Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 97 (2016) 124 – 132

Energy Procedia

European Geosciences Union General Assembly 2016, EGU
Division Energy, Resources & Environment, ERE

Reliable Averages and Risky Extremes – Analysis of spatio-temporal variability in solar irradiance and persistent cloud cover patterns over Switzerland

Annelen Kahl[a*], Juliette Blanchet[b], Bert Kruyt[a,c], Michael Lehning[a,c]

[a]*School of Architecture, Civil and Environmental Energineering, École Polytechnique Fédérale de Lausanne, Switzerland*
[b]*Laboratoire d'Etude des Transferts en Hydrologie et Environnement, Université Grenoble Alpes, Grenoble, France*
[c]*WSL Institute for Snow and Avalanche Research SLF, Davos, Switzerland*

Abstract

With the perspective of Switzerland's phase-out from nuclear energy, solar energy potential may take an important role for the country's future in renewable energy. Based on 12 years of satellite-derived, spatially distributed data of daily average cloud factor, this work analyses cloud cover persistence over all of Switzerland. We used extreme value analysis to study spatial variation of extremal behaviour and to estimate return levels of weeklong cloud cover events throughout Switzerland. Extreme events show different patterns in winter and summer, but overall they occur more frequently and more severely in the northern part of the country and in Ticino.

Peer-review under responsibility of the organizing committee of the General Assembly of the European Geosciences Union (EGU)

Keywords: Solar Power; Renewable Energy, MSG satellite imagery; Extreme Value Analysis; Swiss Energy Strategy 2050

1. Introduction

As the world is waking up to negative impacts of fossil fuel use and dangerous consequences of nuclear energy production, renewable resources are moving into the spotlight of public interest and research. A number of countries

* Corresponding author. Tel.: +41-21-693-80-93; fax: +41-21-693-63-90.
E-mail address: annelen.kahl@epfl.ch

Peer-review under responsibility of the organizing committee of the General Assembly of the European Geosciences Union (EGU)
doi:10.1016/j.egypro.2016.10.035

has partially completed or initiated an energy transition; each of them following their individual path determined by available energy potential as well as by policy making and public opinion.

In 2011 the federal council and government in Switzerland decided to phase out nuclear energy by not renewing existing power plants and to fill the resulting deficit with electricity production from renewable sources by 2050 [1]. As all other countries, Switzerland too has its individual predispositions for a future renewable energy production. Its extraordinary hydropower potential supplies 56% of annual electricity production, but it is almost completely exploited [2]. The amount of available wind potential and the acceptance of the associated infrastructure are largely uncertain. With deep geothermal exploitation on hold [3], solar energy remains the most promising candidate and will have to pick up the lion's share of the new renewable production. Particularly the southern part of the country receives fair amounts of solar radiation and photovoltaic panels are safe and easy to install. But solar power comes with certain well-known complications. First and foremost its periodic variability at daily and seasonal scale has the unfortunate characteristic of following the inverse of the demand, which is high in the morning, evening and in winter and low in summer and at midday. While this can at least be foreseen and may be solved in Switzerland by a combination of novel (Hydrogen) and established (pumped-hydro) storage solutions, the more daunting task will be to adjust to the unpredictable short-term intermittency resulting from cloud cover. Sudden and short fluctuations due to individual clouds stress lines and network components. Long periods of persistent cloud cover challenge the electricity provider to reliably meet the demand. In order to better prepare for this last challenge, it is important to study cloud patterns throughout time and space and to analyse not just the average behaviour but also the extremes. Because it is the extreme conditions that will pose the biggest risk to a reliable supply and a healthy functioning of the electricity system. Previous assessments of Switzerland's future in renewable energies have mostly looked at temporally and/or spatially average conditions or short time series [4][5]. This study uses extreme value analysis of satellite-derived gridded cloud cover information to assess worst-case scenarios for losses in solar power production. We focus on the persistence of cloud cover only and will not address high persistence of sunshine. The latter potentially also threatens the electrical system since more electricity might be produced than can be distributed through the lines.

2. Methods

2.1. Extreme Value Analysis – Model Approach

Extreme Value analysis describes the statistical behaviour of the extremes of a distribution of random variables $X_{1...},X_n$ [6–8]. Let $Z_n = \max(X_1,\ldots X_n)$ denote these extremes. One common approach is to subset the data set into fixed blocks and to pick the highest value from each block. For time series data such a block might be, for example, a month or a year. The resulting extremes are called block-maxima and their distribution follows the Generalized Extreme Value (GEV) distribution:

$$G(z) = \begin{cases} exp\left[-\left(1+\xi\left(\frac{z-\mu}{\sigma}\right)\right)^{-1/\xi}\right], & for\ 1+\xi\left(\frac{z-\mu}{\sigma}\right) > 0 \\ 0 & otherwise \end{cases} \tag{1}$$

The GEV model has three parameters that describe the extreme tail of the distribution. The location parameter μ quantifies the centre of the tail and is thus comparable to the average of the extremes. The scale parameter $\sigma > 0$ measures the spread of the tail, and the shape parameter ξ describes its form. Distinction is being made between heavy tails, with many extreme events, light tails with few extreme values and bound tails that have a maximum upper value. The downside of this approach is that only a very small portion of the existing data can contribute to the analysis and if the existing time series is short, this method might be considered wasteful. Hence another approach is commonly used, the so-called peak over threshold (POT) method [9]. It characterizes every value exceeding a certain threshold u as extreme. The choice of u is critical to the success of these methods. It needs to be high enough to ensure that the extreme value requirements apply for the corresponding exceedances. But it also needs to be low enough to allow for an adequate number of values to contribute to the analysis. There are two representations of POT processes. The first models the magnitude of the extremes, conditional on the fact that they exceed the threshold, with the Generalized Pareto distribution (GPD) [9]. Since these exceedances are a function of the chosen threshold, the resulting model parameters cannot directly be compared to the parameters of the GEV. The second approach adopts a Poisson Point (PP) Process representation, which means that it models the sequence of occurrence and magnitude of the extremes as a 2-dimensional non-homogeneous Poisson process [10]. Model parameters are estimated using maximum likelihood: It selects the set of model parameters that renders the empirical data set most probable. Or differently said: the model parameters that produce a distribution, which agrees most with the measured extremes, will be chosen. The model parameters of the PP representation are identical to those of the GEV distribution and allow for the estimate of return levels z_p, which is very useful to quantify the likelihood of specific extreme events. These methods have been widely applied to hydrologic [11], financial [12] and climatologic [13] questions and more recently also to ecology [14] and snowfall [15,16].

2.2. Return Levels

The return level z_p is the level that is expected to be exceeded on average once every $1/p$ blocks (e.g. years). They represent hence the $(1 - p)th$ quantile of the GEV and can be estimated by inverting the distribution (equation 1):

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi}\left[1 - \{-log(1-p)\}^{-\xi}\right], & for\ \xi \neq 0 \\ \mu - \sigma log\{-log(1-p)\}, & for\ \xi = 0 \end{cases} \tag{2}$$

We compute return levels over different time periods for each pixel in Switzerland, to get an idea of the cloudiest week that will occur on average every 2, 10 and 100 years and we will show their spatial distribution throughout the country. All calculations were carried out using different R packages [ismev, extRemes, evd].

3. Data

3.1. Justification of Data Choice – Why Meteosat Second Generation?

For our analysis we picked data derived from MeteoSat Second Generation (MSG) imagery [17], because it fulfilled all necessary requirements:

- Spatially distributed to study spatial patterns and temporal correlations between locations
- Spatial resolution high enough to capture topography (~1km)
- Time series long enough to allow for reliable extreme value analysis and to cover many weather anomalies
- Temporal resolution high enough to capture daily or sub-daily extremes
- Good snow-cloud discrimination to avoid over-estimation of cloud cover in winter and at high elevations

There is a growing number of satellite derived radiation products that are either freely or commercially available. An overview is given in [18]. Over mountainous terrain the accuracy normally goes down because the complex topography creates additional challenges. Particularly the discrimination between snow and clouds requires an extended use of the electromagnetic spectrum. The spectral profiles of the two surfaces are very similar over large wavelength intervals and in some instances it is even impossible to distinguish the two [19]. Based on a detailed analysis of the MeteoSat imagery, we found that the MeteoSat first generation (MFG) products notoriously misclassify snow as clouds and consequently overestimate cloud cover at high elevations and in winter (Figure 1) while MeteoSat second generation is more realistic. Since radiation is derived from cloud cover information, an overestimate in cloud cover leads to an underestimate in Surface Incoming Shortwave radiation (SIS).

Plotting the difference between the two products as a function of elevation (Figure 1a) and time (Figure 1b) confirms the systematic underestimate in SIS from MFG. Only below 1500m where snow cover is rare, the two products show comparable radiation estimates. And the highest underestimates in SIS are observed during the winter months, when large parts of the country are snow covered. Currently only the so-called HelioMont [20] algorithm, a MSG-based radiation product published by MeteoSwiss is able to capture this important difference.

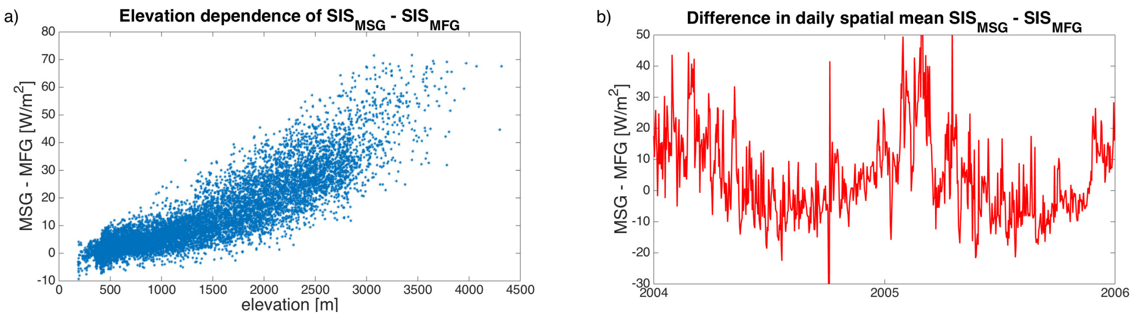

Figure 1: *a) Difference of the temporal mean SIS from MFG and MSG during 2004-2005 as function of elevation. Each dot represents the difference between corresponding pixels in the two dataset. Their respective elevation is derived from a SRTM DEM* [21]. *b) Difference in spatial average SIS from MFG and MSG over all of Switzerland for each day of the period 2004-2006.*

As a result of these considerations, we decided to work with the MSG-derived products provided by MeteoSwiss. For the extreme value analysis we did not use the absolute radiation value in W/m^2 that is given in the SIS product. Instead we analyzed the so-called cloud factor (CF), which is calculated as a pre-processing step to SIS. It quantifies the proportion of theoretical clear-sky radiation that is filtered out by the clouds, which were present at that instance. It is hence a measure for cloudiness that allows for a direct comparison between each and any pixel, independent of the respective pixel's elevation or geographic location. Since this part of our analysis aims at studying the spatial patterns and temporal variability of persistence in cloud cover and not the absolute solar potential, the cloud factor seems to be the most adequate choice.

3.2. Capturing Cloud Cover Persistence

A multi-day moving average was calculated, to capture the persistence of cloud cover rather than the daily cloud cover value itself. But how long is persistence? And how many clouds make it cloudy? Any persistence event is characterized by two parameters: 1) duration (in days) and 2) minimum cloud clover during this period.

Figure 2 gives a quantitative idea of how often events with a certain cloud cover occur and how long they last.

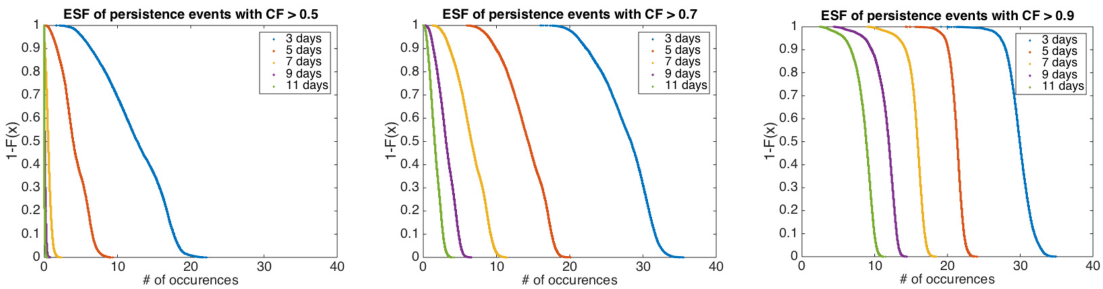

Figure 2*: Experimental survival functions (ESF) of number of occurrences of certain persistence events per year for all pixels in Switzerland.*

As one would expect, there are more 3-day events for any threshold than there are 11-day events. The higher the minimum cloud cover, the fewer the number of events. One interesting finding, which is not obvious, is the change in shape of ESF for high cloud thresholds. Almost all pixels experience about 30 three-day persistence events with a CF threshold of 0.1 per year. But when we look at the ESF of 3-day events with a threshold of CF = 0.5, we see that the slope is less steep. Counts vary from as little as 4 to as much as 20, with about equal fractions of pixels along this spectrum. Having accounted for frequency and magnitude of persistence events, we picked a 7-day window for the extreme value analysis because it is long enough to significantly affect the electricity supply, but it is short enough such that its frequency of occurrence allows for a statistical analysis.

We also looked at how the frequency of persistence events varies throughout the country to find out where long and cloudy periods are most likely to happen. The relative spatial patterns of annual average event counts are similar for various persistence lengths. Hence we focus on the 7-day case for three different thresholds (*Figure 3*). An entire week with a cloud factor of at least 0.5 is quite rare throughout the country as we see in *Figure 3a*. For a threshold of CF > 0.3 the number of average annual occurrences ranges from 1 to 11 (*Figure 3b*), with a clear north south gradient. The occurrence of low cloud cover persistence is more uniform (*Figure 3c*) (as has already become apparent in

Figure 2c). Only southern Ticino and the Rhone valley clearly stand out with low values.

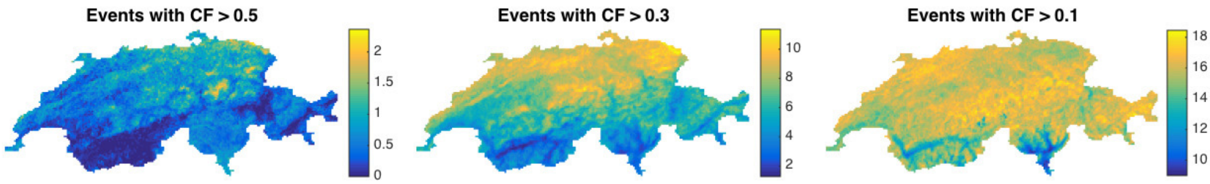

Figure 3: Spatial distribution of average annual occurrence of 7-day persistence events for three thresholds.

Stark difference become apparent, when we subset the year to seasonal averages (*Figure 4*). The spatial patterns in summer show strong elevation dependence, in addition to regionally high values on an east/west stripe across the centre of the country. The lowest values are found in the Rhone valley, the south of Ticino and the western plateau. This is a manifestation of convective cloud formation during the warm season, when a lot of mixing and rising of moist air occurs. In winter this regional trend is almost inverted due to stable stratification and fog formation in the plateau, the entire north of the country and at low elevation in the southern half.

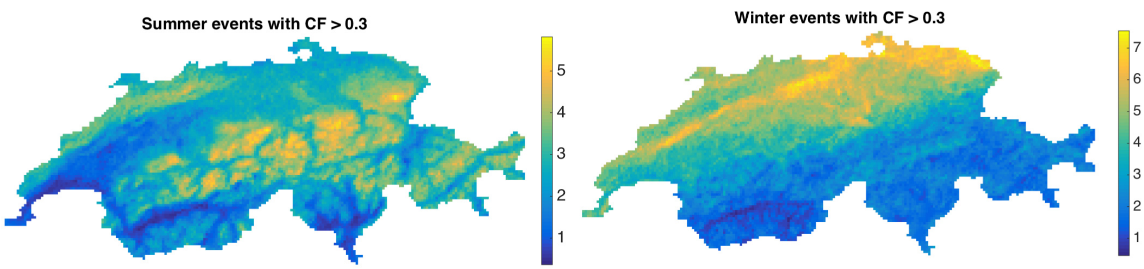

Figure 4: Spatial patterns as in Figure 3 subset into seasonal trends.

3.3. Preprocessing the persistence data for Extreme Value Analysis

7-day cloud cover persistence

Under the considerations of section 3.2 a 7-day moving window seems an appropriate length to study. Since there is no unique definition of cloud cover persistence, we explored two different ways of quantifying it.

Approach 1: We measure extremely cloudy events by the lowest CF value that is reached during each 7-day period. Hence the extreme value analysis was conducted on the upper tail of this measure's distribution (i.e., the maxima of the minimum CF per week). Correspondingly sunny events would be described by the highest cloud factor value of the 7-day period and the extreme value analysis would estimate the lower tail of this distribution (i.e., the minima of the maximum CF).

Approach 2: Persistence in cloud cover is quantified by the 7-day moving average of the daily cloud factor. In this case, the upper tail of the distribution represents persistence of high cloud cover, the lower tail of the same distribution represents low cloud cover.

Both formulations qualify as proxies for cloud cover persistence, but each one carries a slightly different interpretation. Approach 1 adheres to the idea of persistence in a sense of 'without interruption', because the entire 7-day window may never drop below the characteristic CF value. Approach 2 still yields a low or high value if one of the days of the periods is an outlier. Correspondingly the values in the latter representation are more extreme on both ends.
While approach 1 in our opinion is the most precise quantification of persistence in cloud cover, we chose to show and discuss the result of approach 2, for practical reasons. The first formulation is complicated, because it involves the minimum of the distribution of maxima and vice versa. Furthermore, for PV production, a short-duration cloud cover does not pose a significant problem. Finally, a comparison of the two shows that the spatial patterns are similar. So for the sake of clarity we will from now on refer to the 7-day moving average cloud factor as *cloud cover persistence*.

Detrending
In order to make the data set eligible for extreme value analysis, underlying periodic trends and correlation within successive days need to be removed. We fit a linear model including two Fourier terms to the data, subtracted it from the entire time series and added back the overall mean to avoid affecting the location parameter in the extreme value model.

Declustering
Since the extremal data set is the results of a 7-day moving window computation, it has a high degree of clustering within 3 days, which can lead to an overestimate of extreme events if not properly addressed. The extremal coefficient θ is a measure of limiting mean cluster size ([6], chapter 5.2). In the original form, all pixels have θ-values of the order of 0.3 in accordance with the 3.5-day overlap resulting from the moving window operation. For declustering we used the decluster function that is included in the {R extRemes} package [22]. The function detects all threshold exceedences that are separated by fewer than 3 days and sets them to NA. As a result we find θ equals or very close to 1, a necessary (but not sufficient) requirement for an independent series.

Threshold choice
There are several ways to estimate an appropriate threshold for POT approaches, such as parameter stability and mean residual life plot (MRL) [6]. Both give some visual guidance that could assist an informed decision. However our dataset includes about 12000 time series, which have to be fitted individually. It is unreasonable to do this by visual inspection. The distributions vary quite significantly between the pixels and it is also not possible to find an appropriate threshold based on the analysis of a subset of pixels. A threshold that works well for some of the pixels will generate too small or too large a number of exceedances for others. Instead of a universal threshold, we consider the upper five percentile of each individual pixel's dataset as the extreme values.

4. Results

For our extreme value analysis of weeklong cloud cover persistence, we chose peak over threshold modelling, using the Poisson Point representation. Point-wise maximum likelihood estimates of the three GEV parameters and return levels for each pixel in Switzerland give an overall notion of how the distribution of extremes varies throughout the country

4.1. Model Parameters

The location parameter gives the centre of the tail that could be compared to the average of the extremes. The lowest values between 0.5 and 0.6 are found in the Cantons of Wallis and Graubünden. In the Northern part of the country and in Ticino the extremes centre at much higher cloud factor values of about 0.7. The spread of the tails displays a spatial pattern that resembles the elevation profile of the country. Pixels at high elevations tend to have higher scale values than pixels in the valley. The scale parameter is anti-correlated with the location parameter, because the value range of CF has an upper bound. The distribution of shape parameters is concentrated between -0.6 and -0.1. There are only very few positive shape parameters (indicative of heavy tails) and they are spatially correlated with high shape values and thus also with high elevations.

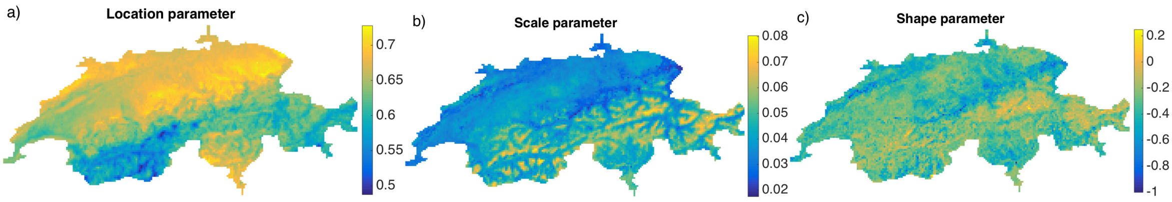

Figure 5: *Model parameters a) location parameter, b) scale parameter, c) shape parameter*

4.2. Return Levels

The return levels indicate the worst, weeklong cloud cover persistence that is to be expected once every *1/p* years. It is important to note that even though we plot the results for each pixel in the same map, they do not necessarily occur simultaneously. So while Zürich might experience its cloudiest week in January, it might well be February for Geneva. The spatial patterns of strong cloud cover are similar to the location parameter, but with increasing return periods the patterns show a slight shift, toward high values at high elevations (with upper limit set to 1). The histograms next to the spatial plots show not just a general shift to higher values, but also a change from a screwed to a more symmetric distribution of the extremes.

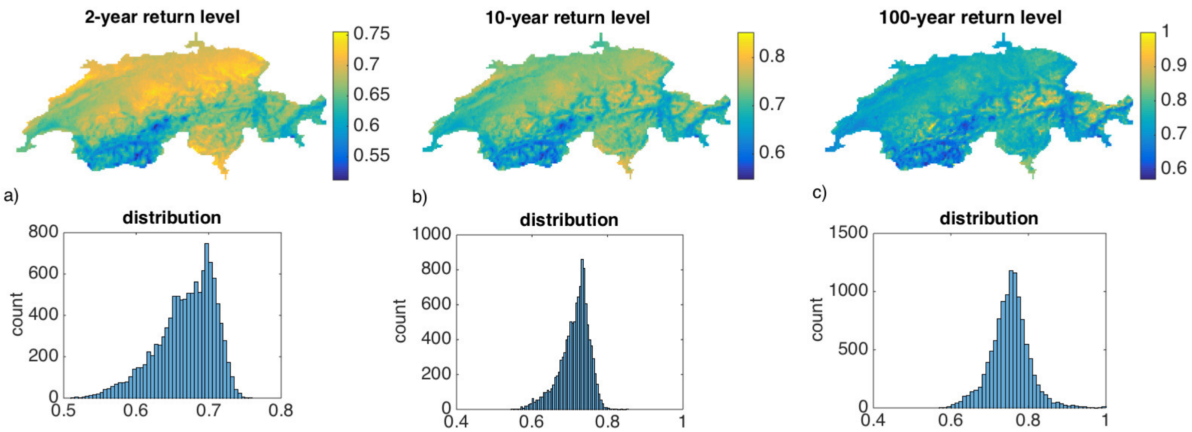

Figure 6: Spatial and statistical distribution of return levels of highest cloud cover persistence for; a) 2-years, b) 10-years, c) 100 years

5. Conclusions

This study shows the application of extreme value analysis to satellite derived gridded cloud cover information with the goal to assess spatial and temporal characteristics of weeklong cloud cover persistence, in order to assess risks and worst-case scenarios of losses in solar power production. This information is useful for grid management and energy production companies, which are concerned about a stable supply mix of energies. For an individual installation, the production potential at a particular site may be more important, though. The actual assessment of productivity of a location however needs to be derived from radiation data. The results may also serve as a safety check to verify whether a location displays extraordinarily strong or frequent extreme conditions.

When analysing cloud cover trends, good input data sets are only available for a limited time span.
Our comparison of two radiation products derived from different satellite sensors (MFG and MSG) show large discrepancies. For Switzerland in particular it is important to work with a product that is able to distinguish cloud and snow, otherwise the radiation that reaches the surface will be underestimated in winter and at high elevations.
Of all modelled GEV parameters the location value is the most informative one in our analysis. It is interesting to see that regions with low location parameter tend to have higher scale and shape parameters than the pixels with high location parameters. It means that in places with long periods of cloudy weather, we will not have many 'really extreme' events (those which make up the extreme end of the tail). This is the case for the yellow regions Figure 5a). However the modelled scale parameters span only 6% of the possible value range, while the location parameter varies from 0.5 to 0.72 (~22% of value range). The shape parameter indicates almost exclusively bound tails for all pixels in Switzerland.
Consistently with the location parameter, the key information that can be gained from the return level analysis is the severity of rare events. If for example we define that a future power system needs to be built such that it can handle the 100-year cloud persistence, we know that it will have to be able to compensate for a one-week loss of a certain percentage (which can be retrieved from the return level maps) of productivity from solar power. The exact conversion from change in cloud factor to resulting change in productivity of a PV panel is beyond the scope of this paper.
Overall the spatial patterns of the location parameter and return values resemble the patterns of *Figure 3*. Even though the two figures do not show the same variable, their respective meaning is comparable; high values correspond to more clouds, and they are primarily located in the Northern part of the country. For most of the country the average statistics (*Figure 3*) reflects the behaviour of the extremes (Figure 5*a,* Figure 6), only Ticino is an exception. Here location parameters as well as return level values are high, while the bulk of the persistence events is characterized by low occurrences at all lengths of persistence. For Switzerland this means that placing a lot of PV panels in Ticino promises high production on average, while simultaneously threating with more frequent extreme events. And for the rest of the world this means that conducting an analysis of extremal behaviour in addition to the standard statistics is important, because it might highlight zones where the risk is higher than one would anticipate from average statistics alone.

Acknowledgements

The authors would like to thank Anthony Davison and Leo Belzile for valuable advice with the extreme value analysis. Reto Stoeckli, Abhay Devasthale and Jörg Trentmann candidly guided our search of satellite products. This project is funded by the EPFL Energy Center and the SCCER SoE.

References

[1] (SFOE) SFO of E. Energiestrategie 2050 n.d. http://www.bfe.admin.ch/energiestrategie2050/index.html (accessed June 3, 2016).
[2] Blumer Y, Martin S. Die Schweizer Wasserkraft, Struktur-und Schwachstellenanalyse basierend of Interviews mit Branchenexperten. ETH-NSSI Arbeitspapier 2012.
[3] Giardini D. Geothermal quake risks must be faced. Nature 2009;462:848–9. doi:10.1038/462848a.
[4] Redondo PD, Van Vliet O. Modelling the Energy Future of Switzerland after the Phase out of Nuclear Power Plants. Energy Procedia, 2015. doi:10.1016/j.egypro.2015.07.843.
[5] Schlecht I, Weigt H. Swissmod - A model of the Swiss Electricity Market. WWZ-Discussion Pap 2014.
[6] Coles S. An Introduction to Statistical Modeling of Extreme Values. London: Springer; 2001.
[7] Haan L de, Ferreira A. Extreme Value Theory: An Introduction. New York: Springer; 2006. doi:10.1007/0-387-34471-3.
[8] Resnick SI. Extreme values, regular variation, and point processses. Springer; 1987. doi:10.1007/978-0-387-75953-1.
[9] Davison AC, Smith RL. Models for exceedances over high thresholds. J R Stat Soc Ser B Methodol 1990;52:393–442. doi:10.2307/2345667.
[10] Leadbetter MR, Lindgren G, Rootzen H. Extremes and related properties of random sequences and processes. New York: Springer; 1983.
[11] Katz R. Statistics of Extremes in Climatology and Hydrology 2002;25:1–48.
[12] Embrechts P, Klueppelberg C, Mikosch T. Modelling extremal events for insurance and finance. Berlin: Springer; 1997.

[13] Brown B, Katz R. Regional Analysis of Temperature Extremes: Spatial Analog for Climate Change? J Clim 1993;8:108–19.

[14] Katz RW, Brush GS, Parlange MB. Statistics of extremes: Modeling ecological disturbances. Ecology 2005;86:1124–34. doi:10.1890/04-0606.

[15] Blanchet J, Marty C, Lehning M. Extreme value statistics of snowfall in the Swiss Alpine region. Water Resour Res 2009;45. doi:10.1029/2009WR007916.

[16] Blanchet J, Lehning M. Mapping snow depth return levels: smooth spatial modeling versus station interpolation. Hydrol Earth Syst Sci 2010;14:2527–44. doi:10.5194/hess-14-2527-2010.

[17] Meteosat — EUMETSAT n.d. http://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Meteosat/index.html (accessed June 3, 2016).

[18] Vernay C, Pitaval S, Blanc P. Review of Satellite-based Surface Solar Irradiation Databases for the Engineering, the Financing and the Operating of Photovoltaic Systems. Energy Procedia 2014;57:1383–91. doi:10.1016/j.egypro.2014.10.129.

[19] Rittger K. Spatial estimates of snow water equivalent in the Sierra Nevada. University of California, Santa Barbara, 2012.

[20] Stoeckli R. The HelioMont surface solar radiation processing. Sci Rep MeteoSwiss 2013:122 pp.

[21] Farr T, Kobrick M. The shuttle radar topography mission. Eos Trans AGU 2001;82:47. doi:10.1029/2005RG000183.1.INTRODUCTION.

[22] Gilleland E, Katz RW. New software to analyze how extremes change over time. Eos (Washington DC) 2011;92:13–4.