Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA

Federal Office of Meteorology and Climatology MeteoSwiss
Resarch & Cooperation

NCCR Climate II: Scientific Results

1 Weather Risk

1.1 Probabilistic Forecasting: August 2005

Limited-area ensemble prediction systems (LEPSs) combine the advantages of probabilistic forecasts with high-resolution numerical weather prediction. The LEPS of the COSMO consortium, referred to as COSMO-LEPS, regionalize and adapt the global-scale EPS of the European Centre for Medium-Range Weather Forecasts (ECMWF) to the regional scale by nesting limited-area forecasts in the global ECMWF EPS members (Montani et al., 2003). COSMO-LEPS is running on a daily basis since November 2002 (fully operational since November 2005). The task "reforecasting extreme events (REXE)" of the NCCR PRECLIM project analyzes COSMO-LEPS forecasts with regard to the climatologic characteristics and investigates the predictability of extreme events.

Probabilistic Forecast of August 2005 Flooding

One example of such an extreme event was the heavy precipitation in August 2005 that caused tremendous floods in Switzerland and in adjacent neighbourhood countries. Subsequently, the COSMO-LEPS forecast for this event is discussed and compared with the corresponding ECMWF EPS forecast.

The heavy precipitation started on 20 August. Within 72 hours the accumulated precipitation exceeded 100 mm in a large area from the western Alps to the north eastern Alpine foreland (see Fig. 1). In central Switzerland more than 150 mm were observed and at some locations even more than 300 mm.

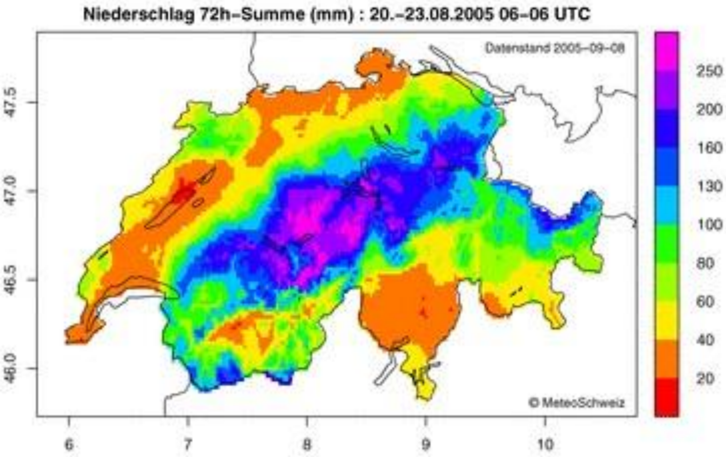


Figure 1: Precipitation sum for 20 August 06 UTC - 23 August 06 UTC based on the Swiss observation network with about 400 stations and a high-resolution precipitation climatology (Courtesy of C. Frei, MeteoSwiss).

Figure 2 (left) shows the COSMO-LEPS forecast from 19 August 2005 12 UTC for the 72-h precipitation sum between 20 August 06 UTC and 23 August 06 UTC (lead-time 18-90h). The four panels show the probabilities to exceed the corresponding thresholds 50, 100, 150 and 250 mm, respectively. The panels show for the entire northern Alpine slopes high probabilities (up to 80%) for accumulated precipitation over 100 mm and high probabilities (over 60%) for precipitation above 150 mm particularly in the Bernese and central Alps. In addition, a scenario with accumulated precipitation over 250 mm for some locations in these regions is predicted with probabilities of 20-30%. The panel for the 100 mm threshold points out a high correlation of regions showing high predicted probabilities with those regions showing precipitation higher than 100 mm. Furthermore, the locations with observed precipitation above 250 mm are in that region where the predicted probabilities are highest for this threshold. Figure 2 (right) presents the corresponding forecast from the ECMWF EPS. The panels indicate only for the 50 mm/72h threshold and only for the eastern Swiss Alps considerable probabilities.

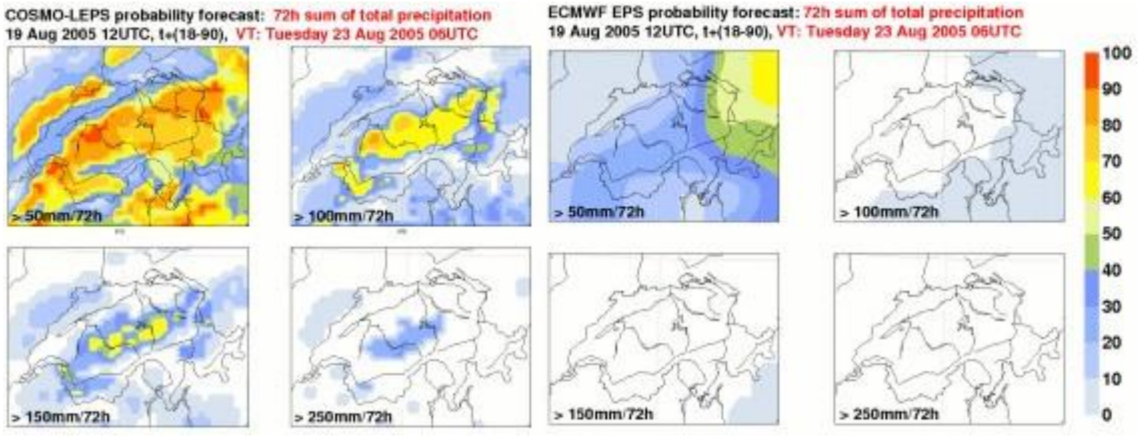


Figure 2: Forecast from 19 Aug 2005 12 UTC for 72-h precipitation sum from (left) COSMO-LEPS and (right) ECMWF EPS. The panels show the probabilities to exceed the given thresholds 50, 100, 150 and 250 mm/72h, respectively, for the period 20 Aug 06 UTC to 23 Aug 06 UTC.

In summary, COSMO-LEPS forecasts provided very appropriate warnings for the extreme precipitation event in August 2005. The system predicted high probabilities for large precipitation amounts for most of the regions hit by the event without giving obvious false alerts for other regions. In order to evaluate the skill of COSMO-LEPS for extreme events objectively, a large sample of reforecasts will be performed and investigated within this project.

References

Montani, A., M. Capaldo, D. Cesari, C. Marsigli, U. Modigliani, F. Nerozzi, T. Paccagnella, P. Patruno, S. Tibaldi, 2003b: Operational limited-area ensemble forecasts based on the Lokal Modell. ECMWF Newsletter Summer 2003, 98, 2-7.

For further information, please contact: André Walser.

1.2 Reforecasting Extreme Events (REXE)

In the framework of this project a warning product for extreme weather events is developed. The product is based on the COSMO LEPS mesoscale ensemble system as this model has shown to be reliable in forecasting extreme events and, as an ensemble model, offers the opportunity to take advantage of its probabilistic nature. Further on the COSMO-LEPS model combines a long forecast time of 132 hours with a high horizontal resolution (10km) which is of importance for regions of complex orography as e.g. Switzerland

To be able to give a warning of extreme events it is necessary to gain knowledge on what to call extreme. A common approach for extreme events is the definition of a physical threshold. A warning will be given when this threshold is crossed. However, this method potentially implies false alarms. Firstly the same threshold might not be applicable for all regions and each season (e.g. a heavy precipitation event at one location might not be remarkable at another place) and secondly the model itself might be biased in a certain direction and permanently under/overestimate a parameter. To work around these difficulties, it is advisable to compare a model forecast to the underlying climatology of the model. Ensemble prediction systems are especially suitable to that approach, as they offer the possibility to directly compare the probability distribution of the forecast to that of the underlying model climatology.

An extreme forecast index (EFI) addressing this problem has been proposed by F. Lalaurette, 2002. Basically it is the weighted difference between the probability distribution of the EPS forecast and the model climatology. In this project an EFI will be used for the COSMO-LEPS, combining the advantages of the model?s high resolution and long forecast time. The rough procedure to achieve the EFI as basis for a warning product is given in Fig. 1.

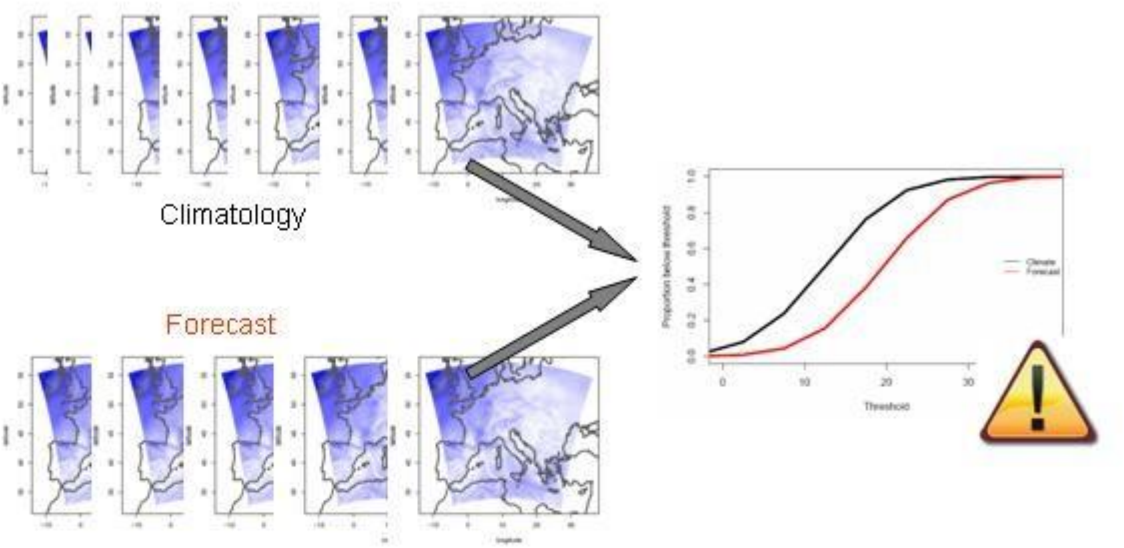


Figure 1: To achieve a warning index for extreme events, in a first step, the model climatology has to be computed. After that the actual probability distribution of the LEPS forecast is compared to the model probability distribution. Depending on the difference between the two distributions a warning can be given.

As the occurrence of extreme events is rather small, a large set of hindcasts is necessary to establish a representative model climatology. In the framework of REXE, re-forecasts of a past period of around 30 years will be calculated at the ECMWF. This should provide a sufficiently large data base to achieve a good statistic, even for rare extreme events.

In a first step the warning index will be developed for extreme precipitation and wind speeds. The extension on further parameters is planned.

NCCR Climate II.docx

References

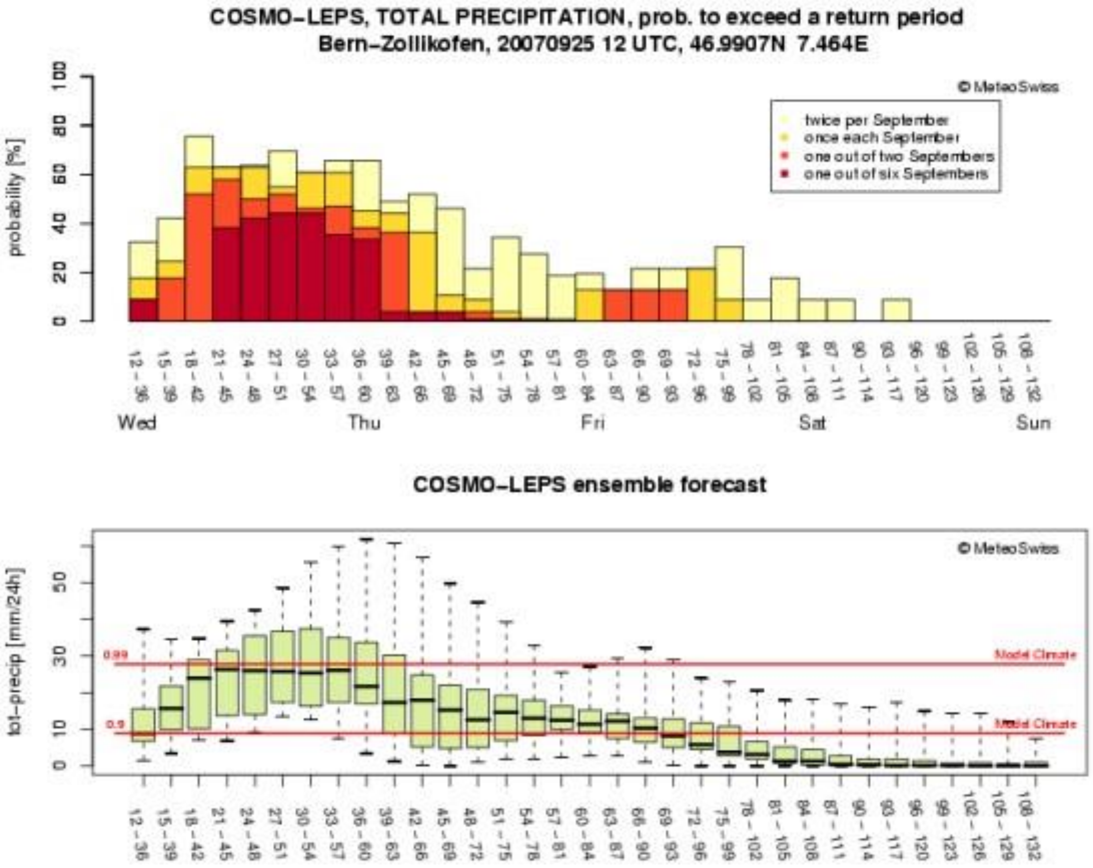
Montani, A., M. Capaldo, D. Cesari, C. Marsigli, U. Modigliani, F. Nerozzi, T. Paccagnella, P. Patruno, S. Tibaldi, 2003b: Operational limited-area ensemble forecasts based on the Lokal Modell. ECMWF Newsletter Summer 2003, 98, 2-7.
Lalaurette, F., 2002, Early Detection of Abnormal Weather Using a probabilistic Extreme Forecast Index. ECMWF Technical Memorandum, 373.

For further questions please contact Mark Liniger.

1.3 Probabilistic reliable warning products

Quasi-operational warning products for rare weather situations such as heavy precipitation and strong wind are developed, using the limited-area ensemble prediction system COSMO-LEPS.

Simultaneous to the forecasts of the COSMO limited area ensemble prediction system (LEPS) a set of reforecasts from 1971 to 2000 is created in real time, forming an extensive model climatology. The COSMO-LEPS forecasts are calibrated with this model climatology with the effect of correcting for local, seasonal depending systematic model biases and errors. Here, a monthly subset of the climatology (+-14 days around the actual forecast date) is used to estimate the frequency of a forecasted event within the model climatology (see figure below). The 16 forecast members allow for a probabilistic forecast of the frequency of events, shown in the figure for 24h rainfall near Bern. With such a calibration it is possible to give warnings in form of a Probability to exceed a given Return Period, referred to as PRP index.



In the upper part, the probability of occurrence and return period of total precipitation is shown relative to the September climate. The colors of the bars indicate the return periods (from events occurring twice per September to events occurring only every sixth September). The height of the bars gives the probability of occurrence. The lower part shows the raw COSMO-LEPS 24h rainfall sums, gliding in steps of 3 hours. The boxes envelop the 25% and 27% quantile, the black bar marks the median and the whiskers show the range of the ensemble forecast. The red lines show the 99% and 90% quantile of the COSMO-LEPS model climatology (1971-2000) in September.

For further information, please contact: Mark Liniger.

1.4 Calibrated COSMO-LEPS ensemble forecasts

Numerical weather forecasts are often subject to systematic errors, so called model biases. This can e.g. manifest in a surplus or deficiency in rainfall or a tendency for too high or low temperatures. Usually, those biases have a spatial and temporal structure. The bias in high altitudes might differ from the bias in the lowland. Here, a bias corrected forecast product derived from the Consortium for Small scale Modelling Limited area Ensemble Prediction System (COSMO-LEPS) is shown.

The temperature bias of COSMO-LEPS in its current model version was found by comparing a set of 20 years hindcasts with observations from the Swiss automated observational network (ANETZ). The found relation of past measured temperature and past forecasted temperature was than used to correct the actual forecast. An example forecast for Zurich is given in Figure 1. It shows the ensemble forecast as boxplot. Additional information about the typical climate for this season is given by the shaded areas in the background. They show different temperature quantiles as observed at this station in order to give the user an idea of how typical the current forecast is with respect to the long-term observations. Further on, to provide information about the current model performance, actual observations are added to the plot as soon as they are available. Also the forecast of the higher resolution deterministic model COSMO-7 is added.

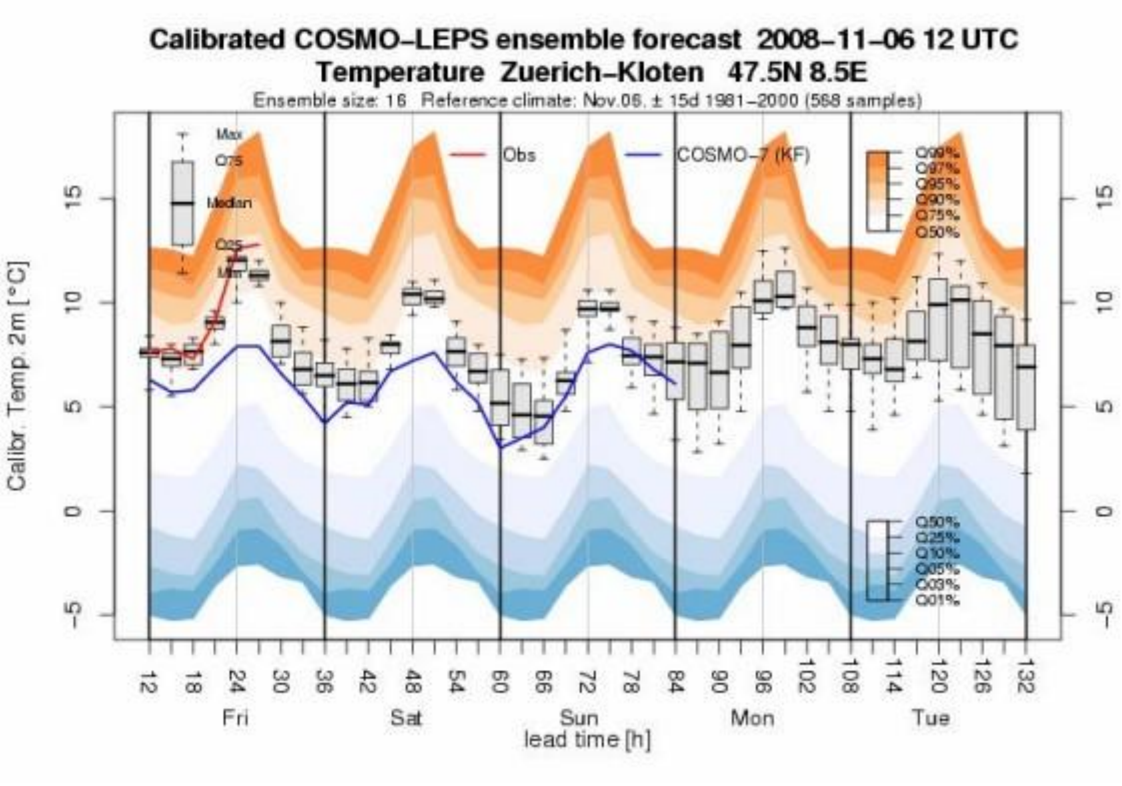


Figure 1: Exemplary COSMO-LEPS 5 days temperature forecast for Zuerich-Kloten including actual observations and the COSMO-7 forecast.

A verification of the usefulness of this product revealed a strong skill improvement in comparison to the uncalibrated COSMO-LEPS output. In case of temperature, a calibration was most effective at high altitude stations during winter months. Also for other seasons and locations significant improvements could be achieved. The all over skill improvement is shown in Figure 2 in form of the ranked probability skill score (RPSS) versus lead time.

NCCR Climate II.docx

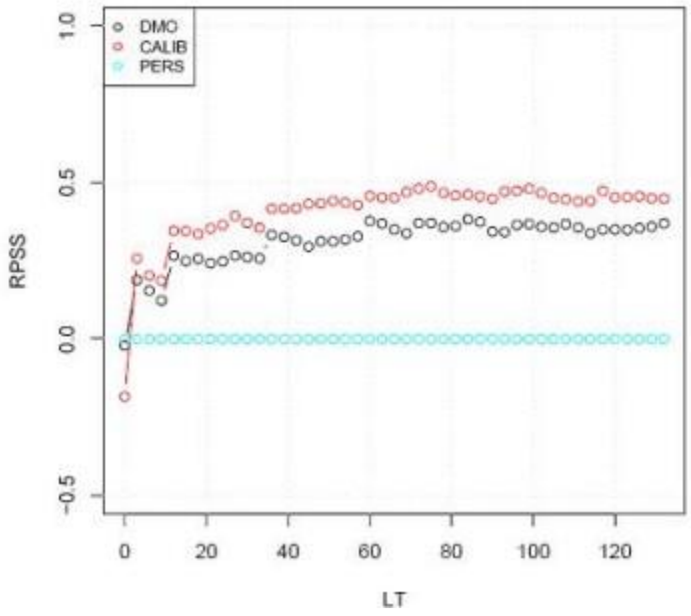


Figure 2: COSMO-LEPS forecast skill for temperature (raw in black and calibrated in red) for the Period Oct. 2006- Nov. 2007. The reference is a presistence forecast (blue) that has no skill by definition.

For further information, please contact Mark Liniger.

NCCR Climate II.docx

2 Climate Risk

2.1 Verification of Ensemble Forecasts

The discrete Brier and ranked probability skill scores

Probabilistic forecasts with ensemble prediction systems (EPS) have found a wide range of applications in weather and climate risk management. The rationale behind the ensemble method used is to approximate the expected probability density function (PDF) of the predictand by a finite set of forecast realizations. However, while probabilistic forecasts contain more information than deterministic ones, it is not trivial to verify them such that the full information content is considered. The problem mainly arises from the fact that probabilities need to be compared to real-valued observations, which, roughly speaking, resembles comparing apples with oranges.

Several approaches have been developed to tackle this problem, but there is no general agreement on which one is the best. A common method is the conversion of the verifying observations into PDFs over a given number of outcome categories (e.g. "too cold", "normal", "too warm"). This is done by assigning the value "1" to the category the observation falls into, and "0" to the other categories. Based on this approach, multi-category scores comparing forecast PDFs to observation PDFs can be defined. Examples are the ranked probability score (RPS), which measures shape and central tendency of the forecast PDF with respect to the observation, and the Brier Score (BS), which is the special case of an RPS for two forecast categories. By relating these scores to a usually simpler reference forecast strategy such as climatological "guessing", so-called **skill scores** (RPSS and BSS, respectively) can be defined. They are formulated such that they assume positive values if the ensemble prediction system outperforms the reference strategy. Zero or even negative indicate no benefit to the reference forecast. A perfect forecast would give a skill score of 1.

The characteristics and properties of the RPSS are investigated in this project. An earlier study has found the RPSS to be negatively biased for small ensemble size, leading to a severe underestimation of predictive skill (Müller et al., 2005). This flaw imposes major problems on the verification of ensemble predictions, especially in the context of seasonal forecasts, where large ensemble sizes are not yet standard.

A new technique is developed to overcome this deficiency. By adequately considering the effects of finite ensemble size in the climatologic reference score, the negative bias can be removed analytically and a "debiased" version of the ranked probability skill score (RPSSD) and Brier skill score (BSSD) can be formulated. This is described in detail in Weigel et al. (2006). As an example, Figure 1 shows a comparison of the classical BSS and the new BSSD-formulation for probabilistic seasonal forecasts over two regions in dependence of ensemble size. It can be seen that the new technique gives reasonable results even for systems with small ensemble sizes. This is of particular importance when multi-model ensembles are evaluated, where the benefit of increasing ensemble size needs to be clearly distinguished from the benefits of multi-model combination.

NCCR Climate II.docx

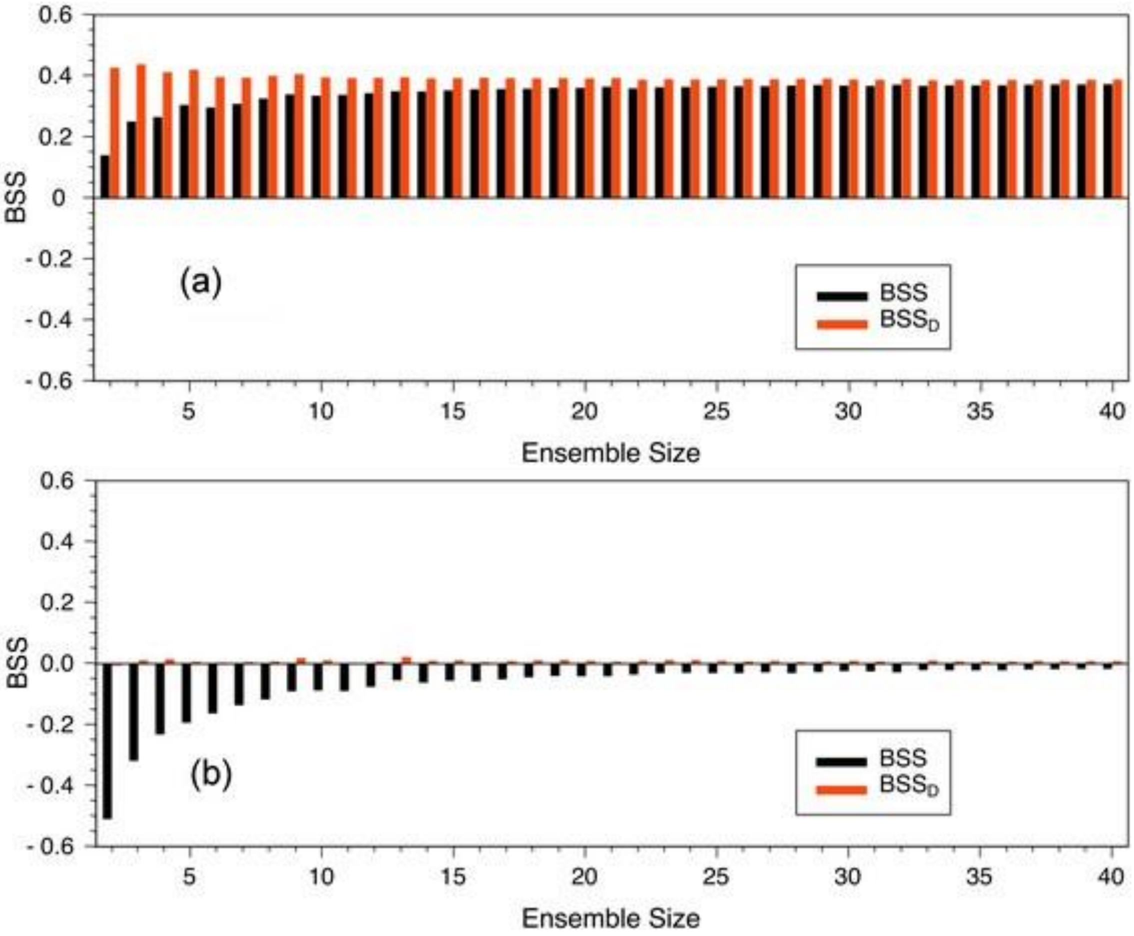


Figure 1. Traditional Brier skill score (BSS, black) and new debiased Brier skill score (BSSD, red) as a function of ensemble size for near-surface temperature predictions for March with a lead time of 4 months. Scores are averaged over 15 years (1988-2002) over (a) the Nino3.4 region in the equatorial Pacific and (b) over Central Europe. The ECMWF System 2 data are verified against ERA40 re-analysis data.

Related references

Müller W. A., C. Appenzeller, F. Doblas-Reyes and M. A. Liniger (2005), A Debiased Ranked Probability Skill Score to Evaluate Probabilisitic Ensemble Forecasts with Small Ensemble Size, J. Clim., 18, 1513-1523

Weigel A. P., M. A. Liniger and C. Appenzeller (2006), The Discrete Brier and Ranked Probability Skill Scores, Mon. Wea. Rev., 2007, 135, 118 -124

For further information, please contact: Mark Liniger.

2.2 Prediction of Winter Storm Risk (PreWiStoR)

European winter wind storms represent one of the major loss potentials for reinsurance companies. Current methods that enable the management of this risk include the coupling of reinsurance loss portfolio models with historical datasets of extreme wind speed. Adequate assessment of risk is currently limited by a lack of high quality wind datasets which extend back in time with sufficient accuracy and spatial coverage. Datasets are typically in the order of 50 to 70 years in length which place limitations on the estimates of wind storm return periods.

Recently, data from dynamical ensemble prediction systems covering seasonal time periods have become available from various forecasting centers. These data are appealing to use in risk assessment since they are based on the dynamics of weather and climate and they provide a direct estimation of the probability density function of a meteorological variable. The meteorological fields derived from a dynamically based ensemble member can be considered as a physically consistent, probable representation of the state of the atmosphere-ocean system. In the project PreWiStoR the dynamical output of seasonal forecasting systems has been linked with a reinsurance loss model to explore the potential predictability of reinsurance losses.

In phase 1 of the PreWiStoR project the seasonal predictability of storminess over the European region has been quantified. The storminess over the European region is summarized by an extreme wind index. This index takes into account the area and magnitude of extreme winds over the European region. When we compare observations of storminess (obtained from ERA-40) with the predictions from s2d models up to 7 months in advance we see that the models only have useful prediction skill in the first month of forecast.

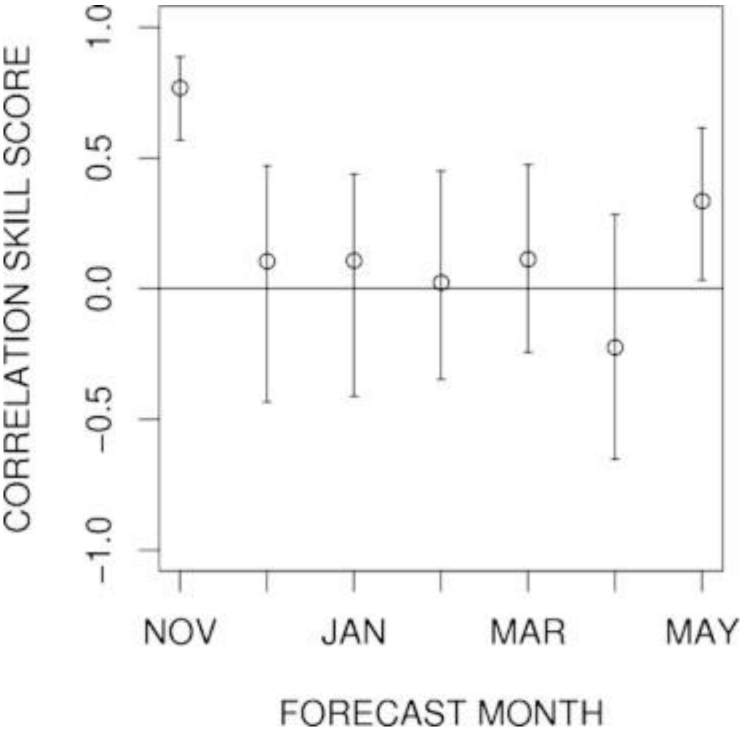


Figure 1: Prediction skill of storminess over the European region as a function of forecast lead time (taken from Della-Marta et al., 2008).

This result encourages the use of monthly time scale climate predictions of European storminess and at the same time justifies the use of s2d models as a random sampe of the climate in risk analysis applications on longer time scales.
In phase 2 of the PreWiStoR project the feasibility of using dynamically based ensemble seasonal to decadal (s2d) climate forecast data to estimate wind storm induced loss has been investigated. A conceptual diagram of this process is shown below.

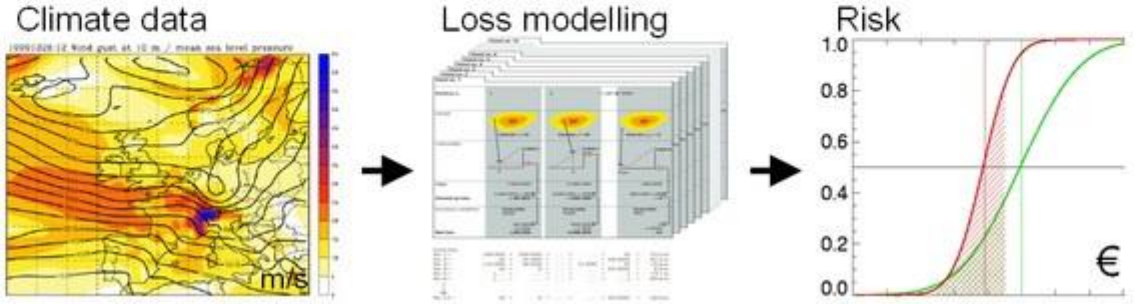


Figure 2: Conceptual diagram of phase one of the PreWiStoR project.

Differences in the magnitude of wind speed between Swiss Re wind storm fields and those identified in ERA-40 and s2d exist due to reanalysis and climate model biases. In order to derive meaningful loss estimates from s2d wind storms requires the application of calibration techniques. When these biases are adequately accounted for the comparison of loss estimates between datasets is possible. An example is shown in the diagram below.

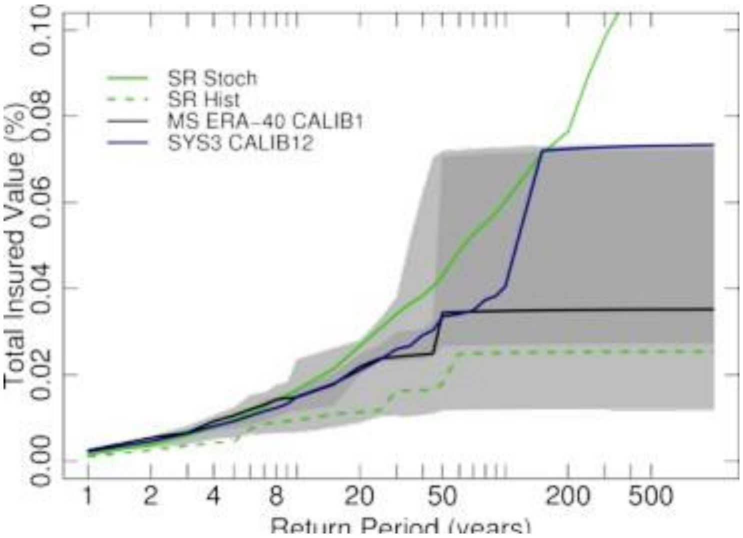


Figure 3: A comparison of loss frequency curves calculated using SwissRe data (green lines), ERA-40 (black line) and an s2d model (blue line) (taken from Della-Marta et al, 2008). The grey shading approximates the uncertainty in the loss estimates based on re-sampling the s2d wind storms.

Figure 3 shows that the uncertainty in loss return periods greater than 50 years is very high when the length of the dataset is in the same order of length. This advocates the use of longer wind storm climatologies such as those based on s2d data which are typically greater than 300 years in length.

Phase 3 of the PreWiStoR project investigates the covariation of wind storm frequency, intensity and loss over Europe with large-scale climate diagnostics. Preliminary results show that both wind storm magnitude and frequency are modulated by large-scale atmospheric flow conditions. Important modulating factors of the intensity of wind storms in the North Atlantic region are found in indices of atmospheric water vapour content.

The PreWiStoR project has been funded by SwissRe through the SNF NCCR Climate and by the EU FP6 project ENSEMBLES.

For further information, please contact: Mark Liniger.

References:

Della-Marta, P. M.; Liniger, M. A.; Appenzeller, C.; Bresch, D. N.; Köllner-Heck, P. & Muccione, V. 2010. Improved estimates of the European winter wind storm climate and the risk of reinsurance loss using climate models. Journal of Applied Meteorology and Climatology, 49, 2092-2120.

NCCR Climate II.docx

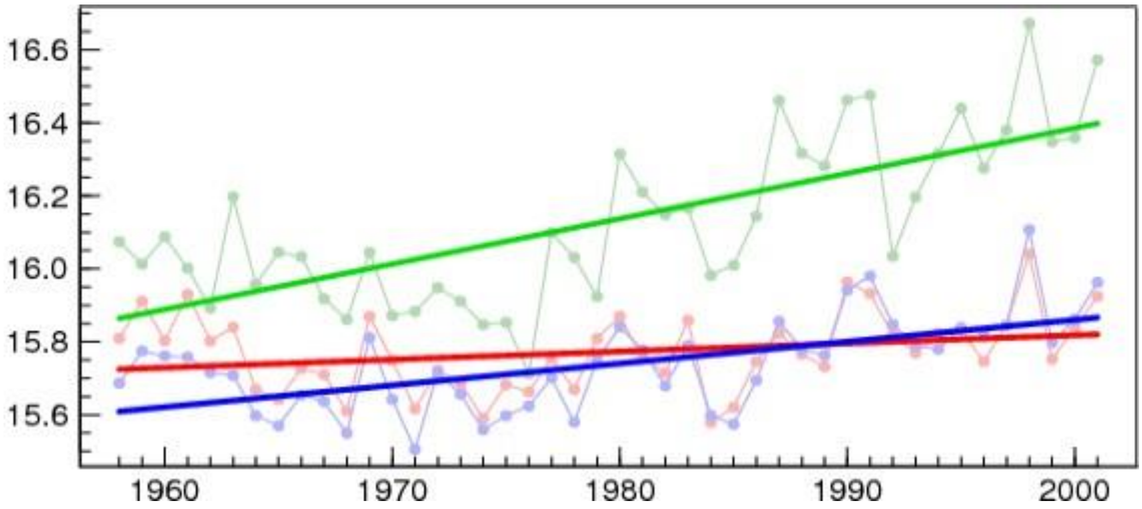
2.3 Realistic Greenhouse Gas Forcing and Seasonal Forecasts

Atmospheric greenhouse gas (GHG) concentrations are a key driver for decadal to centennial climate variability and climate change. For short term climate predictions, e.g. seasonal forecasts, the impact of initial conditions has been thought to dominate over the increasing concentration of GHG in the atmosphere.

The improvement of seasonal forecasts by including realistically varying GHGs was investigated in this study. Forecasts starting every May and November were compared over the period 1958 until 2001. One set has constant GHG concentrations while another one has a realistic GHG trend. The large scale temperature trends derived at different lead times are compared in between the forecast sets and observations over the entire 44 years. An example for the July global mean temperature is shown in the figure. It can be ascertained that the temperature trend of the model with increasing GHG concentrations is closer to the observation.

It has been found that after a few months the anthropogenic climate change signal is lost up to 70% although it was present in the initial conditions. The differences in trends vary with lead times, seasons and regions. Strongest effects are found in the Tropics and the Summer Hemispheres, in particular the Northern one.

The local improvement in trends is strongest in the tropics, in particular the Atlantic and for late boreal summer. This improvement in trends is not widespread and very weak in predicting detrended interannual variability. Thus, the seasonal forecasts mainly benefit from more realistic model climatology and not from an improvement in the actual predictability.



Global mean temperature for July (incl. linear trend) for 1958 - 2001 in °C. The green line stands for the quasi observation, i.e. the ERA40-data. Red and blue denote predictions corresponding to a 2-month lead. Red denotes the data set with the constant GHG concentration while blue stands for the data set with yearly updated GHG concentration.
Global mean temperature for July (incl. linear trend) for 1958 - 2001 in °C. The green line stands for the quasi observation, i.e. the ERA40-data. Red and blue denote predictions corresponding to a 2-month lead. Red denotes the data set with the constant GHG concentration while blue stands for the data set with yearly updated GHG concentration.

Related References:

Liniger, M. A., H. Mathis, C. Appenzeller and F. J. Doblas-Reyes (2007), Realistic Greenhouse Gas Forcing and Seasonal Forecasts,. Geophys. Res. Let., 34, L04705

Mathis, H. (2005), Impact of realistic greenhouse gas forcing on seasonal forecast performance, Veröffentlichungen der MeteoSchweiz, Vol. 72, 79 pp., Zürich, Switzerland.

Veroeff_72.pdf, 11.9 MB

For further information, please contact: Mark A. Liniger

2.4 Statistical Prediction of Western European Summer Temperature

Guidance from the statistical prediction method described here forms one input to our official summer temperature forecast for Switzerland.

Details of the statistical prediction technique

The idea of this statistical method is to use winter and spring precipitation in Southern Europe to predict summer temperature over western Europe. There is scientific evidence presented in Della-Marta et al. (2007) that shows there is useful information in the total precipitation from January till April in the Mediterranean Basin and just north of the Alps to predict summer temperatures (Figure 1a). An additional source of information for the model is the winter sea surface temperatures in the North Atlantic. The physical basis for such a connection can be found in Della-Marta et al. (2007) and the references contained therein.

The prediction scheme is based on a statistical procedure known as canonical correlation analysis. In essence, we take the strength of the pattern in January to April precipitation (Figure 1a) together with the strength of the predictor pattern of December to February sea surface temperatures (Figure 1b) that, in the past, have tended to be associated with warmer than normal summer western European temperature in summer (Figure 1c)

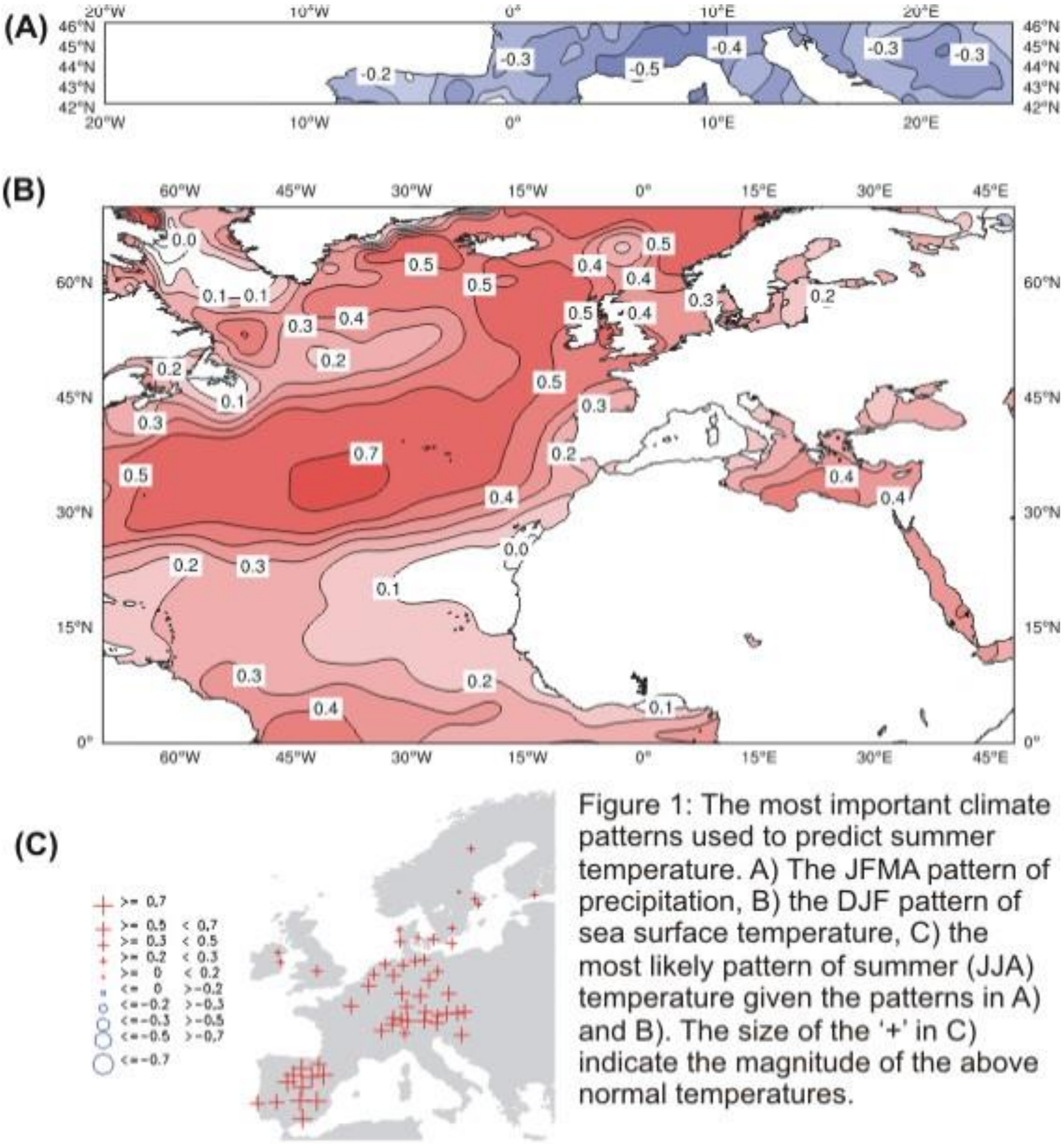


Figure 1: The most important climate patterns used to predict summer temperature. A) The JFMA pattern of precipitation, B) the DJF pattern of sea surface temperature, C) the most likely pattern of summer (JJA) temperature given the patterns in A) and B). The size of the '+' in C) indicate the magnitude of the above normal temperatures.

.

These patterns for the forecast for summer 2008 appear in Figure 2. Slightly drier than normal conditions combined with warmer than normal sea surface temperatures lead to the prediction of warmer than normal temperatures in Switzerland during 2008.

The official MeteoSwiss seasonal forecast.

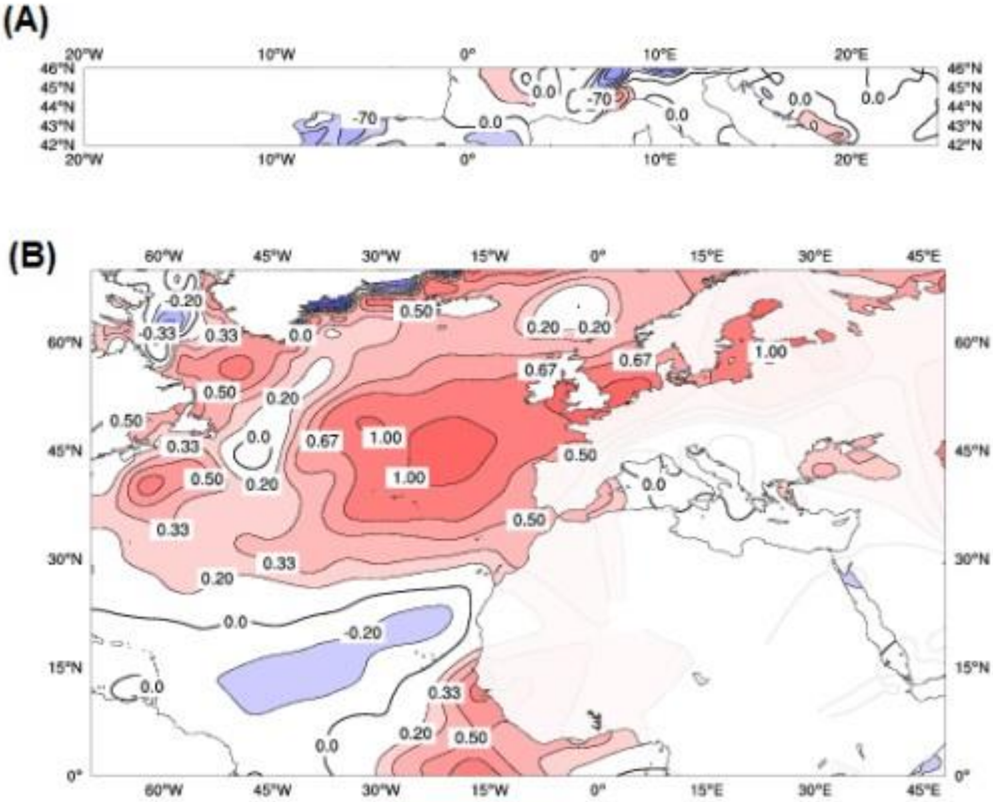


Figure 2: The climate patterns used to predict 2008 summer temperature. (A) the JFMA pattern of precipitation (mm), (B) the DJF pattern of sea surface temperature (degrees Celcius).

References

Della-Marta, P.M. ; Luterbacher, J.; von Weissenfluh, H.; Xoplaki, E.; Brunet, M.; Wanner, H. 2007. Summer heat waves over western Europe 1880-2003, their relationship to large scale forcings and predictability. Climate Dynamics, doi: 10.1007/s00382-007-0233-1.

2.5 Improving prediction skill by multi-model combination

Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?

Multi-model ensemble combination has become a standard technique to improve ensemble forecasts on essentially all time-scales. While the success of multi-model combination has been demonstrated in many studies, the underlying mechanisms have so far not been properly understood. Indeed, given that a multi-model contains information of all participating models, including the less skilful ones, the question remains as to why, and under which conditions, a multi-model can outperform the best participating single model. How can it be that a good forecast can be further improved by adding a poorer model?

In a recent study we have resolved this paradox, applying both synthetic toy model forecasts as well as real seasonal multi-model forecasts (Weigel et al 2008).

The toy model has been designed to allow the generation of perfectly calibrated single model ensembles of any size and skill. Additionally, the degree of ensemble underdispersion (or overconfidence) could be prescribed. Multi-model ensembles were then constructed from both weighted and unweighted averages of these single model ensembles. Applying this toy model, systematic model-combination experiments were carried out (Fig. 1). We evaluated how multi-model performance depends on skill and overconfidence of the participating single models. It turned out that multi-model ensembles can indeed locally outperform a “best model approach”, but only if the single model ensembles are overconfident (black line in Fig. 1). The reason is that multi-model combination reduces overconfidence, i.e. that ensemble spread is widened while the average ensemble mean error is reduced. This implies a net gain in prediction skill, because probabilistic skill scores penalize overconfidence. Under these conditions, even the addition of an objectively poorer model can improve multi-model skill. If the forecast ensembles are totally reliable from the beginning, multi-model combination does not further improve prediction skill (grey line in Fig. 1).

Using seasonal near-surface temperature forecasts from the DEMETER data-set, we showed that the conclusions drawn from the toy model experiments equally hold in a real multi-model ensemble prediction system.

FIGURE 1

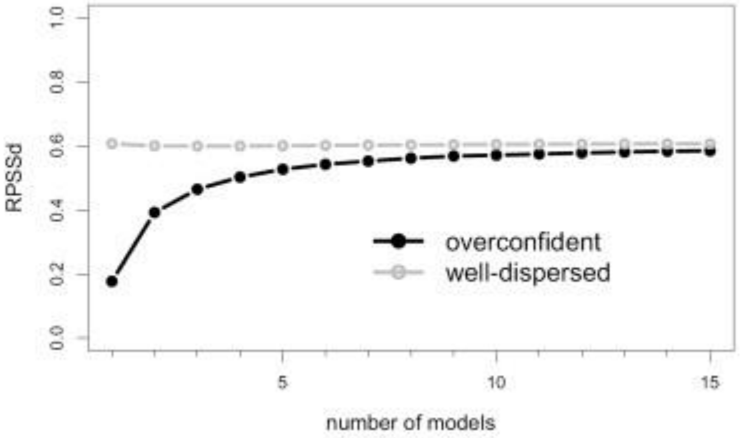


Figure 1. Skill of multi-model ensemble forecasts as a function of the number of participating single model ensembles. The grey line is for well-calibrated reliable ensembles and the black line for highly overconfident ensembles. Only in the latter case does model-combination truly enhance prediction skill.

Related references:

Weigel A. P., M. A. Liniger and C. Appenzeller (2007), Generalization of the discrete Brier and ranked probability skill scores for weighted multi-model ensemble forecasts. Mon. Wea. Rev., 135, 2778-2785
Weigel A. P., M. A. Liniger and C. Appenzeller (2008), Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? Quart. J. Roy. Met. Soc., 134, 241-260

For further information, please contact: Mark Liniger.

NCCR Climate II.docx

2.6 Monthly forecasts

Forecasts within the range of 1 to 4 weeks are important to many applications in the context of weather and climate risk management. The prediction skill of such monthly forecasting system (of ECMWF) in dependency of lead time, sea-son and geographical location has been determined. The results show that monthly forecasts are useful at lead times of 2 to 4 weeks, depending on the region and season considered (Figure 1). Forecasts over the sea remain skilful longer than forecasts over land. This is most notably in the ENSO region and the central Atlantic, where the skill is up to 50% improved against a climate forecast even at a lead time of four weeks. There are also land areas e.g. tropical South America where the model retains significant skill in the fourth week. However, for most extra-tropical landmasses the skill drops sharply after one week (Baggenstos, 2007, Weigel et al. 2008b).

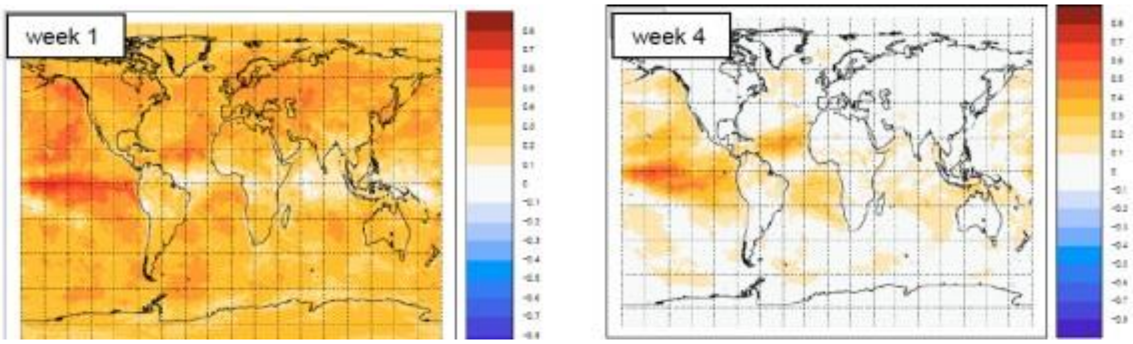


Figure 1) Skill of the ECWMF monthly forecast system for weekly mean near surface temperature for 1 (left) and 4 (right) week lead time.

References:

Weigel, A. P., Baggenstos, D., Liniger, M. A., Vitart, F. and Appenzeller, C. 2008b: Probabilistic verification of monthly temperature forecasts. Mon. Wea. Rev. 136 5162-5182.
Baggenstos, D, 2007: Probabilistic verification of operational monthly temperature forecasts, Veröffentlichung der MeteoSchweiz, Vol. 76, 52 pp, Zürich, Switzerland

For further information, please contact: Mark Liniger

NCCR Climate II.docx

3 Statistical Analysis

3.1 ENSEMBLES WP5.1

Development of daily high-resolution gridded observation datasets for Europe

Introduction

Within the framework of the EU-project ENSEMBLES, the work package 5.1 aims to assemble a daily high-resolution gridded observational dataset for Europe. The underlying observational time series comprehend daily values for temperature (min, max and mean), precipitation, air pressure and snow cover covering at least 45 years. As observational errors and inhomogeneities in time series may have a serious impact on the gridded fields, a special focus must be given on quality control and homogeneity testing.

WP5.1 is under the lead of Albert Klein Tank (KNMI) and the participants are: KNMI (Albert Klein Tank, Lisette Klok), MeteoSwiss (Evelyn Zenklusen, Michael Begert), University of East Anglia (Malcolm Haylock, Phil Jones) and University of Oxford (Mark New, Nynke Hofstra). The contribution of MeteoSwiss focuses on an automated process to detect shift inhomogeneities in the time series and will be described consecutively.

Automated homogeneity testing

The procedure developed for relative homogeneity testing in the framework of ENSEMBLES combines a fully automated method to build a q-series (series of differences between a candidate and several neighbouring reference stations) with a statistical homogeneity test:

1. Vienna Enhanced Resolution Analysis Quality Control (VERAQC, Steinacker et al., 2000) is used to build the q-series and
2. Alexandersson's standard normal homogeneity test (abbr.: SNHT, Alexandersson, 1986) is applied on the q-series to detect the inhomogeneities.

This homogeneity test procedure is henceforth called VERHOM (VERAQC relative homogeneity test method). Figure 1 shows the results of VERHOM applied to the ENSEMBLES observational mean temperature dataset. Most of the series contain 0 or 1 inhomogeneity and very few have 4 or more break points.

NCCR Climate II.docx

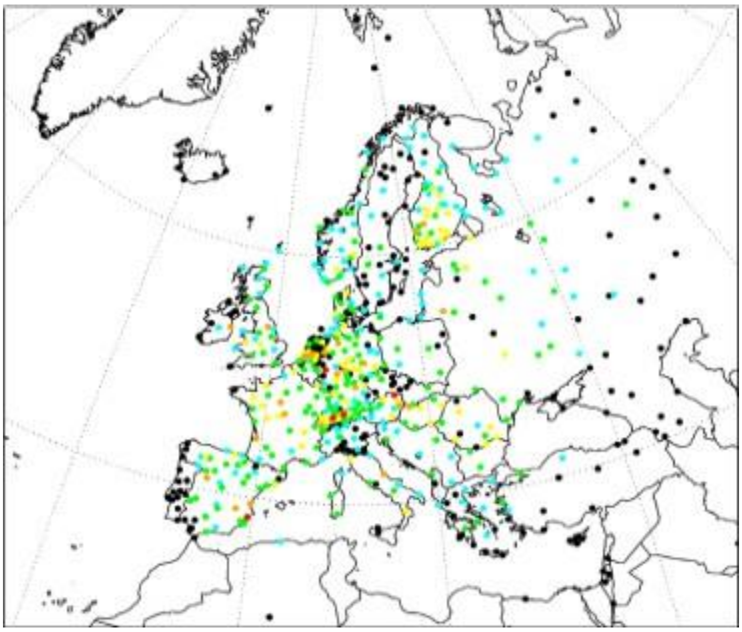


Figure 1: number of break points in mean temperature series detected by VERHOM (period 1960-2000): 0 (blue), 1 (green), 2 (yellow), 3 (orange), >4 (red), undefined (black).

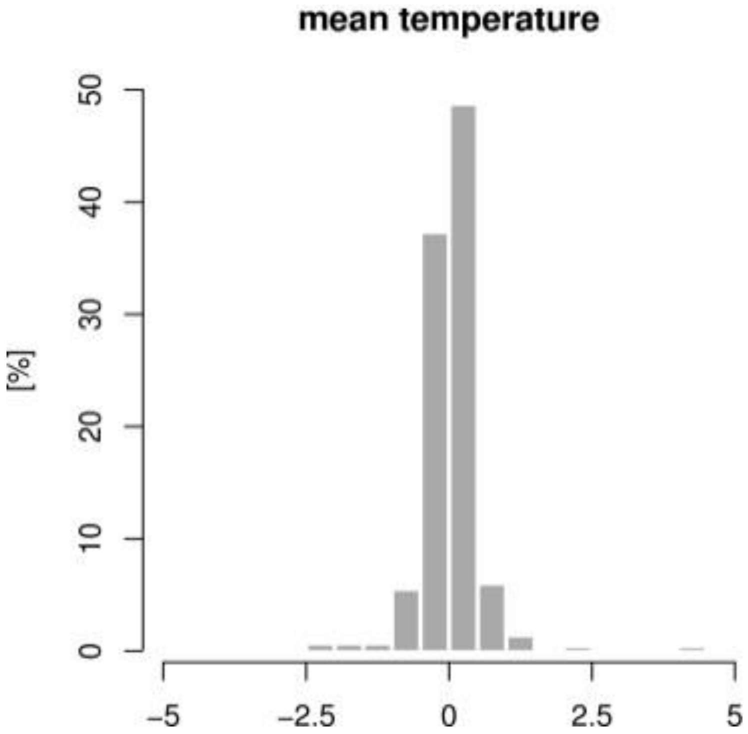


Figure 2: frequency distribution of shift dimensions for mean temperature series of ENSEMBLES data set (1960-200). Classes: 0.5°C.

In Figure 2 the frequency distribution of the shift dimensions is plotted. The shifts vary between -3 and 4.5°C. However, large shifts rarely occur and most of them lie between -1 and 1°C.

Comparing the inhomogeneities detected in the Swiss series to those detected by another semi-automated and station history based homogeneity test procedure developed at MeteoSwiss (THOM-AS, Begert et al., 2003) was used to fine-tune and test VERHOM. This comparison revealed different problems and limitations of VERHOM, such as the influence of a varying network density or problems at the edge of the borders. A further development could improve the performance substantially.

Based on the experience made in the present study the authors are convinced that an automated homogeneity test procedure will not achieve the same quality as thoroughly conducted homogenization based on detailed knowledge of the station history and the manual review of the intermediate results. However, given the large number of time series in the ENSEMBLES dataset a homogeneity assessment based on manual input would be too expensive to realize.

Further information

A detailed report about ENSEMBLES WP5.1 can be downloaded from the ENSEMBLES RT5 website: http://www.knmi.nl/samenw/ensembles_rt5/RT5.html

Related references

Alexandersson H, 1986: A homogeneity test applied to precipitation data. J. Climatol., 6, 661-675.
Begert M, Seiz G, Schlegel T, Musa M, Baudraz G and Moesch M, 2003: Homogenisierung von Klimareihen der Schweiz und Bestimmung der Normwerte 1961-1990. Schlussbericht des Projekts NORM90. Veröffentlichung der MeteoSchweiz, Vol.67. MeteoSchweiz, Zürich.
Steinacker R, Häberli C, and Pöttschacher W, 2000: A transparent method for the analysis and quality evaluation of irregularly distributed and noisy observational data. Monthly Weather Review, 128(7), 2303-2316.

For further information, please contact: Michael Begert, Mark Liniger.

NCCR Climate II.docx

3.2 Climate change in Central Europe: Is there more than pure changes in the mean?

Motivation

Our climate has undoubtedly warmed in the last decades. In central Europe, eight of the ten warmest years in the 1851-2004 temperature record have been observed from 1989-2003. However, climate change is more than changes in mean temperature. Events like the hot summer 2003 or the 2005 summer flooding events in Central Europe demonstrate that also interannual variability could change in the climate system. Such events show that potential changes in mean and variability are crucial for many socio-economic applications related to climate extremes. Although several model studies suggest increasing European summer surface temperature variability in the course of the 21st century (e.g. Schär et al. 2004), it is relatively unclear how variability has changed in the near past or will change in the future.

Results

In a just published study, Scherrer et al. (2005) investigate probability distributions of temperature data from observations and several global climate model scenario runs. The central European temperature distribution is found to change in observational data. Changes are also identified in climate change simulations for the 20th and projections for the 21st century. In order to properly compare observations and the various model results, changes in mean and variability have been scaled relative to each model's and seasonal characteristics.

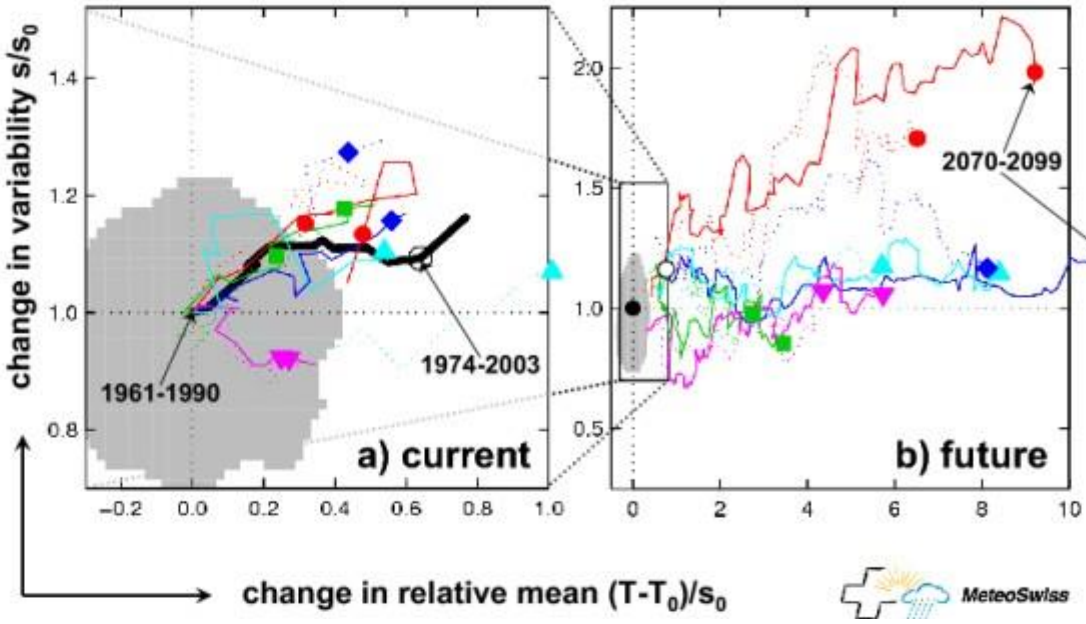


Figure 1. Changes in Central European summer temperature (horizontal axis) and interannual variability (vertical axis) in observations (black line) and IPCC model runs (thin and dotted lines) for (left) the current climate (1961-2004) and (right) the future (2005-2099). 1961-1990 ncertainty is shown as grey area.

For summer, observations show that beside large increases in temperature also the interannual variability has slightly increased. The increase in mean and variability is well reproduced by climate model runs for the current climate summers (Fig.1a). All models and scenario runs show a strong increase in summer means and for most runs there is a tendency for increasing variability in future summers (Fig. 1b). Nevertheless, the model projections for the future show large differences between the individual models.

Winter observations indicate a warming and a decrease of variability which is not found in the present-day model runs. Projections for the future propose a strong increase in the mean, but in contrast to the summer results linked with a tendency for decreasing variability.

In summary, by far the largest future changes in relative mean are found for summer. Although some models propose considerable changes in interannual variability, the dominant changes in seasonal temperature distributions are an increase in the mean value.

References.

Schär, C., P. L. Vidale, D. Lüthi, C. Frei, C. Häberli, M. A. Liniger, and C. Appenzeller, 2004: The role of increasing temperature variability in European summer heatwaves. Nature, 427, 332-336.
Scherrer, S. C., C. Appenzeller, M. A. Liniger, and C. Schär, 2005: European temperature distribution changes in observations and climate change scenarios. Geophys. Res. Lett., 32, L19705, doi:10.1029/2005GL024108.

For further information, please contact Mark Liniger.

NCCR Climate II.docx

3.3 A comparative study of Satellite and ground-based phenology

Long time series of ground based plant phenology, as well as more than two decades of satellite-derived phenological metrics, are currently available to assess the impacts of climate variability and trends on terrestrial vegetation. Traditional plant phenology provides very accurate information on individual plant species but with limited spatial coverage. Satellite phenology allows monitoring of terrestrial vegetation on a global scale and provides an integrative view at the landscape level. Linking the strengths of both methodologies has high potential value for climate impact studies. We focused on Switzerland from 1982-2001 and compared a multispecies index from ground observed spring phases (Studer et al. 2005) with two types (maximum slope and threshold approach) of satellite-derived start-of-season (SOS) metrics.

In general this study (Studer et al. subm.) showed that satellite-derived phenological SOS estimations are comparable to the integrative multispecies spring index based on phenological ground observations. The absolute trend value is smaller for the ground observed index (3 days/decade) than for the satellite NDVI indices (slope: 5.4; threshold: 8 days / decade). Normalized by their interannual variability all methods are very consistent and show a trend close to 1 standard deviation/decade.
From the two tested satellite-derived metrics the threshold method correlated better with the temporal development of the ground observed phenology index and the temporal pattern of the integrated spring temperature.

Snow anomalies may contribute to the deviation of the slope SOS estimate from the temperature driven phenological development. The slope method did not reflect several years with a very early spring onset due to very little snow and warm temperatures (Figure 1). The two satellite-derived SOS measures seem to represent different stages of spring onset with a different year to year behaviour. The threshold method is therefore better able to represent the temperature dependent temporal pattern of observed phenology.

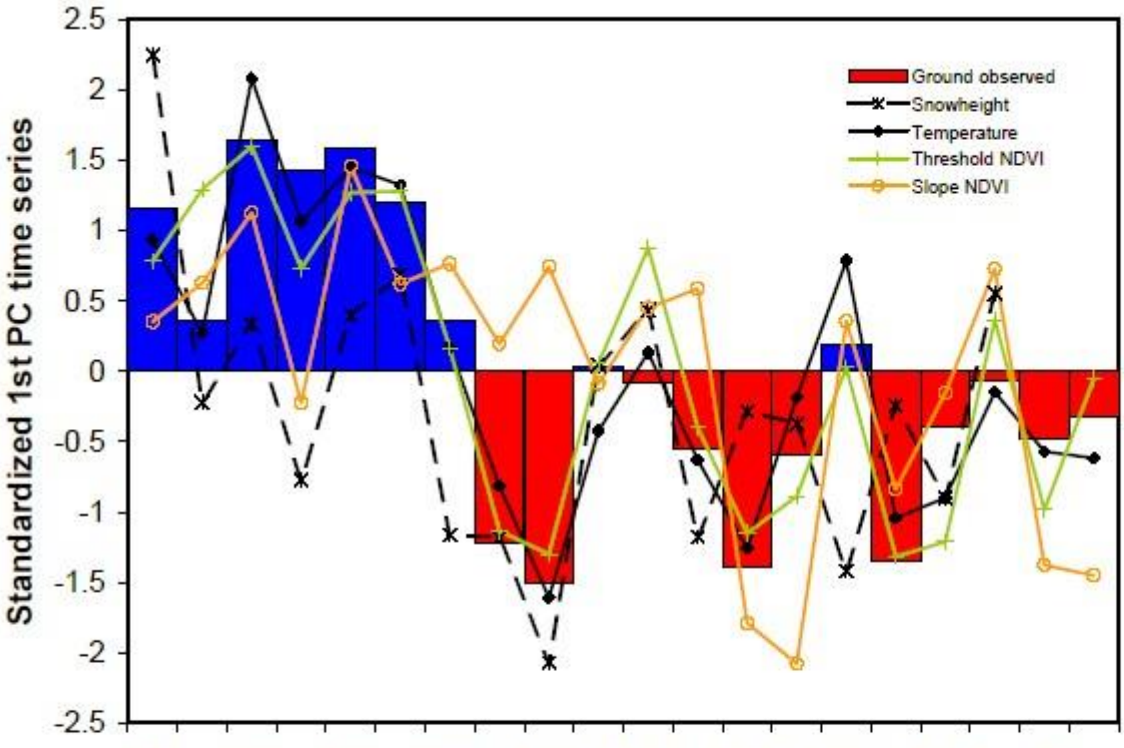


Figure 1: Time evolution of the leading principal components of the different phenology parameters, snowheight and temperature.

NCCR Climate II.docx

The good correspondence of the temporal variability and the spatial patterns of ground and satellite phenology demonstrates that they may complement each other. The precise temporal information of ground phenology can be merged with the continuous spatial coverage of satellite phenology. For instance, satellite phenology may be used to explore the complex ecophysiological responses of plant development at higher altitudes discussed above and is also a promising approach for larger scale studies with unevenly distributed ground based data.

References

Studer, S., Appenzeller, C., Defila, C. 2005: Inter annual variability and decadal trends in Alpine spring phenology. Climatic Change 73: 395-414

Studer, S., Stöckli R., Appenzeller, C., Vidale P.-L. 2007: A comparative study of satellite and ground-based phenology, int. J. Biometeorol, 51, 405-414

3.4 The Quality of ERA-40 over Switzerland

In recent years reanalysis data sets have become available and are gaining interest for climatological studies. The knowledge of their accuracy and limitations is therefore of crucial importance. In this context, the Alpine area and Switzerland therein is ideal to benchmark the reanalysis data: With its position in central Europe, the area has an excellent observational coverage over the entire period of forty years. The density of the station network is higher than the resolution of the reanalysis data. The Alpine area is also highly structured in terms of topography.
The Re-Analysis ERA-40 of the European Centre for Medium-Range Weather Forecasts (ECMWF) has generated consistent global analyses data from 1957 to 2002 with a high temporal (6h) and spatial resolution (~ 120km). Although the data assimilation technique was fixed over the whole time period, quality and quantity of past meteorological observations have changed considerably over the same time period. Observations assimilated in ERA-40 were not rigorously homogenized. This could introduce non climatic trends or low-frequency variations.

In a recent study (Kunz et al., 2007), we have compared long term changes derived from ERA-40 reanalysis with the corresponding changes found in carefully homogenized observational data. Focus is given on daily surface temperature in the Switzerland for the period 1961–2000.

Overall the two temperature data sets agree well. The temperature trend in ERA-40 (0.3°C/decade) is similar to the one in raw observations but underestimates the trend derived from homogenized observations by 0.1° C/decade. Differences between daily ERA-40 and the observations decrease with time from ±2° in the 1960s to ± 1° in the 1990s. Differences show a distinct annual cycle. On average ERA-40 overestimates temperatures in summer by 1.0°C and underestimates them by 0.4°C in winter (Fig. 1, left panel). To compare the temperature values of single Swiss station series directly with ERA-40 data a simple downscaling procedure has been applied. The difference in altitude between the model surface and the actual altitude of the stations has been corrected with a seasonal varying lapse rate, which is calculated as a mean of observed vertical temperature gradients. The correction reduces considerably the seasonal bias found between ERA-40 and the observed series (Fig.1, right panel)

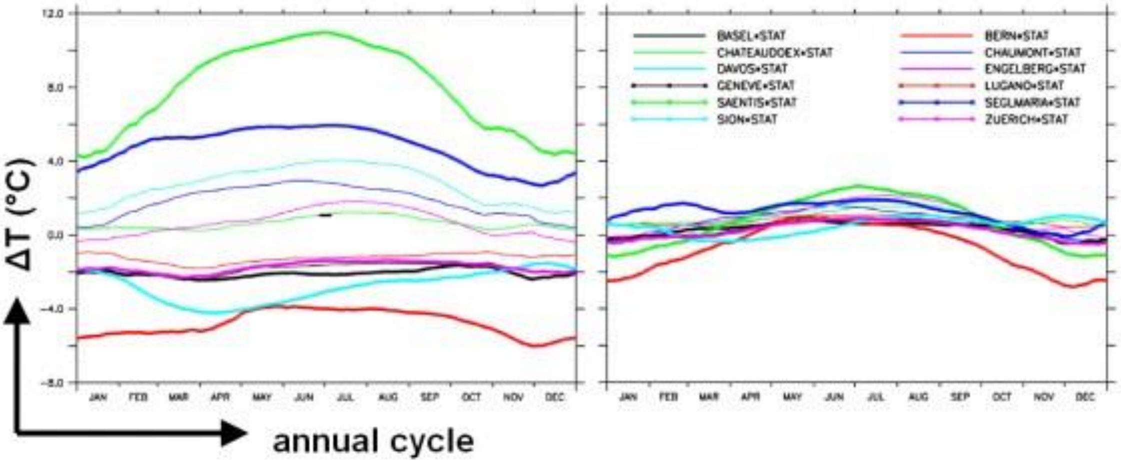


Figure 1. Comparison of ERA-40 surface air temperature with twelve homogeneous Swiss station series. The left panel shows the annual cycle of uncorrected temperature differences between ERA-40 and the corresponding station. Right panel: Statistically downscaled ERA-40 values corrected with a variable lapse rate. Units: °C.

Generally ERA-40 reproduces the main characteristics of the Swiss observational temperature well, which makes it suitable for climatological studies.

NCCR Climate II.docx

References:

Kunz, Heike; Scherrer, Simon C.; Liniger, Mark A.; Appenzeller, Christof, 2007: The evolution of ERA-40 surface temperatures and total ozone compared to observed Swiss time series. Meteorologische Zeitschrift, Volume 16, Number 2, April 2007, pp. 171-181(11), DOI: 10.1127/0941-2948/2007/0183

3.5 Return Period of Wind Storms over Europe

Accurate assessment of the magnitude and frequency of extreme wind speed is of fundamental importance for many safety, engineering and reinsurance applications.

In this study we utilise the spatial and temporal consistency of the European Centre for Medium Range Forecasts ERA-40 reanalysis data to determine the frequency of extreme winds associated with wind storms over the eastern North Atlantic and Europe. Two parameters are investigated: 10m maximum wind gust speed and 10m wind speed analysed every six hours. Wind storm statistics are determined from extreme wind indices that summarise storm magnitude and spatial extent.

We apply classical peak over threshold (POT) extreme value analysis techniques (EVA) to the extreme wind indices in order to determine the return periods of the 200 most prominent European storms which have occurred between 1957 and 2002. The catalogue of storms has been based on the on available literature.

Below is an example extreme wind climatology shown as a return period / return level diagram for a particular extreme wind index. The y-axis denotes the magnitude of the index (non-dimensional units, also called the return level) and the x-axis shows the expected return period in years of wind storm events identified in the ERA-40 dataset. The horizontal and vertical grey lines denote the return level and return period of storms in the storm catalogue respectively.

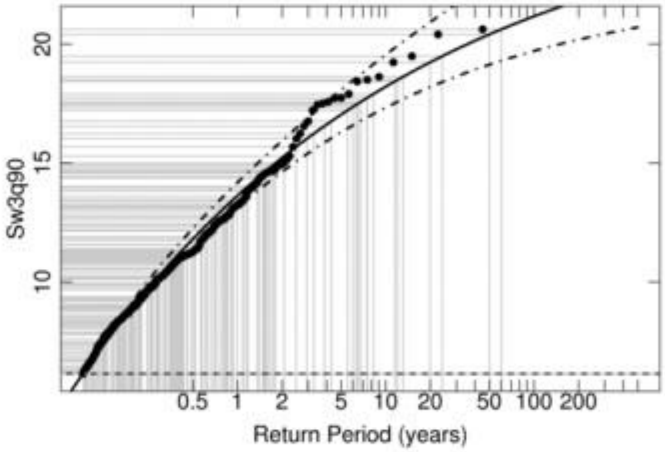


Figure 1: The return period (years) and return level (non-dimensional) of the Generalised Pareto Distribution (black line) of the extreme wind index, Sw3q90, using 10m wind speed calculated over Europe. The black dots represent the maxima of the declustered POT series. Dashed dotted lines show the upper and lower bounds of uncertainty. The horizontal and vertical grey lines denote the return level and return period of the catalogue storms respectively. Note the log scale on the horizontal axis.

Della-Marta et al. (2008) studied five different extreme wind indices each being sensitive to different aspects of the wind ?foot print? associated with each wind storm. Below is a link to figures which show the return period estimates associated with many of the wind storm events in the wind storm catalogue. The figures show the return periods of storms derived from five different indices calculated over two different regions, one considering the whole North Atlantic and European domain of the study and the other considering only the European land area. Users can also choose between which variable the return period estimates are based on, either: 10 m maximum wind gust speed or 10m wind speed analysed every six hours. Users should note that these return period estimates are representative of a return period considering the whole European area. Return period estimates of wind associated with a wind storm calculated over the area of a country or smaller region are likely to be much higher (see Della-Marta et al., 2008 for more details).

The return period (years) of catalogue wind storms for each extreme wind index

Della-Marta, P. M.; Mathis, H.; Frei, C.; Liniger, M. A.; Kleinn, J. & Appenzeller, C.
The return period of wind storms over Europe
International Journal of Climatology, 2008, DOI: 10.1002/joc.1794

The project has partly been funded by PartnerRe and by the EU FP6 project ENSEMBLES.

For more information contact Mark Liniger.

NCCR Climate II.docx

3.6 Spring Phenology Patterns of the Alpine Region and their Relation to Climate Change

According to the Intergovernmental Panel on Climate Change (IPCC AR4, 2007), the warming of the climate system is considerable. Observations form all continents show that many physical and biological systems are affected by changes in regional climate, and especially by temperature increases. Numerous phenological studies have confirmed an earlier onset of spring in response to climate change. However, all these employed studies were confined to single phenological phases at various locations.

Meier et al. (in prep) present a multispecies dataset of fifteen different phenological spring phases of the COST action 725 from 1971 to 2004 covering the Alpine region from 45 to 49°N and from 6 to 17°E. In order to determine the impact of temperature and precipitation on phenological observations principal component analysis (PCA) has been applied.

From 1971 to 1988 phenological spring events occurred on average 2.2 days later than the long term average whereas during the period from 1989 to 2004 earlier appearance dates of 2.4 days were noticed. Therefore, the average beginning of the growing season in the Alpine region has advanced by 1.8 days per decade. Temperature was found to be the main factor for this overall observed change in spring.The respective spatial PCA phenology pattern (Figure 1) was quite homogeneous over the whole Alpine region.

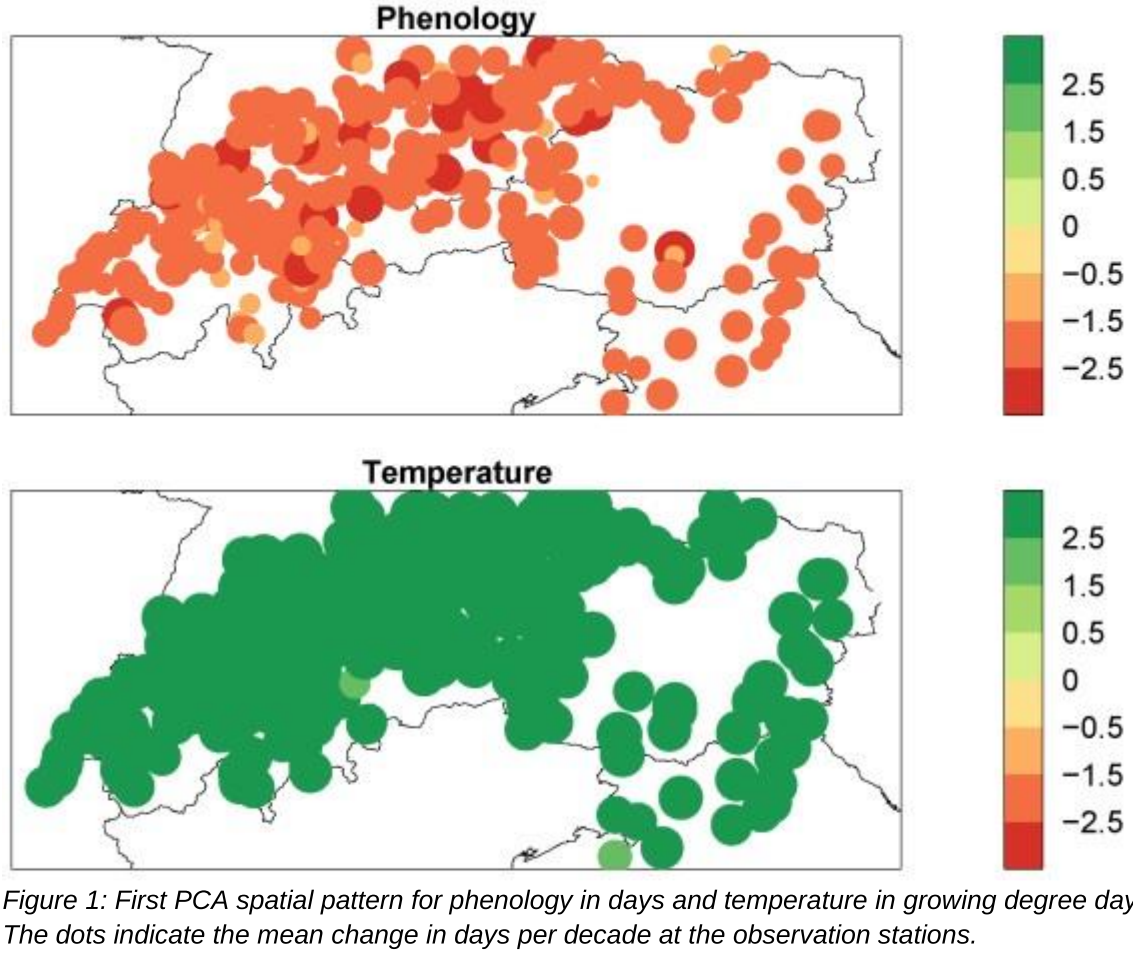


Figure 1: First PCA spatial pattern for phenology in days and temperature in growing degree days. The dots indicate the mean change in days per decade at the observation stations.

Regionally important, the second PCA phenology and temperature pattern (Figure 2) were clearly dominated by altitudinal gradients. Plants in higher elevations are exposed to lower temperatures than plants in the lowlands and thus appear earlier.

An obvious north-south pattern was found in the third PCA precipitation pattern with fewer days of rain in the Alps and south of them and more precipitation in the north of the Alpine region.

Analogous climate change signals were found by Studer et al. (2005) for Switzerland.

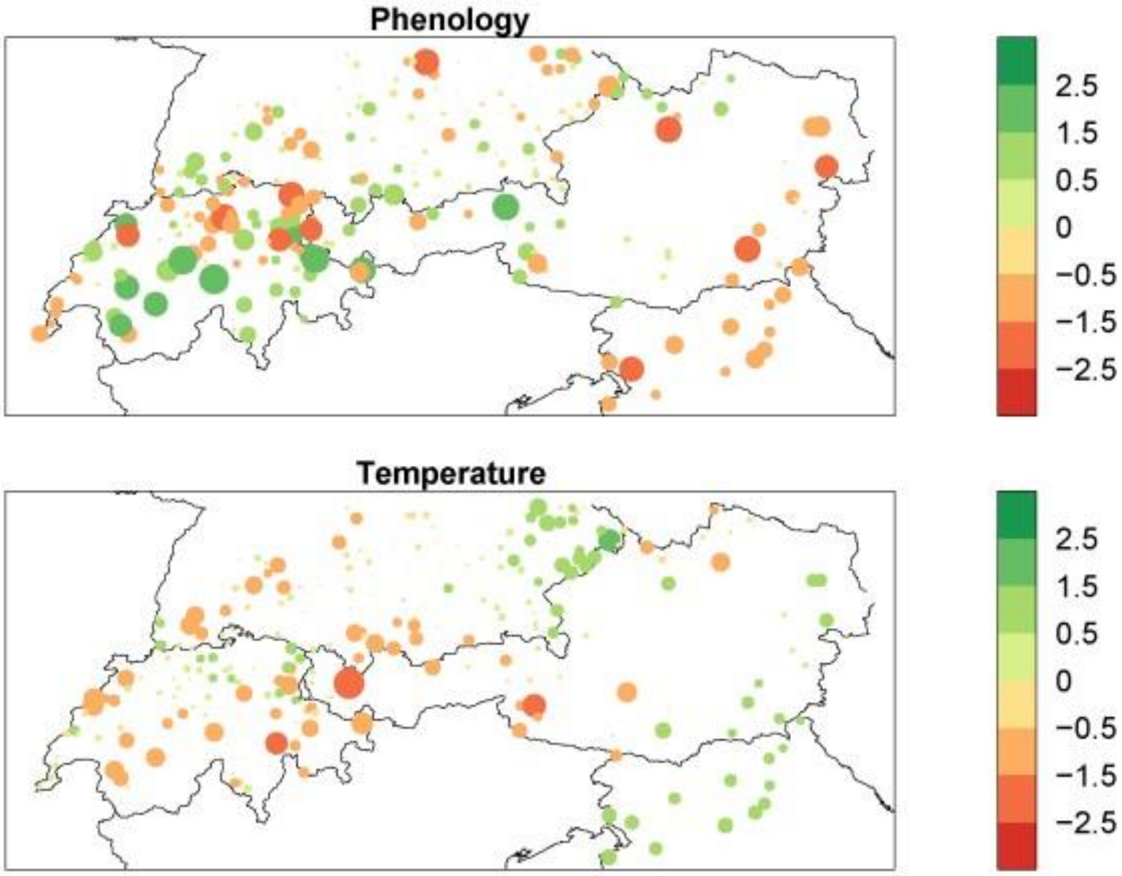


Figure 2: Second PCA spatial pattern for phenology in days and temperature in gro wing degree days. The dots indicate the mean change in days per decade at the observation stations.

References:
IPCC in Climate Change (2007). The Physical Science Basis. Contribution of Working Group I to the fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK.

Meier N., Appenzeller C., Defila C., Liniger M. and M.W. Rotach. Spring Phenology Patterns of the Alpine Region and their Relation to Climate Change. In prep.

Studer S., Appenzeller C. and C. Defila (2005). Inter-annual variability and decadal trends in Alpine spring phenology: a multivariate analysis approach. Climate Change, 73, 395 - 414.

For further information, please contact: Mark Liniger.

NCCR Climate II.docx

3.7 Statistical uncertainty of changes in winter storms

Statistical uncertainty of changes in winter storms over the North Atlantic and Europe in an ensemble of transient climate simulations

Winter storms are among the most important natural hazards affecting Europe. The study quantifies changes in storm frequency and intensity over the North Atlantic and Europe under future climate scenarios in terms of return periods considering uncertainties due to both sampling and methodology.

Return periods of North Atlantic storms' minimum central pressure and maximum vorticity remain unchanged by 2100 for both the A1B and A2 scenarios compared to the present climate. Whereas shortened return periods for maximum vorticity of all intensities are detected for the area between British Isles/ North-Sea / Western Europe as early as 2040. However, the changes in storm maximum vorticity return periods may be unrealistically large: a present day 50 (20) year event becomes approximately a 9 (5.5) year event in both A1B and A2 scenarios by 2100. The detected shortened return periods of storms implies a higher risk of occurrence of damaging wind events over Europe.

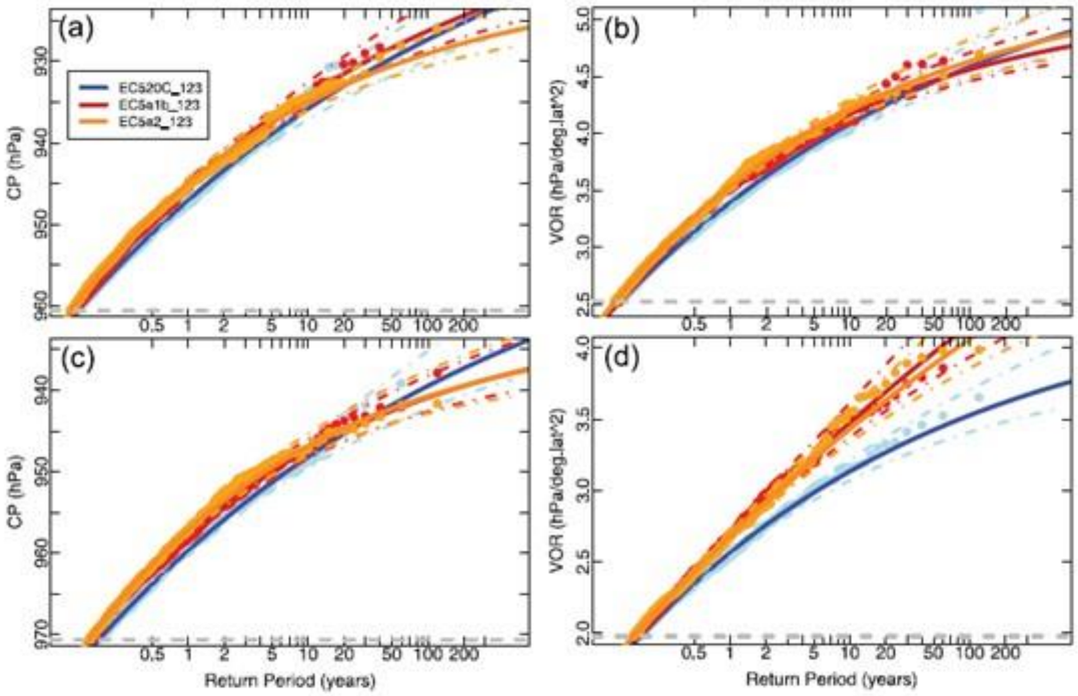


Figure 1: Return periods of storms in terms of (a) central pressure and (b) maximum vorticity for the North Atlantic Region (45°N-65°N, 30°W-10°E), (c) same as Figure 1a (d) same as Figure 1b but for the Region British Isles/North Sea/Western Europe (45°N-60°N, 10°W-30°E) Blue curve corresponds to ECHAM-20C storms, red curve to A1B storms, and orange curve to A2 storms. Individual values are shown as single dots. Dot-dashed lines indicate the 90% confidence interval limits. The units are hPa and hPa/(deg. lat.) for central pressure and maximum vorticity, respectively. Horizontal dashed gray line denotes the GPD (generalised Pareto distribution) threshold. Note the logarithmic x-axis.

This work was partially supported by the ENSEMBLES project, funded by the European Commission?s 6th Framework Programme.

References:

Della-Marta, P. M., and Pinto, J. G. (2009), Statistical uncertainty of changes in winter storms over the North Atlantic and Europe in an ensemble of transient climate simulations, Geophys. Res. Lett., 36, L14703, doi:10.1029/2009GL038557

Contact:

For further information please contact Mark Liniger.