Submitted to *Atmos. Res.*, Manuscript No., as of 18 Mar 2010

Effects of the El Niño – Southern Oscillation (ENSO) on heavy precipitation events and associated losses along the North American Pacific coast

Bernhard Reinhardt[1,*], Nikolai Dotzek[1,2] and Eberhard Faust[3]

[1] Deutsches Zentrum für Luft- und Raumfahrt (DLR) – Institut für Physik der Atmosphäre, Oberpfaffenhofen, 82234 Wessling, Germany

[2] European Severe Storms Laboratory, Münchner Straße 20, 82234 Wessling, Germany

[3] Munich RE, Königinstr. 107, 80802 München, Germany

5th ECSS Special Issue

Received 18 March 2010

[*] *Corresponding Author*: Bernhard Reinhardt, Deutsches Zentrum für Luft- und Raumfahrt (DLR), c/o Meteorologisches Institut München (MIM), Ludwig-Maximilians-Universität (LMU), Theresienstr. 37, 80333 München, Germany. Tel: +49-89-2180-4571, eMail: Bernhard.reinhardt@dlr.de

Abstract

We investigate to what extent it is possible to exploit by statistical means the influence of the El Niño – Southern Oscillation (ENSO) on the rainfall distribution along the North American Pacific coast. Regression with monthly sea surface temperature (SST) anomalies from the central Pacific is tested for predictability of upcoming wet-season extreme precipitation frequency. It is also evaluated if an SST signal can be detected in the weather-related loss data provided by the natural catastrophe data base NatCatSERVICE operated by Munich RE. We show that the summertime SST anomalies in the tropical Pacific can explain 18% of the variance of heavy precipitation frequency in the southwest of North America (USA, Mexico) during the winter season (NDJF). Increased summer SST in the central Pacific is connected to a significantly enhanced risk of a severe winter precipitation season. The probability of heavy precipitation exceeding a certain frequency in the upcoming winter can reasonably be assessed with a lead time of five months.

Keywords: Extreme event, Heavy precipitation, El Niño, ENSO, California, Arizona, Mexico, Pacific, Cox regression model; Seasonal prediction

1 Introduction

Substantial benefits could arise from prior knowledge whether atmospheric conditions will foster heavy precipitation events in an upcoming wet season. One example for this is the insurance industry which could adapt their risk management to account for expected loss levels, but there would also be benefits for society as a whole, since heavy precipitation also contributes largely to annual rainfall (Gershunov and Barnett, 1998). Thus, adaptation of water management would be facilitated as well.

Among others, Gershunov and Barnett (1998), Cayan et al. (1999), or Schubert et al. (2008) have shown that the El Niño – Southern Oscillation (ENSO) has strong influence on rainfall distribution and frequency of heavy precipitation events in America. In this study, we investigate to what extent it is possible to exploit this by statistical means, to what extent regression with monthly sea surface temperature (SST) anomalies can predict upcoming wet-season extreme precipitation frequency, and also if an SST signal can be detected in the weather-related losses provided by the natural catastrophe loss data base NatCatSERVICE operated by Munich RE.

We show that the summer SST anomalies in the tropical Pacific can explain 18% of the variance of heavy precipitation frequency in the southwest of North America (USA, Mexico) during the winter season (NDJF). Increased summer SST in the central Pacific is connected to an enhanced risk of a severe winter precipitation season. This significantly enhanced risk is also found for precipitation-related losses despite a sparse dataset. The paper is organized as follows: Sec. 2 outlines the datasets and definitions. In Sec. 3, we present methods and results and Sec. 4 gives our conclusions.

2 Data and Definitions

2.1 Data

The "US-MEX" dataset from NCEP is our source of precipitation data. It contains rain gauge measurements interpolated to a 1°x1° grid. Daily rain amounts are available in the domain 140W-60W, 10N-60N for the years 1948-2005 (Higgins et al., 1996; Higgins, 2000).

To determine the state of ENSO, we used monthly SST anomalies of the Pacific from three different tropical regions centered on the equator. These anomalies are the Niño3, Niño3.4 and Niño4 indices, according to the designators of the different areas. As shown in Fig. 1, the Niño3 area is the easternmost, the Niño4 area is the westernmost, and the Niño3.4

area overlaps with both of them. The Niño indices for 1950 to 2008 were obtained from the NOAA Climate Prediction Center (www.cpc.ncep.noaa.gov/data/indices/sstoi.indices; downloaded on 12/01/2009).

To highlight the spatial and temporal distribution of the information extractable from the SST and to verify that the information is well-contained in the Niño indices, we use the gridded SST dataset ERSST.v3 (Smith und Reynolds, 2003, 2004; Xue et al., 2003; Smith et al., 2008).

The catastrophe loss database NatCatSERVICE operated by Munich RE was used to assess the socio-economic impacts of the severe weather. In this database, catastrophes are classified into seven categories of severity ranging from "natural event" to "great disaster" (Table 1). While the database contains various natural hazards, only the types "blizzard", "flash flood", "general flooding", "landslide", "tropical cyclone" or "severe storm" were considered here.

2.2 Definitions

Extreme precipitation: We consider precipitation extreme if the local 99.7% quantile of daily rain amount is exceeded. This allows for investigation across different climatic regions since this choice ensures that on long-term average, an extreme event will occur about once (1.09 times) a year everywhere in the domain. For some calculations, a time series of a scalar value representing extreme precipitation rather than an evolving field was necessary. To obtain such, the pixels exhibiting extreme precipitation were counted in the investigation area between 17°N-38°N and 101°W-123°W. The selection of this area is motivated in section 3.2.

State of ENSO: A classification of the state of ENSO – that is, El Niño, neutral or La Niña – was required for some computations. Our classification may be based on one of the three ENSO indices mentioned above. When the selected monthly index is above the 75% quantile, El Niño is diagnosed, when it is below the 25% quantile, La Niña is diagnosed, and otherwise the state is considered neutral.

Damage index: The NatCatSERVICE events were converted to a time series similar to the extreme precipitation field: They were counted over the winter seasons in the investigation area. As a compromise between severity and the available number of events, only catastrophes with a NatCatSERVICE classification of 3 or 4 were used. For the timeframe 1975 to 2007, 64 such events were contained in the database. However, since our analysis was limited to the months NDJF, only 33 events could be used.

Category 3 or "severe catastrophe" denotes a total damage of USD 50-200 million and/or 20-100 fatalities. Category 4 or "major catastrophe" pertains to a total damage of USD 200-500 million and/or 100-500 fatalities (Table 1). To synthesize these events of varying severity by a single quantity, we arrived at a damage index D which weighs category 4 events three times as much as a category 3 event. Hence the formula is:

$$D = N_{Cat3} + 3\,N_{Cat4} \quad , \tag{1}$$

in which N denotes the number of catastrophes of a given category. For example, five category 3 events and two category 4 events will lead to a damage index of eleven.

3 Methods and Results

3.1 Average seasonal distribution of extreme precipitation and losses

The "US-MEX"-dataset includes several climatic regions which have storm seasons at different times in the year or do not have a specific storm season at all. The Pacific coast receives most extreme precipitation during winter, but, for example, northern Mexico has two peaks in extreme rainfall – one due to the winter rainfall and one due to the American Monsoon during summer (e.g., Adams and Comrie, 1997). Fig. 2 shows the recurrence frequency for extreme precipitation as a composite for November to February.

The seasonal distribution of the damage index D in the investigation area marked by the black rectangle in Fig. 1 is shown in Fig. 3. The winter months DJF account for over 50% of the damage. This corresponds well to the fact that extreme precipitation during this time of year is pronounced, too, and making it especially interesting to deal with changes during these months. In the following, we consider the months NDJF as winter, pertaining to similar precipitation patterns during these months rather than to calendrical constraints.

3.2 Change of recurrence intervals in connection with ENSO

The recurrence intervals for extreme precipitation in case of El Niño, La Niña, and for the whole time period were calculated for each month of the year. The ENSO status for this computation was derived from the monthly Niño3.4 index. Generally speaking, during the months November to February, El Niño is associated with a decrease of the recurrence intervals in the north of the dataset domain and with an increase in the south. The change of

sign occurs between 40°N and 45°N. A similar structure becomes apparent in the La Niña anomaly, but with opposite sign.

During the other months of the year, no clear larger-scale anomaly was noticed in the south of the domain. The only signature worth mentioning is an increase of extreme precipitation over Canada during La Niña that is lasting during all months. Due to the similar structure for the months NDJF, we show the recurrence frequency anomalies for El Niño and La Niña together as a difference. The area which our analysis is focusing on was chosen according to the focal areas of response given in Fig. 4. The investigation area contains most of the west coast positive anomaly pole seen in the El Niño – La Niña difference plot. We estimate that at least 55 million people are living in this investigation area.

3.3 Relating ENSO to future wet season intensity

The previous section demonstrated only qualitatively that ENSO can have a great impact on heavy precipitation in the investigation area. Next, we will show which regions of the Pacific have the biggest influence on the extreme precipitation at different lead times. Therefore, we calculated the Pearson correlation r between the number of pixels exhibiting extreme precipitation in the winter season (N, D, J, F) and the monthly SST anomalies from the ERSST dataset. This was repeatedly done for different lead times, defined as time between availability of the monthly SST data and 1 December. The pixels were counted on a daily basis so that each grid point can contribute multiple times per winter. The SST anomalies have been taken from months varying between November, that is, with a lead time of 0 months, and the January before with a lead time of 10 months. The years 1951 to 2004 were considered for this calculation.

In Fig. 5, three of the eleven resulting correlation maps are shown. Significant correlation can be seen in many parts of the Pacific. If we focus on the Niño areas centred on the equator, all three Niño areas contain similarly high correlation. However, if we move on to higher lead times, the correlation vanishes in the eastern Pacific, so that for the June map with a lead time of five months, only the Niño3.4 and the Niño4 areas still display high correlation values. For the highest considered lead time, ten months, the correlation has faded so far that only the Niño4 area contains pixels with substantial correlation.

Fig. 6 shows the square of the Person correlation, averaged over the Niño4 area. This r^2 is the fraction of variance that can be explained by the considered process. We see that for lead times up to five months, the Niño4 SST can explain about 15% of the winterly extreme precipitation variability before the correlation is fading for higher lead times. If we consider

only the SST in the western half of the Niño4 area (dashed line) the figures are slightly higher (up to 18%). How these correlations translate into terms of regression success is covered by the following section.

3.4 Seasonal forecast using Cox regression

In a regression, one or more observables are mapped via an analytical function into a predicted value. Coefficients in the formula need to be determined by a training dataset, so that the outcome fits the observations best. We have used a Niño index as input variable and the percentage of average extreme precipitation frequency in the investigation area in the upcoming winter as output. The Niño index can originate from either of the Niño areas and can have a lead time ranging from zero to ten months.

We then applied a Cox regression model of proportional hazards (Cox, 1972; Andersen and Gill, 1982; Therneau et al., 1990) as it was suggested by Maia and Meinke (2008). With this model, not only the output variable can be computed but also the likelihood for different outcomes. The statistics software R and the included package "survival" provided the required tools for this computation (R Development Core Team, 2009). The base function was estimated with the method by Nelson (1969) and Aalen (1978).

As the values of the Pearson correlation in the maps of Fig. 5a-c are all below 0.5, one would not expect that a perfect regression is possible. Exemplarily, Fig. 7 confirms this by showing that the observed values of extreme precipitation frequency are broadly scattered especially for positive SST anomalies.

To give an evaluation, we determined how often the regression is above or below the 100% line, that is, a prediction is correctly above or below the average seasonal course. A bootstrap method was applied to obtain stable results: 40 years were randomly chosen and used as training data set for the regression. For the remaining 15 years, the results were predicted. As the results are varying for different random choices of the years, the experiment was repeated 50000 times and then results were averaged. In Fig. 8, we see that the regressed values correctly reproduce an above- or below-average course of the season around 70% of the times for lead times up to five months. Note that the variability between different steps Δt is not the result of a special selection of the years but comes indeed out of the data.

The usage of more than one covariate in the regression – for example a second SST anomaly value at a different lead time or also the value of the Pacific North American Oscillation (PNA) index at different lead times – led to no improvement of the results. We

interpret this as an overfitting problem (e.g., Wilks, 2006) and conclude that only one covariate can be fitted with a sample size of 55 cold seasons.

From Fig. 7, it seems that an enhanced SST broadens the spectrum of the extreme precipitation variability in terms of creating favourable conditions for severe winter seasons, but not mandatorily triggers extreme precipitation. For the damage index, a regression did not seem appropriate. Only 33 events were available in 32 winters. Because of this, the seasonal damage index did only occur in five different values. This digital nature cannot be approached appropriately by a regression analysis operating with a continuous independent variable.

3.5 Hazard assessment

To determine the probability of a winter season to exhibit heavy precipitation above a certain threshold, we compared the series of El Niño winters to all winters using the Wilcoxon-Rank-Sum significance test. Calculations were made for a variety of different parameter combinations. Here, we only highlight one result: Winters for which the observed Niño4 index indicated El Niño already in the June before exhibit considerably higher extreme precipitation frequency on average than the whole time series. The Rank-Sum test indicates that the El Niño series is above 130% of the average series on a significance level greater than 90%. This means that there is only a chance of less than 10% that such high values can occur when choosing randomly the same amount of winters. Or in other words: If El Niño conditions prevail in June, the chances are higher than 90% that in the upcoming winter, the extreme precipitation count will be at least 30% higher than normal. The same result applies for the NatCatSERVICE damage index series although only a shorter period of 32 winter seasons could be used.

4 Conclusions

The results we presented did not make full use of the special abilities of the Cox regression model. A simple linear regression could probably provide results of similar quality. We found the data to be too sparse and too scattered to gain reasonably small confidence intervals for the fine-graded risk assessment which is one of the major advantages of the Cox regression model. However, we did exploit the integrated rating of the input variables and in general find the Cox regression model a promising approach. We also encourage other researchers to use methods of the survival analysis, to which the Cox regression model belongs, in the field of meteorology.

Our research revealed that the probability of heavy precipitation along the American Pacific coast in the upcoming winter can reasonably be assessed with a lead time of five months. For example, we found a 90% probability for a 130% severity of the upcoming winter season precipitation if El Niño conditions are present in June. Although the NatCatSERVICE dataset is very sparse in the considered domain, a higher risk for socio-economic impacts could be assessed as well. From the three considered Niño indices, the Niño4 index is best-suited for making estimates at long lead times. By selection of a customized input area in a SST dataset, slight improvements compared to this index appear possible. Nevertheless, an exact quantitative forecast of the extreme precipitation as we defined it is not possible, since El Niño seems not to be the trigger for a severe winter season but only a promoter. However, the regression presented here is beneficial for decision-makers who need estimates of the wet season severity repeatedly every year.

Acknowledgements

The authors are grateful to Simon Unterstrasser for reviewing a draft of this manuscript. This work was primarily funded by Munich RE.

References

Aalen, O., 1978: Nonparametric inference for a family of counting processes. *The Annals of Statistics*, 6(4):701–726.

Adams, D. and Comrie, A., 1997: The North American monsoon. *Bull. Amer. Meteor. Soc.*, 78(10): 2197–2213.

Andersen, P. K. and Gill, R. D., 1982: Cox's regression model for counting processes: A large sample study. *The Annals of Statistics*, 10(4):1100–1120.

Cayan, D. R., Redmond, K. T. and Riddle L. G., 1999: ENSO and hydrologic extremes in the western United States. *Journal of Climate*, 12(9): 2881–2893.

Cox D. R., 1972: Regression models and life-tables. *Journal of the Royal Statistical Society.Series B (Methodological)*, 34(2):187–220.

Gershunov, A. and Barnett, T. P., 1998: ENSO influence on intraseasonal extreme rainfall and temperature frequencies in the contiguous United States: Observations and model results. *Journal of Climate*, 11(7): 1575–1586.

Higgins, R., 2000: Improved US precipitation quality control system and analysis. *NCEP/Climate Prediction Center ATLAS No. 6.*

Higgins, R.W., Janowiak, J. E. and Yao, Y.-P, 1996: A gridded hourly precipitation data base for the United States (1963-1993). *NCEP/Climate Prediction Center Atlas 1, national Centers for Environmental Prediction*, page 46 pp.

Maia, A., and Meinke, H., 2008: Quantifying climate-related risks and uncertainties using Cox regression models. Newsletter of the Climate Variability and Predictability Programme (CLIVAR), 13:23–27.

Nelson,W., 1969: Hazard plotting for incomplete failure data. *Journal of Quality Technology*, 1:27–52.

R Development Core Team, 2009: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Reinhardt, B., 2009: *Einfluss der El Niño – Southern Oscillation auf die Eintrittswahrscheinlichkeit von Starkniederschlägen an der Pazifikküste Nordamerikas (Influence of the El Niño – Southern Oscillation on the likelihood of heavy precipitation along the Pacific coast of North America).* Diploma thesis, Meteorological Institute, Ludwig-Maximilians-University of Munich, 83 pp. [In German]

Schubert, S. D., Chang, Y., Suarez, M. J. and Pegion, P. J., 2008:. ENSO and Wintertime Extreme Precipitation Events over the Contiguous United States, *Journal of Climate*, 21: 22-39

Smith, T. and Reynolds, R., 2003: Extended reconstruction of global sea surface temperatures based on COADS data (1854-1997). *Journal of Climate*, 16(10):1495–1510.

Smith, T. and Reynolds, R., 2004: Improved extended reconstruction of SST (1854-1997). *Journal of Climate*, 17(12):2466–2477.

Smith, T. M., Reynolds, R. W., Peterson, T. C. and Lawrimore, J., 2008: Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006). *Journal of Climate*, 21(10):2283–2296.

Therneau, T. M., Grambsch, P. M. and Fleming, T. R., 1990: Martingale-based residuals for survival models. *Biometrika*, 77(1):147–160.

Wilks, D., 2006: *Statistical Methods in the Atmospheric Sciences.* Academic Press Inc., 2nd ed., Amsterdam, 648 pp.

Xue, Y., Smith, T. and Reynolds, R., 2003: Interdecadal changes of 30-yr SST normals during 1871-2000. *Journal of Climate*, 16(10):1601–1612.

Tables

Table 1: NatCat*SERVICE* natural catastrophe ranking into seven categories (courtesy of Munich RE).

1	**Natural event**	**No property damage (e.g. forest fire with no damage to buildings)**					
1	**Small-scale loss event**	**1-9 fatalities and/or hardly any damage**					
2	**Moderate loss event**	**10-19 fatalities and/or damage to buildings and other property**					
					2000-2005	**1990ies**	**1980ies**
3	**Severe catastrophe**	**20+ fatalities**	**Overall losses**	**US$**	**> 50m**	**> 40m**	**> 25m**
4	**Major catastrophe**	**100+ fatalities**	**Overall losses**	**US$**	**> 200m**	**> 160m**	**> 85m**
5	**Devastating catastrophe**	**500+ fatalities**	**Overall losses**	**US$**	**> 500m**	**>400m**	**> 275m**
6	**Great natural catastrophe "GREAT disaster"**	**Thousands of fatalities, economy severely affected, extreme insured losses (UN definition)**					

Figure captions

Fig. 1: Geographical overview on the Niño areas considered and the investigation area (IA). Niño4: black solid, Niño3.4: gray dashed, Niño3: gray solid.

Fig. 2: Recurrence frequency for extreme precipitation as a composite for November to February in the US-MEX dataset. The investigation area is marked by a black rectangle.

Fig. 3: Distribution of the damage index D inside the investigation area on the four seasons (DJF, MAM, JJA, SON) in percent. Underlying are 64 category 3&4 events during 1975-2007.

Fig. 4: Difference in recurrence frequency anomaly between El Niño- and La Niña-months in a composite for November to February. The investigation area is marked by a black rectangle.

Fig. 5: Correlation coefficient r between SST anomaly in a given month and extreme precipitation frequency in the following winter (NDJF) in the investigation area (IA): (a) November, 0 months lead-time; (b) June, 5 months lead-time; and (c) January, 10 months lead-time. Coloured areas contain correlations on significance levels greater 90%.

Fig. 6: Explained variance r^2 between winter extreme precipitation count and SST anomalies from ERSST-dataset averaged over Niño4 area at different lead times (1951-2004).

Fig. 7: Scatter plot of observed (crosses) relative winter extreme precipitation frequency versus SST anomaly in the Niño4 area at five months lead time. Regression with the full dataset of 55 yeas (1950/51-2004/05) as training data is also shown (circles).

Fig. 8: Exemplary validation of a Cox regression model, showing the percentage of forecasts correctly above or below the average of extreme precipitation frequency at different lead times.

Figures

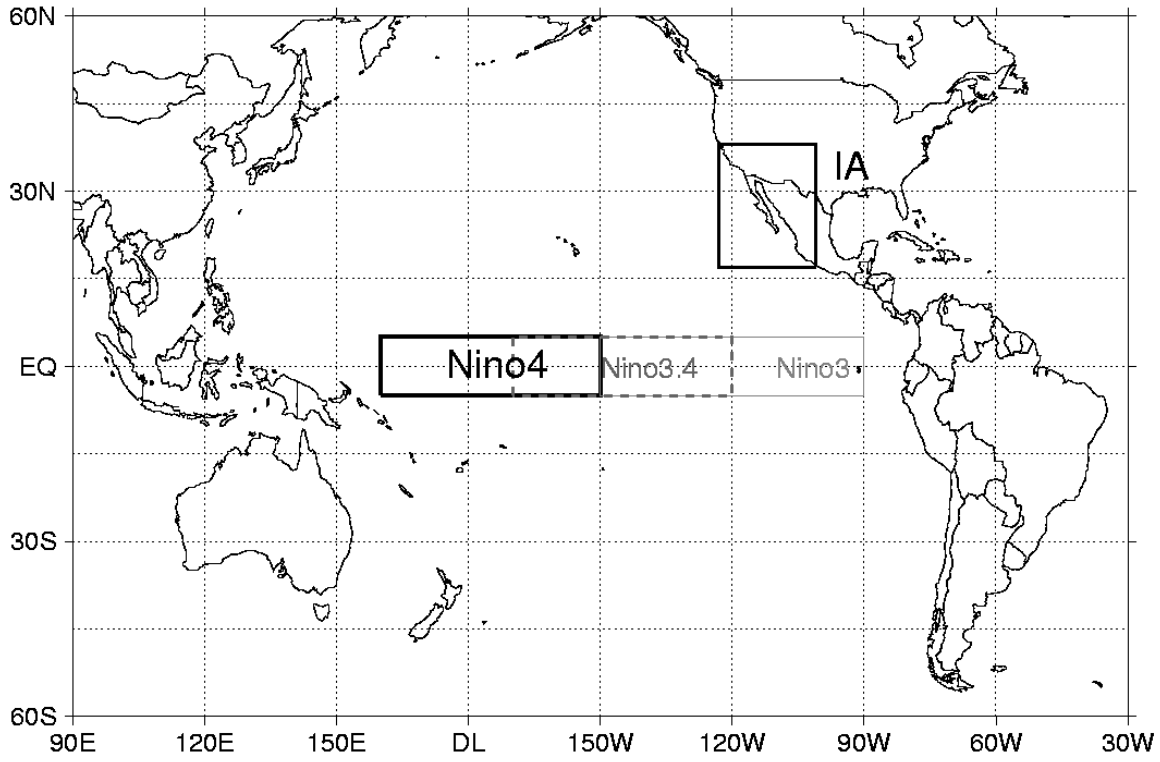


Fig. 1: Geographical overview on the Niño areas considered and the investigation area (IA). Niño4: black solid, Niño3.4: gray dashed, Niño3: gray solid.

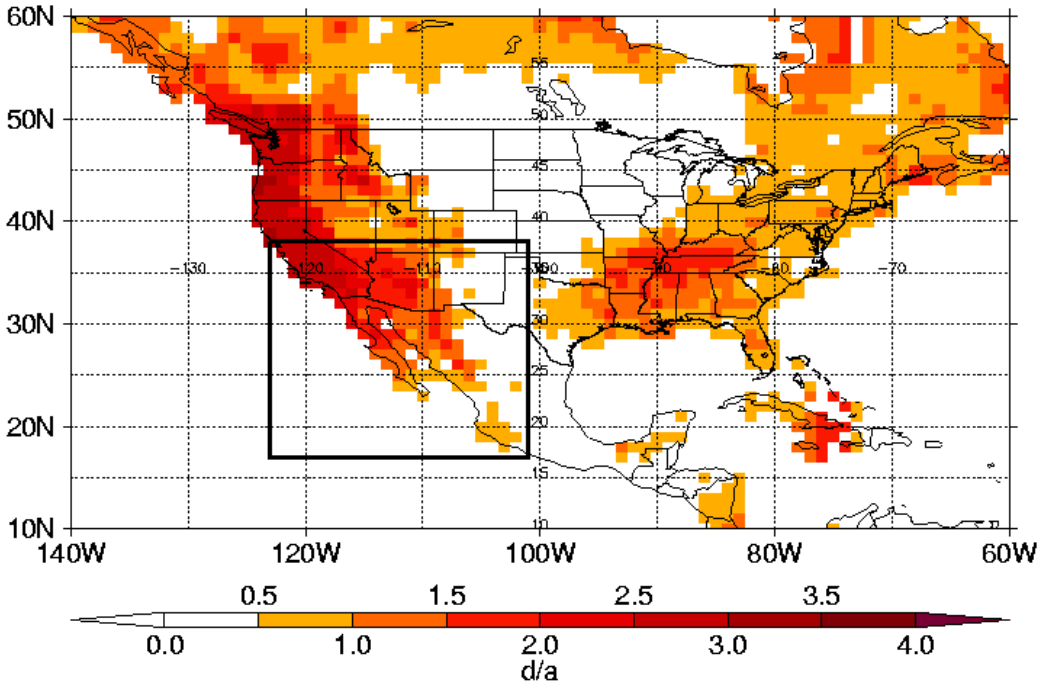


Fig. 2: Recurrence frequency for extreme precipitation as a composite for November to February in the US-MEX dataset. The investigation area is marked by a black rectangle.

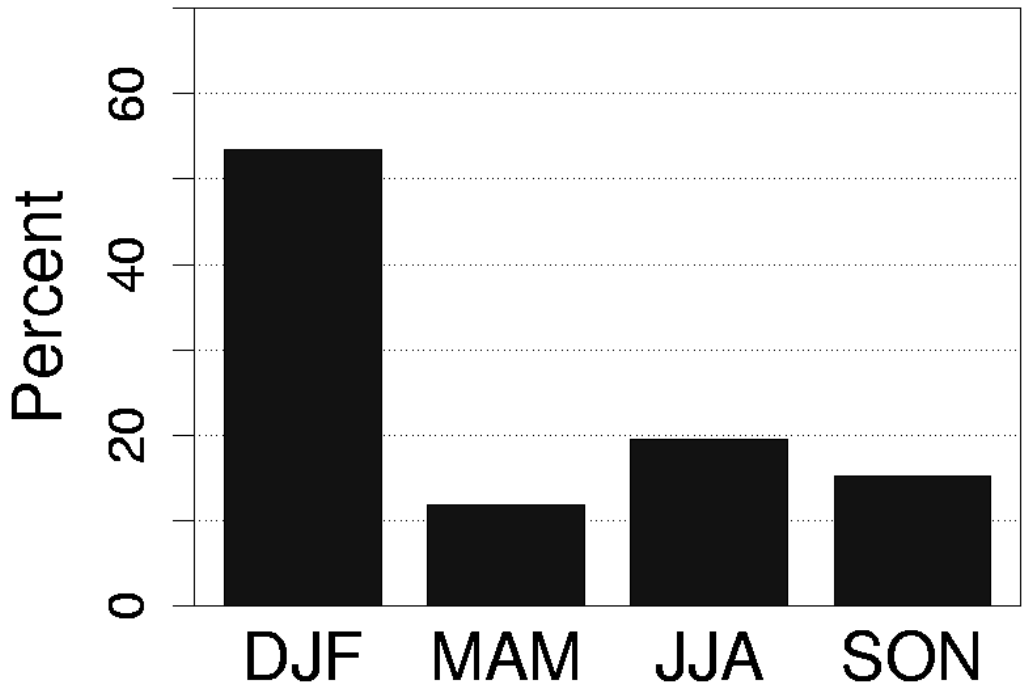


Fig. 3: Distribution of the damage index D inside the investigation area on the four seasons (DJF, MAM, JJA, SON) in percent. Underlying are 64 category 3&4 events during 1975-2007.

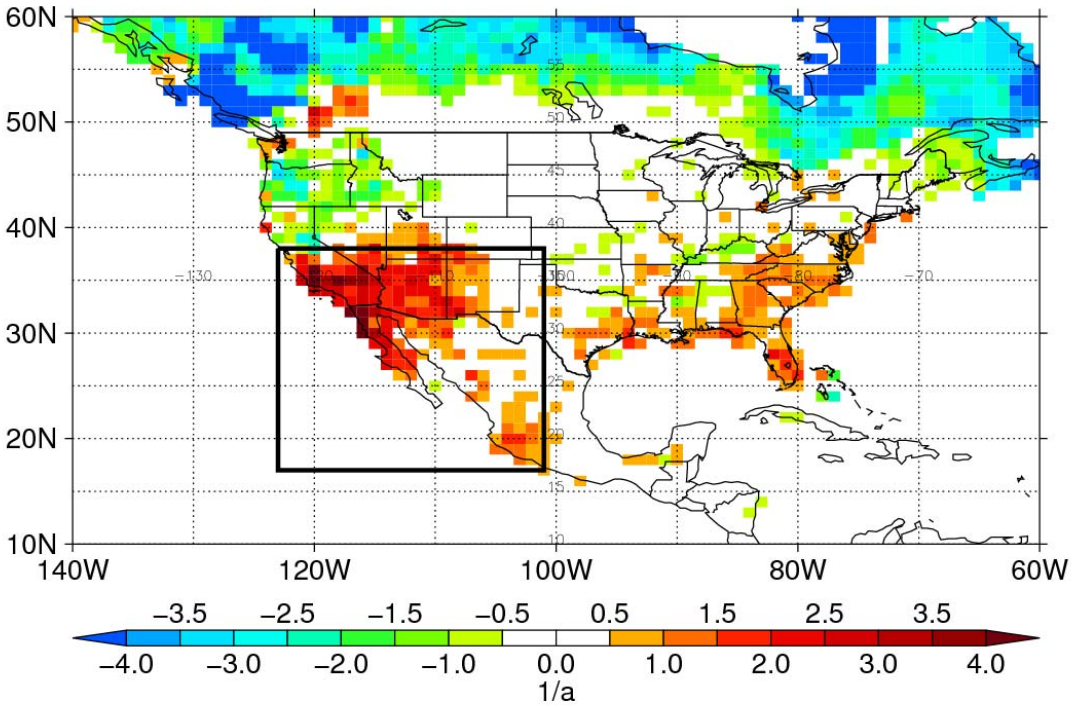
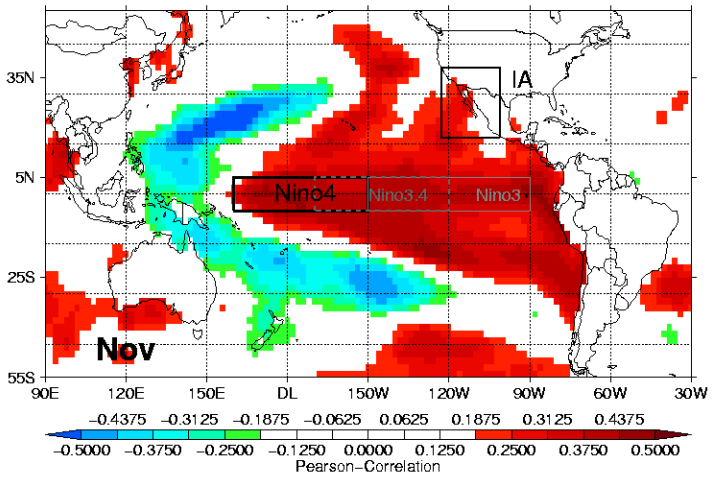
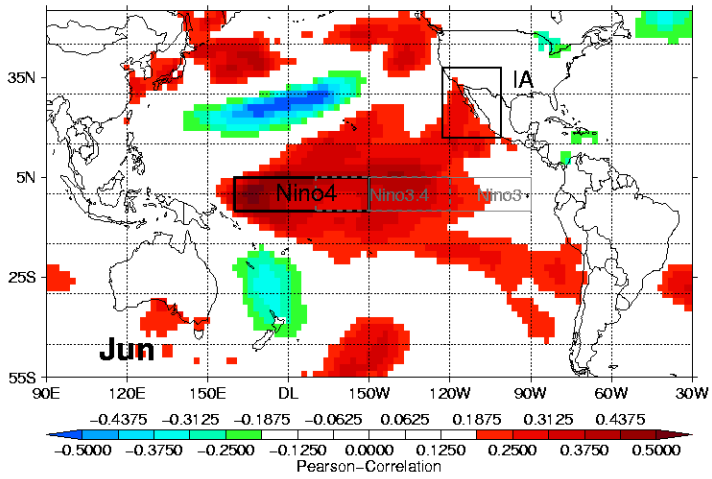


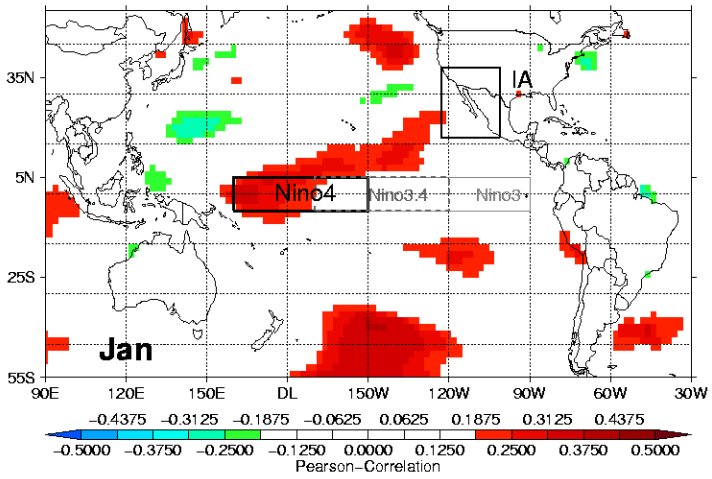
Fig. 4: Difference in recurrence frequency anomaly between El Niño- and La Niña-months in a composite for November to February. The investigation area is marked by a black rectangle.



a)



b)



c)

Fig. 5: Correlation coefficient r between SST anomaly in a given month and extreme precipitation frequency in the following winter (NDJF) in the investigation area (IA): (a) November, 0 months lead-time; (b) June, 5 months lead-time; and (c) January, 10 months lead-time. Coloured areas contain correlations on significance levels greater 90%.

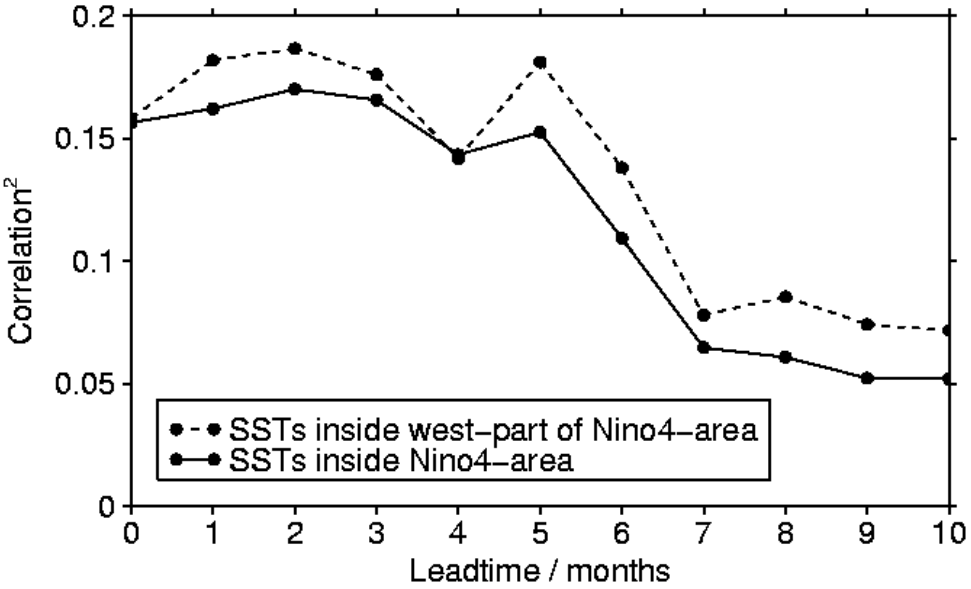


Fig. 6: Explained variance r^2 between winter extreme precipitation count and SST anomalies from ERSST-dataset averaged over Niño4 area at different lead times (1951-2004).

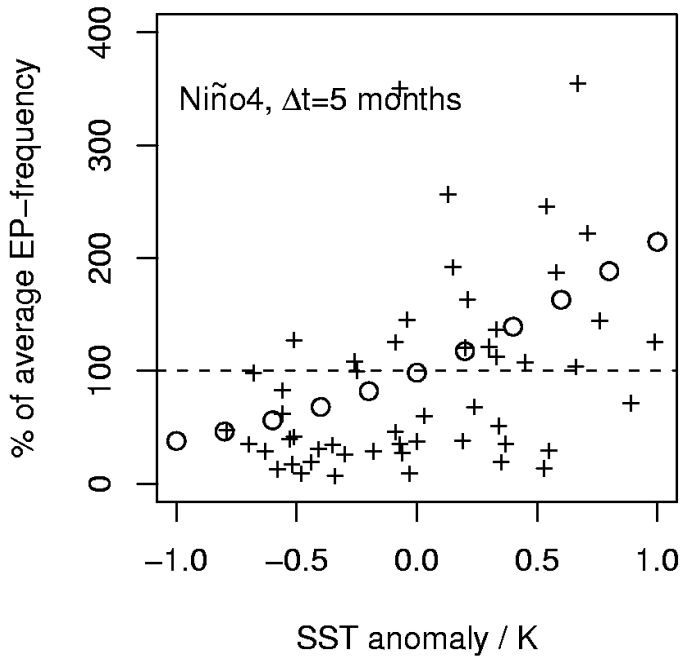


Fig. 7: Scatter plot of observed (crosses) relative winter extreme precipitation frequency versus SST anomaly in the Niño4 area at five months lead time. Regression with the full dataset of 55 yeas (1950/51-2004/05) as training data is also shown (circles).

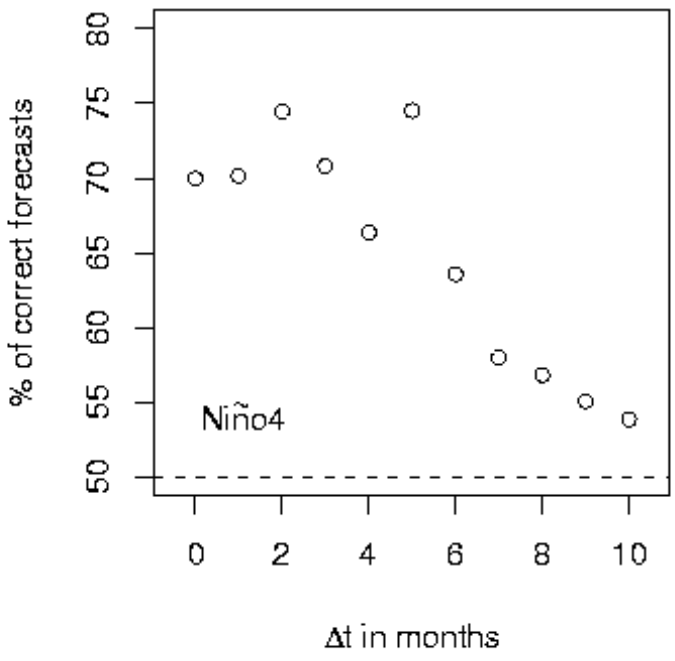


Fig. 8: Exemplary validation of a Cox regression model, showing the percentage of forecasts correctly above or below the average of extreme precipitation frequency at different lead times.