"How many 1 in 100 year extreme weather events can NZ expect to experience over the course of the next decade?"

Abstract:
In this report, we analysed various datasets of recorded extreme weather cases in New Zealand to predict the number of 1 in 100 year extreme weather events in New Zealand over the next decade. Using statistical analysis, we made a prediction that 17 of these events will occur in the next decade.

Introduction:
Due to the implications of climate change, extreme '1 in 100 year weather events' are becoming increasingly common and severe across the world[1]. Climate change is caused by the greenhouse effect, a natural process that warms the Earth's surface. Approximately 30 percent of solar radiation that reaches the earth is reflected back into space, while 70 percent is trapped in the atmosphere by greenhouse gases[2] such as methane and carbon dioxide. As global greenhouse gas emissions increase, more heat is trapped in the atmosphere, contributing to climate change. On average, New Zealand has warmed by approximately 1 degree over the past century[3]. This trend is projected to continue, meaning extreme weather events are expected to become increasingly severe and frequent. Using historical New Zealand weather data and methods of statistical analysis, we will make a prediction for the number of '1 in 100 extreme weather events' that occur in the next decade.

Assumptions:

- We will exclude natural disasters in "extreme weather events", meaning events such as volcanic eruptions, tsunamis and earthquakes are excluded.
- We are assuming that the increase of extreme weather events due to climate change is already accounted for within the data provided.
- We are assuming that the term extreme weather events only contains the following:
    - Flooding
    - Hail
    - Rainfall
    - Wind Storms
    - Lightning Storms
    - Snow/Ice Storms
    - Tornados (and similar variations)
    - Heat/cold waves

[1] "The Sun's impact on the Earth | World Meteorological Organization." https://public.wmo.int/en/sun%E2%80%99s-impact-earth. Accessed 7 Aug. 2021.
[2] "Greenhouse Effect 101 | NRDC." 16 Jul. 2019, https://www.nrdc.org/stories/greenhouse-effect-101. Accessed 7 Aug. 2021.
[3] "Climate change projections: How do scientists predict our ... - NIWA." https://niwa.co.nz/education-and-training/schools/students/climate-change/climate-change-projections. Accessed 7 Aug. 2021.

Data sources:

We have sourced our data from 3 locations:

- MfE (Ministry for the Environment) from 1975 to 2014 (https://data.mfe.govt.nz/table/52586-number-of-extreme-weather-events-identified-by-icnz-19752014/)
- ICNZ (The Insurance Council) from 1966 to 2021 (https://www.icnz.org.nz/natural-disasters/cost-of-natural-disasters)
- NIWA (National Institute of Water and Atmospheric Research) from 1900 to 2020 (https://hwe.niwa.co.nz/)

Definitions:

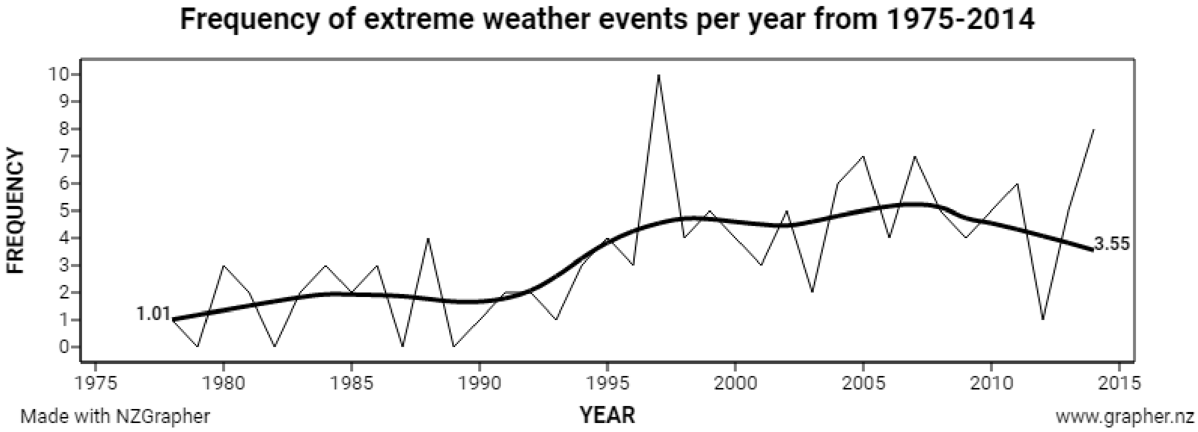
We are defining a “1 in 100 year extreme weather event” as any weather event that has an estimated return period of at least 100 years, that is it occurs, on average, once every hundred years. We define the term “extreme” to mean unlikely, or rare, rather than the alternate definition of damaging or powerful, as that is an incredibly subjective value to determine. We have defined “the next decade” as the time between the start of 2022 to the end of 2031.

Statistical models:

Prediction 1 (using the MfE dataset):

Using a linear model, we made a prediction of the number of extreme weather events in the next decade.

Plotting the frequency of extreme weather events against the year in the data set:



From 1975 to 2014, the long term trend of the model increased from 1.01 to 3.55, meaning there was a consistent increase of 0.065 extreme weather events per year.

Taking the average increase of 0.065 extreme weather events per year as the gradient, we can fit a basic linear model:

$y = 0.065(x - 2014) + 3.55.$

Summing the values of every year in the next decade we get:

$\sum_{x=2022}^{2031} [0.121(x - 2014) + 5.34] = 43.625 \approx 44$extreme weather events.

To calculate a prediction interval, we used the built in forecast feature for NZGrapher:

Time	Min	Prediction	Max
End of 2014	1.8797	5.8093	9.7745
End of 2016	0.66502	5.9309	11.477

$$\Delta prediction\ interval\ =\ \frac{(2016\ Max\ forecast - 2016\ Min\ forecast) - (2015\ Max\ forecast - 2015\ Min\ forecast)}{\Delta years}$$

$$\Delta prediction\ interval\ =\ \frac{(11.477-0.66502) - (9.7745 - 1.8797)}{2}$$

$$\Delta prediction\ interval\ =\ 1.45859$$

Using the change in prediction interval from the end of 2014 to the end of 2016 as the gradient of a linear model, we can extrapolate to get a linear model for a forecast:

$y\ =\ 1.45859(x\ -\ 2014)\ +\ 3.55.$

NZGrapher showed that the prediction interval in 2022 (the beginning of the next decade) was 15.219, and 28.34 in 2031 (the end of the next decade). Averaging these prediction intervals, we get a value of 21.78. We have averaged these values since the number of extreme weather events each year will have their own individual prediction ranges and we are only interested in the overall prediction range.

$$value\ of\ uncertainty\ =\ \pm\frac{Average\ prediction\ interval}{2}$$

$$value\ of\ uncertainty\ =\ \pm\frac{21.78}{2}$$

$$value\ of\ uncertainty\ =\ \pm\ 10.89$$

Accounting for uncertainties, NZ will experience approximately $43.625\ \pm\ 10.89$ extreme weather events over the course of the next decade.

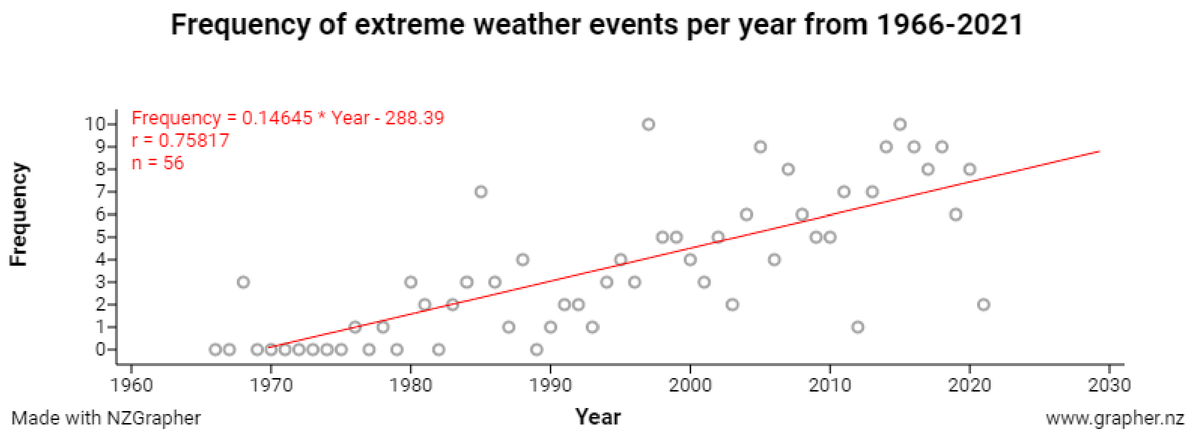
$43.625\ \pm\ 10.89\ =$
$43.625\ +\ 10.89\ =\ 54.5$= 54.5 (1dp)
Or
$43.625\ -\ 10.89\ =\ 32.7$ (1dp)

This gives a maximum value of 54.5≈55 and a minimum value of 32.7≈33 extreme weather events over the course of the next decade.

<u>Prediction 2 (using the ICNZ dataset)</u>:

Using a second linear model, we have made a second prediction of the number of extreme weather events in the next decade. Although both the MfE and ICNZ website data sets state to be sourced from ICNZ, the ICNZ website has additional data from 2014 to 2021 and includes "severe weather events and natural disasters," some of which may not necessarily be considered extreme weather events in the MfE data set used for prediction 1.

Plotting the data using NZGrapher and fitting a linear model:



Equation of the regression line fitted by NZGrapher:

$$y = 0.14645x - 288.39.$$

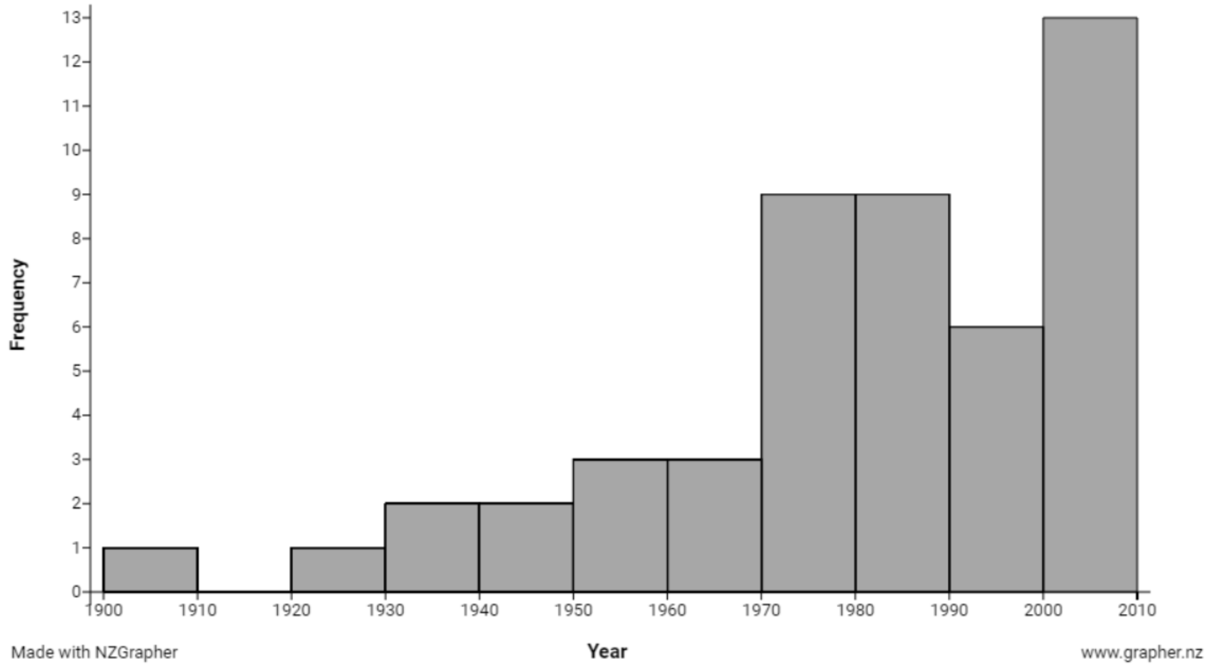
Summing the values of every year in the next decade we get:

$$\sum_{x=2022}^{2031} (0.14645x - 288.39) = 83.9 \approx 84 \text{natural disasters.}$$

The correlation coefficient of 0.75 indicates a strong relationship. This means we can be relatively confident in the above prediction.

The spread of the data in 2020 is roughly +1 to -5, we can assume this spread will continue into the future, giving us an answer range of 78.9 to 84.9 extreme weather events occurring over the course of the next decade.

Prediction 3 (using the NIWA dataset):



The NIWA dataset included many entries that did not fit our definition of "1 in 100 extreme weather events", so we wrote a Python script to download, process, and filter the data to ensure that our results were accurate to the question presented.

Using the information gathered from the NIWA data set, we plotted the information into a histogram separated by decade, we then fitted a power function to this data to approximate the amount of extreme weather events that will occur in any given year, this gives the equation:

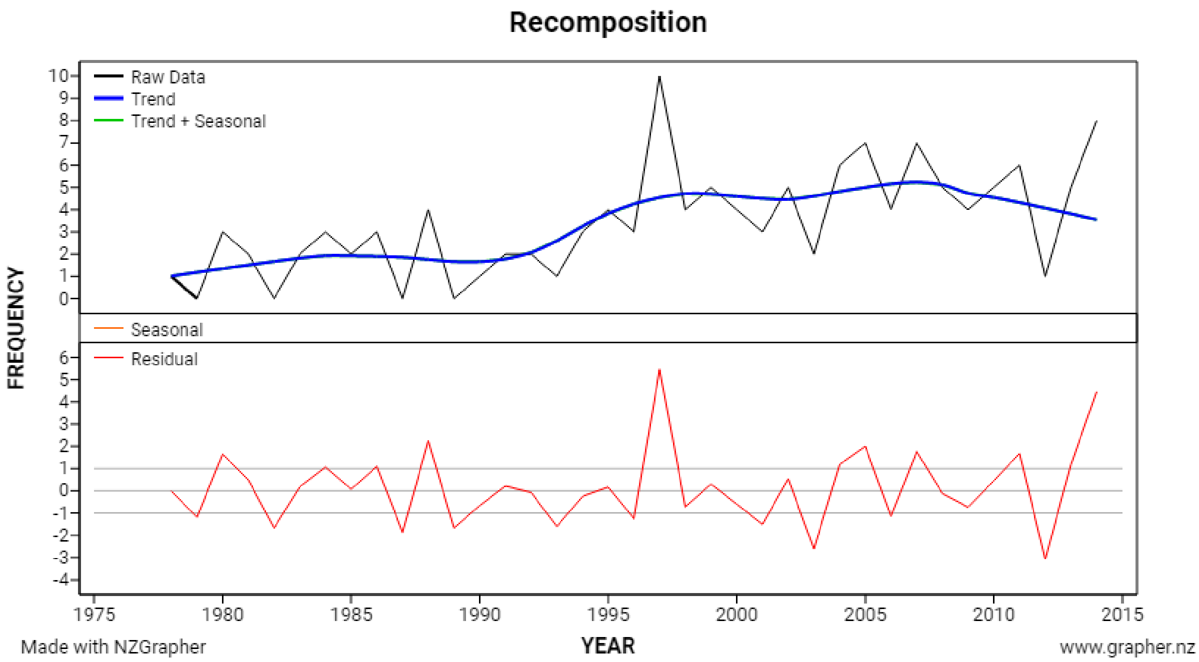
$$f(x) = (0.0000802387(x - 1900)^{2.05191} + 0.040292)$$

We then applied the summation (sigma) function to our equation from 2022 to 2031 to approximate the amount of extreme weather events that will occur over this decade:

$\sum_{x=2022}^{2031} f(x) \approx 16.9197243072 \approx 17$extreme weather events that will occur during this time period.

Confidence in predictions and choosing the final answer:

We are not very confident in our first prediction since the data is from 1975-2014, meaning the predictions made for 2022-2031 will likely be quite inaccurate as we are predicting many years into the future with a small dataset.



Using an NZGrapher residual plot, we can see there are 13 residuals that pass the grey lines. This represents a 10% variation or greater from the raw data. These can be classified as outliers.

For our second prediction, we used a scatter plot to analyse the data. We found that a linear model with an equation $y = 0.14645x - 288.39$ fit the data well with a correlation coefficient of 0.75. This is greater than 0.7 indicating that the relationship is strong, meaning predictions made with this model will be fairly accurate. As the data starts to spread out however, predictions will become increasingly inaccurate. The main reason we are not that confident in this model overall is because the data includes natural disasters such as volcanic activity and earthquakes that may not be 1 in 100 year weather events.

Of all our models, we are most confident in our third model, which uses the NIWA data. This is because we were able to narrow down the data points to what fitted our definition of "extreme 1 in 100 year weather events" most accurately. This allowed us to predict with greater confidence the number of extreme weather events in the next decade. Hence, we believe our result of 17 extreme weather events occurring in the next decade is the most reasonable out of all our predictions.

Limitations:

Many of the limitations of our models are due to the limitations of the data we used. For our data sources, the data on the topic was either incomplete, such as missing event magnitude classifications and expected return periods in the NIWA data. We put a lot of effort into cleaning and filtering data in order to improve the accuracy of our findings, but we accept that there will be inaccuracies in our predictions due to remaining discrepancies. Also, due to the infrequency of an extreme weather event, there are limited amounts of data points available to analyse, thus making it significantly more difficult to fit an accurate model to the data with confidence.

The MfE and ICNZ data is extremely similar, in fact, they are both sourced from ICNZ with minor differences. The MfE data excludes natural disasters and stops at 2014, while the ICNZ data has a wider timeframe, but includes natural disasters as well as extreme weather events. We have decided to use the ICNZ data, as natural disasters are extremely rare and therefore negligible to our final result, and the more recent data will allow for far more accurate results.

The NIWA data better fits our definition of a “1 in 100 year extreme weather event” leading to predictions made using this dataset more useful to answering the question. The NIWA data does contain better distinction between events, as well as a far larger dataset, however the site is incredibly unreliable, and frequently disconnected as we were downloading and processing the data. The way NIWA data specifies the return period is inconsistent and many entries are missing return periods entirely, degrading the accuracy of our filtering of the data. We selected all NIWA events from 1900 to 2021 that either had an upper bound of 100 years or above for their return period, or had a return period category of “Extreme” if no return period was provided.

Weather itself is also unpredictable in nature. As we are predicting many years into the future, it is inevitable that the accuracy of our projections will diminish as we extrapolate further into the future.

Conclusion:

We estimated the number of “1 in 100 year extreme weather events” NZ can expect to experience over the next decade by analysing New Zealand weather data and fitting 3 different models. Extreme weather events are very hard to predict and the accuracy of predictions diminish over time. We are most confident in the predictions made with our third model. We predict that there will be 17 “1 in 100 year extreme weather events” over the next decade (2022-2031). These results are New Zealand specific as all the data sourced is from New Zealand.

Appendix A - Source Code:

NIWA XML dataset downloader

```
import requests
from time import sleep

urlformat = "https://hwe.niwa.co.nz/search/summary/Startdate/01-01-1900/Enddate/31-12-2020/Regions/all/Hazards/all/Impacts/all/Keywords/none/numberOfEvents/20/page/{}/xml"

for x in range(1, 42):
  r = requests.get(urlformat.format(x))
  r.raise_for_status()
  with open("brogle/{:02d}.xml".format(x), "wb") as f:
    f.write(r.content)
  sleep(2)
```

NIWA XML cleanup, filtering and output to CSV

```
import xml.etree.ElementTree as ET
from sys import argv
import csv

def spanparse(s):
  if s is None or not any(c.isdigit() for c in s):
    return 0
  try:
    return int(s)
  except ValueError:
    pass
  #print(s)
  if s.endswith(" years"):
    s = s[:-6]
  elif s.endswith(" year"):
    s = s[:-5]
  max_num = s.split("-", 1)[-1]
  if max_num.endswith("+"):
    return int(max_num[:-1])
  elif max_num.startswith("Up to "):
    return int(max_num[6:])
  else:
    return int(max_num)

def findcategory(e):
  hs = set(e.findall(".//{http://hwe.niwa.co.nz/schema/2011}Hazard"))
  if len(hs) > 1:
    return ":".join(j.attrib["type"] for j in hs if j.attrib["type"] is not None)
  elif len(hs) == 1:
    return hs[0].attrib["type"]
  else:
    return "None"

def dat(l):
  return int(l.split("-", 2)[0])

decades = { 10 * n: 0 for n in range(190, 203) }
with open('eggs.csv', 'w', newline='') as csvfile:
  c = csv.writer(csvfile)
  c.writerow(["Year", "period", "rankthing"])
  for n in range(1, 39):
    tree = ET.parse("brogle/{:02d}.xml".format(n))
    events = tree.getroot()[0][1:]
    for e in events:
      ret = e.find("./{http://hwe.niwa.co.nz/schema/2011}ReturnPeriod")
      rgood = ret is not None and ret.text is not None
      cat = e.find("./{http://hwe.niwa.co.nz/schema/2011}ReturnPeriodCategory")
      cgood = cat is not None and cat.text is not None

      if len(argv) > 1 and argv[1] == "exclude":
        if (rgood and spanparse(ret.text) >= 100) or (cgood and (cat.text == "Extreme" or cat.text == "Severe")):
          yr = dat(e.find("./{http://hwe.niwa.co.nz/schema/2011}StartDate").text)

          pd = spanparse(ret.text)
          decade = yr // 10
          decades[10 * decade] += 1
          c.writerow([yr, pd])
      else:
        yr = dat(e.find("./{http://hwe.niwa.co.nz/schema/2011}StartDate").text)

        if rgood:
          pd = spanparse(ret.text)
        else:
          pd = 0
        decade = yr // 10
        decades[10 * decade] += 1
        if cat is not None:
          cg = cat.text
        else:
          cg = None
        c.writerow([yr, pd, cg])


print(decades)
```

ICNZ website CSV processor

```
import csv

decades = { n: 0 for n in range(1966, 2024) }
with open("../Cost of natural disasters - ICNZ.csv", "r") as csvf, open("br.csv", "w") as fb:
  reader = csv.reader(csvf)
  writer = csv.writer(fb)
  for l in reader:
    if l[0] != "Date":
      decades[int(l[0][:4])] += 1
    writer.writerow([l[0][:4], l[1], l[2]])

print("Year,Frequency")
for k,v in decades.items():
  print(str(k) + "," + str(v))
```