Brönnimann, S. (2017) Weather Extremes in an Ensemble of Historical Reanalyses. In: Brönnimann, S. (Ed.) *Historical Weather Extremes in Reanalyses.* Geographica Bernensia G92, p. 7-22, DOI: 10.4480/GB2017.G92.01.

Weather Extremes in an Ensemble of Historical Reanalyses

Stefan Brönnimann*

Oeschger Centre for Climate Change Research and Institute of Geography, University of Bern, Switzerland

Abstract

Since the first historical reanalysis product, the "Twentieth Century Reanalysis" (20CR), has been published in 2011, numerous studies have made use of this data sets for a wide range of applications. In the meantime several new reanalysis data sets have been produced. Currently, four reanalyses reach back to at least 1900, additional research products have been generated, and more reanalyses will likely be produced in the future. As a means of evaluation, ten individual extreme weather events are studied in this volume in one or several of these data products. Together, they demonstrate the usefulness and limitations of historical reanalyses for studying past weather extremes. Supplementing our first volume of case studies, these cases also allow comparisons across data sets. This introductory paper gives an overview of the cases presented and the data sets used. Furthermore, the paper briefly addresses the challenges and opportunities of analysing extremes in a "multi-reanalysis ensemble".

1. Introduction

Extreme weather events have long entered the spotlight of climate science. Even more than changes in the mean climate state, changes in the frequency or intensity of extreme events are relevant for many climate impacts and thus for society. However, even for the present climate, little is still known about extreme weather events, their frequency, and their decadal variability. This is because extreme events are rare and thus long records are required. Such records exist from weather stations and they are invaluable for studying extremes, particularly when combined with documentary information (for an example, see EUROCLIMHIST, www.euroclimhist.unibe.ch/) that are rich on extreme events. However, these local series often do not allow a quantitative, physical interpretation of the weather events.

* Corresponding author: Stefan Brönnimann, University of Bern, Institute of Geography, Hallerstr. 12, CH-3012 Bern, Switzerland. E-mail: Stefan.broennimann@giub.unibe.ch

From the entire body of historical observations, however, a comprehensive, physically consistent depiction of the atmosphere can be generated using data assimilation. This was achieved in historical reanalyses such as the "Twentieth Century Reanalysis" (20CR, Compo et al., 2011) produced by CIRES/NOAA or ERA-20C (Poli et al., 2016) and the coupled reanalysis CERA-20C (Laloyaux et al., 2017) produced by the European Centre for Medium Range Weather Forecasts (ECMWF). These data sets extended the time range for studying extreme events from 40-50 years or so to now 130-150 years. However, the suitability of historical reanalysis for studying extremes remains unclear. For instance, discussions emerged on their suitability for determining trends in storminess (Donat et al., 2011; Brönnimann et al., 2012; Krüger et al., 2013; Wang et al., 2016), concluding that a careful assessment is required. While suitability must eventually be addressed for each analysis, such an assessment is ideally based on the study of many individual cases that expose the strengths and weaknesses of the data sets. Our collection of papers contributes towards that aim.

Many studies have recently used 20CR or ERA-20C to study long-term changes in extremes (*e.g.,* Donat et al., 2016a; Jones et al., 2016; Matthews et al., 2016; Pérez-Zanón et al., 2016). Other studies use historical cases as a reference or for comparison with present cases (*e.g.,* Kendon and MacCarthy, 2015; Dangendorf et al., 2016; Donat et al., 2016b) or for a chronology of storms (Cornes, 2014). Numerous case studies of extreme weather and climate have analysed and compared historical reanalyses. For instance, an extreme winter in Iceland (Moore and Babij, 2017) showed that the different sea-surface temperature and sea ice forcing of different reanalysis matters. Useful atmospheric diagnostics could be derived for flood events in northeastern Iberia (Pino et al., 2016). 20CR has been shown to be suitable for numerical downscaling for a blizzard in New York (Michaelis and Lackman, 2013), storms in Switzerland (Stucki et al., 2015) and Iberia (Fortunato et al., 2017), a snow fall event (Brugnara et al., 2016) and a flash flood in Italy (Parodi et al., 2017). A new winter storm hazard map for Switzerland was generated based on downscaling ca. 100 storms from 20CR (Dierer et al., 2013). Other studies have even attempted to dynamically downscale 20CR for the entire period (Ishida and Kavvas, 2017). The list could easily be extended. Given this immense popularity of historical reanalyses for diverse applications, case studies of individual extreme events are useful to learn strengths and weaknesses of the data sets.

In a previous volume we have analysed extreme weather events in the "Twentieth Century Reanalysis" version 2 (20CRv2) (Brönnimann and Martius, 2013). The papers were written by students of a seminar that is taught annually. They could show that most of the extreme weather events analysed were indeed well represented in that reanalysis. However, there were also cases where 20CRv2 did not reproduce the event well or not at all. Further, we also demonstrated that the ensemble mean is not always suitable (although it mostly was) and that the entire ensemble (20CRv2 consists of 56 members) should be considered.

In the past four years, new historical reanalyses have become available. The European products ERA-20C and CERA-20C were already mentioned. Specific land-surface products have been or are generated for the latter two reanalyses. CIRES/NOAA produced an updated version of 20CR, termed 20CRv2c, which reaches back to 1851 and also is an ensemble of 56 members. Furthermore, several test reanalyses were produced. For this volume we now have the situation that we can study extreme events in an ensemble of historical reanalyses. In fact, many of the cases in this volume are covered in several reanalysis data sets.

Having several products, some of which are ensembles, allows more detailed analyses, but raises additional questions. The first relates to feasibility. Studying a case in 56 members of a reanalysis is already asking a lot, but studying it in all members of many reanalysis individually is hardly ever possible. So, are there statistical frameworks that allow us to condense the combined information of all reanalyses? Should we, similar to many climate model analyses, use "multi-reanalyses means"? And what can be learned from the ensemble spread? This introductory paper gives a short preview of the volume, and it briefly touches upon the question of challenges and opportunities of having multiple reanalyses. The last part of this paper then gives a glimpse at the possible future development of historical reanalyses.

2. The data sets

All papers in this volume are self-contained and include a brief description of the data sets. Since all papers use the same data sets, this introductory paper provides some additional details on the data sets used. An overview of all reanalysis data sets including some more specific information is compiled in Table 1.

One data set is used by every paper: Version 2c of the "Twentieth Century Reanalysis" (20CRv2c). This is a global three-dimensional atmospheric reanalysis data set reaching back to 1851. It builds on version 2, which reaches back to 1871 (Compo et al., 2011). Both reanalyses provide ensembles of 56 members based on the assimilation of surface and sea-level pressure observations, *i.e.,* the distribution of atmospheric mass, taken from the International Surface Pressure Databank (ISPD). The data assimilation was performed using an Ensemble Kalman Filter technique, with first guess fields generated by a 2008 experimental version of the US National Center for Environmental Prediction Global Forecast System atmosphere/land model (NCEP/GFS, see Saha et al., 2010). The GFS model was integrated at a resolution of T62 in the horizontal, which corresponds to a spatial resolution of ca. 2° x 2°, and 28 hybrid sigma-pressure levels in the vertical. The analysis is performed every six hours, but 3-hourly forecasts of some variables are also available, allowing an even more detailed view of the temporal development of some of the extreme events.

One main difference between versions 2 and 2c concerns the sea-surface temperature (SST) and sea ice distributions. Version 2 used the HadISST1.1 data set (Rayner et al., 2003). An error in the specification of sea ice led to erroneous heat fluxes along the Arctic coast, which affected winter temperatures in the lower troposphere in that region. In version 2c this is fixed. SSTs are now taken from the Simple Ocean Data Assimilation system with sparse observational input (SODAsi, Giese et al., 2016), which uses an atmospheric reanalysis as boundary conditions. At high latitudes (>60°) SSTs were corrected to COBE-SST2 (Hirahara et al., 2014). From 2013 onwards, NOAA OI SST V2 data were used (Reynolds et al., 2007).

Another difference between versions 2 and 2c concerns the data assimilated. Pressure observations in 20CR version 2 were from ISPD version 2, 20CR version 2c used ISPD version 3.2.9 (Cram et al., 2015). Marine data were from the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) version 2.5.1 (Woodruff et al., 2011).

Further historical reanalysis data sets are used in many of the papers. They comprise three historical reanalyses from the ECMWF, namely ERA-20C, CERA-20C, and ERA-

PreSAT. All of them are based on 4-dimensional variational data assimilation (4D-Var), and all of them use ECWMF's Integrated Forecast System (IFS) model.

ERA-20C has a set-up comparable to 20CRv2c. In addition to the assimilation scheme and model used, one main difference between to 20CR is that ERA-20C assimilates marine winds, in addition to surface and sea-level pressure (which are from ISPD version 3.2.6 and ICOADS version 2.5.1; see Table 1). ERA-20C spans the years 1900 to 2010 with a temporal resolution of three hours and a spatial resolution of 1.125°x1.125° (Poli et al., 2016). A preliminary 10-member assimilation (using different SST realisations from HadISST2.1.0.0) was used to estimate the background error in the final (one member) assimilation. From ERA-20C, a land-surface product was generated by driving ECMWF's land-surface scheme HTESSEL off-line.

The reanalysis CERA-20C (Laloyaux et al., 2017) assimilated the same atmospheric observations, but was coupled to the community ocean model NEMO (Nucleus for European Modelling of the Ocean, Madec et al., 2012). In this set-up, the model integration is performed coupled. Both models then have their individual variational assimilation scheme, but with one additional iteration that allows the ocean to "see" the atmospheric update and vice versa.

Finally, a test reanalysis ERA-PreSAT was performed for the years 1939-1967 that used the same set-up as ERA-20C, but additionally assimilated upper-air observations from the CHUAN data set (Stickler et al., 2014). Details on this product can be found in Hersbach et al. (2017). The historical reanalyses are supplemented by conventional reanalyses such as NCEP/NCAR reanalysis (Kistler et al., 2001) in some of the papers of this book. An overview of all products is given in Figure 1.

Together the data sets provide rich information on global weather during the past 150 years. They constitute a "multi-reanalysis ensemble". Several reanalyses (20CRv2, 20CRv2c and CERA-20C) are ensemble products. However, there are differences in the meaning of the ensemble spread. 20CR (v2 and v2c) are atmospheric reanalyses, the ensemble is an initial condition ensemble. ERA-20C used a 10-member ensemble, which also accounts for variations in SSTs, for the determination of the background error. CERA-20C is a coupled reanalysis and its 10 members thus cover a coupled phase space. ERA-PreSAT is one member only. While some of the cases studied in this book are only analysed in one or two of the products, some are analysed in four or five data sets.

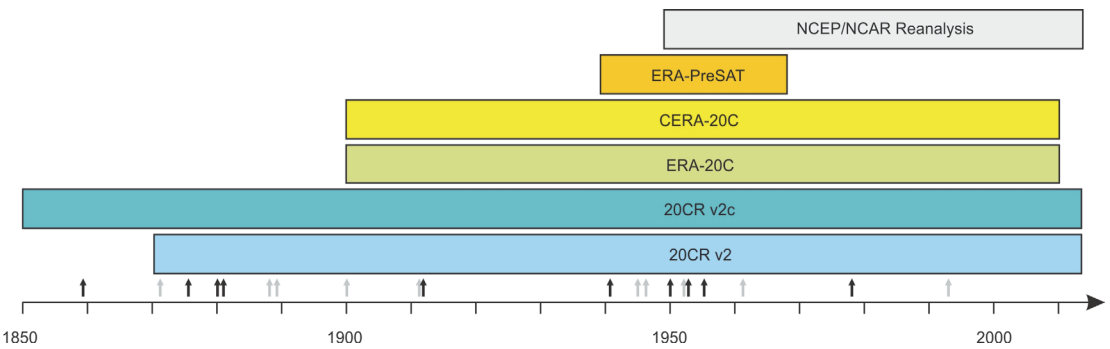

Figure 1. Time covered by the different reanalysis products used in this volume. The short black arrows indicate the timing of the ten case studies, grey arrows denote the cases published in Brönnimann and Martius (2013).

Table 1. Overview of the historical global dynamical reanalysis datasets used in this book. Res. = Resolution at equator, *n* is the number of members. 20CR: Twentieth Century Reanalysis, EnKF: Ensemble Kalman Filter, 4D-Var: Four-dimensional variational data assimilation, SI = Statistical Interpolation (3D-Var), p = pressure

Reanalysis	Model	Observ.	ISPD	Scheme	Period	Res.	n	Reference
20CRv2	NCEP/GFS, T62/L28	p	2	EnKF	1871–2014	320 km	56	Compo et al. (2011)
20CRv2c	NCEP/GFS, T62/L28	p	3.2.9	EnKF	1851–2014	320 km	56	Compo et al. (2011)
20CR-1816*	NCEP/GFS, T62/L28	p	3.2.9	EnKF	1815–1817	320 km	56	Brohan et al. (2016)
ERA-20C	IFS, Cy38r1 T159/L91	p, wind	3.2.6	4D-Var	1901-2010	125 km	1+10[+]	Poli et al. (2016)
CERA-20C	IFS, Cy41r2 T159/L137 NEMO 1°, 41 levels	p, wind	3.2.6	IFS 4D-Var NEMO 3D-Var	1901-2010	125 km	10	Laloyaux et al. (2017)
ERA-PreSAT*	IFS, Cy38r1 T159/L91	p, wind, upper-air	3.2.6	4D-Var	1939-1967	125 km	1	Hersbach et al. (2017)
NCEP/NCAR		all	-	SI	1948-2017		1	Kalnay et al. (1996)

*research versions

[+]one deterministic member, 10 members were used to estimate the background error

In all of the studied cases, also other data sets than reanalyses were consulted, including instrumental observations and derived products (*e.g.,* historical weather charts). Moreover, contemporary literature was available in all cases to put the results found with 20CR into context. These data sets and sources are discussed in the individual papers.

3. Selection of events

Ten extreme weather events were chosen for this book. Table 2 provides a list and references to the papers; Figure 2 gives a geographical overview of the locations. The ten events comprise storms, floodings, and cold surges. The earliest of these events was the storm on 25-26 October 1859 (Villiger et al., 2017) that sank the "Royal Charter" and other ships, killing around 800 people. That storm played a decisive role in reinforcing attempts to set up a warning system (Moore, 2015), which led to the first operational weather forecasts two years later, issued by Admiral FitzRoy. Because it occurred in the first decade of 20CRv2c, when observations were very sparse, it is an interesting case to test the limits of historical reanalyses

Table 2. List of events in this compilation in chronological order

Event	Type	Location	Year	Paper
"Royal Charter" Storm	Storm	UK	1859	Villiger et al. (2017)
"Märzorkan"	Storm	Europe	1876	Ernst et al. (2017)
Sitka Hurricane	Storm	USA	1880	Franke et al. (2017)
Great Gale	Storm	UK	1881	Meyer et al. (2017)
"White Hurricane"	Storm	USA	1913	Gassner et al. (2017)
Iberian Storm	Storm	Portugal/Spain	1941	Baumann and Reichen (2017)
Appalachian Storm	Storm	USA	1950	Burkart and Wyss (2017)
1953 friagem	Cold Surge	Brasil	1953	Zamuriano et al. (2017)
1956 coldwave	Cold Surge	Europe	1956	Dizerens et al. (2017)
Rhine Flooding	Flooding	Germany	1978	Stucki et al. (2017)

and perhaps provides a glimpse at the quality we might expect when extending reanalyses even further back than presently.

The "Märzorkan" of 1876 (Ernst et al., 2017) also is a challenging case, as this storm was of a relatively small spatial scale, though crossing regions with good observational coverage. Franke et al. (2017) analyse an unsusually strong storm that hit Sitka, Alaska, in October 1880. The storm is named "Sitka Hurricane" even though it was not a hurricane but an extratropical cyclone. Incidentally, measurements from a ship are available that were not assimilated into 20CRv2c. So, even more than the "Royal Charter" storm, this storm tests the limits of historical reanalyses in data sparse regions.

Although occurring only one year later than the "Sitka hurricane", the "Great Gale" of 14 October 1881, a severe storm that hit the UK, demonstrates the effects of good data coverage. The storm can be tracked across the Atlantic and is well reproduced in 20CRv2c over Western Europe (Meyer et al., 2017). The "White Hurricane" or "Great White Storm" of November 1913 (Gassner et al., 2017) also was not a hurricane (though hurricane-force winds were observed). It was one of the most deadly storms in the Great Lakes region, killing 250 people. It originated from the merging of two storm systems, possibly fuelled by the warm water surface of the Great Lakes. Occurring in a region with a dense observing system, we expect surface-based reanalyses to capture the storm, but whether the products also represent snow cover is a more challenging question.

In the middle of the Second World War, in February 1941, a severe storm struck Iberia. This event is studied by Baumann and Reichen (2017). Another storm studied in this book is the Great Appalachian Storm of 1950 that hit the eastern USA on Thanksgiving and caused 353 fatalities, power outages, and large economic losses (Burkart and Wyss, 2017). Both cases concern well-observed regions and time periods. The latter case is analysed in four different products and thus allows a comparison.

Figure 2. Map showing the locations of case studies comprised in this volume (red stars). Gray dots refer to case studies in Brönnimann and Martius (2013).

Two cases in the 1950s concern cold surges. In 1953, a cold surge hit Brasil (Zamuriano et al., 2017). The frost had large effects on coffee plants. This translated to changes in world market prices and eventually affected Brasil's production policy. Three years later, southern Europe was affected by a cold wave that lasted one month and let fountains in Marseille freeze (Dizerens et al., 2017). These cases are also covered by six reanalysis data sets and thus allow a comparison. The most recent case concerns one of the largest flooding events of the Rhine river. Occurring in 1978 (Stucki et al. 2017), this event just narrowly predates the availability of the current satellite based reanalyses and is therefore not well studied.

4. Analyses

Analysing multiple reanalysis data sets, some of which are ensembles, provides new challenges. While the climate model community has had this problem for a long time, this is new to atmospheric science community. In the following we touch upon several possible directions of how to use the information in a multi-reanalysis ensemble. The approaches encompass (1) descriptive and statistical ways of assessing the entire ensemble information, (2) analysis (and selection) of individual members, (3) arguments of physical consistency, and (4) application of model output statistics or similar techniques.

The example we use is the same as in the introductory paper for the last volume – winter wind storms. Using the case of a severe Föhn storm in Zurich on 5 January 1919 (see Brönnimann et al., 2012) we outline several issues related to "multi-reanalysis ensembles".

4.1. Statistical analysis of the ensemble

The information in a multi-reanalysis ensemble can be used to characterise (visualize or quantify) the uncertainty. But how? A straight forward way would be to analyse the distribution or (in a parametric sense) the ensemble mean and spread. Most papers in this book analyse the spread of individual reanalyses; this is a useful diagnostic. If this is valuable also for multiple reanalyses, this would make the analysis rather simple as 20CR, for example, provides ensemble mean and spread on the website. However, the spread refers to 6-hourly states. For any derived quantity, the covariance structure needs to be accounted for. For obtaining ensemble information in a trend analysis, trends need to be computed for each member and only from that result, an ensemble distribution can be obtained (in Brönnimann et al., 2012, we did this for trends in wind extremes in Zurich).

For a multi-reanalysis ensemble, things become even more complicated. As a first step, one may wish to visualise the raw data. For the storm case in Zurich (Fig. 3) we show mean sea-level pressure and 10-m wind speed (which in 20CR products is not from the analysis) on 5 January 1919, 6 UTC at the grid point closest to Zurich. The distributions of the ensembles 20CRv2 and 20CRv2c (n = 56) are shown as smoothed kernel densities. The CERA-20C ensemble (n = 10) is added as lines, the deterministic ERA-20C reanalysis as bar. Such plots might be instructive, but the interpretation is difficult. In this case we find large differences between the products (note that possible displacements in time and space are not considered), against which the spread within a given product dwarfs. No realisation reproduces the observations, which is however not unexpected: Föhn storms are strongly dependent on topography, which is far too coarsly resolved in all reanalyses. The figure clearly illustrates that a multi-reanalysis mean is not meaningful.

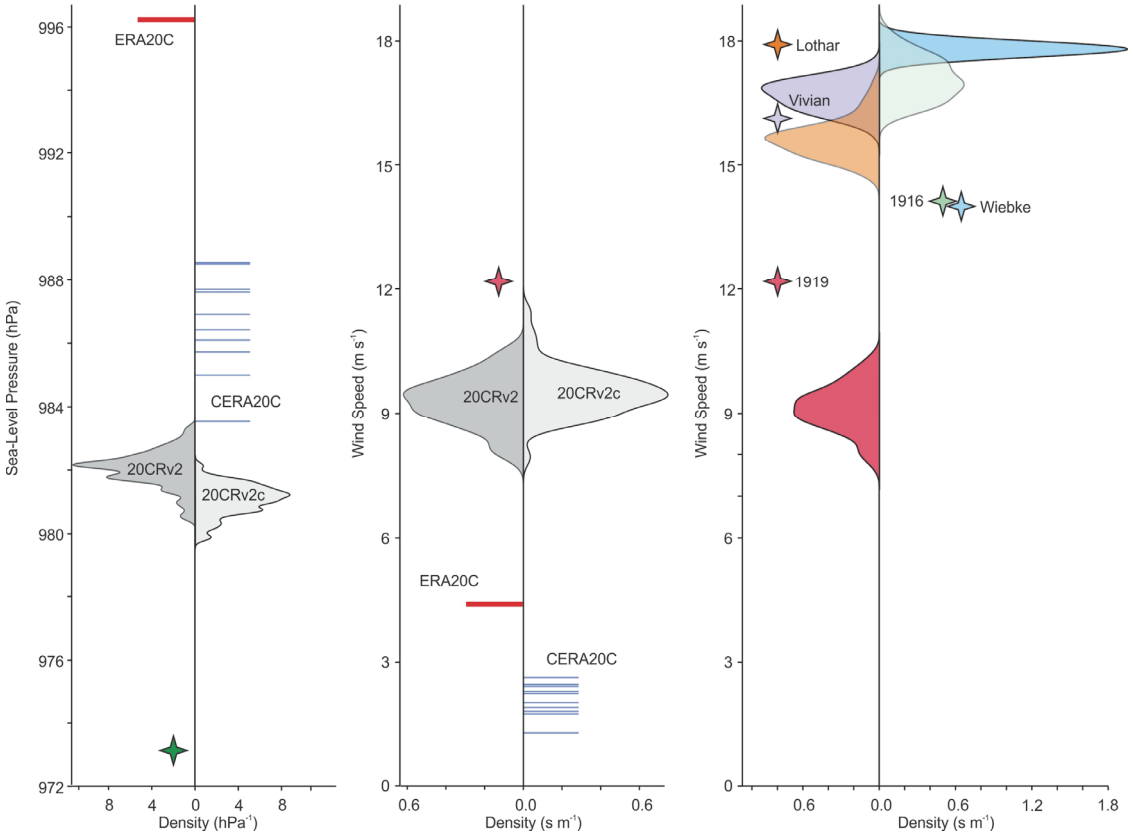

Figure 3. Sea-level pressure (left) and wind speed (middle) in Zurich on 5 January 1919, 6 UTC, in different reanalysis ensembles and observations (star). The variable is plotted on the vertical axis, the horizontal axis shows the density (the bar length has no meaning on the horizontal axes). For the two 20CR reanalysis the shading gives a kernel density. Right: Wind speed during five major wind storms in Zurich from quantile-mapped 20CRv2 (densities) and observations (stars).

4.2. Selection of members

We can also analyse the full output of each member of all reanalyses, *e.g.,* sea-level pressure and wind fields (Fig. 4). For reasons of presentation, only every 8th member of 20CRv2 and 20CRv2c is plotted, plus all members of CERA-20C and ERA-20C. This figure shows 25 possibilities of how the approaching cyclone, with a strong pressure gradient in the southeast, might have looked like. In this book, individual ensemble members are studied in some detail in three early cases (Royal Charter Storm, Märzorkan, Sitka Hurricane), where individual cases, but not the ensemble mean, might capture the reality most closely. For most of the later cases, only the ensemble mean and spread are analysed. Additionally, most of the studies compare different data sets.

If processing all members is not possible, one might wish to analyse only the "best" member or a representative subset. An example could be dynamical downscaling, which is computationally expensive. A "Best Ensemble Member" selection can be achieved by minimizing a cost function according to some target variable. For instance, all members displayed in Figure 4 show the pronounced low pressure system, but there are difference, both within the data set as well as between data sets. If we have only computing resources for downscaling one member, which one should be downscaled to obtain the best results? Should we pick just the one with the highest wind speed in Zurich on 5 January 1919?

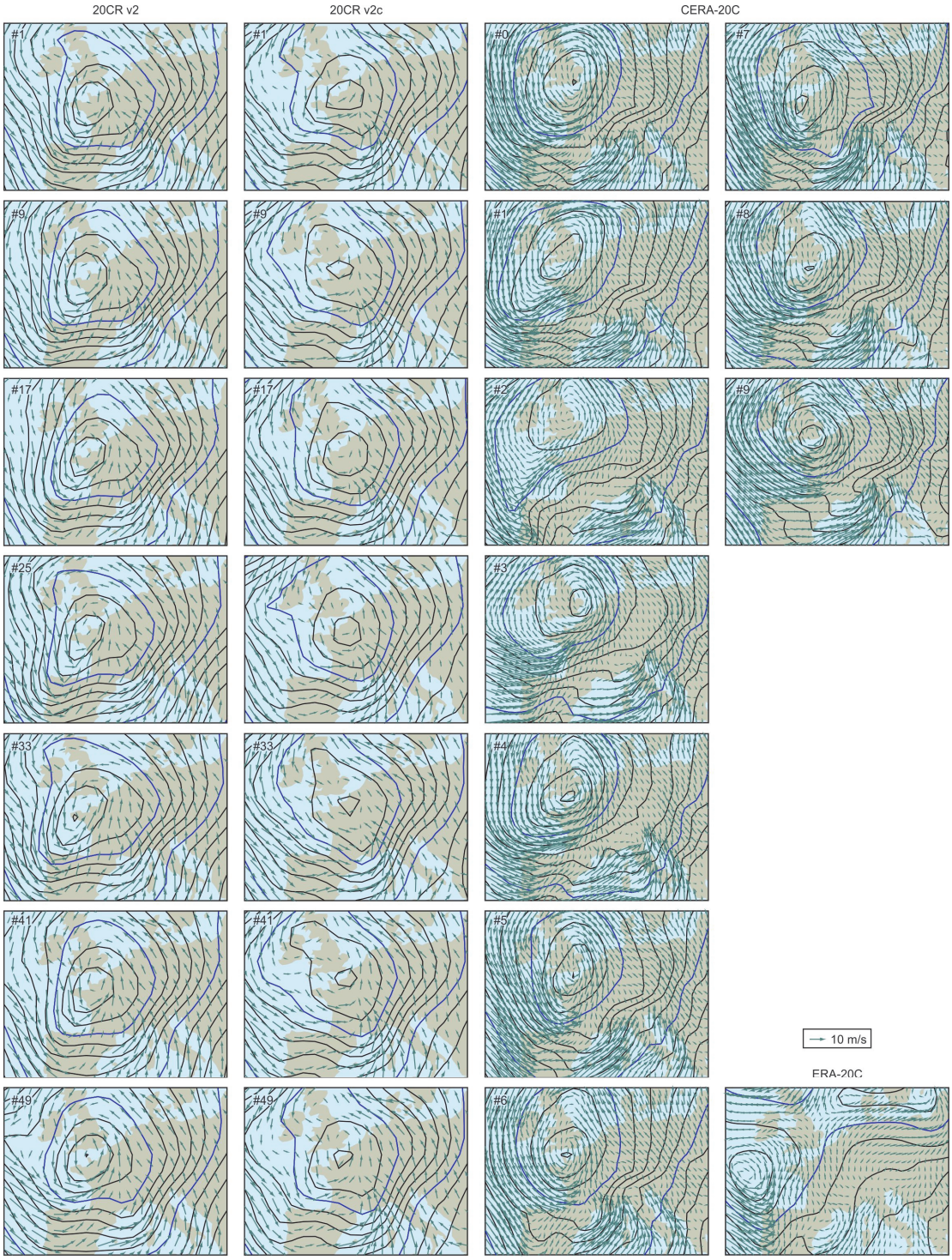

Figure 4. Mean sea-level pressure and 10-m wind speed on 5 January 1919, 6 UTC, in every 8th member of 20CRv2 (first column), 20CRv2c (second column), all 10 members of CERA-20C (third and fourth column) and the deterministic reanalysis ERA-20C (bottom right). Blue contours indicate the 980 and 1000 hPa isolines.

Stucki et al. (2016) dynamically downscaled all members of 20CRv2 for two windstorm cases in Switzerland and then analysed a posteriori whether a subsampling of members might have been possible based on 20CRv2. Such strategies are important to reduce computations, which otherwise would be impossible. In that case, however, only a limited predictability was found. Still, reducing the sample size might be possible in some cases.

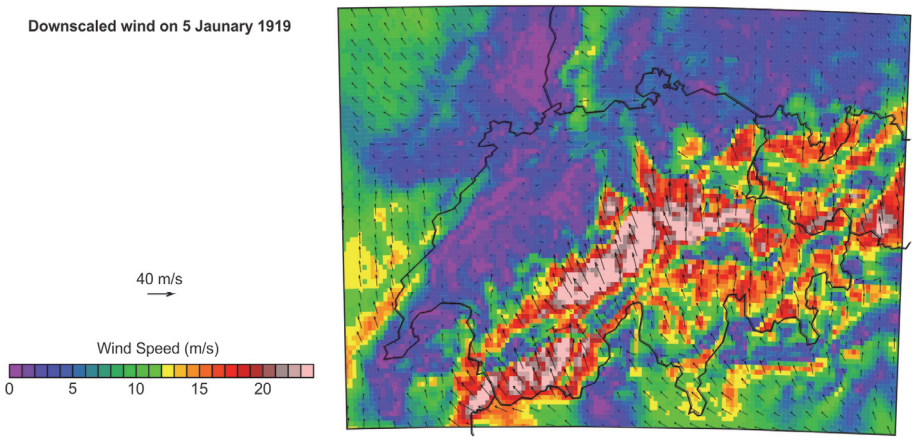

Figure 5. 10-m wind speed during the peak of the storm on 5 January 1919 in the WRF downscaling (see Stucki et al., 2016).

4.3. Physical considerations

The choice of a product can be motivated by physical considerations. For instance, due to differences in model orography, not all reanalyses may be able to depict storms in complex terrain equally well. This holds particularly true for Föhn storms such as 1919. Given a correct depiction of the synoptic-scale situation, dynamical downscaling can be used to increase the resolution and thus better capture the influence of orography and land surface. Stucki et al. (2016) downscaled the Föhn storm 1919 from the 20CRv2 ensemble mean using the Weather Research and Forecasting model (WRF). The resulting wind field (Fig. 5) is realistic and shows high wind speeds in Föhn valleys. The high observed wind speed in Zurich is still not captured, but 20 m s^{-1} are modelled only 15 km south of Zurich.

In addition to orography or model resolution, there may peculiarities specific to the assimilation system. An example is tropical cyclones. All historical reanalysis products assimilate tropical cyclone data from the BestTrack archive. However, when presented to the model, the deviation of the central pressure of a hurricane from the model state is very large. Typically such an observation would be rejected as unplausible. To avoid rejection, hurricane track data were "white-listed" in 20CRv2 and 20CRv2c. Therefore, rather strong hurricanes appear where track data are assimilated (see also the 1938 Long Island hurricane in 20CRv2c on the cover page of this book). In ERA-20C, CERA-20C, and ERA-PreSAT, many track data are rejected (Poli et al., 2016; Hersbach et al., 2017) and hence many hurricanes are missing or are too weak ("white-listing" here would lead to unrealistically large systems and thus would not solve the problem, Patrick Laloyaux, pers. comm.).

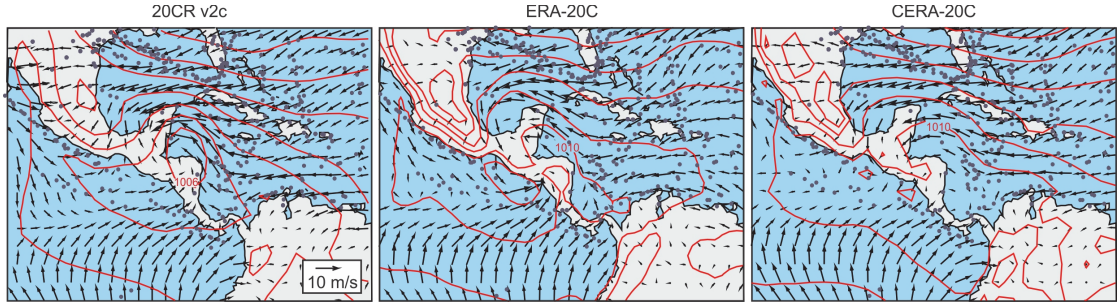

Figure 6. Sea-level pressure and 10-m wind on 27 September 1955, 12 UTC, during hurricane Janet in three reanalyses. Blue dots indicate the location of the assimilated observations.

As an example, Figure 6 shows hurricane Janet, which hit Belize in 1955. While the flow is similar in all three reanalyses, the system is strongest in 20CRv2c which assimilates the track, but weaker in ERA-20C and CERA-20C. This puts constraints on using reanalyses for certain purposes. It also shows that reproducing hurricanes in reanalyses is challenging.

4.4. Model output statistics, quantile mapping or statistical downscaling

Forecasters and climate modellers working with ensemble output of multiple regional models often use statistical techniques to make models comparable among each other and with observations. These techniques are often termed model output statistics and in essence fit the model statistics to the observation statistics. A popular technique is quantile mapping, which adjusts the distribution of modelled variables to that of observed variables.

We performed a quantile mapping of daily maximum wind speeds in 20CRv2 (maximum of 6-hourly data) to observations in Zurich (daily maximum of hourly wind measurements at 10 m). The quantile mapping was performed over the 1981-2008 period for each member separately (linear option in fitQmapQUANT, R-package qmap in steps of 0.01, Gudmundsson et al., 2012). Results are added to Fig. 3 (right). The 1919 storm is not improved in the quantile-mapped data – hence the fact that it is not well depicted is not due to a bias in the entire distribution. The two strongest westerly wind storms in the Zurich observations, Vivian (1990) and Lothar (1999), are much better reproduced, though the latter is underestimated. The right side of the *x*-axis shows the two strongest storms in 20CR (Wiebke, 1990, and an unnamed storm in 1916). Here we find an overestimation, *i.e.,* 20CRv2 shows too high wind speeds. While westerly wind storms are now well depicted in general, other storms (such as Föhn) are not, which shows that part of the errors, such as those due to boundary conditions and model physics, remain. Conversely, the quantile mapping adds further uncertainties and assumptions (*e.g.,* extrapolation of the scaling of the 99th percentile, seasonality).

5. Future development

Historical reanalyses have quickly become widely used and very powerful tools. How will these products develop? In this Section I sketch several different developments that have already started: (a) extension further back in time, (b) assimilation of further variables, and (c) ocean-atmosphere coupling.

5.1. Extension back in time

How far back can this approach be extended? 20CR reaches back more or less to the start of national weather services in the second half of the 19th century. This is a reasonable choice, although many regions of the globe are not covered at all in the 19th century. Conversely, for Europe (and perhaps New England), early instrumental data would possibly allow a further backward extension. For the bicentenary of the Tambora eruption in 2015, a 20CR version was produced that covered the eruption (Brohan et al., 2016). Figure 7 shows ensemble mean fields of 500 hPa geopotential height and 850 hPa temperature (left) as well as 10-m wind, precipitation and 500 hPa vertical motion on 29 July 1816. The figure shows a frontal system stretching across Europe, with high precipitation amounts over France and western Switzerland, promoted by synoptic-scale as well as (near the Alps) orographic uplift. In fact, according to Swiss diary entries it rained almost every day in July 1816, including on 29 July.

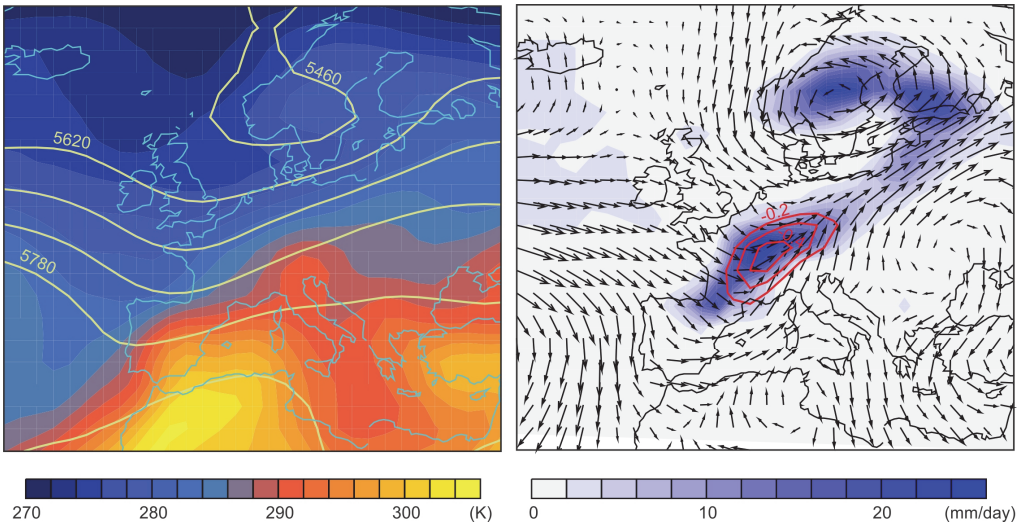

Figure 7. (left) 500 hPa geopotential height (contours, gpm) and 850 hPa temperature and (right) 10-m wind (vectors), precipitation (colours) and 500 hPa vertical motion (contours, only ascent is shown) on 29 July 1816 in 20CR-1816 (Brohan et al., 2016).

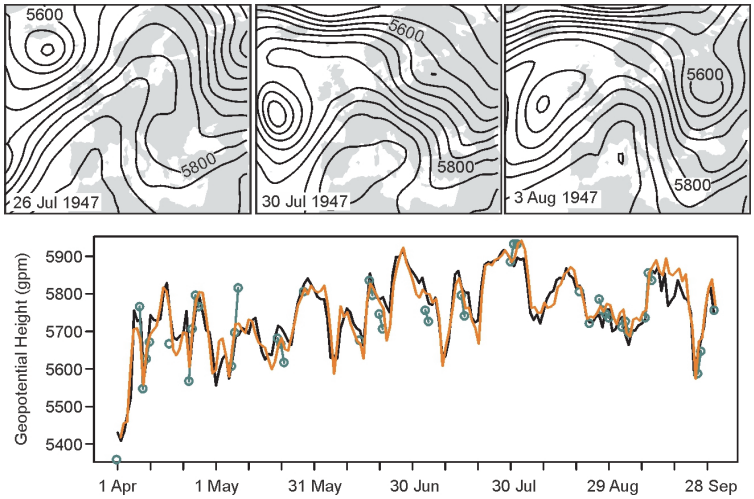

Figure 8. (top) 500 hPa geopotential height from 20CRv2c on 26 Jul, 30 Jul and 3 Aug 1947, during one the most intense heat wave of the 20th century in Switzerland. (bottom) 500 hPa geopotential height at the location of Payerne, Switzerland from 20CRv2c (black), ERA-PreSAT (orange) and observations (green; Grütter et al., 2013).

Two days later, it was so cold that one had to heat the houses. In fact, the cold air is visible two days earlier over the North Sea, approaching rapidly southward with strong northerly flow (see Brönnimann and Krämer, 2016).

This reanalysis was based on observations from some 30 stations and few ships. This is only a very small fraction of what could be made available, although coverage will remain limited to Europe, eastern North America, and some scattered ships. A similar coverage could perhaps be reached back to the 1760s. A 250-year dynamical reanalysis thus seems possible, although meaningful results will only be reached in Europe and small parts of North America.

5.2. Assimilation of further variables

The historical reanalyses 20CRv2 and 20CRv2c assimilate only pressure and yield very good results. ERA-20C and CERA-20C additionally assimilate marine wind. The wide spread wind observations provide additional information from which future reanalyses might benefit.

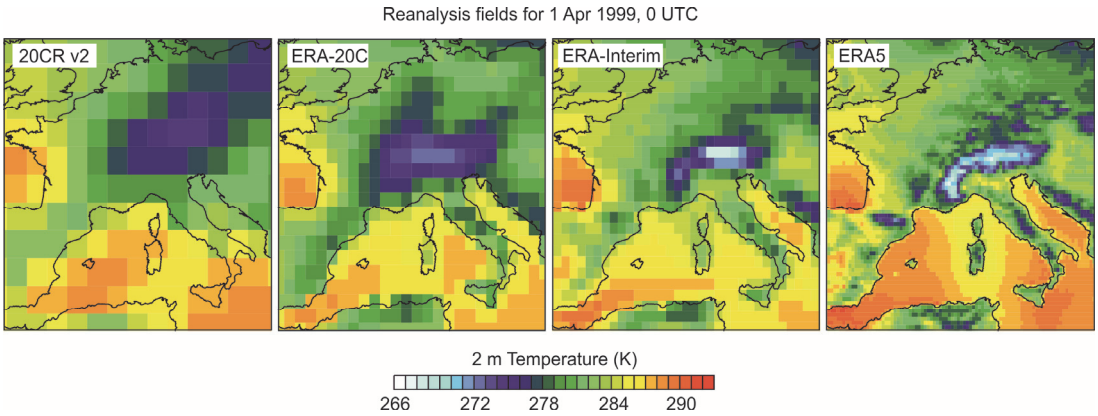

Figure 9. Reanalysis fields for 2-m temperature on 1 April 1999, 0 UTC in four different reanalysis products.

Further variables could be assimilated. For instance, upper-air data have been digitised back to the late 19th century (Stickler et al., 2014). From about 1918 onward, their number might be sufficient to add value to a global product, thus a 100-yr "full reanalysis" seems possible. A test was performed for the period 1939-1967, termed ERA-PreSAT (Hersbach et al. 2017), which showed a clear improvement, though bias issues are still present. As an example, Figure 8 shows results for the summer of 1947, when intense heatwaves struck central Europe. Fields of geopotential height at 500 hPa for the peak phase show extended high-pressure ridges over central Europe. For the gridpoint closest to Payerne (Fig. 8, bottom), ERA-PreSAT shows a higher correlation with radiosoundings from Payerne than 20CRv2c (0.94 instead of 0.92). This is due to the fact that the ERA-PreSAT system assimilates the soundings, while at the same time it is not affected by outliers.

5.3. Ocean-atmosphere land coupling and resolution

The difference between CERA-20C and the other reanalyses is its coupling with an ocean model. In that case only the forecast is fully coupled, while an additional iteration in the assimilation step assures that ocean and atmosphere can react to each other's updates. However, coupled products might become more widespread in the future. For instance, decadal prediction systems need to be tested based on long data sets.

Another approach is used by 20CRv2c. Since gridded SSTs products, the most important boundary condition of the model, are not available far back into the 19th century, an iterative approach is used (Giese et al., 2016) in which a version of the 20CR reanalysis is first generated from climatological SSTs. Fluxes and winds from that reanalysis are then assimilated, together with oceanic data, into an ocean assimilation system, termed SODAsi (Simple Ocean Data Assimilation with sparse input). Sea-surface temperatures are then used to generate a new atmospheric reanalysis version. Feeding atmospheric forcing to the ocean system and returning oceanic boundary conditions to the next iteration of the atmosphere is another way of coupling ocean and atmosphere in a reanalysis system.

The land surface is still relatively crude in most reanalysis products. From the European Reanalyses (ERA-Interim, ERA-20C, CERA-20C), land-surface products are generated off-line (no observations are assimilated) by driving a state-of-the-art land-surface model with reanalysis input. Not only for land-surface applications, but also for analysing exteme events, resolution is a key issue. The latest European reanalysis, ERA-5 (Hersbach and Dee, 2016),

has a resolution of 31 km globally. As an example, Fig. 9 shows temperature over Europe on 1 April 1999 in four reanalyses. The change in resolution from 20CRv2c (2° x 2°) to ERA-5 is impressive. While ERA-5 is still in production, a backward extension to 1950 using historical upper-air data is planned. Although this does not rival the historical reanalyses for most of the cases presented in this book, some (such as the Appalachian storm of 1950 or the cold waves in Brasil 1953 and Europe 1956) could then be studied in even greater detail.

6. Conclusions

The papers compiled in this second collection analyse a selection of extreme events during the past 160 years in several historical reanalyses. In this introductory chapter we introduce the cases and the data sets used and discuss issues related to the fact that we now have several products, some of which are ensemble products.

Together, the papers in this book show that historical reanalyses are generally suitable for studying past extremes. Most of the events appear in 20CRv2c, but in two early cases the ensemble mean is of little use and individual members need to be studied. Magnitudes are mostly underestimated in the ensemble mean. Cases in well-observed regions such as Europe and North America are well represented. Also, sea-level pressure and wind are typically well represented, while temperature and especially precipitation are less well captured.

Having several reanalysis products is an opportunity and provides additional information, but there is no framework for analysing "multi-reanalysis ensembles". Knowing the strengths and weaknesses of each individual product is key to a careful analysis in this case. In the future, there will be even more historical products, perhaps reaching further back, with a higher resolution and a coupled ocean.

Acknowledgements

The author acknowledges funding by European projects ERA-CLIM2 and EUSTACE and the Swiss National Science Foundation project EXTRA-LARGE. Support for the Twentieth Century Reanalysis Project is provided by the U.S. Department of Energy, Office of Science Innovative and Novel Computational Impact on Theory and Experiment (DOE INCITE) program, and Office of Biological and Environmental Research (BER), and by the National Oceanic and Atmospheric Administration Climate Program Office.

References

Baumann, T. and L. Reichen (2017) The Iberian Storm of 1941 in the Twentieth Century Reanalysis. In: Brönnimann, S. (Ed.) *Historical Weather Extremes in Reanalyses.* Geographica Bernensia G92, p. 113-122, DOI: 10.4480/GB2017.G92.10.

Brönnimann, S. and D. Krämer (2016) *Tambora and the „Year Without a Summer" of 1816. A Perspective on Earth and Human Systems Science.* Geographica Bernensia G90, 48 pp., DOI: 104480/GB2016.G90.02.

Brönnimann, S. and O. Martius (2013) *Weather extremes during the past 140 years.* Geographica Bernensia G89, p. 59-68, DOI: 104480/GB2013.G89.07.

Brönnimann, S., O. Martius, H. von Waldow, C. Welker, J. Luterbacher, G. P. Compo, P. D. Sardeshmukh, and T. Usbeck (2012) Extreme winds at northern mid-latitudes since 1871. *Meteorol. Z.,* **21,** 13–27.

Brohan, P., G. P. Compo, S. Brönnimann, R. J. Allan, R. Auchmann, Y. Brugnara, P. D. Sardeshmukh, and J. S. Whitaker (1816) The 1816 "year without a summer" in an atmospheric reanalysis. *Clim. Past Discuss.,* doi:10.5194/cp-2016-78.

Brugnara, Y., S. Brönnimann, M. Zamuriano, J. Schild, C. Rohr, and D. M. Segesser (2016) December 1916: Deadly Wartime Weather. *Geographica Bernensia,* G91, 8 pp., doi:10.4480/GB2016.G91.01.

Burkart, K. and A.-M. Wyss (2017) The "Great Appalachian Storm" of 1950 in Reanalyses and Historical Weather Charts. In: Brönnimann, S. (Ed.) *Historical Weather Extremes in Reanalyses.* Geographica Bernensia G92, p. 47-58, DOI: 10.4480/GB2017.G92.04.

Compo, G. P., J. S. Whitaker, P. D. Sardeshmukh, N. Matsui, R. J. Allan, X. Yin, B. E. Gleason, R. S. Vose, G. Rutledge, P. Bessemoulin, S. Brönnimann, M. Brunet, R. I. Crouthamel, A. N. Grant, P. Y. Groisman, P. D. Jones, M. Kruk, A. C. Kruger, G. J. Marshall, M. Maugeri, H. Y. Mok, Ø. Nordli, T. F. Ross, R. M. Trigo, X. Wang, S. D. Woodruff, and S. J. Worley (2011) The Twentieth Century Reanalysis Project. *Q. J. R. Meteorol. Soc.,* **137,** 1-28.

Cornes, R. C. (2014) Historic storms of the northeast Atlantic since circa 1700: a brief review of recent research. *Weather,* **69,** 121-125.

Cram, T. A., G. P. Compo, X. Yin, R. J. Allan, C. McColl, R. S. Vose, J.S. Whitaker, N. Matsui, L. Ashcroft, R. Auchmann, P. Bessemoulin, T. Brandsma, P. Brohan, M. Brunet, J. Comeaux, R. Crouthamel, B. E. Gleason, Jr., P. Y. Groisman, H. Hersbach, P. D. Jones, T. Jonsson, S. Jourdain, G. Kelly, K. R. Knapp, A. Kruger, H. Kubota, G. Lentini, A. Lorrey, N. Lott, S. J. Lubker, J. Luterbacher, G. J. Marshall, M. Maugeri, C. J. Mock, H. Y. Mok, O. Nordli, M. J. Rodwell, T. F. Ross, D. Schuster, L. Srnec, M. A. Valente, Z. Vizi, X. L. Wang, N. Westcott, J. S. Woollen, and S. J. Worley (2015) The International Surface Pressure Databank version 2. *Geoscience Data Journal,* **2,** 31-46.

Dangedorf, S., A. Arns, J. G. Pinto, P. Ludwig, and J. Jensen (2016) The exceptional influence of storm 'Xaver' on design water levels in the German Bight. *Environ. Res. Lett.,* **11,** 054001.

Dizerens, C., S. Lenggenhager, M. Schwander, A. Buck, and S. Foffa (2017) The 1956 Cold Wave in Western Europe. In: Brönnimann, S. (Ed.) *Historical Weather Extremes in Reanalyses.* Geographica Bernensia G92, p. 101-111, DOI: 10.4480/GB2017.G92.09.

Donat, M. G., D. Renggli, S. Wild, L. V. Alexander, G. C. Leckebusch, and U. Ulbrich (2011) Reanalysis suggests long-term upward trends in European storminess since 1871. *Geophys. Res. Lett.,* **38,** L14703.

Donat, M. G., L. V. Alexander, N. Herold, and A. J. Dittus (2016a) Temperature and precipitation extremes in century-long gridded observations, reanalyses, and atmospheric model simulations. *J. Geophys. Res.,* **121,** 11174-11189, doi:10.1002/2016JD025480.

Donat, M. G., A. D. King, J. T. Overpeck, L. V. Alexander, I. Durre, and D. J. Karoly (2016b) Extraordinary heat during the 1930s US Dust Bowl and associated large-scale conditions. *Clim. Dynam.,* **46,** 413-426.

Ernst, J., N. Glaus, M. Schwander, and M. Graf (2017) Reanalysis of the "Märzorkan" of 1876. In: Brönnimann, S. (Ed.) *Historical Weather Extremes in Reanalyses.* Geographica Bernensia G92, p. 23-34, DOI: 10.4480/GB2017.G92.02

Franke J., L. Di Marco, M. Pickl, C. Mock, K. Wood, and S. Brönnimann (2017) The "Sitka Hurricane" in south-eastern Alaska in the year 1880. In: Brönnimann, S. (Ed.) *Historical Weather Extremes in Reanalyses.* Geographica Bernensia G92, p. 59-67, DOI: 10.4480/GB2017.G92.05.

Gassner, S., L. Frigerio, M. Schwander, and C. Dizerens (2017) The "Great White Storm" of 1913. In: Brönnimann, S. (Ed.) *Historical Weather Extremes in Reanalyses.* Geographica Bernensia G92, p. 123-131, DOI: 10.4480/GB2017.G92.11

Giese, B. S., H. F. Seidel, G. P. Compo, and P. D. Sardeshmukh (2016) An ensemble of ocean reanalyses for 1815-2013 with sparse observational input. *J. Geophys. Res. Ocean.,* **121,** 6891–6910.

Grütter, J., S. Lehmann, R. Auchmann, O. Martius, and S. Brönnimann (2013) The heatwaves in Switzerland in summer 1947. In: Brönnimann, S. and O. Martius (Eds.) *Weather extremes during the past 140 years.* Geographica Bernensia G89, p.69-80, DOI: 10.4480/GB2013.G89.08.

Gudmundsson, L., J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen (2012) Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations - a comparison of methods. *Hydrology and Earth System Sciences,* **16,** 3383-3390.

Hersbach, H. and D. Dee (2016) ERA5 reanalysis is in production. *ECMWF Newsletter,* **147,** 7.

Hersbach, H., S. Brönnimann, L. Haimberger, M. Mayer, L. Villiger, J. Comeaux, A. Simmons, D. Dee, S. Jourdain, C. Peubey, P. Poli, N. Rayner, A. Sterin, A. Stickler, M. A. Valente, and S. Worley (2017) The potential value of early (1939-1967) upper-air data in atmospheric climate reanalysis. *Q. J. R. Meteorol. Soc.,* **143,** 1197-1210.

Hirahara S., I. Masayoshi, and Y. Fukuda (2014) Centennial-scale sea surface temperature analysis and its uncertainty. *J. Climate,* **27,** 57–75.

Ishida, K. and M. L. Kavvas (2017) Climate change analysis on historical watershed-scale precipitation by means of long-term dynamical downscaling. *Hydrol. Process.,* **31,** 35–50.

Jones, P. D., C. Harpham, and D. Lister (2016) Long-term trends in gale days and storminess for the Falkland Islands. *Int. J. Climatol.,* **36,** 1413-1427.

Kendon, M. and M. McCarthy (2015) The UK's Wet and Stormy Winter of 2013/2014. *Weather,* **70,** 40-47.

Kistler, R., E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, H. van den Dool, R. Jenne, M. Fiorino (2001) The NCEP-NCAR 50-year reanalysis: monthly means CD-ROM and documentation. *Bull. Amer. Meteorol. Soc.,* **82,** 247-267.

Krüger, O., F. Schenk, F. Feser, and R. Weisse (2013) Inconsistencies between Long-Term Trends in Storminess Derived from the 20CR Reanalysis and Observations. *J. Climate,* **26,** 868-874.

Laloyaux, P., E. de Boisséson, and P. Dahlgren (2017) CERA-20C: An Earth system approach to climate reanalysis. *ECMWF Newsletter,* **150,** 25–30.

Madec, G. and the NEMO team (2012) *NEMO ocean engine.* Note du Pole de modélisation de l'Institut Pierre-Simon Laplace, France, No 27 ISSN No 1288-1619.

Matthews, T., D. Mullan, R.L. Wilby, C. Broderick, and C. Murphy (2016) Past and future climate change in the context of memorable seasonal extremes. *Climate Risk Management,* **11,** 37-52.

Meyer, L., R. Hunziker, J. Weber, and A. Zürcher (2017) An Analysis of the "Great Gale of October 1881" using the Twentieth Century Reanalysis. In: Brönnimann, S. (Ed.) *Historical Weather Extremes in Reanalyses.* Geographica Bernensia G92, p. 91-100, DOI: 10.4480/GB2017.G92.10.

Michaelis, A. C. and G. M. Lackmann (2013) Numerical modeling of a historic storm: Simulating the Blizzard of 1888. *Geophys. Res. Lett.,* **40,** 4092–4097.

Moore, G. W. K. and M. Babij (2017) Iceland's Great Frost Winter of 1917/1918 and its representation in reanalyses of the twentieth century. *Q. J. R. Meteorol. Soc.,* **143,** 508–520.

Moore, P. (2015) *The weather experiment.* Chatto & Windus, London.

Parodi, A., L. Ferraris, W. Gallus, M. Maugeri, L. Molini, F. Siccardi, and G. Boni (2017) Ensemble cloud-resolving modelling of a historic back-building mesoscale convective system over Liguria: The San Fruttuoso case of 1915. *Clim. Past.,* **13,** 455-472.

Pérez-Zanón, N., M. C. Casas-Castillo, R. Rodríguez-Solà, J. C. Peña, J. G. Solé, and A. Redaño (2016) Analysis of extreme rainfall in the Ebre Observatory (Spain). *Theor. Appl. Climatol,* **124,** 935-944.

Pino, D., J. L. Ruiz-Bellet, J. C. Balasch, L. Romero-León, J. Tuset, M. Barriendos, J. Mazon, and X. Castelltort (2016) Meteorological and hydrological analysis of major floods in NE Iberian Peninsula. *J. Hydrol.,* **541,** 63-89.

Poli, P., H. Hersbach, D. Dee, P. Berrisford, A. Simmons, F. Vitart, P. Laloyaux, D. Tan, C. Peubey, J.-N. Thépaut, Y. Trémolet, E. Holm, M. Bonavita, L. Isaksen, and M. Fisher (2016) ERA-20C: An Atmospheric Reanalysis of the Twentieth Century. *J. Clim.,* **29,** 4085–4097.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late Nineteenth Century. *J. Geophys. Res.,* **108,** 4407, DOI: 10.1029/2002JD002670.

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax (2007) Daily High-Resolution-Blended Analyses for Sea Surface Temperature. *J. Climate,* **20,** 5473-5496.

Saha, S., S. Moorthi, H.-L. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, D. Behringer, H. Liu, D. Stokes, R. Grumbine, G. Gayno, J. Wang, Y.-T. Hou, H.-Y. Chuang, H.-M. H. Juang, J. Sela, M. Iredell, R. Treadon, D. Kleist, P. Van Delst, D. Keyser, J. Derber, M. Ek, J. Meng, H. Wei, R. Yang, S. Lord, H. Van Den Dool, A. Kumar, W. Wang, C. Long, M. Chelliah, Y. Xue, B. Huang, J.-K. Schemm, W. Ebisuzaki, R. Lin, P. Xie, M. Chen, S. Zhou, W. Higgins, C.-Z. Zou, Qu. Liu, Y. Chen, Y. Han, L. Cucurull, R. W. Reynolds, G. Rutledge, and M. Goldberg (2010) The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteorol. Soc.,* **91,** 1015-1057.

Stickler, A., S. Brönnimann, M. A. Valente, J. Bethke, A. Sterin, S. Jourdain, E. Roucaute, M. V. Vasquez, D. A. Reyes, R. Allan, and D. Dee (2014) ERA-CLIM: Historical Surface and Upper-Air Data for Future Reanalyses. *B. Amer. Meteorol. Soc.,* **95,** 1419–1430.

Stucki, P., S. Brönnimann, O. Martius, C. Welker, R. Rickli, S. Dierer, D. Bresch, G. Compo, and P. Sardeshmukh (2015) Dynamical downscaling and loss modeling for the reconstruction of historical weather extremes and their impacts - A severe foehn storm in 1925. *Bull. Amer. Meteor. Soc.,* **96,** 1233-1241.

Stucki, P., S. Dierer, C. Welker, J. Gómez-Navarro, C. Raible, O. Martius, and S. Brönnimann (2016). Evaluation of downscaled wind speeds and parameterised gusts for recent and historical windstorms in Switzerland. *Tellus A,* **68,** 31820.

Stucki, P., A. Baumann, and D. Bucher (2017) The 1978 heavy precipitation and flood event in Switzerland. In: Brönnimann, S. (Ed.) *Historical Weather Extremes in Reanalyses.* Geographica Bernensia G92, p. 69-80, DOI: 10.4480/GB2017.G92.06.

Villiger, L., M. Schwander, L. Schürch, L. Stanisic, and S. Brönnimann (2017) The "Royal Charter" Storm of 1859. In: Brönnimann, S. (Ed.) *Historical Weather Extremes in Reanalyses.* Geographica Bernensia G92, p. 35-45, DOI: 10.4480/GB2017.G92.03.

Wang, X. L., Y. Feng, R. Chan, and V. Isaac (2016) Inter-comparison of extra-tropical cyclone activity in nine reanalysis datasets. *Atmospheric Research,* **181,** 133-153.

Woodruff, S. D., S. J. Worley, S. J. Lubker, Z. Ji, J. E. Freeman, D. I. Berry, P. Brohan, E. C. Kent, R. W. Reynolds, S. R. Smith, and C. Wilkinson (2011) ICOADS release 2.5: Extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.,* **31,** 951-967.

Zamuriano, M., N. Imfeld, S. Hunziker, R. Peier, and G. Santi (2017) Reanalysis of a Cold Surge in Brazil in 1953. In: Brönnimann, S. (Ed.) *Historical Weather Extremes in Reanalyses.* Geographica Bernensia G92, p. 81-89, DOI: 10.4480/GB2017.G92.07.