Appendix I – From National Flood Vulnerability to Risk

1. INTRODUCTION

The aim of this section was to explore the processes needed to develop a National Flood Risk Index based on the information gathered from the insurance companies about the costs and claims related with flood events. It was considered, in this approach, that a flood event is the **number of occurrences, or in this case insurance claims and losses (€), likely to have been affected by the same climate conditions that originated the flood.**

2. METHODS

2.1. Datasets

A query involving the major insurance companies in Portugal was made in order to assess flood risk at local and national scale. In this appendix it was used the following daily information: (i) number of claims and (ii) claims cost, grouped by the first four digits zip code between the 1st of January 2000 and 27th of October 2011.

2.2. Data analysis

To assess the costs and claims per flood event it was developed a Visual Basic script that searches the database for flood occurrences likely to be affected by the same climatic conditions. This analysis was developed in three steps:

1. All occurrences with less than 2 claims that didn't had any occurrence in the contiguous zip codes areas were excluded;
2. All occurrences per zip code in a six day time frame were grouped;
3. All occurrences in the contiguous zip code areas in a six day time frame since the event started were grouped.

The six days window was agreed based on the duration of the climate conditions that can originate floods (from 6h to 3 days) and the time needed for insurance companies to register the claims (usually 3 business days).

With the approximately 12 years of data, which is a limited time frame, it was fitted probability densities functions **just as an example** to calculate the probability of non-exceedance of the: (i) number of flood events per year; the (ii) maximum number of claims per flood event and the (iii) the maximum cost per flood event. The same approach can be used for the (iv) total number of claims per year; and the (v) annual cost on floods. Ideally its needed thirty years of data to capture the population variability but the goal with this exercise is to explore the information that can be extracted at the national scale and use it in the future in a national risk analysis approach.

Since this is just an exercise it was assumed that the variable “number of flood events per year” had normal distribution and for the variables “maximum number of claims per flood event” and “maximum cost per flood event” an EV type I distribution was applied.

3. RESULTS

The analysis of figure 1 and 2 show that the numbers of claims and the estimated flood events have identical annual trends reaching, on average per year, up to 821 claims and 91 flood events just during the winter months. The historical analysis show that February 2008 had the highest insurance losses caused by floods reaching about 6 million euros. In this month, between the 17th and the 23rd, a flood event affected almost all the coastal municipalities from Alcobaça to Grândola (Fig. 3) and had 895 claims with insurance losses of 5.387.333 euros.

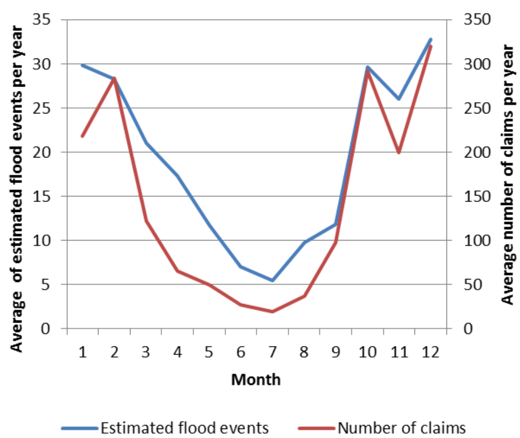

Fig. 1 – Estimated number of flood events and claims per year

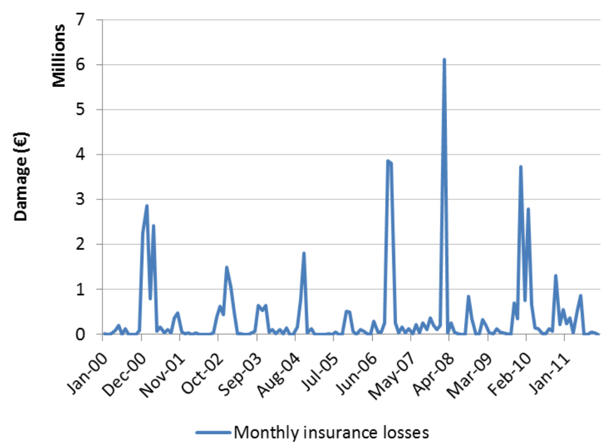

Fig. 2 – Monthly flood insurance losses

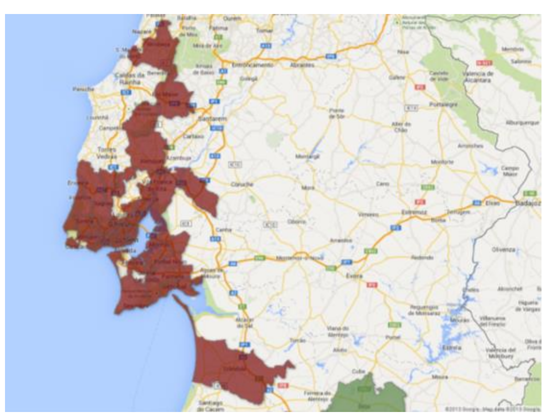

Fig. 3 – Affected zip codes areas, in red, of a flood event between the 17th and the 23rd of February 2008

Figure 4, 5 and table 1 shows the type of results that is possible to obtain assuming a normal distribution of the number of flood events per year.

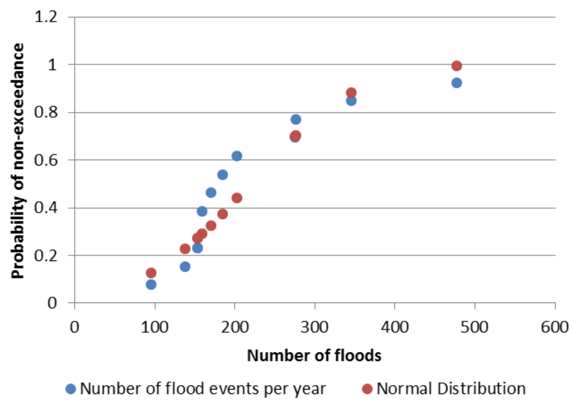

Fig. 4 – Comparison of the EV I and the ranking of the number of flood events per year.

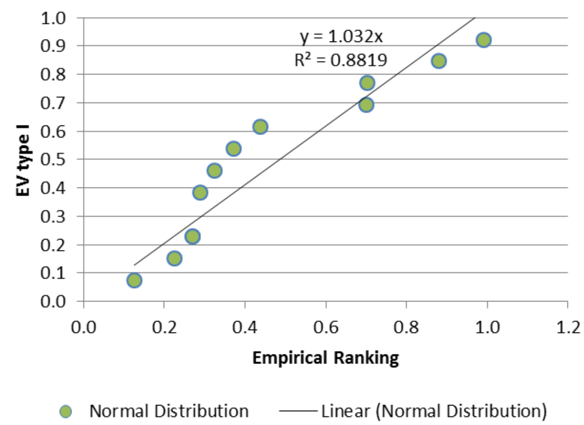

Fig. 5 - Linear fit between EV I density function and the number of flood events per year ranking.

Table 1 - Return Period of the estimated flood events.

Return Period	Estimated flood events per year
1.5	174
5	310
10	358
15	381
20	397
30	417
40	431
50	441
60	449
80	461
100	470
500	529

These results are just an example and cannot be used.

Figure 6, 7 and table 2 shows the same approach for the maximum number of claims per flood event assuming an EV I distribution.

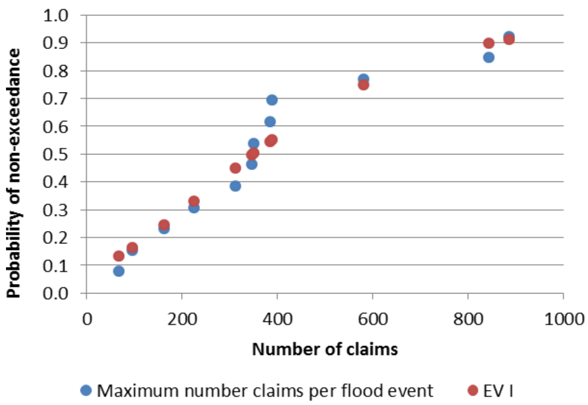

Fig. 6 - Comparison of the EV I and the ranking of the maximum number of claims per flood event.

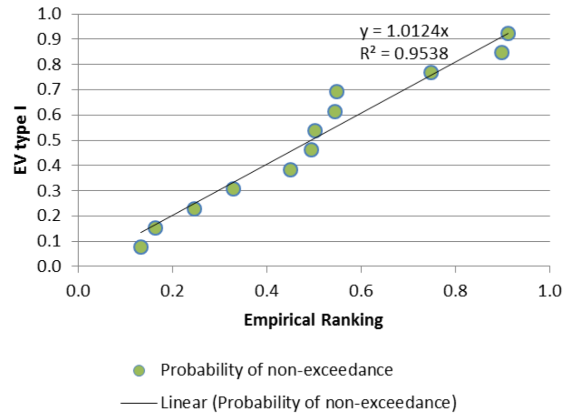

Fig. 7 - Linear fit between EV I density function and maximum number of claims per flood event ranking.

Table 2 – Return period of the maximum number of claims per flood event.

Return Period	Maximum number of claims per flood event
1.5	229
5	651
10	849
15	961
20	1040
30	1149
40	1227
50	1286
60	1335
80	1412
100	1471
500	1898

Figure 8, 9 and table 3 shows the same approach for the estimated insurance losses per flood event assuming an EV I distribution.

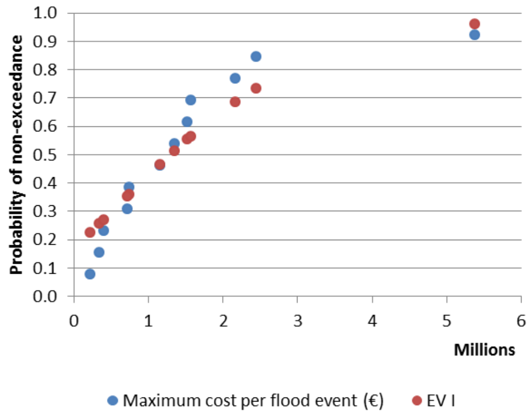

Fig. 8 - Comparison of the EV I and the ranking of the maximum cost per flood event.

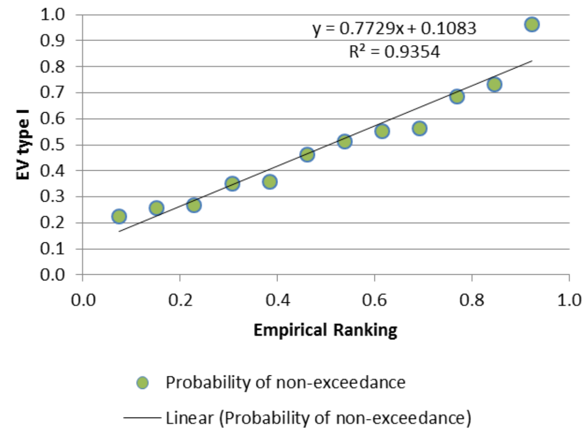

Fig. 9 - Linear fit between EV I density function and the maximum cost per flood event ranking.

Table 3 – Return period of the estimated maximum insurance losses per flood event

Return Period	Estimated maximum cost per flood event (Million euros)
1.5	0.65
5	2.90
10	3.97
15	4.56
20	4.98
30	5.57
40	5.98
50	6.30
60	6.56
80	6.97
100	7.29
500	9.57

4. CONCLUSIONS

The methodology presented showed that is possible to aggregate flood occurrences into flood events which also means that is possible to have a flood frequency and the respective insurance losses at the national scale. For this to be possible all flood events should be compared with the synoptic climate conditions in a clustering analysis, identifying and grouping weather patterns responsible for flood events. The analysis of the selected weather patterns in climate change scenarios can point out the flood trends frequency, intensity and expected damage.

It is very important to address the fact that the work presented is an exploratory analysis based on almost 12 years of data which is clearly insufficient to take any final conclusions. The ideal scenario is to have at least 30 years of data to be able to accommodate the climatic variability. It is also important to take into account that the results presented are biased by the representativeness of the insurance policies in each zip code. Including other databases of occurrences can improve the results.