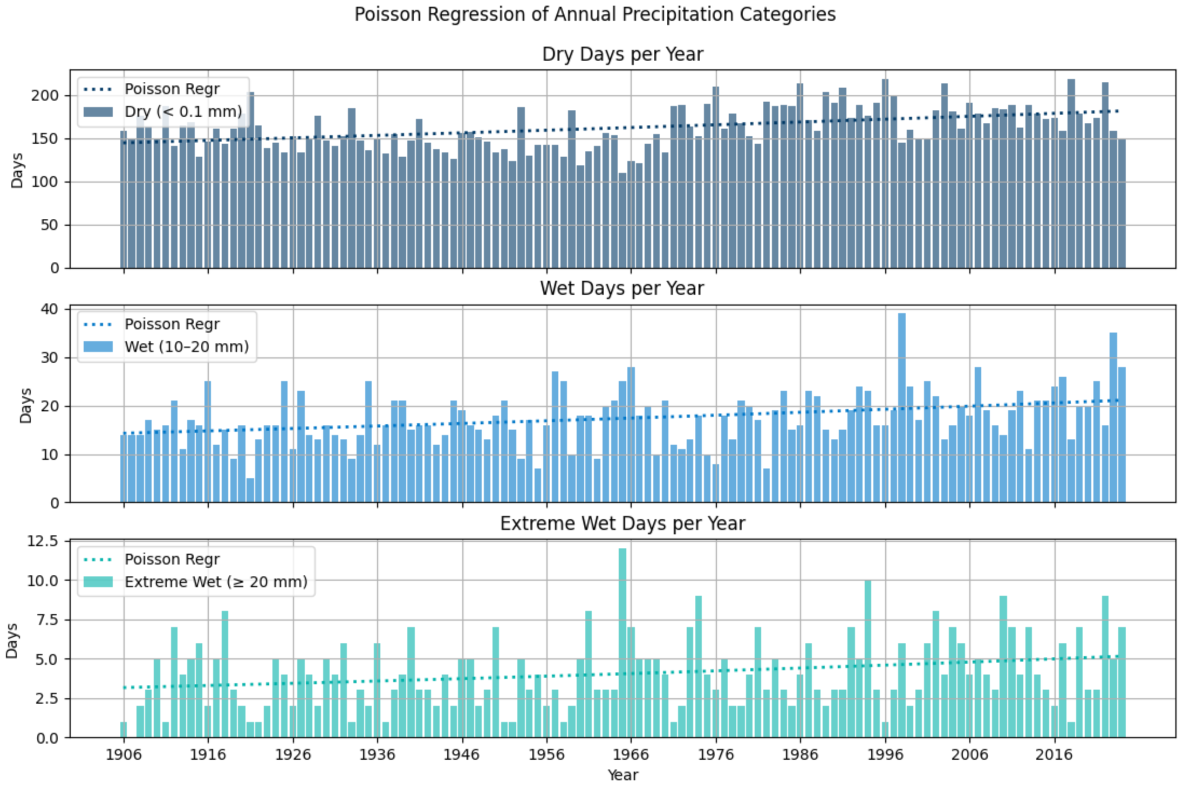

Figure 3.6: Poisson regression preformed on annual number of days in De Bilt categorized as dry, wet, or extreme wet.

Visualizations of annual counts of dry, wet, and extreme wet days from 1906 onwards can be found in Figure 3.6. The separated bar plots here offer a clearer and more detailed visualization of annual trends in daily precipitation categories compared to a stacked bar chart in Figure 2.6. This format makes it easier to observe year-to-year fluctuations and long-term patterns. Each subplot shows the observed counts as bars and the fitted Poisson regression trends. The dry and wet day trends exhibit somewhat steeper slopes over time compared to the extreme wet days, indicating a stronger temporal change in those categories. However, since Poisson regression assumes the mean equals the variance, it is important to check for overdispersion in the count data. This is further discussed in Chapter 4 and 5.

3.4. Isotonic Regression for Count Data

Classical isotonic regression (see Section 3.2) provides a non-parametric method for estimating monotonic trends in continuous data. However, because it relies on squared error loss, it is not well suited for count data. In cases where the response variable takes only non-negative integer values and is better described by a Poisson process, using a Gaussian error model can result in biased estimates and an inaccurate representation of the data's variability. Therefore, we adopt a slightly different approach that combines the monotonicity constraint of isotonic regression with the likelihood-based framework of Poisson regression (see Section 3.3), which is more appropriate for modelling counts.

The Isotonic regression for count data used here, assumes that the observed counts $y_1, y_2, \ldots, y_n$ follow independent Poisson distributions

$$y_i \sim \text{Poisson}(\lambda_i),$$

where the intensity parameters λ_i are constrained to follow a monotonic trend, such that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.

Instead of minimizing squared errors, the model minimizes the Poisson negative log-likelihood,

$$\hat{\lambda} = \arg \min_{\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n} \sum_{i=1}^{n} \left(\lambda_i - y_i \log \lambda_i \right),$$

which corresponds to maximizing the likelihood under the Poisson assumption. This formulation ensures that the discrete nature of the response is respected, and that the variance is modelled appropriately as a function of the mean, as opposed to being held constant. More details can be found in *Generalized isotonic regression* [10].

In this study, we used the Python library `cvxpy` to fit the Poisson isotonic regression model. The model was estimated by minimizing the negative log-likelihood under monotonicity constraints, using the ECOS solver. We applied these constraints either directly to the rate parameters λ_i or to their log values for better stability. This approach works well for modelling monotonic trends in count data, especially in settings where the data are discrete and follow a natural order, such as event counts over time.

As with other isotonic methods, a decreasing form of Poisson isotonic regression is obtained by reversing the inequality constraints to enforce a non-increasing trend in the fitted intensity parameters.

3.4.1. Application on Counts of Extreme Weather Events

For climate indicators expressed as annual counts, such as the number of tropical days or ice days, we apply isotonic regression within a Poisson modelling framework. We assume the count response Y follows a Poisson distribution with mean $\lambda(t)$, which varies monotonically over time:

$$Y \sim \text{Poisson}(\lambda(t)), \quad \text{with} \quad \log(\lambda(t)) = f(t),$$

where f is a monotonic function to be estimated.
This approach is used for variables such as $Y \in \{N_t^{\text{trop}}, N_t^{\text{ice}}, N_t^{\text{dry}}, N_t^{\text{wet}}, N_t^{\text{ext}}\}$, which represent event counts per year. For example, N_t^{trop} (number of tropical days) is expected to increase over time, while N_t^{ice} is expected to decrease. By fitting f under a monotonicity constraint, we preserve the expected direction of change while avoiding assumptions about the exact rate or form of the trend.

The isotonic Poisson regression framework is well-suited for detecting long-term trends in climatological count data, such as the frequency of extreme events. It does not assume a specific functional form for the trend, making it flexible and robust. Because isotonic regression makes minimal assumptions, it can be applied directly to count data.

In Figure 3.7, isotonic regression has been applied to both the number of tropical and ice days. The tropical days trend shows a clear and consistent upward pattern, with the line rising in almost constant upward steps over time. This indicates a steady increase in the number of tropical days throughout the observed period. In contrast, the number of ice days follows a downward trend, with the steps starting to decline around 1970. This suggests that the frequency of ice days has gradually decreased since then, with fewer years experiencing a significant number of ice days. These trends are consistent with the results obtained from the Poisson regression analysis that is visualized in Figure 3.5.

However, it is important to note that interpreting the trends at the start and end of the data can be somewhat misleading, like mentioned before. At these points, the regression is influenced by data from only one side, which may not fully represent the underlying long-term trend. Caution is needed when drawing conclusions from the boundaries of the dataset.

In Figure 3.8, for each category, an isotonic regression line is included, which provides a smoothed, non-decreasing trend based on the data. These regression lines show small step-like increases, indicating subtle upward trends over the long term. Although the number of days changes from year to year, the isotonic trends suggest that the number of dry, wet, and very wet days has generally increased slightly since the early 1900s. This may reflect shifts in precipitation intensity or frequency over time. Overall, this pattern aligns well with previous Poisson regression analyses in Figure 3.6, suggesting positive trends in precipitation regimes.

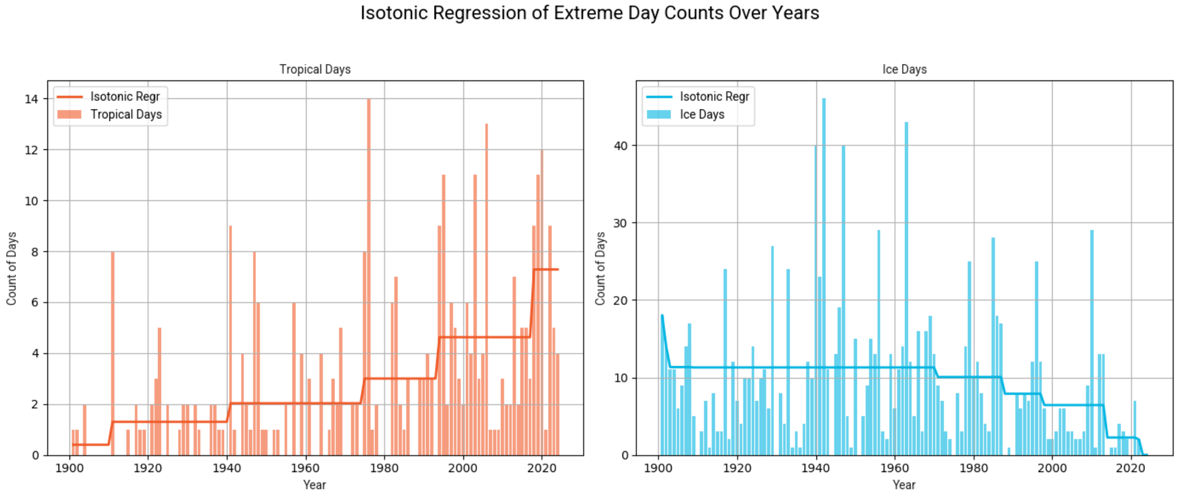

Figure 3.7: Isotonic regression trends of Tropical and Ice Days per year. The left plot shows the number of Tropical Days (maximum temperature ≥ 30°C) and the right plot shows the number of Ice Days (maximum temperature < 0°C), the maximum temperature is measured in the Bilt.

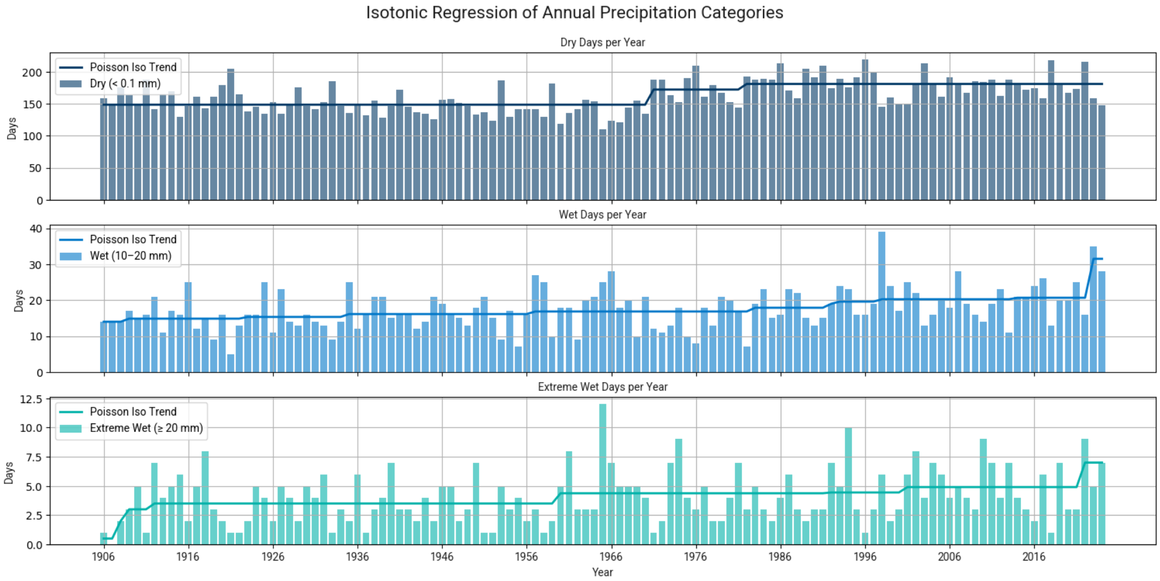

Figure 3.8: Annual number of days in De Bilt categorized as dry, wet, or extreme wet based on daily precipitation totals (1906–2024) in De Bilt. Isotonic regression lines highlight trends in each category.

3.5. Negative Binomial Regression

Negative Binomial regression is a count regression model used to analyze outcome variables that represent counts and exhibit overdispersion. This occurs when the variance of the response variable exceeds the mean, violating the equidispersion assumption of the Poisson regression model and motivating the use of a more flexible alternative [7].

In our case, we consider a single predictor variable x_i. The model expresses the expected count $\mu_i = \mathbb{E}[y_i]$ as

$$\log(\mu_i) = \beta_0 + \beta_1 x_i.$$

The variance of the outcome y_i is then modelled as

$$\mathrm{Var}(y_i) = \mu_i + \theta\mu_i^2,$$

where $\theta > 0$ is the dispersion parameter that quantifies the degree of overdispersion. When $\theta = 0$, the model reduces to standard Poisson regression.

The standard approach to estimating parameters in the Negative Binomial model is to maximize the full log-likelihood over both $\boldsymbol{\beta} = (\beta_0, \beta_1)$ and θ simultaneously. For observation i, the log-likelihood contribution is

$$\ell_i(\boldsymbol{\beta}, \theta) = \log\Gamma(y_i + \theta^{-1}) - \log\Gamma(y_i + 1) - \log\Gamma(\theta^{-1}) + y_i \log(\theta^{-1}\mu_i) - (y_i + \theta^{-1})\log(1 + \theta^{-1}\mu_i).$$

The total log-likelihood is then

$$\ell(\boldsymbol{\beta}, \theta) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\beta}, \theta).$$

Because the likelihood function for the Negative Binomial model does not have a closed-form solution, it must be maximized using numerical methods. Although this allows for joint estimation of both the regression coefficients $\boldsymbol{\beta}$ and the dispersion parameter θ, the procedure can be unstable in practice. It is often sensitive to poor starting values or to covariates that are on very different scales, which can lead to convergence issues or unreliable estimates. To avoid these difficulties, a two-step estimation procedure is often used as a more stable and practical alternative.

3.5.1. Two-Step Estimation

This approach first fits a Poisson regression model to estimate $\boldsymbol{\beta}$, and then estimates θ separately using the residual variation captured by the Pearson chi-squared statistic. Note overdispersion is ignored to obtain initial fitted means

$$\hat{\mu}_i = \exp(\hat{\beta}_0 + \hat{\beta}_1 x_i).$$

These fitted values serve as estimates for the true conditional means μ_i.

The key idea is to estimate the dispersion parameter θ by comparing the observed variance to the variance implied by the Poisson model using the Pearson chi-squared statistic

$$X^2 = \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i}.$$

This statistic is used because it provides a standardized measure of the discrepancy between observed counts y_i and their expected values $\hat{\mu}_i$ under the fitted Poisson model. Specifically, it sums the squared residuals scaled by the model-implied variance. Since the Poisson model assumes the variance equals the mean, the denominator $\hat{\mu}_i$ standardizes each squared difference, making the contributions from all observations comparable regardless of their mean level.

Under the Poisson assumption, the variance equals the mean, so $\mathbb{E}[X^2] \approx n - p$, where p is the number of estimated parameters in the model (degrees of freedom). When overdispersion is present, the observed variability in the residuals will be larger than expected under the Poisson model, causing the Pearson chi-squared statistic to be systematically larger than $n - p$. Therefore, comparing the observed X^2 to its expected value under the Poisson assumption allows us to quantify the extra variability in the data. This extra variability can then be attributed to the dispersion parameter θ in the Negative Binomial model, which models the variance as

$$\mathrm{Var}(y_i) = \mu_i + \theta\mu_i^2.$$

Thus, the Pearson chi-squared statistic serves as a natural tool to estimate θ by capturing how much the observed variance deviates from the Poisson variance assumption.

We derive an estimator for θ by equating the observed Pearson statistic to its expectation under the Negative Binomial model

$$\begin{aligned}\mathbb{E}[X^2] &= \sum_{i=1}^{n} \frac{\mathrm{Var}(y_i)}{\hat{\mu}_i} \\ &= \sum_{i=1}^{n} \frac{\hat{\mu}_i + \theta\hat{\mu}_i^2}{\hat{\mu}_i} \\ &= \sum_{i=1}^{n} (1 + \theta\hat{\mu}_i) \\ &= n + \theta \sum_{i=1}^{n} \hat{\mu}_i.\end{aligned}$$

Now, we approximate the expectation with the observed value X^2, treating the fitted means $\hat{\mu}_i$ as plug-in estimates for μ_i. This is a form of moment matching, where the sample-based Pearson statistic is assumed to be close to its theoretical expectation under the Negative Binomial model. More detail can be found in the book *Statistical Inference* [5]. Equating the two expressions gives

$$X^2 \approx n + \theta \sum_{i=1}^{n} \hat{\mu}_i.$$

Solving for θ, we obtain the following estimator:

$$\hat{\theta} = \frac{X^2 - n}{\sum_{i=1}^{n} \hat{\mu}_i} = \frac{\sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i} - n}{\sum_{i=1}^{n} \hat{\mu}_i}.$$

Using the estimated $\hat{\theta}$ from above, we then refit the regression coefficients $\boldsymbol{\beta}$ by maximizing the Negative Binomial likelihood while keeping θ fixed. This two-step procedure typically improves numerical stability and convergence properties compared to joint estimation, especially with moderate sample sizes or complex models.

We implemented this two-step method using Python's `statsmodels` library. The Poisson regression provides fitted means $\hat{\mu}_i$, which are used to estimate $\hat{\theta}$. This estimate is passed as a fixed parameter to the `NegativeBinomial` family in the final model to estimate $\boldsymbol{\beta}$ via maximum likelihood.

3.5.2. Application to Temperature and Precipitation

Each of the annual event counts defined above (e.g., $N_t^{\text{trop}}, N_t^{\text{ice}}, N_t^{\text{dry}}, N_t^{\text{wet}}, N_t^{\text{ext}}$) is modelled separately using a negative binomial regression. As an example, for the number of tropical days N^{trop}, we assume

$$N_t^{\text{trop}} \sim \mathrm{NB}(\mu_t, \theta)$$

with a log-linear model for the expected value:

$$\log(\mu_t) = \beta_0 + \beta_1 t$$

Here, μ_t is the expected number of tropical days in a given year, t is the year (used as a continuous predictor), and θ is the dispersion parameter that accounts for overdispersion.

The variance is given by

$$\mathrm{Var}(N_t^{\text{trop}}) = \mu_t + \theta\mu_t^2$$

Models for the other event counts $N_t^{\text{ice}}, N_t^{\text{dry}}, N_t^{\text{wet}}, N_t^{\text{ext}}$ are specified in the same way, with time as predictor.

Figure 3.9 shows the Negative Binomial regression fits for the counts of tropical days and ice days. Visually, there is a clear increasing trend in tropical days and a clear decreasing trend in ice days over time. These patterns are consistent with earlier observations from Isotonic and Poisson regression and suggest notable changes in temperature extremes over the study period.

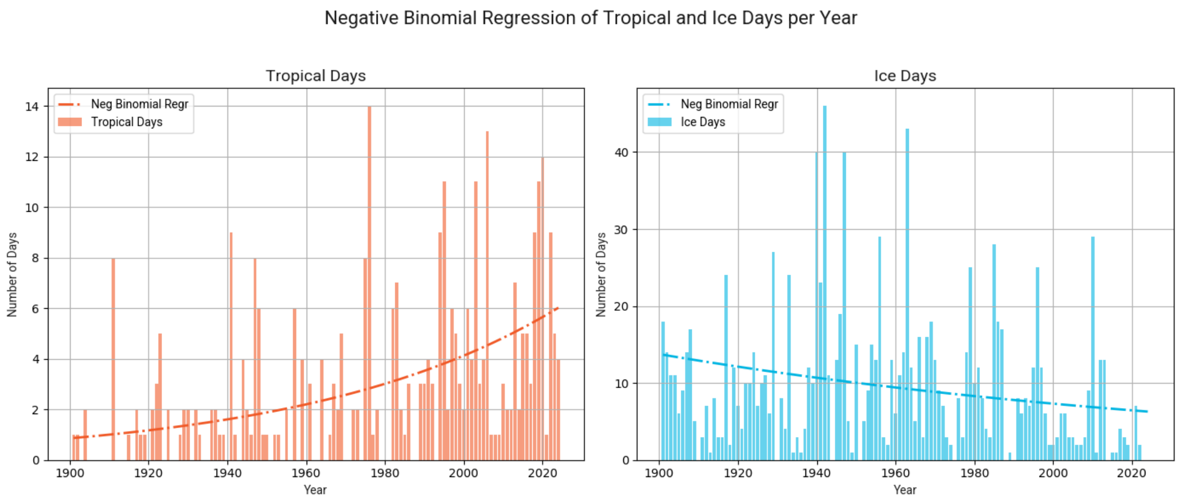

Figure 3.9: Negative Binomial regression of extreme weather events per year. The left plot shows the number of Tropical Days (maximum temperature ≥ 30°C) per year with the Negative Binomial regression fit, and the right plot shows the number of Ice Days (maximum temperature < 0°C) per year with the Negative Binomial regression fit.

Lastly, Figure 3.10 shows the Negative Binomial regression fits for the counts of dry days, wet days, and extreme wet days. All three variables display a slight increasing trend over the study period. These visual observations also align with the trends identified by the other models discussed earlier. Formal hypothesis testing of these models for Negative Binomial regression will be addressed in Chapter 4 and 5.

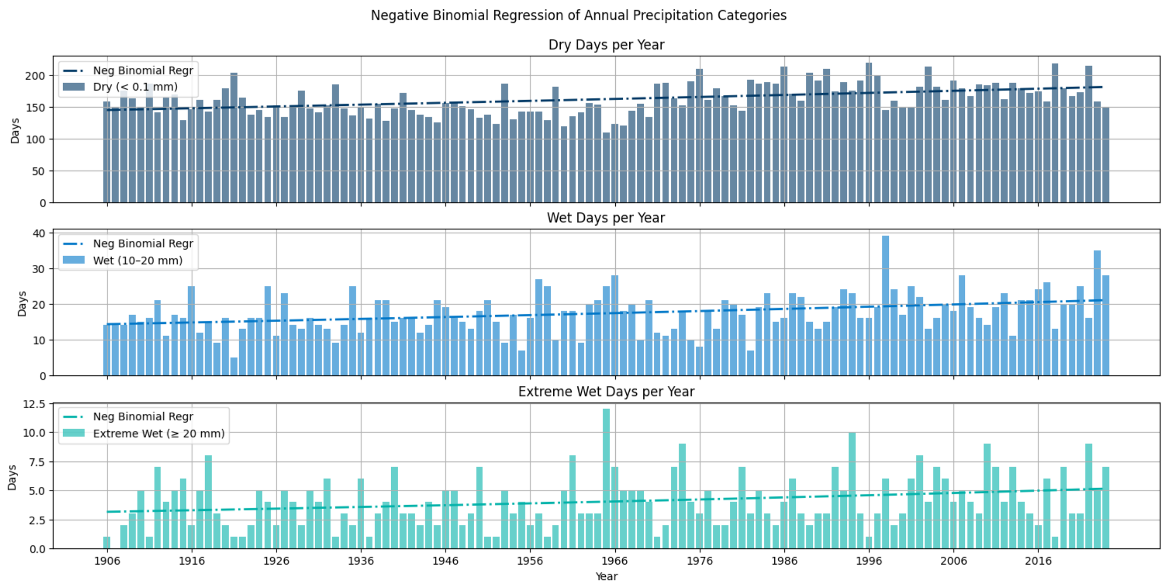

Figure 3.10: Negative Binomial regression preformed on annual number of days in De Bilt categorized as dry, wet, or extreme wet.

4

Hypothesis Testing for Trend Detection

This chapter focuses on hypothesis testing for detecting trends in temperature, precipitation, and climate-related count data such as tropical and ice days. Based on the regression models introduced in previous chapters, namely isotonic, linear, Poisson, and negative binomial, we present the appropriate testing frameworks to evaluate the presence of statistically significant trends. In particular, we are interested in detecting whether, in most cases, there is a *positive trend* or, in the case of ice and dry days, *negative trend* over time. The Python implementation can be found in Appendix A.3.

4.1. Linear Regression

As introduced in Section 3.1, linear regression models the relationship between a continuous response variable and time:

$$Y_i = a + bx_i + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2).$$

In this study, separate linear models are fitted to the yearly, summer, and winter subsets of the data to assess seasonal differences in trends, where t denotes the year (or year index).

4.1.1. Estimated Models for Temperature and Precipitation

The fitted linear regression models for the yearly average temperature over time are:

$$\begin{aligned}
\widehat{T}_t^{\text{year}} &= -23.5830 + 0.0169\, t, \\
\widehat{T}_t^{\text{summer}} &= -20.7640 + 0.0189\, t, \\
\widehat{T}_t^{\text{winter}} &= -30.7065 + 0.0171\, t,
\end{aligned}$$

The fitted models for total precipitation over time are:

$$\begin{aligned}
\widehat{P}_t^{\text{year}} &= -122.4626 + 0.0965\, t, \\
\widehat{P}_t^{\text{summer}} &= 64.0527 + 0.0056\, t, \\
\widehat{P}_t^{\text{winter}} &= -223.2284 + 0.1473\, t,
\end{aligned}$$

Here, T_t represents the average temperature (in °C) and P_t the total precipitation (in millimetres) for year t. The slope coefficients represent the estimated average change per year.

All temperature models show positive slopes, indicating increasing trends over time. Specifically, the yearly temperature increases by approximately 0.0965°C per year, while summer and winter show smaller but still positive increases of about 0.0189°C and 0.0171°C per year, respectively. This suggests a stronger overall warming trend on an annual basis. For precipitation, the positive slopes indicate increasing trends as well, with the strongest increase in winter precipitation (0.1473 mm/year), a modest increase in summer (0.0056 mm/year), and a smaller increase annually (0.0169 mm/year).

4.1.2. Slope Testing

To test whether a statistically significant positive trend exists over the years (i.e., whether the slope parameter b is greater than zero), we consider the one-sided hypotheses:

$$H_0 : b = 0 \quad \text{(no trend, i.e., no change in expected response)}$$
$$H_1 : b > 0 \quad \text{(positive trend in the response variable)}$$

The test statistic is defined as:

$$t = \frac{\hat{b}}{\mathrm{SE}(\hat{b})},$$

where $\hat{b}$ is the estimated slope and $\mathrm{SE}(\hat{b})$ its standard error. Under H_0, t follows a t-distribution with $n-2$ degrees of freedom. The corresponding p-value assesses statistical significance.

We reject H_0 in favour of H_1 if the test statistic exceeds the critical value $t_{1-\alpha,n-2}$ from the t-distribution at significance level α, or equivalently if the one-sided p-value is less than α.

$$P(T > t_{1-\alpha,n-2}) = \alpha, \quad T \sim t_{n-2}.$$

Additionally, a $(1-\alpha) \times 100\%$ confidence interval for the slope can be used; if the entire interval lies above zero, this further supports a positive trend. This testing procedure is applied to different seasonal and annual temperature and precipitation data subsets. Test statistics, p-values, critical values at $\alpha = 0.05$, and confidence intervals for slopes are computed for each model. The results of these tests are presented in Chapter 5.

4.2. Classic Isotonic Regression

Isotonic regression is, as introduced in Section 3.2, a non-parametric technique for estimating a monotonic (non-decreasing or non-increasing) relationship between an ordered explanatory variable x and a response variable y, without assuming a specific functional form. Given data points $(x_1, y_1), \ldots, (x_n, y_n)$ with $x_1 < x_2 < \cdots < x_n$, isotonic regression estimates a monotonic function f by solving

$$\min_{f(x_1) \le f(x_2) \le \cdots \le f(x_n)} \sum_{i=1}^{n} \left(y_i - f(x_i)\right)^2.$$

The goal is to test whether there is evidence of a monotonic trend in the data. This can be formalized as a hypothesis test on the expected values $\mu_i = \mathbb{E}[Y_i]$ at each ordered x_i:

$$H_0 : \quad \mu_1 = \mu_2 = \cdots = \mu_n \quad \text{(no trend)}$$
$$H_1 : \quad \mu_1 \le \mu_2 \le \cdots \le \mu_n \quad \text{(with at least one strict inequality)}$$

Because classical parametric tests are not suitable in this setting, a permutation-based hypothesis test based on the reduction in residual sum of squares (RSS) from fitting an isotonic regression versus a flat (constant mean) model as the test statistic is used. This approach is well-suited for detecting monotonic trends without assuming linearity or normality.

4.2.1. Permutation Test

To quantify how much better isotonic regression fits the data compared to a null model with no trend, we compute the test statistic

$$T_{\text{obs}} = \mathrm{RSS}_{\text{constant}} - \mathrm{RSS}_{\text{isotonic}},$$

where

$$\mathrm{RSS}_{\text{constant}} = \sum_{i=1}^{n} (y_i - \bar{y})^2,$$
$$\mathrm{RSS}_{\text{isotonic}} = \sum_{i=1}^{n} (y_i - \hat{f}(x_i))^2.$$

Here, $\bar{y}$ is the sample mean and $\hat{f}(x_i)$ is the fitted value from the isotonic regression. The larger the value of T_{obs}, the greater the improvement in fit by the isotonic model, compared to the constant model, indicating stronger monotonic structure.

To give an example, the original isotonic regression fits were plotted alongside the mean temperature values for the yearly, summer, and winter average temperature in Figure 4.1. These plots show clear monotonic trends in the data, with the isotonic regression lines deviating from the constant mean lines, indicating potential temporal changes in temperature averages.

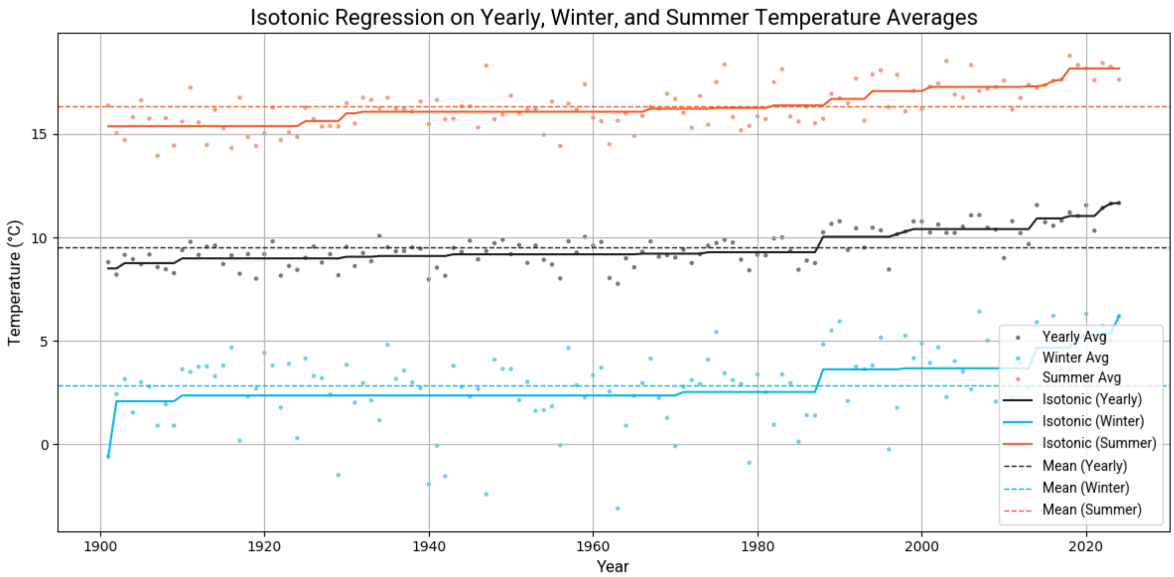

Figure 4.1: Isotonic regression trends of average yearly, summer, and winter temperatures in De Bilt, shown alongside the corresponding mean temperatures.

To assess the significance of the observed test statistic, we generate a draw from the null distribution, given the individual values of y_i, by permuting the response values y_i while keeping x_i fixed. For each permutation b, the isotonic regression is refitted and the test statistic T_b is recomputed

$$T_b = \text{RSS}_{\text{constant}} - \text{RSS}_{\text{isotonic}}^{(b)}.$$

An example of a permutation result can be found in Figure 4.2. When comparing the fitted isotonic regression line from the permuted data to the mean, it is often relatively flat and closely follows the average, indicating an absence of a clear monotonic trend. This contrasts with the original fit (in Figure 4.1), where the isotonic regression captures a more pronounced upward trend.

This process is repeated B times (in this research, $B = 10{,}000$) to generate an empirical null distribution under the assumption that there is no monotonic relationship between x and y.

The one-sided p-value is then computed as

$$p = \frac{1 + \sum_{b=1}^{B} \mathbb{I}(T_b \geq T_{\text{obs}})}{B + 1},$$

where $\mathbb{I}(\cdot)$ is the indicator function. A small p-value indicates that the observed improvement in fit is unlikely under the null hypothesis, supporting the presence of a statistically significant monotonic trend.

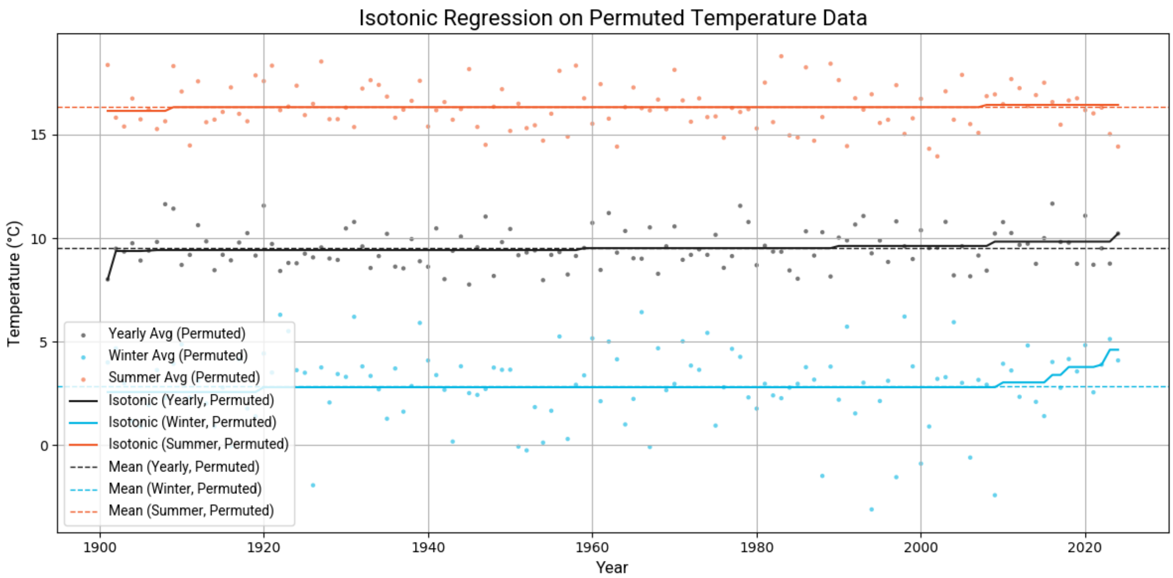

Figure 4.2: Example of a permuted time series of average temperatures (yearly, summer, and winter). The isotonic regression lines fitted to the permuted data are solid lines. Horizontal lines indicate the mean temperature for each season.

4.3. Poisson Regression

To model count data, such as the number of tropical or ice days per year, Poisson regression is used (see Section 3.3). This approach assumes that the dependent variable follows a Poisson distribution, with a mean that varies as a function of time:

$$Y_i \sim \text{Poisson}(\lambda_i), \quad \text{with} \quad \log(\lambda_i) = a + bx_i.$$

The logarithmic link function ensures that the expected count λ_i remains strictly positive. A positive value of the slope b implies an increasing expected count over time, while a negative value indicates a decreasing trend.

4.3.1. Estimated Models for Extreme Temperatures and Precipitation Counts

To illustrate, the fitted Poisson regression models for tropical and ice days take the following forms:

$$\widehat{\log \mathbb{E}[N_t^{\text{trop}}]} = -30.2804 + 0.0159\,t,$$
$$\widehat{\log \mathbb{E}[N_t^{\text{ice}}]} = 12.1281 - 0.0050\,t.$$

Similarly, the fitted models for dry, wet, and extremely wet days are:

$$\widehat{\log \mathbb{E}[N_t^{\text{dry}}]} = 1.3010 + 0.0019\,t,$$
$$\widehat{\log \mathbb{E}[N_t^{\text{wet}}]} = -3.6373 + 0.0033\,t,$$
$$\widehat{\log \mathbb{E}[N_t^{\text{ext}}]} = -6.7669 + 0.0042\,t.$$

These equations describe the logarithm of the expected number of days of each type as a linear function of year t. The slope coefficient b in each model reflects the annual rate of change in the log-expected count. For example, a slope of $b = 0.0159$ for tropical days implies that the expected number of tropical days increases by a factor of $\exp(0.0159) \approx 1.016$ per year, or roughly 1.6% growth per year. Conversely, the negative slope of -0.0050 for ice days indicates a declining trend, with expected counts decreasing by about 0.5% per year.

The other models for precipitation show small positive slopes, suggesting slightly increasing trends over time. A formal test of statistical significance for these slopes is presented in the next section.

4.3.2. Slope Testing

To determine whether there is a statistically significant trend in the count data, we test the slope parameter b of the Poisson regression model.

For variables with expected increases (e.g., tropical days and the precipitation models), the hypotheses are:

$$H_0 : b = 0 \quad \text{(no trend)}$$
$$H_1 : b > 0 \quad \text{(positive trend)}$$

For variables with expected decreases (e.g., ice days), the alternative hypothesis is reversed

$$H_0 : b = 0 \quad \text{(no trend)}$$
$$H_1 : b < 0 \quad \text{(negative trend)}$$

Unlike linear regression, where slope inference is based on a t-distribution and normally distributed errors, Poisson regression assumes a count response with a Poisson distribution. This makes classical t-tests invalid for testing coefficients. Therefore different methods suited for generalized linear models are needed.

The first approach to hypothesis testing is the Wald test, which uses the estimated slope coefficient $\hat{b}$ and its estimated standard error $\widehat{\mathrm{SE}}(\hat{b})$. The standard error is defined as the square root of the estimated variance of $\hat{b}$, typically obtained from the inverse Fisher information matrix. The test statistic is given by

$$z = \frac{\hat{b}}{\widehat{\mathrm{SE}}(\hat{b})}.$$

Under the null hypothesis $H_0: b = 0$, this statistic approximately follows a standard normal distribution,

$$z \sim \mathcal{N}(0,1).$$

This allows one to assess whether the slope significantly differs from zero by comparing z to critical values from the normal distribution.

The second approach is the Likelihood Ratio Test (LRT), which compares the goodness of fit between the full model (including the slope term) and a reduced model (excluding it). The test statistic is

$$\Lambda = -2\left[\ell_{\text{reduced}} - \ell_{\text{full}}\right],$$

where ℓ denotes the log-likelihood. Under the null hypothesis, $\Lambda \sim \chi^2_1$. A significant value of Λ indicates that the model with the time trend fits the data significantly better.

A statistically significant result from either test suggests the presence of a systematic change in the count data over time.

4.4. Isotonic Regression for Count Data

Classical isotonic regression (see Section 4.2) works well for continuous data with constant variance. It typically assumes Gaussian errors and minimizes the residual sum of squares (RSS). However, these assumptions do not hold for count data, where observations are discrete and their variability scales with the mean.

To address this, we use a generalized isotonic regression framework that replaces the RSS with a likelihood-based loss function appropriate for the data's distribution. For count data, this leads to a Poisson-based model, where each observed count y_i is assumed to follow a Poisson distribution with mean λ_i, subject to monotonicity

$$y_i \sim \text{Poisson}(\lambda_i), \quad \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n.$$

Instead of minimizing squared errors, the model minimizes the Poisson negative log-likelihood

$$\min_{\lambda_1 \le \cdots \le \lambda_n} \sum_{i=1}^{n} \left(\lambda_i - y_i \log \lambda_i\right),$$

which properly accounts for the mean-variance relationship in count data [10]. This approach generalizes isotonic regression to settings where distributional assumptions other than Gaussian are more appropriate.

4.4.1. Permutation Test with Poisson Likelihood

Similar to the classical isotonic regression permutation test (Section 4.2), we test for a monotonic trend by comparing the isotonic model for count data against a null model where the rate is constant. The test statistic is

$$T_{\text{obs}} = \mathscr{L}_{\text{constant}} - \mathscr{L}_{\text{isotonic}},$$

where

$$\mathscr{L}_{\text{constant}} = \sum_{i=1}^{n} \left(\bar{y} - y_i \log \bar{y}\right),$$
$$\mathscr{L}_{\text{isotonic}} = \sum_{i=1}^{n} \left(\hat{\lambda}_i - y_i \log \hat{\lambda}_i\right).$$

Here, $\bar{y}$ is the overall mean count and $\hat{\lambda}_i$ are the fitted values from the isotonic regression.
To understand if the observed test statistic is unusual, we create a null distribution by permuting the count data many times, fitting the generalized isotonic model to each permuted set, and calculating

$$T_b = \mathscr{L}_{\text{constant}} - \mathscr{L}^{(b)}_{\text{isotonic}} \quad \text{for } b = 1, \ldots, B.$$

The p-value is the proportion of permuted statistics T_b that are as large or larger than T_{obs}:

$$p = \frac{1 + \sum_{b=1}^{B} \mathbb{I}(T_b \geq T_{\text{obs}})}{B+1}.$$

This method respects the discrete and overdispersed nature of count data, making it a more appropriate test for monotonic trends in counts than classical isotonic regression [10].

4.5. Negative Binomial Regression

When overdispersion is present, the Poisson model's assumptions are violated. In such cases, the Negative Binomial (NB) regression offers a more flexible alternative by introducing a dispersion parameter. The model, as discussed in Section 3.5, is specified as

$$Y_i \sim \text{NB}(\mu_i, \theta), \quad \log(\mu_i) = a + b x_i,$$

where μ_i is the expected count, and θ is the dispersion parameter that accounts for extra variability.

4.5.1. Estimated Models for Extreme Temperatures and Precipitation Counts

The fitted NB regression models for tropical and ice days are:

$$\widehat{\log \mathbb{E}\left[N_t^{\text{trop}}\right]} = -30.0591 + 0.0157\,t \quad \text{with } \theta = 0.7568,$$
$$\widehat{\log \mathbb{E}\left[N_t^{\text{ice}}\right]} = \quad 14.6205 - 0.0063\,t \quad \text{with } \theta = 0.8425,$$

These functions describe the logarithm of the expected count of extreme temperature days as a linear function of time. The positive slope in the tropical day model indicates a rising trend over the years, while the negative slope in the ice day model suggests a decline, both in line with expectations under climate change.

Similarly, the models for dry, wet, and extremely wet days are

$$\widehat{\log \mathbb{E}\left[N_t^{\text{dry}}\right]} = \quad 1.3650 + 0.0019\,t \quad \text{with } \theta = 0.0119,$$
$$\widehat{\log \mathbb{E}\left[N_t^{\text{wet}}\right]} = -3.5656 + 0.0033\,t \quad \text{with } \theta = 0.0321,$$
$$\widehat{\log \mathbb{E}\left[N_t^{\text{ext}}\right]} = -6.7666 + 0.0042\,t \quad \text{with } \theta = 0.0393.$$

These equations also suggest slight upward trends in precipitation-related extremes. For example, the increasing slope for extremely wet days implies a rise in the frequency of such events over time.

Notably, the Negative Binomial regression estimates are numerically quite similar to those obtained from the Poisson regression in Section 4.3.1. This similarity indicates that the primary trend patterns are consistent across models, while the NB formulation provides better accommodation for overdispersion and excess zeros.

4.5.2. Slope Testing

As with Poisson regression, we test for the presence of a time trend by evaluating the slope parameter b. The null and alternative hypotheses follow the same one-sided logic

$$H_0 : b = 0, \qquad H_1 : b > 0 \quad (\text{or } H_1 : b < 0 \text{ for expected negative trends}).$$

For negative binomial regression, the same two hypothesis testing approaches, described in Section 4.3.2 can be applied: the Wald test and the Likelihood Ratio Test. These tests retain their general structure, but account for the distinct variance specification of the negative binomial model.

The first approach is the Wald test, which has the same mathematical form as in the Poisson case:

$$z = \frac{\hat{b}}{\widehat{\mathrm{SE}}(\hat{b})},$$

where $\hat{b}$ is the estimated slope coefficient and $\widehat{\mathrm{SE}}(\hat{b})$ is its standard error. In the Negative Binomial setting, this standard error is derived from the inverse of the observed Fisher information matrix based on the Negative Binomial likelihood, which includes the estimation of a dispersion parameter α that allows the variance to exceed the mean. Under the null hypothesis $H_0 : b = 0$, the test statistic approximately follows a standard normal distribution,

$$z \sim \mathcal{N}(0, 1).$$

Although the formula is identical to that used in Poisson regression, the standard error in Negative Binomial regression is typically larger due to the extra variability captured by the dispersion parameter, making the test more conservative in the presence of overdispersion. The second approach is the Likelihood Ratio Test (LRT), which compares the fit of the full model (including the slope term) to that of a reduced model (excluding it). The test statistic is given by

$$\Lambda = -2\,[\ell_{\text{reduced}} - \ell_{\text{full}}]\,,$$

where ℓ denotes the log-likelihood under the Negative Binomial model. As in Poisson regression, the test statistic approximately follows a chi-squared distribution with one degree of freedom under the null hypothesis,

$$\Lambda \sim \chi^2_1.$$

A key consideration is that the dispersion parameter θ must be estimated under both the full and reduced models to ensure a valid comparison of likelihoods. A statistically significant value of either the Wald test or the LRT statistic provides evidence that the slope coefficient differs significantly from zero, indicating a systematic change in the count data over time. Compared to the Poisson framework, Negative Binomial regression yields more reliable inference when the data exhibit overdispersion, by adjusting both standard errors and likelihoods to better reflect the underlying variance structure.

5

Results of Trend Tests

This chapter evaluates whether Dutch climate indicators show statistically significant increasing or decreasing trends over time by testing the trend under a one-sided alternative. For each variable, we estimate the trend using different regression models discussed in the chapter before, e.g. linear, two types of isotonic, Poisson, and negative binomial, depending on its distribution and scale. We report the estimated slopes, confidence intervals, and one-sided p-values to assess the strength and direction of these trends across temperature and precipitation metrics.

5.1. Linear Regression

The regression analysis results for temperature and precipitation in Table 5.1 reveal several interesting patterns in the trends of temperature and precipitation over time. The yearly temperature exhibits a highly significant positive trend, with an estimated slope of 0.0169 and a one-sided p-value effectively equal to zero, indicating strong evidence of warming across the entire period analysed. This warming is consistent across seasons, as reflected by the winter temperature trend, which also shows a statistically significant positive slope of 0.0171 and a correspondingly low p-value, further confirming that the observed increase is not limited to the annual average but persists during colder months as well. In both cases, the absolute t-statistics exceed the one-sided critical value of 1.645 at the 5% significance level, providing strong support for the alternative hypothesis of a positive trend.

Table 5.1: Results of linear trend analysis for temperature and precipitation trends. Temperature models use 124 observations (df = 122), precipitation models use 118 observations (df = 116).

Model	Coef.	Std. Err.	t-stat	p-value (1-sided)	95% CI
Yearly Temperature	0.0169	0.0017	10.183	0.0000×10^{0}	[0.0136, 0.0201]
Summer Temperature	0.0189	0.0021	9.111	9.9920×10^{-16}	[0.0148, 0.0230]
Winter Temperature	0.0171	0.0044	3.907	7.6956×10^{-5}	[0.0084, 0.0258]
Yearly Precipitation	0.0964	0.0303	3.188	9.2027×10^{-4}	[0.0365, 0.1564]
Summer Precipitation	0.0057	0.0634	0.089	4.6463×10^{-1}	[-0.1199, 0.1312]
Winter Precipitation	0.1472	0.0554	2.659	4.4746×10^{-3}	[0.0376, 0.2569]

In contrast, precipitation trends demonstrate greater variability across seasons. While the overall yearly precipitation displays a significant upward trend with a slope of 0.0964 and a p-value below 0.001, this pattern is not uniform throughout the year. Specifically, the summer precipitation trend is small and statistically insignificant, with a slope near zero and a p-value around 0.46. The corresponding t-statistic falls far below the critical threshold, suggesting no clear evidence of increased precipitation during summer months. On the other hand, winter precipitation shows a significant positive trend, with a slope of 0.1472 and a p-value of approximately 0.0045. Here, the t-statistic again exceeds the critical value of 1.645, indicating that precipitation has increased notably in the winter season.

5.2. Classic Isotonic Regression

To assess the presence of significant monotonic trends in the Dutch climate data, we compared the fit of a constant model to that of an isotonic regression using a permutation test as described in Section 4.2. The difference in residual sum of squares (RSS) between the two models served as the test statistic, and its significance was evaluated based on 10,000 permutations under the null hypothesis of no trend.

Figure 5.1 shows the permutation distributions of RSS differences for average yearly, summer, and winter temperatures. All three distributions are sharply peaked near zero, as expected under the null hypothesis, with frequencies rapidly decreasing as the RSS difference increases. In each subplot, the observed RSS difference (marked by a vertical line) lies far to the right of the distribution, indicating that the isotonic regression captures a strong monotonic trend in temperature that would be highly unlikely to happen by chance.

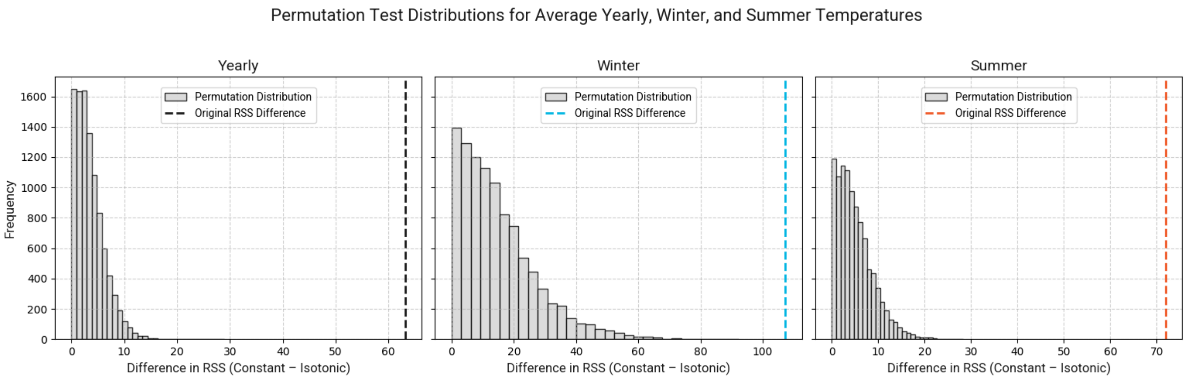

Figure 5.1: Permutation distributions of RSS differences between constant and isotonic regression models for Dutch average yearly, winter, and summer temperatures. The distributions are based on 10,000 permutations under the null hypothesis of no trend. The vertical line in each plot indicates the observed RSS difference from the original data.

Figure 5.2 displays the results for average precipitation, split by season and year. Significant monotonic trends are observed in both yearly and winter rainfall, where the observed RSS values lie near the upper tail of the permutation distributions. Summer precipitation, by contrast, shows no such trend. The observed RSS difference is near the peak of the null distribution, suggesting that any trend is likely due to random variation.

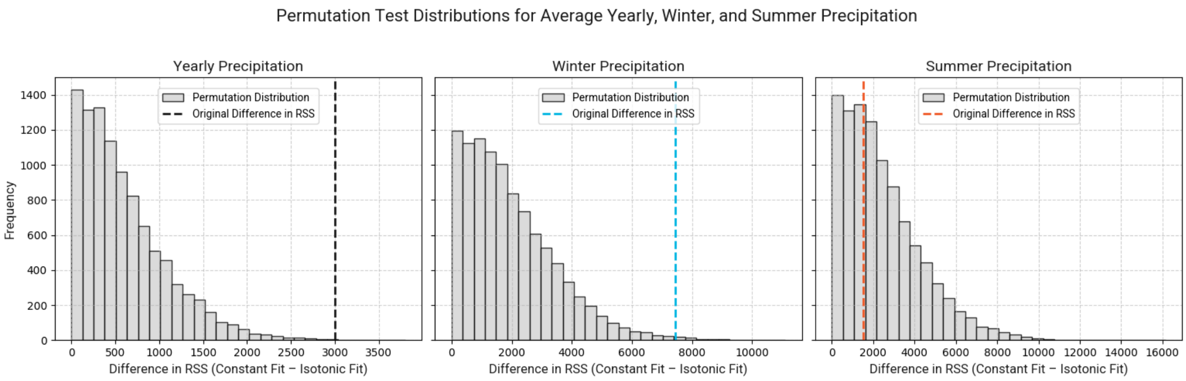

Figure 5.2: Permutation distributions of RSS differences between constant and isotonic regression models for average precipitation: yearly, winter, and summer. Each distribution is based on 10,000 permutations. Vertical lines indicate the observed RSS differences.

Table 5.4 summarizes the permutation test outcomes. The observed test statistics consistently exceed the maximum values from the permutation distributions for most variables, particularly for temperature, confirming that the trends seen in the figures are not due to random variation. The extremely low p-values (often near or below 10^{-3}) provide strong statistical evidence for these trends.

Table 5.2: Permutation test results for isotonic regression on temperature and precipitation data.

Model	Test Statistic	Permutation p-value	Permuted Stats Min	Permuted Stats Max
Yearly Temperature	63.169	9.9900×10^{-4}	0.0000	16.219
Summer Temperature	71.919	9.9900×10^{-4}	0.0000	26.638
Winter Temperature	107.26	9.9900×10^{-4}	0.0000	66.741
Yearly precipitation	3003.3	1.1999×10^{-3}	0.0000	3798.8
Summer precipitation	1518.8	6.1904×10^{-1}	0.0000	16150
Winter precipitation	7436.8	5.3995×10^{-3}	0.0000	11109

In summary, the isotonic regression analysis reveals strong and statistically significant monotonic trends in Dutch climate data, particularly for temperature and dry extremes. These trends are evident across most seasons and variables, although some, like summer precipitation and Ice Days, show more subdued signals. The results underline a consistent pattern of warming and drying, with notable seasonal and variable-specific variation.

5.3. Poisson Regression

The Poisson regression models examined the relationship between time and counts of various climate-related events: Ice Days, Tropical Days, and precipitation extremes categorized as Dry Days, Wet Days, and Extreme Wet Days. The results can be found in Table 5.3.

Table 5.3: Results of Poisson regression one-sided tests on the effect of time. Models for Ice and Tropical Days use 124 observations (df = 122), and precipitation models (Dry, Wet, Extreme Wet Days) use 119 observations (df = 117).

Model	Coef.	Std. Err.	z-stat	p-value (1-sided)	95% CI	Dispersion
Ice Days	-0.0050	0.0008	-6.1228	4.5974×10^{-10}	[-0.0067, -0.0034]	8.96
Tropical Days	0.0159	0.0017	9.4203	0.0000×10^{0}	[0.0126, 0.0191]	3.07
Dry Days	0.0019	0.0002	9.1892	0.0000×10^{0}	[0.0015, 0.0023]	3.00
Wet Days	0.0033	0.0006	5.1527	1.2837×10^{-7}	[0.0020, 0.0046]	1.57
Extreme Wet Days	0.0042	0.0013	3.1191	9.0689×10^{-4}	[0.0015, 0.0068]	1.20

Among the trends that are analysed, the decline in Ice Days stands out as particularly strong and statistically robust. With a slope of -0.0050 and a z-value of -6.12, the result is well below the one-sided critical threshold of -1.6449, supporting a significant decrease over time ($p < 5 \times 10^{-10}$). This decline is consistent with a warming climate, where fewer days meet the freezing requirements.

Equally notable is the sharp rise in Tropical Days, with a steep slope of 0.0159 and a z-value of 9.42, far exceeding the critical value of 1.6449. This indicates a highly significant increase ($p \approx 0$), pointing to a growing prevalence of extremely hot days.

Dry Days also exhibit a very strong upward trend (z = 9.19), suggesting a marked shift toward more frequent dry conditions. This is supported by an extremely low p-value ($< 1 \times 10^{-20}$), reinforcing concerns over increasing dryness in the observed region.

However, it is important to note that most Poisson models exhibited overdispersion, where the variance of the response variable exceeds its mean. This violates the key Poisson assumption of equidispersion and may reflect underlying heterogeneity or temporal clustering not captured in the model. As a result, while all trends here are statistically significant, the standard errors may be underestimated, and caution is warranted when interpreting these results quantitatively.

5.4. Isotonic Count Regression

We assess monotonic trends in climate-related count data using the Poisson-based permutation test described in Section 4.4.1. The test compares a constant-rate model to a Poisson isotonic model, with the difference in log-likelihoods serving as the test statistic. Significance is determined by comparing the observed value to a null distribution generated from 10,000 random permutations under the assumption of no trend.

Permutation results for extreme temperature indicators are presented in Figure 5.3. While the distribution for Tropical Days mirrors the patterns seen in temperature, since the with the observed test statistic is clearly separated from the permutation values, the result for Ice Days is more modest. The observed value for Ice Days is near the tail of the distribution but not as far removed, suggesting a weaker though still detectable trend. This contrast highlights the relative strength of warming trends compared to reductions in cold extremes.

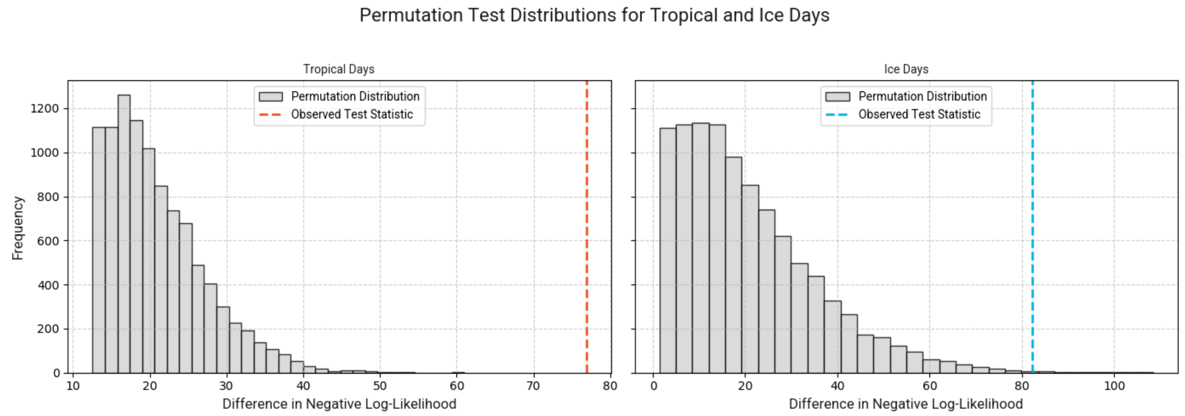

Figure 5.3: Permutation distributions of test statistics quantifying the improvement in fit between a constant-rate model and an isotonic count trend model for Tropical Days and Ice Days. The vertical lines indicate the observed Test statistic from the original data.

Figure 5.4 concerns extreme precipitation day counts. Here, Dry Days exhibit a particularly strong trend: the observed test statistic far exceeds the values seen in any of the permutations. Wet Days also show a substantial deviation from the null, though less pronounced. The trend for Very Wet Days is weaker, with the observed value located at the edge of the permutation distribution.

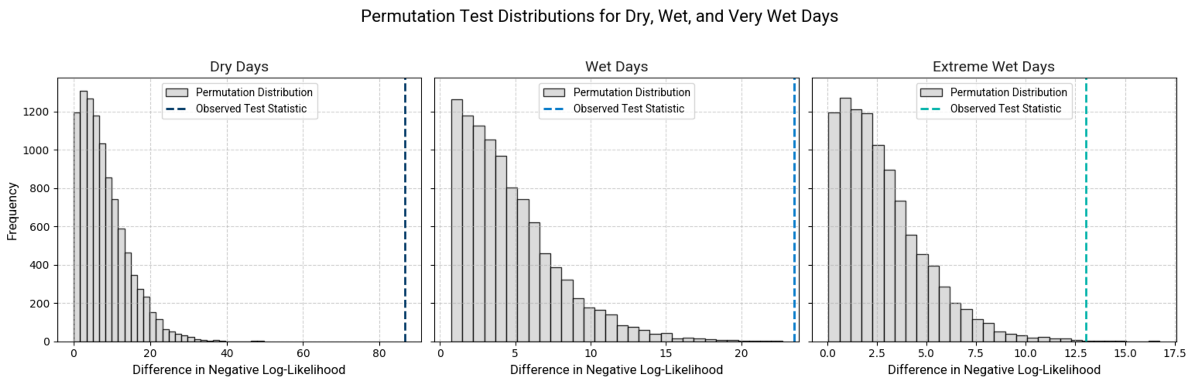

Figure 5.4: Permutation distributions of likelihood-based test statistics comparing constant and isotonic count models for Dry, Wet, and Extreme Wet Days. The vertical dashed lines shows the observed test statistics.

Table 5.4 summarizes the results of permutation tests evaluating the strength of monotonic trends in various count-based climate indicators using isotonic count modeling. For each category, the test statistic reflects the improvement in fit when replacing a constant model with an isotonic model that allows for a non-decreasing or non-increasing trend over time. The permutation p-values quantify the probability of observing such improvements under the null hypothesis of no trend.

Table 5.4: Permutation test results for isotonic regression on various count data categories.

Model	Test Statistic	Permutation p-value	Permuted Stats Min	Permuted Stats Max
Tropical Days	76.866	9.9990×10^{-5}	12.556	60.961
Ice Days	82.289	2.0998×10^{-3}	1.432	108.416
Dry Days	86.617	9.9990×10^{-5}	0.0460	49.707
Wet Days	23.475	9.9990×10^{-5}	0.7410	22.706
Extreme Wet Days	13.027	1.0999×10^{-3}	0.0669	16.737

The low p-values (all below 0.01) across categories suggest strong evidence for temporal trends in the data. For Tropical and Ice Days, the maximum of most permuted distributions exceed the test statistic from the original regression, indicating highly significant departures from constancy. Among the precipitation-related counts, Dry Days show the strongest trend, followed by Wet and Extreme Wet Days, which also demonstrate significant but more modest effects. Overall, the results support the presence of systematic changes over time in both temperature and precipitation related extremes.

5.5. Negative Binomial Regression

Table 5.5 summarizes the results of one-sided Negative Binomial regressions examining the effect of year on several climatic indices. All models account for overdispersion relative to the Poisson distribution, with corresponding dispersion values reported in the last column of Table 5.3. It is important to note that not all models exhibit strong overdispersion; in particular, the model for Extreme Wet Days shows an overdispersion value of 1.20, suggesting that a Poisson model might already provide a reasonable fit in that case.

Table 5.5: Results of Negative Binomial regression one-sided tests on the effect of year. Models for Ice and Tropical Days use 124 observations (df = 122), and precipitation models (Dry, Wet, Extreme Wet Days) use 119 observations (df = 117).

Model	Coef.	Std. Err.	z-stat	p-value (1-sided)	95% CI	Dispersion
Ice Days	-0.0063	0.0027	-2.2983	1.0774×10^{-2}	[-0.0117, -0.0009]	1.03
Tropical Days	0.0157	0.0028	5.5494	1.4334×10^{-8}	[0.0102, 0.0213]	1.15
Dry Days	0.0019	0.0004	5.3192	5.2117×10^{-8}	[0.0012, 0.0026]	1.02
Wet Days	0.0033	0.0008	4.1240	1.8615×10^{-5}	[0.0017, 0.0048]	1.01
Extreme Wet Days	0.0042	0.0014	2.8961	1.9000×10^{-3}	[0.0013, 0.0070]	1.03

The model for Tropical Days required a two-step estimation approach due to instability in maximum likelihood optimization. First, a Poisson model was fit to the data, and the dispersion parameter θ was estimated using the Pearson chi-squared statistic. This estimated θ was then used to fit a Negative Binomial model with a fixed dispersion parameter, like discussed in Section 3.5.1. For all other models, the optimization procedure successfully converged, allowing all parameters, including θ, to be estimated jointly.

The regression results reveal several notable trends. The number of Tropical Days shows a strong and highly significant positive trend, with a coefficient of 0.0157 and a one-sided p-value of 1.43×10^{-8}, indicating an accelerating increase in hot days over time. Conversely, Ice Days show a significant decreasing trend, with a negative coefficient of -0.0063 and a p-value of 0.0108. All precipitation based indices display significant positive associations with year. This suggests an intensification of both dry days and extreme precipitation events. The overdispersion statistics are all close to 1, indicating that the Negative Binomial models adequately capture extra-Poisson variability across all outcomes.

6
Conclusion and Discussion

This thesis investigated long-term trends in Dutch climate data using more than a century of observations from the De Bilt weather station. The analysis focused on changes in temperature, precipitation, and the frequency of extreme weather events, using a variety of statistical models suited to different types of data. Continuous variables such as average temperature and total precipitation were examined with both linear and isotonic regression, while count-based variables such as tropical days and extreme precipitation events were analysed using isotonic, Poisson, and Negative Binomial regression.

The central research question posed in this thesis was: *To what extent do historical weather observations from De Bilt reveal statistically significant trends in temperature, precipitation, and extreme weather events over the past century?* Based on the results presented in Chapters 4 and 5, this question can be answered clearly: the long-term data from De Bilt show strong and statistically significant evidence of warming, along with changing patterns in both average and extreme precipitation.

Table 6.1: Summary of trend directions and significance across climate indicators and regression models. Arrows indicate direction of trend; checkmarks denote significance at the 5% level.

Category	**Variable**	**Linear**	**Isotonic (2 types)**	**Poisson**	**NegBin**
Average Temperature	*Yearly*	↑ ✓	↑ ✓	–	–
	Summer	↑ ✓	↑ ✓	–	–
	Winter	↑ ✓	↑ ✓	–	–
Average Precipitation	*Yearly*	↑ ✓	↑ ✓	–	–
	Summer	↑ ×	↑ ×	–	–
	Winter	↑ ✓	↑ ✓	–	–
Extreme Temperature	*Tropical Days*	–	↑ ✓	↑ ✓	↑ ✓
	Ice Days	–	↓ ✓	↓ ✓	↓ ✓
Extreme Precipitation	*Dry Days*	–	↑ ✓	↑ ✓	↑ ✓
	Wet Days	–	↑ ✓	↑ ✓	↑ ✓
	Extreme Wet Days	–	↑ ✓	↑ ✓	↑ ✓

All temperature indicators, including yearly, summer, and winter averages, show statistically significant upward trends. These results are consistent across both linear and isotonic models, which strengthens confidence in the observed patterns. Similarly, the number of tropical days has increased and the number of ice days has decreased, with isotonic, Poisson, and Negative Binomial models confirming these changes as statistically significant.

Precipitation trends show a more varied picture. Yearly and winter precipitation averages have increased significantly, while summer precipitation shows no statistically significant trend. In contrast, the number of dry, wet, and extremely wet days has increased slightly but consistently, with dry days showing the clearest upward trend. These changes suggest a gradual shift toward more frequent precipitation extremes, in line with broader expectations under climate change.

An overview of the trend directions and their statistical significance is provided in Table 6.1. The table shows clearly which trends are significant and which are not, based on the chosen regression models. All observed trends are statistically significant at the five percent level, with the exception of summer precipitation. The consistency across different methods reinforces the reliability of the findings.

Although the analysis reveals clear climate trends at the De Bilt station, there are some limitations. The study is based on data from a single location and assumes independence between years. Future research could expand the spatial scope to include more weather stations across the Netherlands, apply time series models that account for autocorrelation, or make use of regional climate model data to explore attribution. It would also be valuable to investigate compound events, such as hot and dry summers, which pose particular risks under climate change.

Taken together, these results paint a clear picture: the Dutch climate, as observed at De Bilt, is undergoing measurable and statistically significant changes. Rising temperatures, shifts in the frequency of extreme events, and evolving precipitation patterns all point to a climate in transition. Still, more research is needed, especially covering larger areas, these findings highlight the importance of ongoing climate monitoring and help guide practical plans to adapt to changing conditions.

A

Modelling

A.1. Data Processing Code

Data preprocessing for all 4 indicators.

Temperature Data Preprocessing

```
df = pd.read_csv("data/temperature.txt", sep=r"\s+", skiprows=27, names=["STN", "DATE", "TG_hom"])
df["DATE"] = pd.to_datetime(df["DATE"], format="%Y%m%d")
df["TG_hom"] = pd.to_numeric(df["TG_hom"], errors="coerce")
df = df.drop(columns="STN")

df["Year"] = df["DATE"].dt.year
df["Month"] = df["DATE"].dt.month
```

Seasonal Temperature Averages

```
yearly_avg = df.groupby("Year")["TG_hom"].mean()
summer_avg = df[df["Month"].isin([6, 7, 8])].groupby("Year")["TG_hom"].mean()

df["WinterYear"] = df["Year"]
df.loc[df["Month"] == 12, "WinterYear"] += 1
winter_avg = df[df["Month"].isin([12, 1, 2])].groupby("WinterYear")["TG_hom"].mean()
```

Precipitation Data Preprocessing

```
df_rain = pd.read_csv("data/rainfall.txt", sep=r"\s+", skiprows=15)
df_rain = df_rain.dropna(subset=["YEAR"])

months = ['JAN', 'FEB', 'MAR', 'APR', 'MAY', 'JUN',
          'JUL', 'AUG', 'SEP', 'OCT', 'NOV', 'DEC']
monthly = df_rain.melt(id_vars=["YYYY"], value_vars=months,
                       var_name="Month", value_name="Rainfall")

month_number = {month: i for i, month in enumerate(months, 1)}
monthly["Month_Num"] = monthly["Month"].map(month_number)
monthly["Date"] = pd.to_datetime(dict(year=monthly["YYYY"],
                                      month=monthly["Month_Num"],
                                      day=1))
```

Tropical and Ice Days Classification

```
df = pd.read_csv("data/weather_data.txt", comment='#', header=None,
                 names=["STN", "DATE", "TG", "TN", "TX", "SQ", "RH"])
df = df.drop(columns="STN")
df["DATE"] = pd.to_datetime(df["DATE"], format="%Y%m%d")

for col in ["TG", "TN", "TX", "SQ", "RH"]:
    df[col] = pd.to_numeric(df[col], errors='coerce') / 10
df[['SQ', 'RH']] = df[['SQ', 'RH']].replace(-1, 0)
df = df[df["DATE"].dt.year != 2025]

df["TD"] = (df["TX"] >= 30).astype(int)
df["ID"] = (df["TX"] < 0).astype(int)
```

Precipitation Day Classification

```
df = pd.read_csv("data/weather_data.txt", comment='#', header=None,
                 names=["STN", "DATE", "TG", "TN", "TX", "SQ", "RH"])
df = df.drop(columns="STN")
df["DATE"] = pd.to_datetime(df["DATE"], format="%Y%m%d")

for col in ["TG", "TN", "TX", "SQ", "RH"]:
    df[col] = pd.to_numeric(df[col], errors='coerce') / 10
df[['SQ', 'RH']] = df[['SQ', 'RH']].replace(-1, 0)
df = df[df["DATE"].dt.year != 2025]

df["dry"] = (df["RH"] < 0.1).astype(int)
df["wet"] = ((df["RH"] >= 10) & (df["RH"] < 20.0)).astype(int)
df["very_wet"] = (df["RH"] >= 20.0).astype(int)

df["year"] = df["DATE"].dt.year
df = df[df["year"] >= 1906]

rain_counts = df.groupby("year")[["dry", "wet", "very_wet"]].sum()
```

A.2. Regression Models

Examples of fitting different regression models.

Linear Regression

```
ols_yearly = sm.OLS(Y_yearly, X_const).fit()
ols_summer = sm.OLS(Y_summer, X_const).fit()
ols_winter = sm.OLS(Y_winter, X_const).fit()
```

Classic Isotonic Regression

```
    iso = IsotonicRegression(increasing=False)
yearly_ir = iso.fit_transform(x_flat, yearly)
summer_ir = iso.fit_transform(x_flat, summer)
winter_ir = iso.fit_transform(x_flat, winter)
```

Isotonic Count Regression

```
def poisson_isotonic_regression(x, y, increasing=True, log_scale=True):
"""
Fit a Poisson isotonic regression model using convex optimization.

Args:
    x (array-like): 1D array of predictors (used only for sorting).
    y (array-like): 1D array of count responses.
    increasing (bool): Whether to enforce increasing trend. If False, decreasing.
    log_scale (bool): Whether to model   = exp(  ) to ensure positivity.

Returns:
    lambda_hat: Fitted Poisson rates (same shape as y).
"""
# Ensure inputs are numpy arrays
x = np.asarray(x)
y = np.asarray(y)

# Sort x and reorder y accordingly
sorted_indices = np.argsort(x)
x_sorted = x[sorted_indices]
y_sorted = y[sorted_indices]
n = len(y_sorted)

# Define variables
if log_scale:
    theta = cp.Variable(n)  # log(  )
    lam = cp.exp(theta)
else:
    lam = cp.Variable(n)
    theta = cp.log(lam)

# Define monotonicity constraints
diffs = theta[1:] - theta[:-1] if log_scale else lam[1:] - lam[:-1]
if increasing:
    constraints = [diffs >= 0]
else:
    constraints = [diffs <= 0]

# Poisson negative log-likelihood: sum( _i  - y_i * log( _i ))
objective = cp.sum(lam - cp.multiply(y_sorted, theta))
prob = cp.Problem(cp.Minimize(objective), constraints)
prob.solve()

if prob.status not in ["optimal", "optimal_inaccurate"]:
    raise RuntimeError("Optimization failed.")

# Return fitted values in original x order
fitted = lam.value
unsorted_fitted = np.empty_like(fitted)
unsorted_fitted[sorted_indices] = fitted

return unsorted_fitted
```

Poisson Regression

```
poisson_td = smf.glm(formula='TD_sum ~ YEAR', data=extreme_year,
```

```
    family=sm.families.Poisson()).fit()

def check_dispersion(model):
    pearson_chi2 = sum(model.resid_pearson**2)
    df_resid = model.df_resid
    dispersion = pearson_chi2 / df_resid
    print(f"Dispersion statistic: {dispersion:.2f}")
    if dispersion > 1.5:
        print("    Evidence of overdispersion.")
    elif dispersion < 0.8:
        print("    Evidence of underdispersion.")
    else:
        print("    Dispersion is approximately acceptable.")
```

Negative Binomial Regression

```
    nb_td = smf.glm(
    formula='TD_sum ~ YEAR',
    data=extreme_year,
    family=NegativeBinomial()
).fit()
```

Two step alpha

```
# Tropical Days
poisson_td = smf.glm('TD_sum ~ YEAR', data=extreme_year, family=sm.families.Poisson()).fit()
mu_td = poisson_td.fittedvalues

n = len(extreme_year)
pearson_td = np.sum((extreme_year['TD_sum'] - mu_td)**2 / mu_td)
alpha_td = (pearson_td - n) / np.sum(mu_td)

nb_td = smf.glm('TD_sum ~ YEAR', data=extreme_year,
family=NegativeBinomial(alpha=alpha_td)).fit()
extreme_year['TD_nb_pred'] = nb_td.predict(extreme_year)
```

A.3. Trend Testing

Functions used to preform one sided and permutation tests.

Linear Regression

```
    def test_trend_one_sided(ols_result, alpha=0.05, verbose=True):
    """
    Performs a one-sided test of the slope coefficient from an OLS regression.

    H0: beta = 0
    H1: beta > 0

    Returns test statistic, critical value, one-sided p-value,
    confidence interval, and standard error.
    """
    coef = ols_result.params[1]
    stderr = ols_result.bse[1]
    t_stat = ols_result.tvalues[1]
    df = int(ols_result.df_resid)
```

```
# One-sided p-value
if coef > 0:
    p_one_sided = 1 - stats.t.cdf(t_stat, df)
else:
    p_one_sided = 1.0

# Critical t-value
t_crit = stats.t.ppf(1 - alpha, df)

# Confidence interval (two-sided)
ci_low, ci_high = ols_result.conf_int()[1]

# Decision
reject = t_stat > t_crit

if verbose:
    print("=== One-sided Test for Trend ===")
    print(f"Toetsingsgrootheid (t-statistic): {t_stat:.4f}")
    print(f"Standard Error: {stderr:.4f}")
    print(f"Critical value (alpha = {alpha}): {t_crit:.4f}")
    print(f"One-sided p-value: {p_one_sided:.4e}")
    print(f"95% Confidence Interval (slope): [{ci_low:.4f}, {ci_high:.4f}]")
    print(f"Reject H0? {'Yes' if reject else 'No'}\n")

return {
    "coefficient": coef,
    "standard_error": stderr,
    "t_statistic": t_stat,
    "t_critical": t_crit,
    "p_value_one_sided": p_one_sided,
    "confidence_interval": (ci_low, ci_high),
    "reject_null": reject
}
```

Permutation testing

Function used to test the classic isotonic regression models

```
def isotonic_permutation_test(y, x=None, B=10000, random_state=None):
rng = np.random.default_rng(random_state)
y = np.asarray(y)

if x is None:
    x = np.arange(len(y))
else:
    x = np.asarray(x)

# Remove NaNs
mask = ~np.isnan(y)
y = y[mask]
x = x[mask]

# Fit isotonic regression on real data
ir = IsotonicRegression(increasing=True)
y_iso = ir.fit_transform(x, y)
```

```
rss_null = np.sum((y - np.mean(y))**2)
rss_iso = np.sum((y - y_iso)**2)
T_obs = rss_null - rss_iso

T_perm = []
for _ in range(B):
    y_perm = rng.permutation(y)
    y_iso_perm = ir.fit_transform(x, y_perm)
    rss_null_perm = np.sum((y_perm - np.mean(y_perm))**2)
    rss_iso_perm = np.sum((y_perm - y_iso_perm)**2)
    T_perm.append(rss_null_perm - rss_iso_perm)

T_perm = np.array(T_perm)
p_value_raw = (np.sum(T_perm >= T_obs) + 1) / (B + 1)
p_value_str = f"{p_value_raw:.4e}"  # format as scientific notation with 4 decimals

return T_obs, p_value_str, T_perm
```

This is for testing the continuous isotonic regression

```
def isotonic_permutation_test(y, x=None, B=10000, random_state=None):
rng = np.random.default_rng(random_state)
y = np.asarray(y)

if x is None:
    x = np.arange(len(y))
else:
    x = np.asarray(x)

# Remove NaNs
mask = ~np.isnan(y)
y = y[mask]
x = x[mask]

# Fit isotonic regression on real data
ir = IsotonicRegression(increasing=True)
y_iso = ir.fit_transform(x, y)

rss_null = np.sum((y - np.mean(y))**2)
rss_iso = np.sum((y - y_iso)**2)
T_obs = rss_null - rss_iso

T_perm = []
for _ in range(B):
    y_perm = rng.permutation(y)
    y_iso_perm = ir.fit_transform(x, y_perm)
    rss_null_perm = np.sum((y_perm - np.mean(y_perm))**2)
    rss_iso_perm = np.sum((y_perm - y_iso_perm)**2)
    T_perm.append(rss_null_perm - rss_iso_perm)

T_perm = np.array(T_perm)
p_value_raw = (np.sum(T_perm >= T_obs) + 1) / (B + 1)
p_value_str = f"{p_value_raw:.4e}"

return T_obs, p_value_str, T_perm
```

This is for testing the count data

```
def poisson_isotonic_regression(x, y, increasing=True, log_scale=True):
    """
    Fit a Poisson isotonic regression model using convex optimization.

    Args:
        x (array-like): 1D array of predictors (used only for sorting).
        y (array-like): 1D array of count responses.
        increasing (bool): Whether to enforce increasing trend. If False, decreasing.
        log_scale (bool): Whether to model   = exp(  ) to ensure positivity.

    Returns:
        lambda_hat: Fitted Poisson rates (same shape as y).
    """
    # Ensure inputs are numpy arrays
    x = np.asarray(x)
    y = np.asarray(y)

    # Sort x and reorder y accordingly
    sorted_indices = np.argsort(x)
    x_sorted = x[sorted_indices]
    y_sorted = y[sorted_indices]
    n = len(y_sorted)

    # Define variables
    if log_scale:
        theta = cp.Variable(n)  # log(  )
        lam = cp.exp(theta)
    else:
        lam = cp.Variable(n)
        theta = cp.log(lam)

    # Define monotonicity constraints
    diffs = theta[1:] - theta[:-1] if log_scale else lam[1:] - lam[:-1]
    if increasing:
        constraints = [diffs >= 0]
    else:
        constraints = [diffs <= 0]

    # Poisson negative log-likelihood: sum( _i  - y_i * log( _i ))
    objective = cp.sum(lam - cp.multiply(y_sorted, theta))
    prob = cp.Problem(cp.Minimize(objective), constraints)
    prob.solve(solver=cp.ECOS, verbose=False, max_iters=10000)

    # Return fitted values in original x order
    fitted = lam.value
    unsorted_fitted = np.empty_like(fitted)
    unsorted_fitted[sorted_indices] = fitted

    return unsorted_fitted

# Example usage with your extreme_year DataFrame:
# Assuming extreme_year has columns: 'YEAR', 'TD_sum', 'ID_sum'

x_td = extreme_year['YEAR'].values
y_td = extreme_year['TD_sum'].values
td_poisson_iso_pred = poisson_isotonic_regression(x_td, y_td, increasing=True, log_scale=True)
```

```
extreme_year['TD_poisson_iso_pred'] = td_poisson_iso_pred

x_id = extreme_year['YEAR'].values
y_id = extreme_year['ID_sum'].values
id_poisson_iso_pred = poisson_isotonic_regression(x_id, y_id, increasing=False, log_scale=True)
extreme_year['ID_poisson_iso_pred'] = id_poisson_iso_pred


def poisson_neg_log_likelihood(y, lam):
    """
    Compute Poisson negative log-likelihood (up to constants):
    sum(lam_i - y_i * log(lam_i))
    """
    return np.sum(lam - y * np.log(lam))

def poisson_isotonic_permutation_test(y, x=None, increasing=True, B=10000, random_state=None):
    rng = np.random.default_rng(random_state)
    y = np.asarray(y)

    if x is None:
        x = np.arange(len(y))
    else:
        x = np.asarray(x)

    mask = ~np.isnan(y)
    y = y[mask]
    x = x[mask]

    lam_iso = poisson_isotonic_regression(x, y, increasing=increasing)  # pass increasing here

    ll_null = poisson_neg_log_likelihood(y, np.full_like(y, y.mean()))
    ll_iso = poisson_neg_log_likelihood(y, lam_iso)
    T_obs = ll_null - ll_iso

    T_perm = []
    for _ in range(B):
        y_perm = rng.permutation(y)
        lam_perm = poisson_isotonic_regression(x, y_perm,
        increasing=increasing)  # pass here too
        ll_null_perm = poisson_neg_log_likelihood(y_perm, np.full_like(y_perm, y_perm.mean()))
        ll_iso_perm = poisson_neg_log_likelihood(y_perm, lam_perm)
        T_perm.append(ll_null_perm - ll_iso_perm)

    T_perm = np.array(T_perm)
    p_value_raw = (np.sum(T_perm >= T_obs) + 1) / (B + 1)
    p_value_str = f"{p_value_raw:.4e}"

    return T_obs, p_value_str, T_perm

test_results = {}
labels = {
    "Tropical Days": ('TD_sum', True),   # increasing isotonic regression
    "Ice Days": ('ID_sum', False)         # decreasing isotonic regression
}

for label, (column, increasing) in labels.items():
```

```
    T_obs, p_val_str, T_perm = poisson_isotonic_permutation_test(
        extreme_year[column].values,
        extreme_year['YEAR'].values,
        increasing=increasing,
        B=10000,
        random_state=42
    )

    test_results[label] = {
        "T_obs": T_obs,
        "p_value_str": p_val_str,
        "T_perm": T_perm
    }

    print(f"{label}:\n"
          f"  Observed Test Statistic = {T_obs:.4f}\n"
          f"  p-value = {p_val_str}\n"
          f"  Permuted stats min: {T_perm.min():.4f}, max: {T_perm.max():.4f}\n")
```

Poisson Regression

```
    def format_p_value(p):
    dec_p = Decimal(str(p))
    if dec_p == 0:
        return "{:.4e}".format(p if p > 0 else 1e-20)
    else:
        return "{:.4e}".format(p)

def poisson_one_sided_test(model, param_name='year', alternative='less', model_name='Model'):
    coef = model.params[param_name]
    se = model.bse[param_name]
    z_stat = coef / se
    z_95 = norm.ppf(0.975)
    ci_lower = coef - z_95 * se
    ci_upper = coef + z_95 * se

    if alternative == 'less':
        p_val = norm.cdf(z_stat)
        alt_text = "less than 0"
    elif alternative == 'greater':
        p_val = 1 - norm.cdf(z_stat)
        alt_text = "greater than 0"
    else:
        raise ValueError("alternative must be 'less' or 'greater'")

    p_val_str = format_p_value(p_val)

    print(f"\n{model_name} - Poisson Regression One-Sided Test on '{param_name}'")
    print("-------------------------------------------------------")
    print(f"Coefficient (slope): {coef:.4f}")
    print(f"Standard Error: {se:.4f}")
    print(f"95% Confidence Interval: ({ci_lower:.4f}, {ci_upper:.4f})")
    print(f"Z Statistic: {z_stat:.4f}")
    print(f"One-sided p-value (H1: coef {alt_text}): {p_val_str}")

    check_dispersion(model, model_name)
```

Negative Binomial Regression

```
def negbinom_one_sided_test(model, param_name='YEAR', alternative='less',
model_name='Negative Binomial Model'):

coef = model.params[param_name]
se = model.bse[param_name]
z_stat = coef / se
z_95 = norm.ppf(0.975)
ci_lower = coef - z_95 * se
ci_upper = coef + z_95 * se

if alternative == 'less':
    p_val = norm.cdf(z_stat)
    alt_text = "less than 0"
elif alternative == 'greater':
    p_val = 1 - norm.cdf(z_stat)
    alt_text = "greater than 0"
else:
    raise ValueError("alternative must be 'less' or 'greater'")

p_val_str = format_p_value(p_val)

print(f"\n{model_name} - Negative Binomial Regression One-Sided Test on '{param_name}'")
print("------------------------------------------------------------")
print(f"Coefficient (slope): {coef:.4f}")
print(f"Standard Error: {se:.4f}")
print(f"95% Confidence Interval: ({ci_lower:.4f}, {ci_upper:.4f})")
print(f"Z Statistic: {z_stat:.4f}")
print(f"One-sided p-value (H1: coef {alt_text}): {p_val_str}")
```

Bibliography

[1] Alan Agresti. *Foundations of Linear and Generalized Linear Models*. John Wiley & Sons, 2015. ISBN 9781118730034.

[2] Alexander Bakker. Time series transformation tool version 3.1: Description of the program to generate time series consistent with the knmi'14 climate scenarios. Technical Report TR-349, Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands, 2015. URL `https://cdn.knmi.nl/knmi/pdf/bibliotheek/knmipubTR/TR349.pdf`. Accessed: 2025-05-26.

[3] F. Bijma, M. Jonker, and A. Van der Vaart. *An Introduction to Mathematical Statistics*. Amsterdam University Press, Amsterdam, 2017.

[4] A. Colin Cameron and Pravin K. Trivedi. *Regression Analysis of Count Data*. Cambridge University Press, Cambridge, UK, 1998. ISBN 9780521655418.

[5] George Casella and Roger L. Berger. *Statistical Inference*. Duxbury, 2nd edition, 2002.

[6] Python Software Foundation. Python programming language, version 3.11, 2022. URL `https://www.python.org`.

[7] Joseph M. Hilbe. *Negative Binomial Regression*. Cambridge University Press, 2nd edition, 2011.

[8] Vincent Humphrey, Lukas Gudmundsson, and Sonia I. Seneviratne. Assessing global water storage variability from grace: Trends, seasonal cycle, and extremes. *Surveys in Geophysics*, 42(4):1173–1202, 2021. doi: 10.1007/s10712-021-09638-4.

[9] Koninklijk Nederlands Meteorologisch Instituut. Maandgegevens van het weer in nederland, 2024. URL `https://www.knmi.nl/nederland-nu/klimatologie/maandgegevens`. Accessed: 2025-04-29.

[10] Ronny Luss and Saharon Rosset. Generalized isotonic regression. *arXiv preprint arXiv:1104.1779*, 2011.

[11] NOAA National Centers for Environmental Information. Meteorological versus astronomical seasons, 2015. URL `https://www.ncei.noaa.gov/news/meteorological-versus-astronomical-seasons`. Accessed: 2025-04-29.

[12] Intergovernmental Panel on Climate Change (IPCC). *Climate Change 2021: The Physical Science Basis*. Cambridge University Press, 2021. doi: 10.1017/9781009157896. URL `https://www.ipcc.ch/report/ar6/wg1/`.

[13] World Meteorological Organization. Centennial observing stations: De bilt (netherlands), 2017. URL `https://wmo.int/activities/centennial-observing-stations/centennial-observing-stations`. Accessed: 2025-04-29.

[14] Xuebin Zhang, Lisa V Alexander, Gabriele C Hegerl, Phil Jones, Albert M G Klein Tank, Thomas C Peterson, et al. Indices for monitoring changes in extremes based on daily temperature and precipitation data. *WIREs Climate Change*, 2(6):851–870, 2011. doi: 10.1002/wcc.147. URL `https://doi.org/10.1002/wcc.147`.