Can Perils Reasonably Be Assumed To Occur Independently In Time?

Disclaimer

This report has been produced by Ed Wheatcroft for Lloyd's for general information purposes only. While care has been taken in gathering the data and preparing the report, Lloyd's does not make any representations or warranties as to its accuracy or completeness and expressly excludes to the maximum extent permitted by law all those that might otherwise be implied.

Lloyd's accepts no responsibility or liability for any loss or damage of any nature occasioned to any person as a result of acting or refraining from acting as a result of, or in reliance on, any statement, fact, figure or expression of opinion or belief contained in this report. This report does not constitute advice of any kind.

Contact details

Ed Wheatcroft
Independent consultant
edward.wheatcroft1@gmail.com

Contents

1 Introduction

In the United Kingdom, insurance companies are legally required to hold reserves such that they are able to survive a 1 in 200 year event. It is the insurer's responsibility to demonstrate to the regulator that this is the case. For the Corporation of Lloyd's, calculations are usually made with the assumption that certain weather events occur independently of each other. The aim of this short project was to examine this assumption. In this report, this question is considered from a number of different angles, specifically whether the assumption of independence is reasonable between perils of the same type, whether different types of perils are associated and whether there is evidence to reject the hypothesis of independence when considering all perils simultaneously. The assumption of whether this concept holds for extreme events and for correlation between annual counts of each type of peril, was also considered.

2 Data

16 global perils were identified as being of particular interest to the study. Data was collected on each of these perils from a number of different sources, resulting in data sets for further analysis (*see table 1, below*). The number in the left column is the peril number which, at times, is used as shorthand for each peril type.

Table 1: The sample size and date range for data corresponding to each peril type.

Number	Peril Type	Count	Start Year	End Year
1	Tropical Cyclones - Gulf and Florida	620	1851	2012
2	Tropical Cyclones - NE USA and Canada	1,102	1851	2012
3	US Floods	410	1985	2015
4	US Tornadoes	58,885	1950	2014
5	European Windstorms	50	1981	2012
6	EU Floods	297	1985	2015
7	North West Pacific Tropical Cyclones	1,930	1945	2011
8	Australia Windstorms	734	1907	2015
9	Australia Floods	123	1985	2015
10	Australia Wildfires	58	1851	2015
11	South Africa Floods	51	1985	2015
12	Indian Ocean Tropical Cyclones	2,542	1945	2012
13	Thailand/Malaysia Floods	122	1985	2015
14	Mexico Floods	80	1985	2015
15	Brazil Floods	97	1985	2015
16	China Floods	344	1985	2015

The numbers in the left column are sometimes used as shorthand to refer to each peril type.

For some perils, inclusion of events is somewhat arbitrary. For example, for European windstorms, the source of the data consists of ‘50 of the worst windstorms to hit Europe since 1981’, ‘worst’ being defined somewhat loosely as those that resulted in particularly high insurance losses. This means that, for this study, only very extreme European windstorms are considered. A similar situation exists for Australian wildfires in which only a small subset of events, those that are noteworthy for being particularly destructive, are included in the data set. Conversely, for some perils, there are likely to be many events in the data set that do not result in high insurance losses. This is particularly true of tornadoes, a very large number of which occur each year and cause relatively low damage. When referring to an occurrence of a peril, unless otherwise stated, this covers those that are included in the study data set.

3 Testing independence within each peril type

If the study perils do not occur independently in time, a likely cause is that this assumption fails within peril types, ie either there is a tendency for periods of high activity or that a second event is less likely given the occurrence of a first. Either of these effects seems plausible physically. For example, there could be periods of time in which the meteorological conditions are favourable for a certain type of peril. It is therefore sensible to start by testing for independence within each peril type.

3.1 Methodology

If perils truly occur independently of each other in time, the probability of an occurrence is the same regardless of whether another has recently occurred. Statistically, this means that the marginal distribution of the occurrence of a new event is equivalent to the conditional distribution given the recent occurrence of some other event. To test whether this is the case, it is reasonable to estimate the distribution of elapsed times between events under the assumption of independence and compare this distribution with that of the observed times. If the two distributions are significantly different, there is evidence that the events are not independent.

Define a set of waiting times to be the time in days between *n* occurrences of a particular peril considered in chronological order. To estimate the distribution of waiting times under the assumption of independence, we would like to be able to take random draws from the distribution of calendar days (integers between 1 and 365) on which events occur and calculate the elapsed time between them. To do this, one method is to simply sample from the set of observed calendar days which can be considered draws from this distribution. The number of events in each year is kept the same as the observed number and calendar days were resample without replacement. This means that each observed calendar day is resampled precisely once and assigned to a calendar year. It is then possible to calculate waiting times between the resampled events. Resample occurred 128 times and hence there are 128 times as many draws from the expected distribution of waiting times than the observed distribution.

Having drawn from the distribution of waiting times under the assumption of independence, we then require some methodology to test whether the observed waiting times can be considered to be drawn from the same distribution. To do this, we use a Kolmogorov-Smirnov test. To conduct this test, the empirical cumulative density functions of the expected and observed waiting times are compared and the maximum distance between the two CDF's is calculated. Since we discretise waiting times (ie we measure this in discrete days rather than continuous time), the test is not exact but tends towards an exact test as the sample size tends to infinity. For large enough samples, this effect is small. We apply this test to each of the perils individually and note the p value for each one.

3.2 Results

The results of applying the Kolmogorov-Smirnov test to each of the perils are shown in Table 2 (*see below*). In most cases, there is little evidence to reject the null hypothesis that events occur independently of each other in time. For US tornadoes, however, the p value implies extremely strong evidence to reject it. This is actually not surprising as we will discuss shortly. There is also enough evidence to reject the null hypothesis in two other cases, namely North West Pacific and Indian Ocean Tropical Cyclones. It should be noted that the probability of observing a significant result when the null hypothesis is true is much increased when multiple tests are applied and therefore the likelihood of a type one error should be considered.

Another very important issue to consider is the impact of small sample sizes. In some cases, most notably for floods, Australian wildfires and European windstorms, the sample size in this study is low. This has the effect of significantly reducing the power of the test and hence, whilst there may actually be dependence between events, the sample size is simply too low to show significant evidence of this. In all such cases, non-significant results should be interpreted as showing that there is not enough evidence to reject the null hypothesis rather than suggesting strong evidence of independence.

Following on from the results (*see Table 2, p7*), we now demonstrate a number of them in more detail. Whilst we are particularly interested in those with a low p value, we first demonstrate one in which the p value is relatively high to illustrate a case in which there is no significant evidence to reject the null hypothesis of independence.

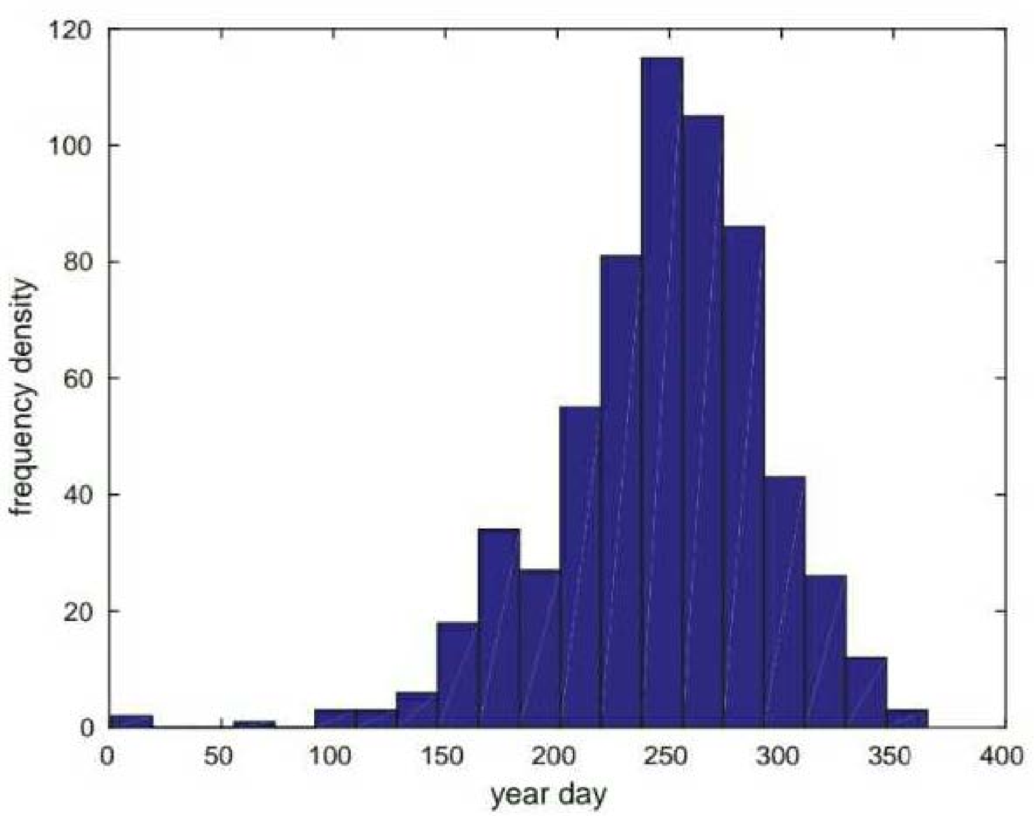
Table 2: p values and sample sizes for each peril type

Peril Type	Sample Size	p-value
Atlantic Tropical Cyclones - Gulf and Florida	620	0.4517
Atlantic Tropical Cyclones - North East and Canada	1102	0.2034
US Floods	410	0.8626
US Tornadoes	58885	0.0000
European Windstorms	50	0.1535
EU Floods	297	0.2706
North West Pacific Tropical Cyclones	1930	0.0000
Australia Windstorms	734	0.9735
Australia Floods	123	0.8544
Australia Wildfires	58	0.8427
South Africa Floods	51	0.9833
Indian Ocean Tropical Cyclones	2542	0.0052
Thailand/Malaysia Floods	122	0.3938
Mexico Floods	80	0.9077
Brazil Floods	97	0.7947
China Floods	344	0.6571

3.3 Atlantic Tropical Cyclones - North East and Canada

In this example, we focus on Atlantic Tropical Cyclones in the North East USA and Canada area. The data consist of 1,102 events which occurred between 1851 and 2012. For this type of peril, there is a well-defined distribution over a calendar year which is shown as a histogram (*see Figure 1, p8*). In this case, the test did not give a significant result ($p = 0.2034$) and so there is little evidence to reject the null hypothesis that such events occur independently of each other.

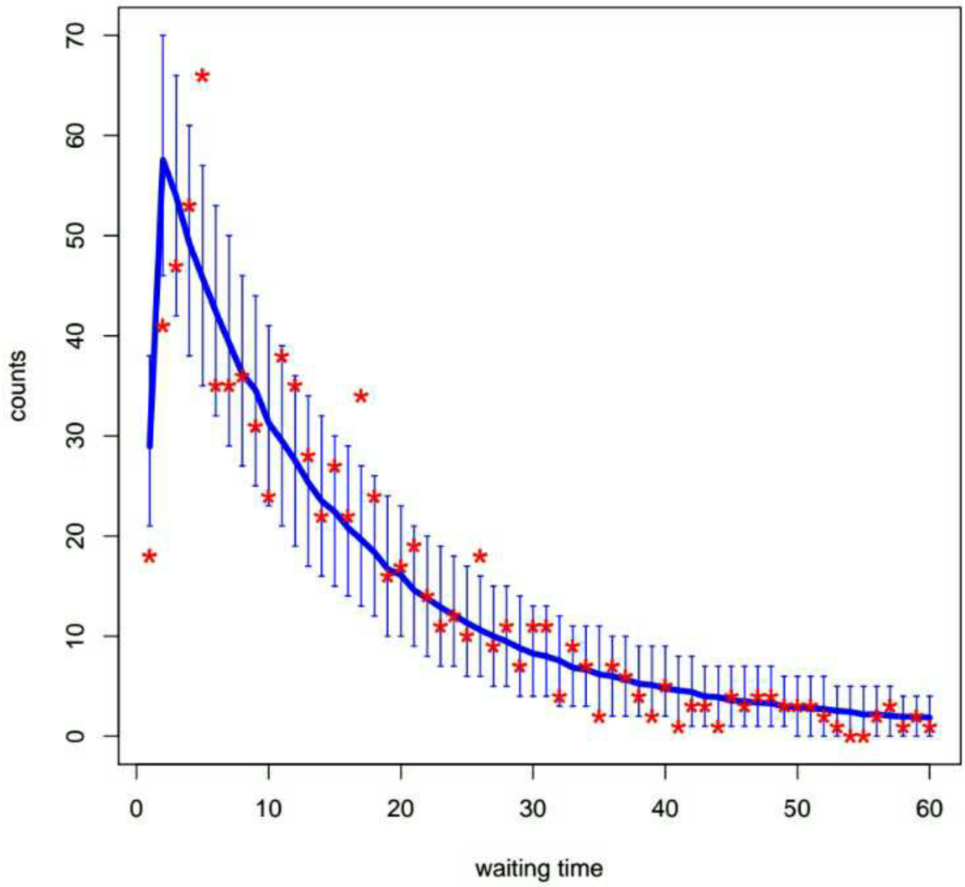
Figure 1: Tropical cyclones in North East US and Canada



A histogram of days of the year in which Tropical Cyclones occurred in the North East US and Canada area over the period of the dataset.

To illustrate our approach we can compare the expected count of each waiting time with the counts that were actually observed. These are illustrated in Figure 2 (*see p9*) in which the blue line represents the former and the red stars the latter. The error bars represent the range in which the observed counts are expected to fall with 90% probability when the null hypothesis of independence is true. Consistent with the p value, there appears to be very little evidence that the distributions are different besides normal sample variation.

Figure 2: Atlantic tropical cyclones in the North East US and Canada – expected counts and observed counts.

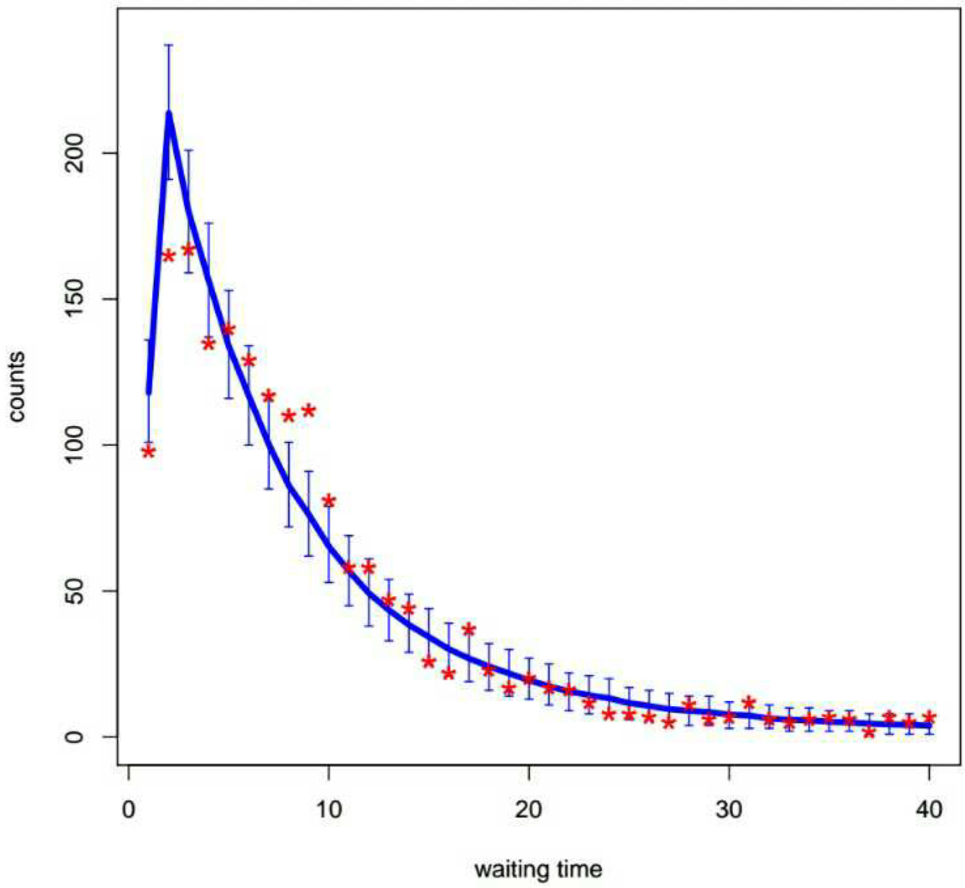


Expected (blue line) and observed (red stars) counts of each waiting time in days for Atlantic tropical cyclones in the North East US and Canada area. The error bars represent the region in which there is a 90% probability of the observed counts falling under the null hypothesis of independence. There appears to be little evidence to reject the null hypothesis that events occur independently of each other.

3.4 North West Pacific Tropical Cyclones

From the test in the previous section, there is evidence to reject the null hypothesis ($p < 0.0001$) that tropical cyclones in the North West Pacific occur independently of each other. The expected counts of each waiting time under the assumption of independence and the observed number of counts (*see Figure 3, below*). In this case, the significant result appears not to result from clustering but rather the opposite effect that the occurrence of an event tends to reduce the likelihood of another. This is evident from the fact that there are a smaller number of observed waiting times of zero, one, two and three days than expected.

Figure 3: North West Pacific tropical cyclones – expected counts and observed counts.

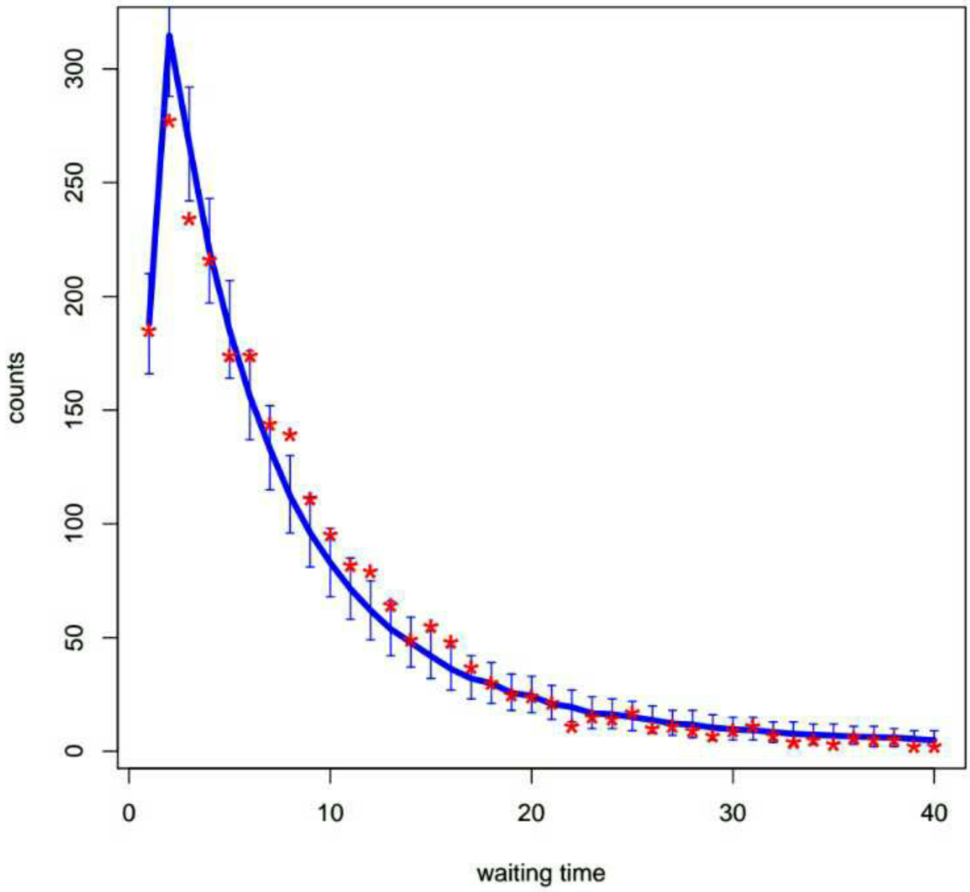


Expected counts with 90% error bars (blue lines) and observed counts (red stars) of each waiting time in days for North West Pacific tropical cyclones. There is some evidence that the occurrence of a tropical cyclone in this area results in a lower probability of another event in quick succession.

3.5 Indian Ocean Tropical Cyclones

In our analysis, we found significant evidence (p=0.0052) to reject the null hypothesis that Indian Ocean tropical cyclones occur independently of each other. Again, we investigate whether an occurrence of an event makes another in quick succession more or less likely. The observed and expected number of counts for each waiting time are compared (*see Figure 4, below*). Here, as with North West Pacific tropical cyclones, it appears that the occurrence of an event tends to decrease the likelihood of another in quick succession.

Figure 4: Indian Ocean tropical cyclones – expected counts and observed counts.

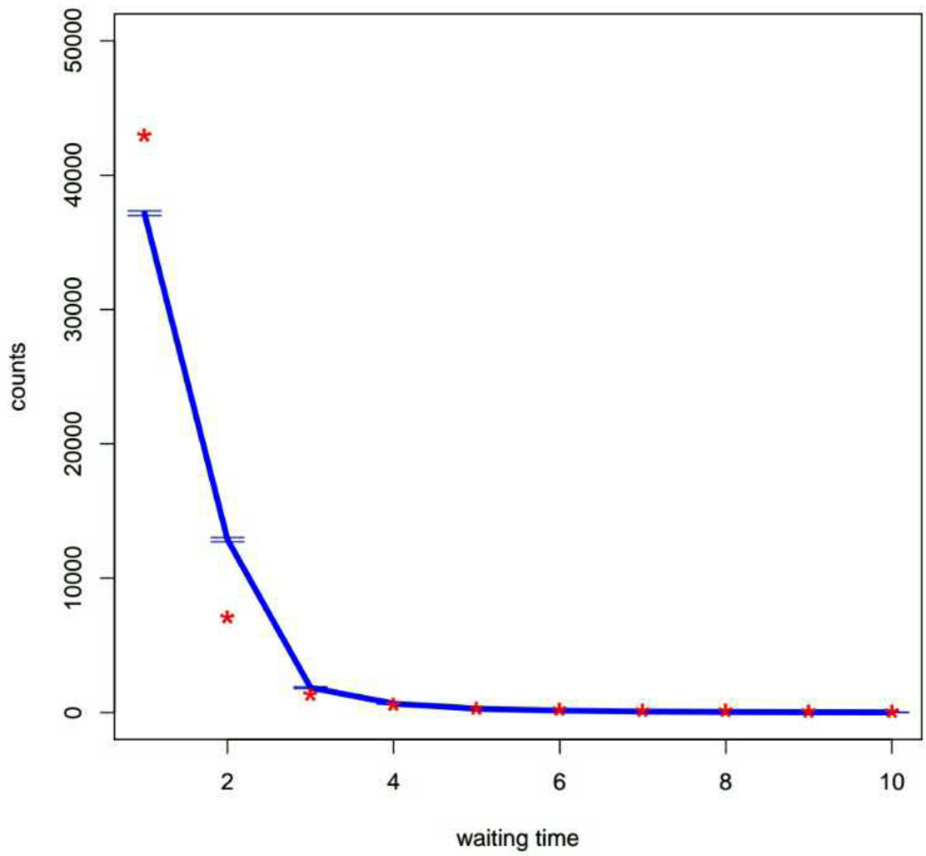


Expected counts with 90% error bars (blue lines) and observed counts (red stars) of each waiting time in days for Indian Ocean tropical cyclones. There is some evidence that the occurrence of a tropical cyclone in this area results in a lower probability of another subsequent event.

3.6 US Tornadoes

Previously, we found that the p value for the case of US tornadoes was very small suggesting extremely strong evidence to reject the null hypothesis that they occur independently. In fact, the phenomenon of clustering is well known to occur in US tornadoes since, commonly, more than one will occur in the same locality in very quick succession as a result of the same favourable conditions. We therefore see this effect in the analysis of this peril. The expected and observed counts of waiting times are shown in Figure 5 (*see below*). Here, it is clear that there are many more observed waiting times of zero days (ie events occurring on the same day) than would be expected were tornadoes to occur independently.

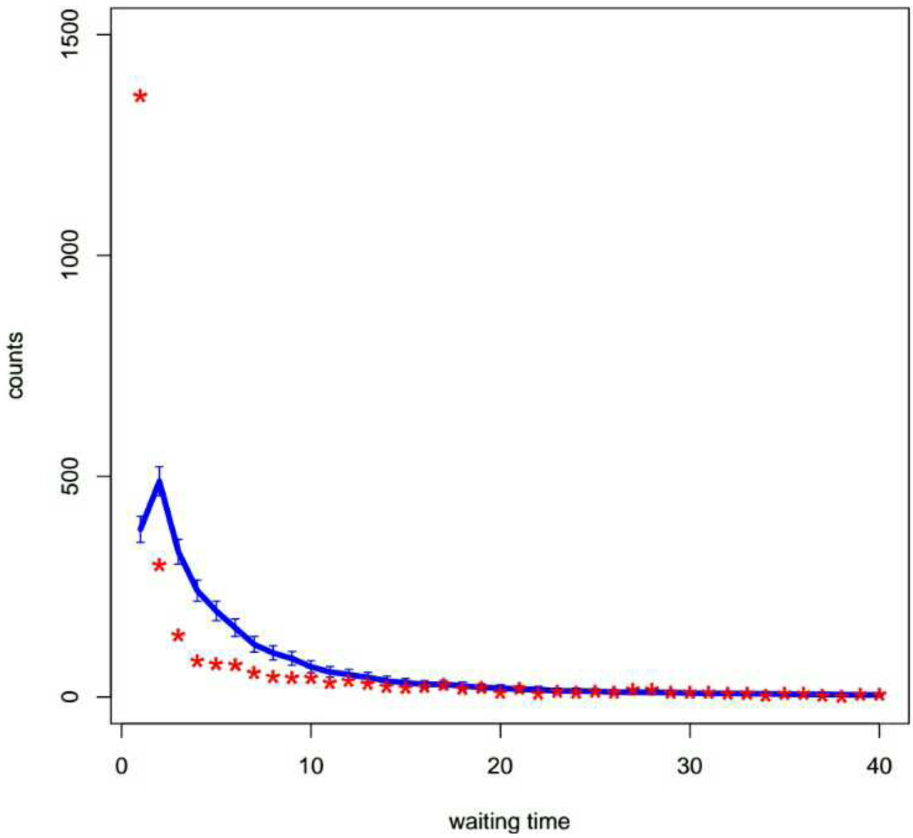
Figure 5: US tornadoes – expected counts and observed counts.



Expected counts with 90% error bars (blue lines) and observed counts (red stars) of each waiting time in days for US tornadoes. Here, there appears to be far more tornadoes occurring on the same day than would be expected under the assumption that events occur independently.

Many tornadoes occur each year in the USA, many of which are not severe in nature and are unlikely to have a significant cost to Lloyd's. Such events, therefore, are not as relevant as other perils. To test whether more destructive tornadoes are clustered, we now consider only the more intense tornadoes. In our data set, the intensity of tornadoes is measured on the Fujita scale which ranges from 0 to 5. Around 6% of tornadoes are considered to have an intensity of 3 or higher on this scale. Tornadoes are assigned 3 on this scale if there is considered to be 'severe' damage whilst 4 and 5 describe 'devastating' and 'incredible' damage respectively. When applying the Kolmogorov-Smirnov test to only tornadoes of these intensities, gain, we get a p value of virtually zero. The observed and expected counts in this case are shown in Figure 6 (*see below*). Again, it is clear that there are many more tornadoes that occur on the same day than would be expected if these events were independent.

Figure 6: US tornadoes that register as 3 or higher on the Fujita scale – expected counts and observed counts.



Expected counts with 90% error bars (blue lines) and observed counts (red stars) of each waiting time in days for US tornadoes that register as 3 or higher on the Fujita scale. There appears to be far more tornadoes of this intensity occurring on the same day as would be expected under the assumption that events occur independently.

4 Pairwise Tests of Independence

In the previous section, we investigated whether events within the same peril type can be considered to occur independently of each other. In this section, we test for independence between peril types, ie we ask whether the occurrence of an event within one peril type changes the likelihood of a subsequent event within another peril type.

4.1 Methodology

We take a similar approach to that used in the previous section. For each occurrence of a given peril type, which we refer to as peril one, we look at the waiting time until the next occurrence of some other peril type, which we refer to as peril two (If an occurrence of peril one and peril two are observed on the same day we treat this as a waiting time of 0). To sample from the distribution of waiting times under the assumption of independence, we resample without replacement from observed year days of peril one and peril two, again preserving the annual count for each and find the time between each occurrence of peril one and the next occurrence of peril two. We create 128 sets of waiting times from the expected distribution such that there are 128 times more expected waiting times than observed waiting times. As before, we then use a Kolmogorov-Smirnov test to test if there is evidence to reject the null hypothesis that the expected and observed times come from the same distribution. We repeat the test for all possible peril pairs.

4.2 Results

The p values of each test are shown where row numbers correspond to the peril type of peril one and the columns to that of peril two (*see table 3, below*) – for clarity, peril types are given as the peril numbers defined in Table 1 (*see p4*). Any p values less than 0.01 are starred and treated as significant and we find that there are a large number of cases in which the null hypothesis of independence is rejected. Since the number of tests is high (a total of 240), however, we should be cautious in how we interpret our results since we expect 2.4 tests on average to be significant at this level even if the null hypothesis is true in all cases.

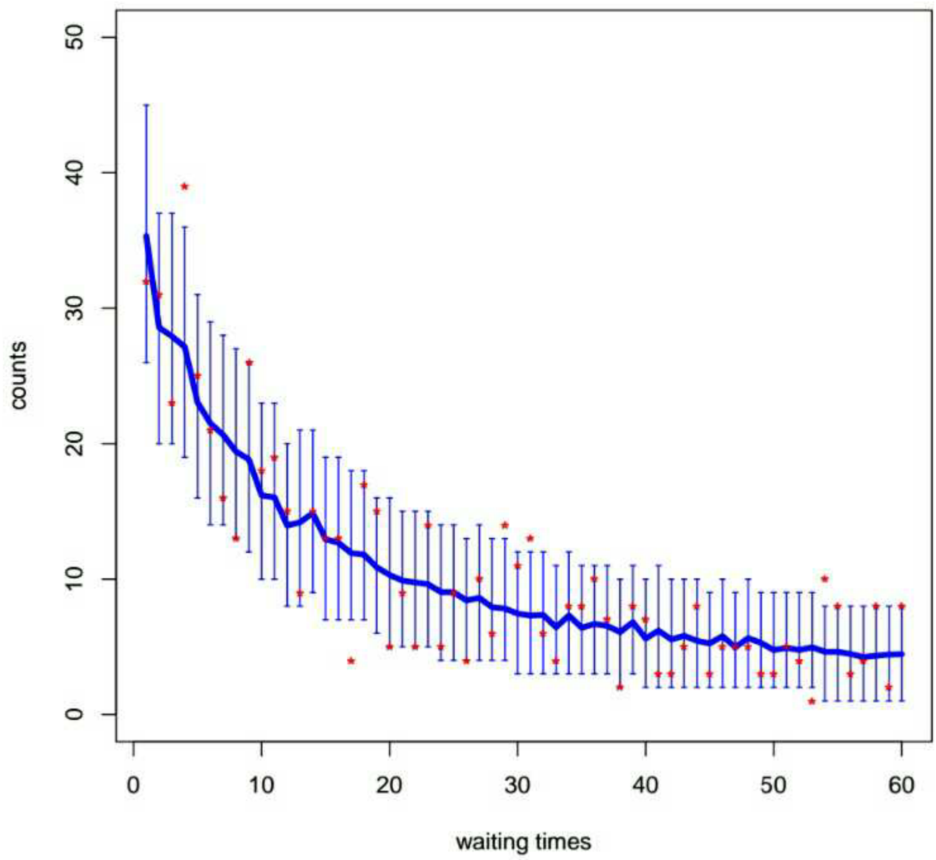
Table 3: p values of pairwise tests with rows corresponding to peril one and columns to peril two.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		0.3766	0.3178	0.2930	0.4635	0.2153	0.9064	0.1186	0.0589	0.9166	0.0989	0.3734	0.7162	0.3197	0.8764	0.6724
2	0.5897		0.2305	0.0077*	0.5646	0.0393	0.4856	0.2877	0.0044*	0.1305	0.3159	0.3258	0.0192	0.6485	0.0049*	0.1458
3	0.3393	0.9428		0.0014*	0.0201	0.0655	0.8121	0.9334	0.0520	0.0009*	0.0010*	0.7353	0.2194	0.5462	0.8465	0.6671
4	0.3478	0.1581	0.0001*		0.2605	0.0000*	0.3665	0.2774	0.1406	0.1170	0.5350	0.0979	0.0652	0.0290	0.0013*	0.6523
5	0.7946	0.8168	0.3427	0.0063*		0.7259	0.6229	0.4650	0.0988	0.3306	0.9732	0.9040	0.1669	0.0514	0.2125	0.5138
6	0.9213	0.3466	0.5957	0.0007*	0.2851		0.6798	0.9113	0.5042	0.0105	0.9341	0.1161	0.5677	0.6041	0.3056	0.3761
7	0.7379	0.4124	0.7108	0.0014*	0.0003*	0.0113		0.7651	0.0000*	0.0002*	0.3417	0.9994	0.6189	0.4537	0.0892	0.9342
8	0.1670	0.2612	0.6247	0.0000*	0.0545	0.0701	0.3057		0.0009*	0.8009	0.1218	0.0000*	0.0003*	0.3774	0.8698	0.9397
9	0.6577	0.3937	0.8327	0.0000*	0.7659	0.3013	0.6096	0.8777		0.8405	0.5153	0.6966	0.1304	0.8808	0.0451	0.5067
10	0.9908	0.6169	0.7945	0.8944	0.5852	0.2505	0.1310	0.3868	0.5769		0.1844	0.4530	0.9554	0.2613	0.5907	0.6771
11	0.7042	0.9088	0.6674	0.0138	0.4530	0.0636	0.9279	0.6122	0.7896	0.0383		0.7859	0.8417	0.7188	0.8126	0.9541
12	0.5114	0.8782	0.0237	0.0000*	0.1607	0.0003*	0.1698	0.0000*	0.0010*	0.0010*	0.0172		0.0003*	0.0133	0.3732	0.7108
13	0.9937	0.5613	0.7957	0.0048*	0.5426	0.2194	0.2922	0.7758	0.2420	0.0347	0.6239	0.9574		0.6676	0.7024	0.9098
14	0.1357	0.7988	0.1748	0.0227	0.7541	0.4159	0.7774	0.3494	0.4858	0.8704	0.5570	0.7736	0.8096		0.5761	0.9484
15	0.8997	0.7871	0.2709	0.0282	0.4876	0.2138	0.1625	0.2789	0.4886	0.6410	0.6366	0.9344	0.4643	0.9702		0.9582
16	0.3954	0.7347	0.3715	0.6750	0.0173	0.4965	0.5016	0.4977	0.0003*	0.0042*	0.1373	0.3315	0.4816	0.6196	0.0000*	

Values are starred if they are significant at the 1% level.

There are far too many significant results from these tests to show each in more detail. Instead, we simply take two examples, the first in which there is no significant evidence to reject independence between the pairs and the second in which there is. In the first example, we look more closely at the case in which peril one consists of North West Pacific tropical cyclones and peril two consists of Indian Ocean tropical cyclones. The p value of this test is 0.9994 and thus there is very little evidence to reject independence. This is illustrated further in Figure 7 (*see below*) in which, as with earlier plots, the blue line represents the expected counts of each waiting time under independence and the red stars the observed waiting times. Here, there appears to be little difference between the counts beyond normal sample variation.

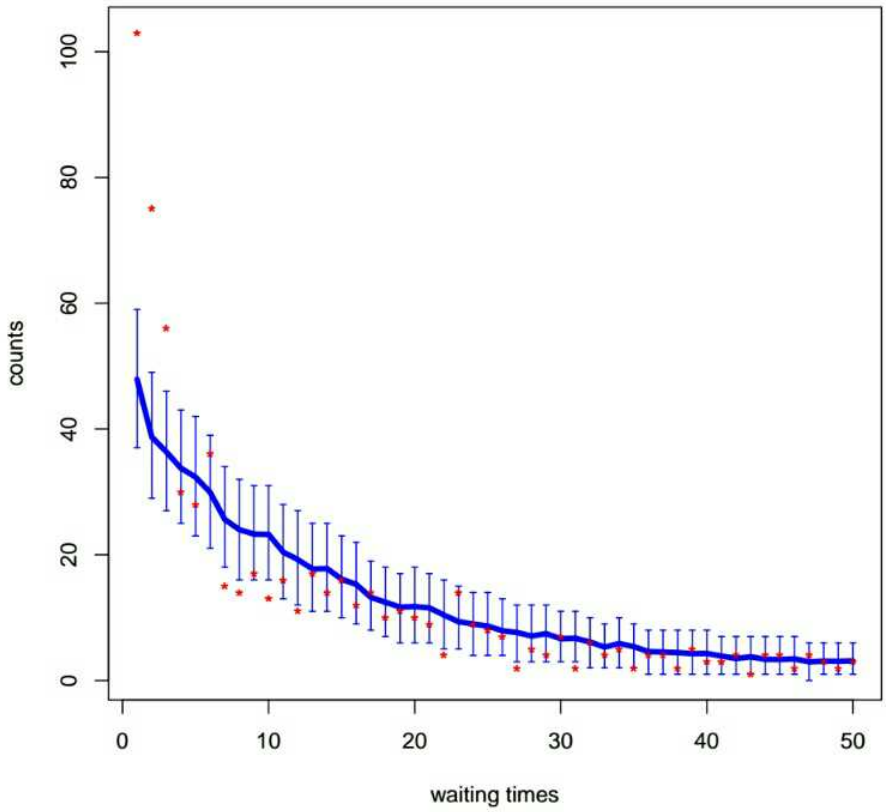
Figure 7: North West Pacific tropical cyclones and Indian Ocean tropical cyclones – expected counts and observed counts.



Expected counts with 90% error bars (blue lines) and observed counts (red stars) of each waiting time in days between occurrences of North West Pacific tropical cyclones and Indian Ocean tropical cyclones. In this case, there appears to be little evidence to reject the null hypothesis that such these events occur independently of each other.

In the second example, we look at the case in which peril one consists of Indian Ocean tropical cyclones and peril two Australian windstorms. In this case, as illustrated in Figure 8 (*see p17*), the observed numbers of very short waiting times are much higher than the expected number. This suggests that the occurrence of an Indian Ocean tropical cyclone increases the likelihood of an Australian windstorm shortly after.

Figure 8: Indian Ocean tropical cyclones and Australian windstorms – expected counts and observed counts.



Expected counts with 90% error bars (blue lines) and observed counts (red stars) of each waiting time in days between Indian Ocean tropical cyclones and Australian windstorms. In this case, the occurrence of the former appears to be increase the likelihood of an occurrence of the latter in quick succession.

Whilst we do not show any more examples, on further inspection, there are very few cases in which conclusive evidence about the nature of the relationship between two peril types can be found. Whilst in the example shown above, there is strong evidence to suggest that the occurrence of an Indian Ocean Tropical Cyclone tends to increase the likelihood of an Australian windstorm in quick succession, in no other case is there such a clear pattern and thus the only conclusion we can draw is that the observed distribution of waiting times is different to the expected distribution under the assumption of independence. Due to the complex nature of Earth's climate, this seems like a reasonable conclusion. It should be further noted that if an occurrence of peril one changes the likelihood of an occurrence of peril two, the reverse must also be true, ie the occurrence of peril two changes the likelihood of an occurrence of peril one. When this is not evident in the results of the tests, this is likely due to differences in the sample sizes of each peril type.

5 Simultaneous Testing of All Perils

So far, we have tested for independence both within and between peril types. Given that we have found significant results in a non-trivial number of cases, we have very strong evidence to suggest that Lloyd's perils, in fact, do not occur independently of each other. This is not particularly surprising given the complex nature of the world's climate and, even if all perils were truly independent, it would be difficult to prove in any case. This suggests a slight change in our aims. Instead of asking whether events are truly independent of each other, we can instead consider whether the assumption of independence provides a reasonable approximation to their behaviour. Given that we can never expect to have a perfect model of something as complex as the relationship between Lloyd's perils, if the null hypothesis can not be rejected, it could be argued that the independence assumption is a reasonable enough approximation to be informative. In this section, we test for independence when all perils are considered simultaneously to investigate whether this is the case.

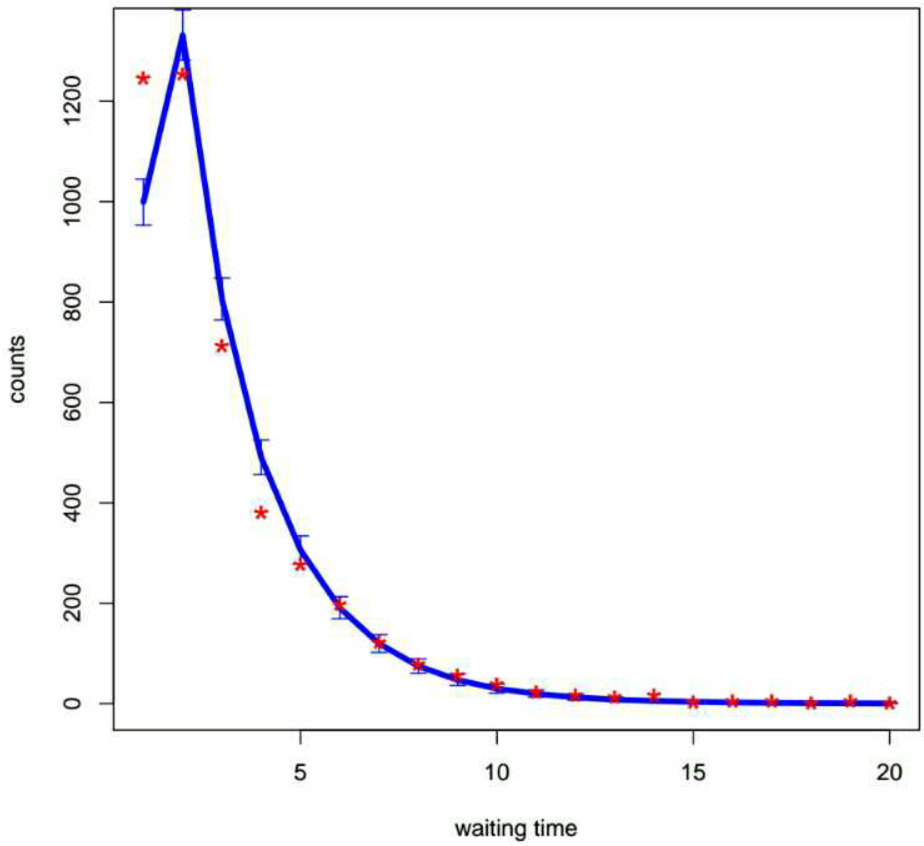
5.1 Methodology

The methodology used in this section is very similar to that of the previous section. However, instead of comparing waiting times between occurrences of each individual peril, we compare waiting times between multiple types of peril. To do this, for each peril, we draw, without replacement, the same number of events as was observed in each calendar year from the set of all observed year days (for that peril). The times between consecutive events are then treated as draws from the expected distribution of waiting times under the assumption of independence. We resample 128 times such that the number of samples from the expected distribution of waiting times is 128 times larger than the number of observed waiting times. As in the previous section, we compare the distributions using the Kolmogorov-Smirnov test and report the p value.

5.2 Results

In the first case, we test the hypothesis that all perils can be considered to occur independently of each other. In the case of tornadoes, we only consider those categorised as 3 or higher on the Fujita scale. Applying the methodology described above, we obtain a p value very close to zero. The number of observed counts of each waiting time are compared with their expected numbers (*see Figure 9, p19*). Here, there are many more events that occur on the same day than would be expected were the events to occur independently of each other. This, in fact, is entirely in line with what we have discovered so far since we have concluded that tornadoes are highly likely to be clustered in time.

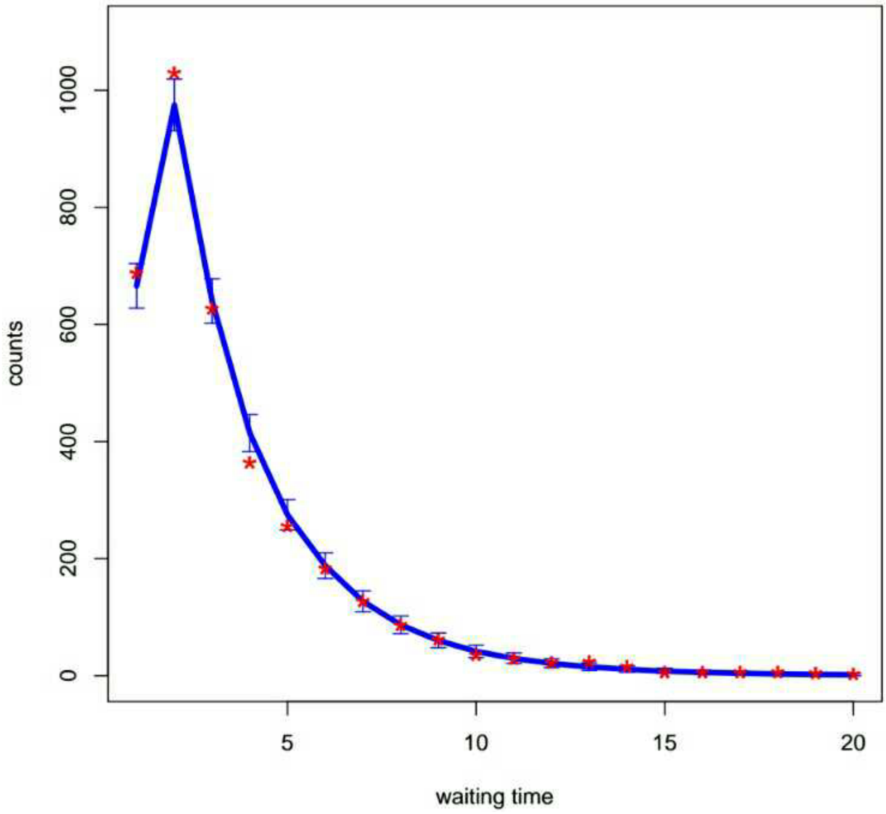
Figure 9: All study perils – expected counts and observed counts.



Expected counts with 90% error bars (blue lines) and observed counts (red stars) of each waiting time in days for all study perils. Consistent with the significant result, the observed and expected distributions appear to be different.

To investigate whether events can be considered to be independent aside from between US tornadoes, we repeat the experiment omitting this peril. In this case, we obtain a p value of 0.0845, a weakly significant result. There is thus weak evidence to reject the null hypothesis that perils occur independently once tornadoes have been excluded. The observed and expected counts of waiting times for this case are shown in Figure 10 (*see p20*).

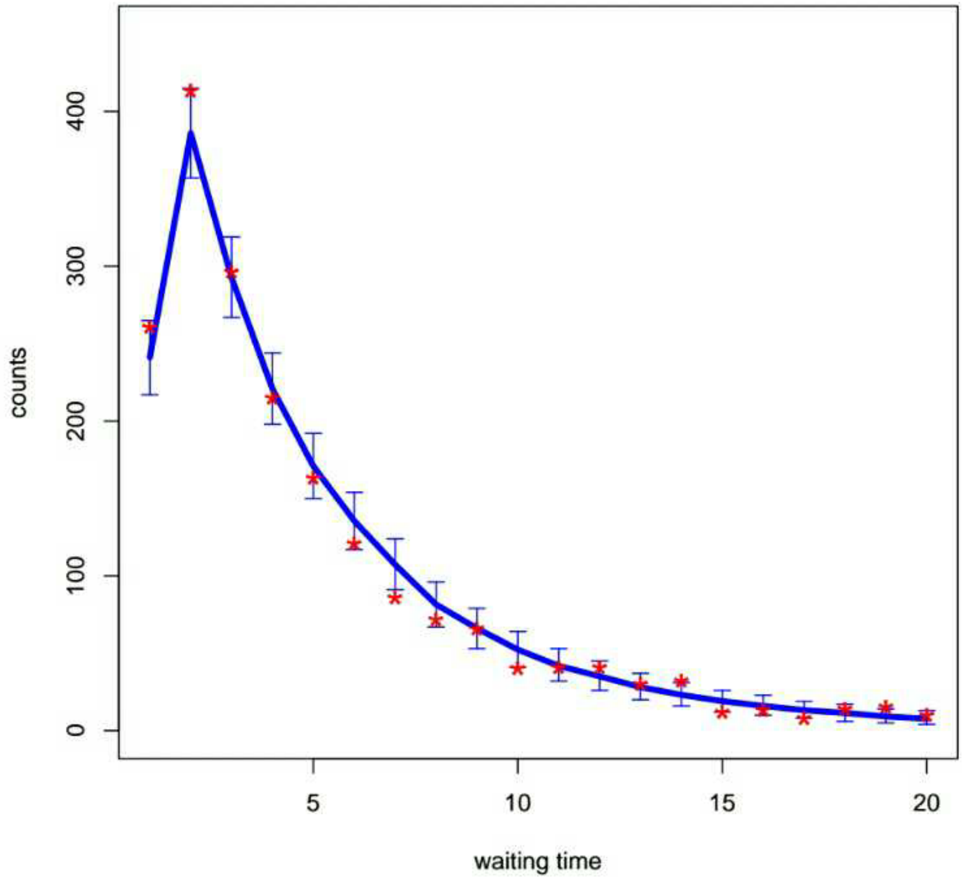
Figure 10: All study perils, except tornadoes – expected counts and observed counts.



Expected counts with 90% error bars (blue lines) and observed counts (red stars) of each waiting time in days for all study perils except tornadoes. Consistent with the non-significant result of the Kolmogorov- Smirnov test, once tornadoes are excluded, the distributions do not appear to be very different.

Here, consistent with the result of the test, there does not appear to be much difference between the observed and expected counts beyond normal sample variation. Previously, we also found significant evidence to reject the null hypothesis that North West Pacific and Indian Ocean Tropical Cyclones occur independently of each other. Applying our methodology without these two perils as well as US tornadoes, we obtain a p value of 0.1583, a non-significant result according to standard test levels. The observed and expected counts are shown in Figure 11 (*see p21*) and, again, consistent with the test result, there is little sign of any significant difference between the observed and expected counts.

Figure 11: All study perils, except tornadoes and Indian Ocean and North West Pacific tropical cyclones – expected counts and observed counts.



Expected counts with 90% error bars (blue lines) and observed counts (red stars) of each waiting time in days for all study perils except tornadoes and Indian Ocean and North West Pacific tropical cyclones. Consistent with the non-significant result of the Kolmogorov-Smirnov test, once these perils are excluded, the distributions do not appear to be very different.

6 Tests for Independence of Extreme Events

Lloyd's and other major insurers can be hit particularly hard by the occurrence of extreme events. So far, we have spent little time considering the occurrences of those events that would be considered particularly extreme or unusual. In this section, we single out those events that have the potential for particularly high insurance losses and test whether they can be considered to occur independently of each other. In some cases, only the most extreme occurrences are included in the data set. For example, in the case of European Windstorms, only the 50 most extreme in terms of insurance losses were included. In these cases, we simply treat each event as extreme. In some cases, notably Indian Ocean Tropical Cyclones, there are a number of missing values for intensity. These events are simply omitted. Missing values often result from a lack of reporting equipment in certain areas and we thus assume that they do not result in large insurance losses. The definition of extreme events used for the study is summarised below (*see Table 4, below*).

Table 4: Definitions of events treated as 'extreme'.

Peril Type	Threshold	Proportion qualifying
Tropical Cyclones	Windspeed more than 90 knots	20%
Floods	Magnitude greater than 6	30%
US Tornadoes	4 or 5 on Fujita Scale	1%
European Windstorms	all	100%
Australia Wildfires	all	100%

6.1 Results

Applying the same methodology as in the previous section, we get a p value of 0.0230 suggesting some evidence that the null hypothesis can be rejected. As with our previous analyses, however, we expect to see some clustering due to the inclusion of tornadoes. Repeating the experiment without tornadoes gives us a p value of 0.9834 suggesting very little evidence to reject the null hypothesis of independence. It should be noted that, when only the most extreme events are analysed, the power of the test is lower due to the smaller sample size and thus we are less likely to reject the null hypothesis if it is untrue. Even so, the number of data points considered is still quite high (1,056 events) and thus the test can still be considered to have high power.

7 Comparing Annual Counts of Perils

So far, we have tested whether the occurrence of a Lloyd's peril changes the likelihood of another in quick succession. Whilst this is an interesting and useful approach, it tells us nothing about the overall frequency of events. In this section, we investigate whether there is any association between annual counts of different perils.

To test for association between annual counts of each type of peril, we calculate Pearson's correlation coefficient for each possible pairing of perils. These are shown in Table 5 (*see p26*) in which the columns and rows correspond to different peril numbers. Whilst it is difficult to interpret the results of applying many significance tests, we can still gain some insight into possible relationships. In total, we find that 6 out of 120 pairs produce a p value of less than 0.01 or 1%, all of which result from a positive correlation coefficient. We note that we expect, on average, to obtain 1.2 such significant results when applying this number of significance tests even when each of the null hypotheses are true. The probability of observing 6 or more such results is around 0.0002 and so it is highly likely that, in at least some cases, the null hypothesis that the counts are truly uncorrelated is in fact false.

Table 5: Pairwise correlation coefficients for each pairing of perils.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	0.23	-0.25	0.11	0.08	0.17	-0.21	-0.1	0.29	0.18	0.03	-0.3	0.01	0.24	0.16	-0.22
2		1	0.13	0.4	-0.28	0.03	-0.27	-0.25	0.33	0.19	0.05	-0.35	0.23	-0.03	0.35	0.09
3			1	0.19	-0.12	0.46	-0.23	-0.02	0.49	0.04	0.32	0.18	0.36	0.17	0.3	0.42
4				1	-0.03	0.2	0	-0.27	0.61	0.15	0.32	-0.33	0.49	0.26	0.29	0.07
5					1	-0.06	0.16	0.2	-0.32	-0.12	-0.19	0.07	-0.42	0.07	-0.32	-0.15
6						1	-0.17	0.04	0.44	0.17	0.39	0.09	0.36	0.31	0.32	0.55
7							1	-0.09	-0.45	-0.2	0.08	0.34	-0.13	-0.3	-0.21	-0.18
8								1	-0.1	0.24	-0.44	0.4	-0.26	0.19	0.11	0.16
9									1	0.18	0.36	-0.2	0.61	0.52	0.42	0.3
10										1	0.03	-0.35	0.33	0.02	0.31	0.13
11											1	-0.11	0.63	0.14	0.2	0.17
12												1	-0.33	-0.14	-0.29	0.18
13													1	0.11	0.38	0.13
14														1	0.05	0.34
15															1	0.24
16																1

The numbers in the left column and the top row correspond to the peril numbers defined in Table 1.

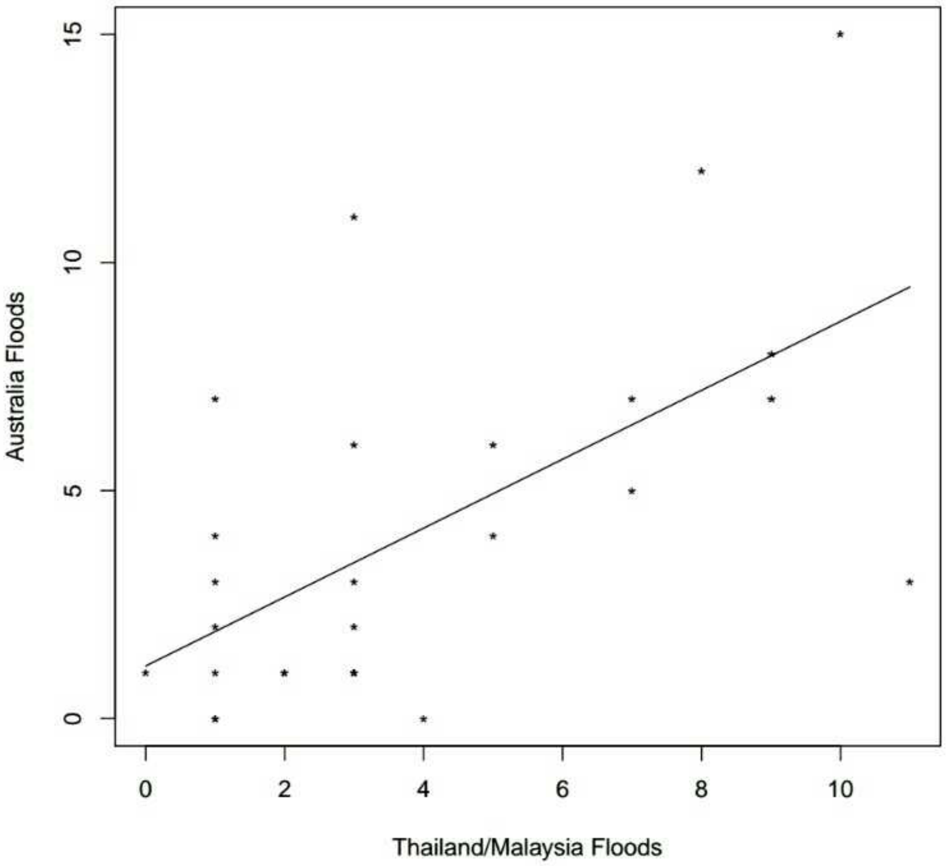
The pairs of events that indicate significant correlation are shown in table 6 (*see below*). For at least some of these, it seems plausible that there may be some genuine relationship. However, extreme caution should be taken and on further inspection we can speculate as to what causes the result to be significant.

Table 6: Pairs of events with significant correlation coefficients at the 1% level.

Peril one	Peril two	P Value
Australia Floods	US Floods	0.0096
Australia Floods	US Tornadoes	0.0008
EU Floods	China Floods	0.0027
Australia Floods	Thailand/Malaysia Floods	0.0007
Australia Floods	Mexico Floods	0.0053
South Africa Floods	Thailand/Malaysia Floods	0.0004

A scatter plot of yearly counts of floods in Australia and in Thailand/Malaysia is shown in Figure 12 (*see p28*) along with the least squares regression line.

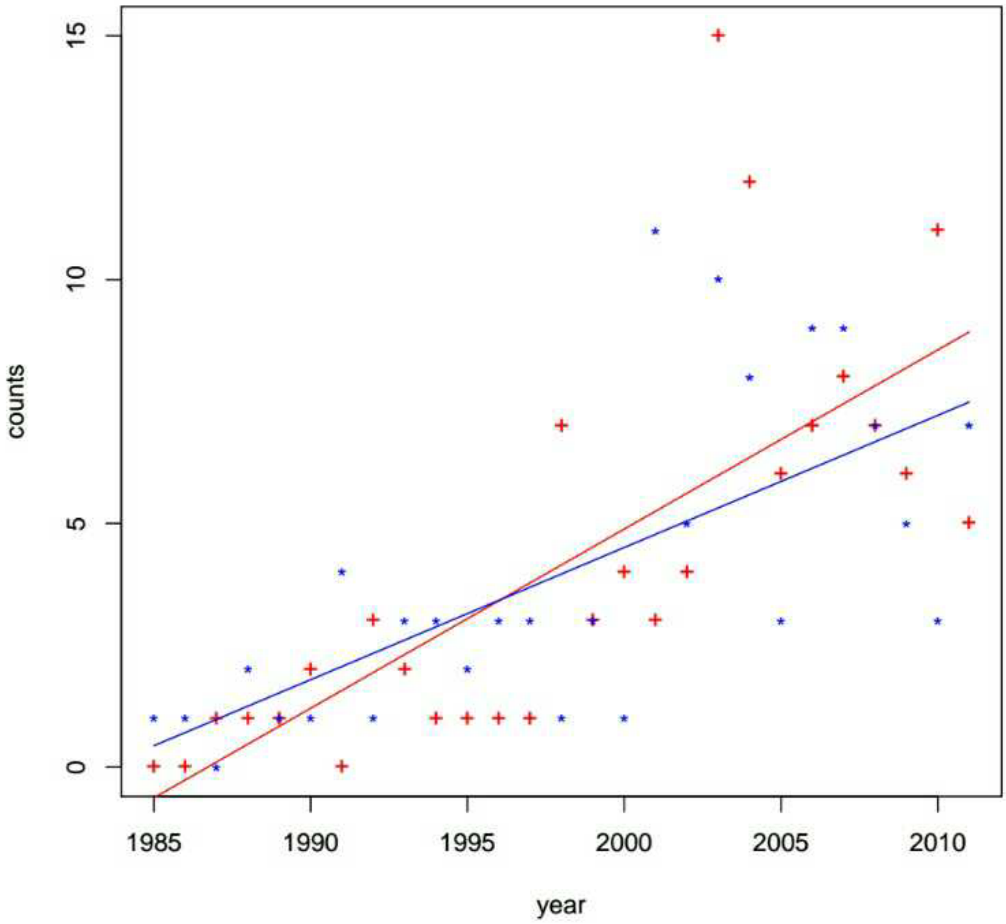
Figure 12: Frequency of floods in Australia and Thailand/Malaysia.



A scatterplot of the frequency of floods in Australia and Thailand/Malaysia with least squares regression line.

Here, there appears to be a strong positive correlation as indicated by the small p-value. With further investigation, however, we find that the number of reported events in both cases increased over the time period considered (*see Figure 13, p29*). Given that both types of peril increased over this time period, we expect to see a positive correlation and thus, whilst it is quite possible that there is an association beyond this mutual growth in frequency, it is hard to draw this conclusion from this test.

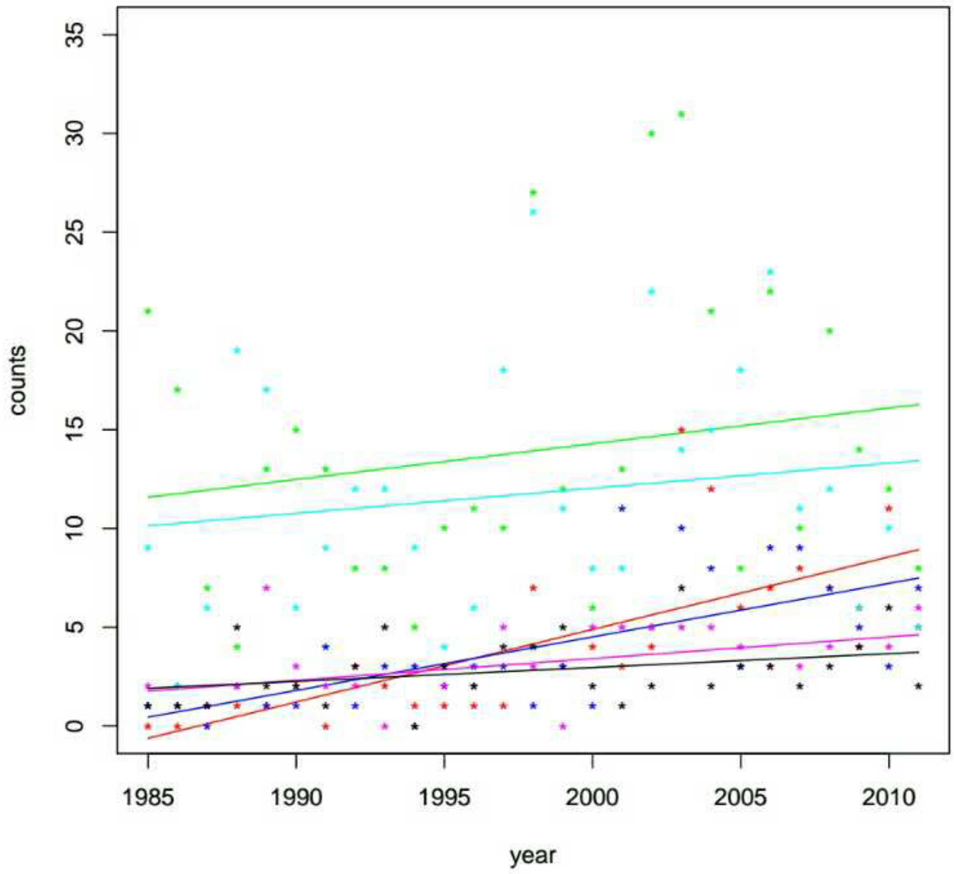
Figure 13: Annual counts of floods in Australia and Thailand/Malaysia.



Annual counts of floods in Australia (blue stars) and Thailand/Malaysia (red crosses) with least squares regression lines in each case.

To investigate further, counts of each type of flood that occur in a pair with a significant result are mapped out (*see Figure 14, p30*). It appears that the number of floods recorded by the flood observatory is increasing over time and thus is likely to be the cause of the correlations we have observed.

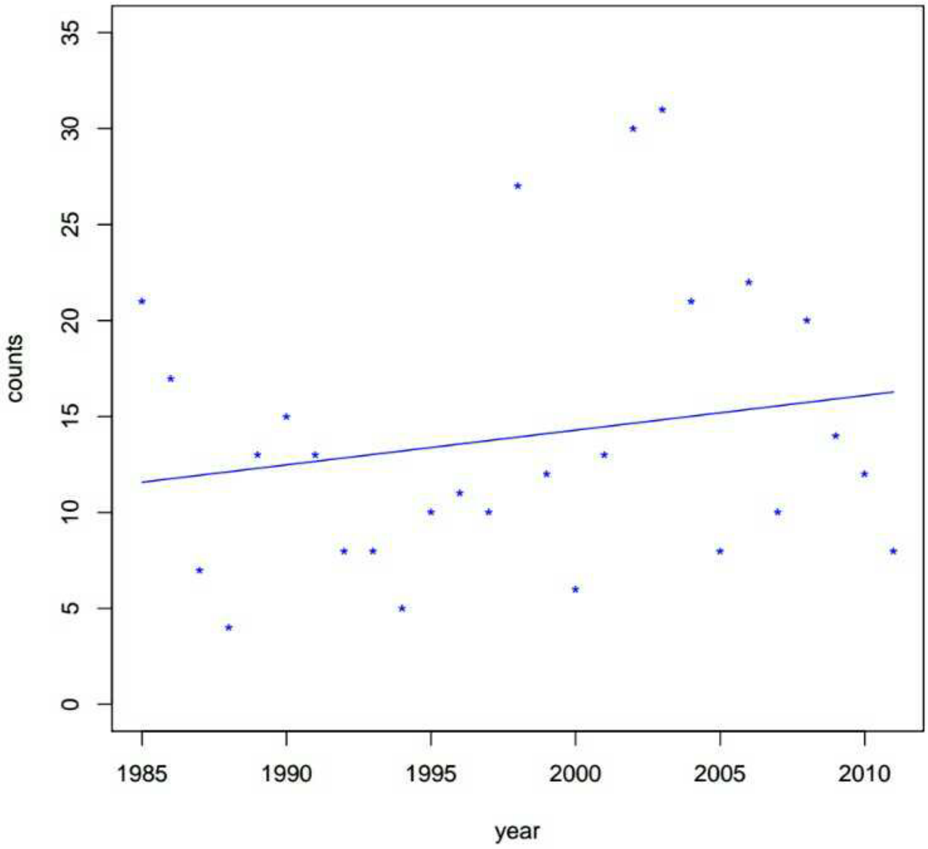
Figure 14: Annual counts of floods in the USA, Australia, Thailand/Malaysia, Brazil, China and Mexico.



Annual counts of floods in the USA (green), Australia (red), Thailand/Malaysia (blue), Brazil (magenta), China (cyan) and Mexico (black).

Similarly, the number of tornadoes over the period in which the flood data correspond to also appear to have increased though less markedly (*see Figure 15, p33*). From the data available, it is thus difficult to determine whether there is a genuine relationship beyond their common increase in frequency over time. Note also that this does not necessarily mean that the number of each type of flood is increasing. This effect could simply result from a change in the likelihood of a flood being recorded by the flood observatory from which the data is obtained. Further investigation into this would be interesting but it is not the purpose of this project to investigate whether the number of observed perils is changing. More investigation could be made to attempt to remove the effect of the trend though it is not clear how to do this in a satisfactory way.

Figure 15: Annual counts of US tornadoes.



Annual counts of US tornadoes with the least squares regression line.

8 Conclusions

In this report, we have considered the question of whether the study perils can reasonably be assumed to occur independently of each other. We first tested this assumption within each type of peril. We found that, in most cases, there is little evidence to reject the null hypothesis. In three cases, however, the observed waiting times were found to be significantly different from the expected times. Investigating further we found that for US tornadoes, this is caused by clustering, ie the likelihood of observing two tornadoes close together is higher than would be expected were they to occur independently of each other. In the case of Indian Ocean and North West Pacific tropical cyclones on the other hand, a significant result was found but this was shown to result from the opposite effect, ie that a new event is less likely to occur given the recent occurrence of another. Given the extremely strong evidence of clustering between US tornadoes, it is clear that the assumption of independence of Lloyd's perils is, in fact, violated.

Having tested the independence assumption of each individual type of peril, we then moved on to consider whether perils of different types can be considered to occur independently of each other. In many cases, we found significant evidence to reject the null hypothesis of independence. Although it can be difficult to interpret a large number of test results due to the inevitability of type I errors, the large number of significant results adds significant evidence to suggest that Lloyd's perils do not always occur independently of each other.

Having accepted that Lloyd's perils almost certainly do not occur independently of each other, a slight change of philosophy can be made. Instead of requiring that perils occur truly independently of each other, we can simply consider whether making the assumption of independence provides a reasonable approximation to their true behaviour. We tested whether the observed distribution of waiting times between all types of peril differs significantly from the expected distribution under independence. Consistent with our earlier findings, we found significant evidence to reject the null hypothesis. We then repeated the test omitting US tornadoes and found that doing this yielded only a weakly significant result. Next, we repeated this also omitting Indian Ocean and North West Pacific tropical cyclones and found a higher p value and thus little evidence to reject the null hypothesis of independence with standard significance levels. This means that, although we are almost certain that events are not independent, this assumption may actually provide a good approximation of the true behaviour once certain perils are removed. Although we have not attempted to test this, treating multiple tornadoes that occur in quick succession in the same vicinity as a single event may help improve the validity of the independence approximation when considering all perils simultaneously.

We then turned our attention to the most extreme events to test whether their behaviour is different to that of more common events. We thus repeated the analysis on events that we defined to be particularly severe. Considering all perils simultaneously, we found a highly significant result to reject independence. Removing US tornadoes from the analysis, however, resulted in a non-significant p value and hence there appears to be little difference in the conclusions obtained from considering perils of all intensities.

Finally, we tested for correlation in the frequency of different perils over a calendar year. Although we found several significant results, on further inspection, we found that these results are likely to be caused by mutual increases in reported events. To test this further would therefore require further thought as to how to alleviate this problem.