



Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

This discussion paper is/has been under review for the journal Hydrology and Earth System Sciences (HESS). Please refer to the corresponding final paper in HESS if available.

A quantitative analysis to objectively appraise drought indicators and model drought impacts

S. Bachmair¹, C. Svensson², J. Hannaford², L. J. Barker², and K. Stahl¹

¹Hydrology, Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany

²Centre for Ecology and Hydrology, Wallingford, UK

Received: 20 August 2015 – Accepted: 30 August 2015 – Published: 18 September 2015

Correspondence to: S. Bachmair (sophie.bachmair@hydrology.uni-freiburg.de)

Published by Copernicus Publications on behalf of the European Geosciences Union.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Abstract

Drought monitoring and early warning is an important measure to enhance resilience towards drought. While there are numerous operational systems using different drought indicators, there is no consensus on which indicator best represents drought impact occurrence for any given sector. Furthermore, thresholds are widely applied in these indicators but, to date, little empirical evidence exists as to which indicator thresholds trigger impacts on society, the economy, and ecosystems. The main obstacle for evaluating commonly used drought indicators is a lack of information on drought impacts. Our aim was therefore to exploit text-based data from the European Drought Impact report Inventory (EDII) to identify indicators which are meaningful for region-, sector-, and season-specific impact occurrence, and to empirically determine indicator thresholds. In addition, we tested the predictability of impact occurrence based on the best performing indicators. To achieve these aims we applied a correlation analysis and an ensemble regression tree approach (“random forest”), using Germany and the UK (the most data-rich countries in the EDII) as a testbed. As candidate indicators we chose two meteorological indicators (Standardized Precipitation Index (SPI) and Standardized Precipitation Evaporation Index (SPEI)) and two hydrological indicators. The analysis revealed that accumulation periods of SPI and SPEI best linked to impact occurrence are longer for the UK compared with Germany, but there is variability within each country, among impact categories and, to some degree, seasons. The median of regression tree splitting values, which we regard as estimates of thresholds of impact occurrence, was around -1 for SPI and SPEI in the UK; distinct differences between northern/northeastern vs. southern/central regions were found for Germany. Predictions with the ensemble regression tree approach yielded reasonable results for regions with good impact data coverage. The predictions also provided insights into the EDII, in particular highlighting drought events where missing impact reports reflect a lack of recording rather than true absence of impacts. Overall, the presented quantitative framework proved to be a useful tool for evaluating drought indicators, and

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



to model impact occurrence. In summary, this study demonstrates the information gain for drought monitoring and early warning through impact data collection and analysis, and highlights the important role that quantitative analysis with impacts data can have in providing “ground truth” for drought indicators alongside more traditional stakeholder-led approaches.

1 Introduction

Drought is less tangible than other natural hazards, such as earthquakes or floods, due to its slow onset, “creeping” nature, and complex, often non-structural impacts (Gillette, 1950; Wilhite et al., 2007). Nonetheless, drought is known to affect more people than any other hazard, and to cause high economic loss (Loayza et al., 2012; Wilhite et al., 2007). While droughts cannot be prevented, societal vulnerability can be reduced, with monitoring and early warning (hereafter, M&EW) being one important measure to enhance drought resilience. The aim of M&EW is to provide adequate and timely information on drought conditions to enable people and organizations to be better prepared and react accordingly (Svoboda et al., 2002; Wilhite and Svoboda, 2000). Such systems are usually based on several drought indicators representing different domains of the hydrological cycle, i.e. indicators for meteorological drought, soil moisture drought and vegetation stress, hydrological drought, and groundwater drought.

A recent trend has been the design of “combined” or “multivariate” indicators consisting of a blend of individual ones. The rationale behind the construction of blended indicators is that a single indicator is not sufficient to adequately capture different types of drought, and the corresponding multiplicity of drought impacts that differ markedly in response time (Hao and Singh, 2015). There have been several studies assessing the link between indicators of different types of droughts, e.g. between meteorological drought and streamflow, soil moisture, or remotely sensed vegetation stress indicators (Haslinger et al., 2014; Ji and Peters, 2003;

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Martínez-Fernández et al., 2015; Vicente-Serrano and López-Moreno, 2005; Vicente-Serrano et al., 2012). These are useful when there is an assumption that the lag between, say, meteorological and hydrological drought represents the response time for impact occurrence in, say, riverine ecosystems. Drought indicator choices can be substantiated by stakeholder consultation or expert judgement, as has been implemented for the operational US Drought Monitor (Svoboda et al., 2002). Similar initiatives have been developed in research project settings in southwest Germany (Stölzle and Stahl, 2011) and Switzerland (Kruse et al., 2010).

However, while indicators representing different types of drought are commonly used as proxies for impact occurrence, there is, to date, little empirical evidence as to which indicator best represents drought impact occurrence for any given sector. Lackstrom et al. (2013) identified an impact-driven perspective as the “missing piece” of drought monitoring; what is of ultimate interest is knowledge of when and where a precipitation shortfall or low streamflow or groundwater level will translate into impacts on society, the economy, and ecosystems. A direct, empirical evaluation of drought indicators with impact information would obviate the need for assumptions based on intercomparing different drought indicators.

Aside from identifying indicators important for drought impacts, there is a need for a better understanding of the meaning of indicator thresholds used for drought declaration and as triggers for management actions in drought plans. Such thresholds are mostly based on hazard intensity classes corresponding to a certain frequency of occurrence, e.g. following the widely accepted Standardized Precipitation Index scheme, with classes ranging from 0 to -0.99 (mild drought), -1 to -1.49 (moderate drought), -1.5 to -2 (severe drought), and < -2 (extreme drought) (McKee et al., 1993). The US Drought Monitor (USDM) differentiates between five drought severity classes based on several indicators and corresponding thresholds (Svoboda et al., 2002). Different thresholds again are used for delineating alert classes of the Combined Drought indicator of the European Drought observatory (European Drought observatory, 2013).

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

HESSD

12, 9437–9488, 2015

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Common to all thresholds is that they are arbitrary cut-off points (e.g. McKee et al., 1993; Svoboda et al., 2002). A survey among drought managers in the US on drought plans and respective indicators and triggers revealed that there is large uncertainty in the selection of thresholds, with one survey reply uncovering that most states selected their indicators “out of a hat” without knowing whether they “worked” (Steinemann, 2014). There is currently no consensus on appropriate drought indicators and thresholds meaningful for practitioners of different sectors.

Regarding drought prediction, a substantial body of research has been dedicated to forecasting drought indicators with sufficient lead time (e.g. Dutra et al., 2014; Mehta et al., 2014; Trambauer et al., 2014; Wetterhall et al., 2015). However, while the models used for forecasting may propagate the climate signal into soils and hydrology, they do not include a further link to the tangible negative environmental and socio-economic impacts of a particular drought. Models bridging the gap between drought indicators and impacts are rare. While predictions of crop yield are more common (e.g. Hlavinka et al., 2009; Mavromatis, 2007; Qiring and Papakryiakou, 2003), very few studies have tested approaches for modeling other types of drought impacts (e.g. Blauhut et al., 2015; Stagge et al., 2014; Gudmundsson et al., 2014; and Vicente-Serrano et al., 2012). The complexity of processes and the interconnectedness of the multitude of drought impacts, which may occur with much delay and even outside of the hazard affected area (Logar and van den Bergh, 2013; Wilhite et al., 2007), may be one reason why few drought impact models have been presented.

The most important obstacle, however, is a paucity of information on drought impacts. Initiatives to rectify this include the US Drought Impact Reporter (DIR) (Wilhite et al., 2007), and the more recently developed European Drought Impact report Inventory (EDII) (Stahl et al., 2015a). Both provide text-based, categorized information on reported drought impacts. The majority of impacts of the US DIR stem from online media clipping (Wilhite et al., 2007), meaning that it can be used as a real-time monitoring tool. In contrast, the EDII is designed as a research database with a focus on past drought events. Other potential sources of drought impact data are

reported crop yields, or losses assembled in the Emergency Events Database EM-DAT (www.emdat.be) or by re-insurance companies. Nevertheless, crop yield reductions may not necessarily be due to drought and loss data mostly provides aggregated information on large events without details on the temporal and spatial evolution of impacts, which is essential for empirically validating indicators and developing drought impact models.

Only very few studies to date have exploited text-based impact datasets. Dieker et al. (2010) qualitatively and quantitatively compared the USDM to impact data from the US DIR. Stagge et al. (2014) and Blauhut et al. (2015) both worked with EDII data at the country- or macro-region-scale across Europe, with impacts coded as a binary response variable (impact vs. no impact) to determine the likelihood of impact occurrence for different impact types. Bachmair et al. (2015) also used EDII data to test the feasibility of evaluating drought indicators with impacts at smaller spatial scales in Germany. As an extension to Stagge et al. (2014) and Blauhut et al. (2015), they replaced the binary data with the number of impact occurrences, thus providing a measure of impact severity. A correlation analysis and extraction of indicator values concurrent with past impact onset showed variability in indicator performance and onset thresholds at the sub-country scale and between drought events. The effect of different impact categories or types was not assessed (Bachmair et al., 2015).

Building on these previous efforts, the aim of this study is to exploit the EDII to link drought indicators to impacts using quantitative methodologies. Germany (DE) and the UK were selected as a test-bed, since they represent the countries with most impact data in the EDII database, but the aim is to develop methods that can be extended to other geographical areas in future applications. Specifically, the aims are to

- evaluate different drought indicators using text-based impact information to identify indicators that are meaningful for region-, sector-, and season-specific impact occurrence,

HESSD

12, 9437–9488, 2015

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- to empirically determine indicator thresholds representative for impact occurrence, as an alternative to using the default, arbitrarily selected hazard class thresholds intrinsic to indicators such as the SPI,
- to model impact occurrence via machine learning to assess the potential for predictive purposes (i.e. predicting impacts based on indicators alone), and exploit the relationships between indicators and text-based impact data to “backwards learn” about the nature of the impact data itself.

2 Data

2.1 Spatial and temporal resolution

As temporal and spatial resolution of the drought indicator and impact data we selected monthly time series for the period 1970–2012, aggregated at the NUTS1 level (level 1 of the Nomenclature of Units for Territorial Statistics, a spatial unit used in the European Union). NUTS1 regions represent major socio-economic regions. In Germany they correspond to the federal states. In the UK there are 12 NUTS1 regions, in Germany 16 (see Table 1 for a list of NUTS1 regions considered for analysis and abbreviations used in this study). Note that two NUTS1 regions in the UK and three in Germany were excluded from the analysis due to having insufficient impact data (see Sect. 2.3 for details).

2.2 Drought indicators

As drought indicators we selected the Standardized Precipitation Index (SPI) (McKee et al., 1993), the Standardized Precipitation Evaporation Index (SPEI) (Vicente-Serrano et al., 2010), and streamflow percentiles (Q). In addition, groundwater level percentiles (G) were included for Germany. For the SPI and SPEI, accumulation periods of 1–8, 12, and 24 months were chosen. Gridded SPI and SPEI data were

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

calculated based on E-OBS gridded data (version 9.0; 0.25° regular spatial grid, Haylock et al., 2008) using the R Package “SCI” (Stagge et al., 2015). The gamma distribution was used for the computation of the SPIs and the generalized logistic distribution for the SPEIs (reference period: 1971–2010). Potential evapotranspiration for the SPEI was estimated using the Hargreaves method (Hargreaves, 1994). For each NUTS1 region, regional averages of mean monthly SPI- n or SPEI- n were calculated. Here, n denotes the accumulation period. The mean was chosen since Bachmair et al. (2015) found little differences between the performance of different indicator metrics per spatial unit (e.g. mean vs. minimum, or 10th percentile vs. percent area with SPI or SPEI below a threshold). The reference period for calculation of streamflow percentiles is 1960–2012 in the UK, and 1970–2011 in Germany (also for groundwater).

The monthly streamflow percentiles are based on monthly mean streamflows. In Germany these are calculated from daily streamflow records for several gauging stations per federal state; monthly groundwater percentiles come from weekly to monthly readings of groundwater levels or spring discharge for several monitoring stations per state (data provision by different agencies of the German federal states, see Kohn et al., 2014). Many of these stations are used for the federal states’ hydrological forecasting systems and thus represent stations with good data quality. Monthly streamflow records for the UK were taken from daily river flow records held on the UK National River Flow Archive (NRFA) (www.ceh.ac.uk/data/nrfa/index.html). The UK Benchmark Network (Bradford and Marsh, 2003) of near-natural catchments was used, alongside the network of sites used in the National Hydrological Monitoring Programme (NHMP: <http://www.ceh.ac.uk/data/nrfa/nhmp/nhmp.html>). No groundwater measurements were used from the UK due to the limited number of NHMP borehole records available in many NUTS regions, reflecting the concentration of productive aquifers in the south and east of the country.

The streamflow gauging stations in the UK and Germany encompass both natural and anthropogenically influenced catchments. Figure 1 displays the spatial location

of Q and G measurement stations and the boundaries of the NUTS1 regions in the UK and Germany. The number of stations per NUTS1 region is displayed in Table 1. Regional average mean monthly Q and G values were calculated for each NUTS1 region, provided there was at least one station with non-missing observations in the region. As further predictors that may modify the drought indicators' power to explain drought impact occurrence we also selected the month of impact occurrence (M) and the year of impact occurrence (Y). For this purpose the series of months (1–12) was transformed into a sinusoidal curve shifted by four months (peak in July and lowest value in January).

2.3 Drought impacts

Drought impact data come from the European Drought Impact report Inventory (EDII) (Stahl et al., 2015a), which can be viewed online at <http://www.geo.uio.no/edc/droughtdb/> (data extraction for this study: October 2014). The EDII defines a “drought impact” as a negative environmental, economic or social effect experienced under drought conditions. Examples of drought impacts are crop losses, water supply shortages and hosepipe bans, increased mortality of aquatic species, reduced production at thermal or nuclear power plants due to a lack of cooling water, or impaired navigability of streams, to name a few. Drought conditions themselves (anomaly in precipitation, soil moisture, streamflow, groundwater levels etc.), without a negative consequence or at least evoking serious concerns, are not considered an impact. The source of EDII entries is text-based reports on drought impacts, e.g. governmental or NGO reports, books, newspapers, digital media or scientific papers. Each impact report in the EDII contains the following information: (1) a spatial reference (different levels of geographical regionalization, including the European Union NUTS regions standard), (2) a temporal reference (at least the year of occurrence), and (3) an assigned impact category. The 15 categories, e.g. agriculture, water supply, etc., are shown in Fig. 2. Each category subsumes several impact type subcategories (see Stahl et al., 2015a for details).

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



For the analysis the qualitative information on drought impacts was transformed into monthly time series of number of drought impact occurrences per NUTS1 region. The same methodology as in Bachmair et al. (2015) was applied during the conversion of a “drought impact report” (EDII entry) into “drought impact occurrence” (hereafter termed I). In short, this entails the following (see Bachmair et al., 2015 for details):

- Each impact report was assigned to a NUTS1 region. Impact reports with country-level information only were omitted from the analysis. An impact report was converted into several I if (1) the impact report stated impact occurrence in several NUTS1 regions or (2) an impact fell into several impact subtypes.
- Each I is temporally referenced by specifying a start and end month. Impact reports only stating the year of occurrence were omitted from the analysis. In case only the season was provided in the report, we assumed the drought impact occurred during each month of this season (winter = DJF, spring = MAM, summer = JJA, fall = SON).

For each NUTS1 region and month the total number of I was determined, hereafter termed N_I . Table 1 shows the N_I per NUTS1 region included in the analysis, which sum up to 4551 N_I (UK) and 1534 N_I (DE) in total for each country. Some analyses were undertaken for impacts separated into the 15 impact categories. However, a different kind of split of the data was also made, into two larger groups:

- hydrological drought impacts (I_h), i.e. impacts resulting from drought conditions of surface waters or groundwater,
- impacts due to other types of drought (I_o), i.e. impacts associated with meteorological and soil moisture drought and concurrent extremes (e.g. heat waves).

The differentiation between I_h and I_o is based on a keyword search of the impact description field in the database and therefore does not strictly follow any impact

category or impact subtype. Examples of I_h include impaired navigability of streams, increased temperature in surface waters negatively affecting aquatic species, drying up of reservoirs, or reduced fishery production. I_o comprises most agricultural and forestry impacts, impacts on recreation or human health, soil subsidence, or wildfire. Figure 2 shows the total number of I , I_h and I_o per NUTS1 region and season, as well as their categorical distributions.

2.4 Selection of years for analysis

For each NUTS1 region separately, a subset of years within 1970–2012 were selected for analysis based on drought impact occurrence. Years with at least one impact occurrence in the region were selected. All months of the selected years were included in this censored time series. The censoring was undertaken to exclude years with drought conditions yet no impact reports in the EDII, similar to Bachmair et al. (2015). The search for impact reports in both countries focused on known drought events; the absence of impact reports in the EDII for years with drought conditions may therefore be attributable to either a lack of impact occurrence or simply a lack of drought impact reports, whether through not being discovered or not being published in the first place. Table 1 shows the length of time series per region and the percentage of months with impact occurrence in this censored time series. Despite the above-described censoring approach a considerable percentage of months with zero impact occurrence remained. The data analysis was only applied to regions with at least 10 months with impact occurrence, which led to the exclusion of Northern Ireland and Scotland (UK), and the Hanseatic City of Bremen, Hanseatic City of Hamburg, and Thuringia (DE).

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



3 Methods linking indicators and impacts

3.1 Correlation analysis

First, we carried out a cross-correlation analysis between different drought indicators and the number of impacts, accounting for temporal autocorrelation in the indicator and/or impact time series. Spearman rank correlation coefficients (ρ) were calculated between time series of drought indicators and number of impact occurrences, for each NUTS1 region separately. Rank correlation was chosen over Pearson correlation since the counts of the impact data are not normally distributed. Correlations were undertaken between time series of different indicators on the one hand (mean SPI and SPEI for 1–8, 12, and 24 months; Q ; G (DE only); month (M) and year (Y) of impact occurrence), and time series of number of impact occurrences for different impact subsets on the other:

- total impacts (N_I)
- hydrological drought impacts (N_{Ih})
- impacts due to other types of drought (N_{Io})
- impacts per impact category, and
- impacts per season (DJF, MAM, JJA, SON).

A subset of impact data was only included in the analysis if there were at least 10 months with impact occurrence. Since there was temporal autocorrelation present in the time series of SPI and SPEI of longer accumulation periods, in time series of Q and G , and in the impact time series for most UK and some German NUTS1 regions, significance levels of the cross-correlation analysis had to be corrected. Temporal autocorrelation of time series used in cross-correlation analysis violates the assumption of serial independence and increases the likelihood of type I error

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



(Hurlbert, 1984; Jenkins, 2005). We applied the “Modified Chelton method” by Pyper and Peterman (1998), which adjusts the “effective” number of degrees of freedom used for determining significance levels. While we use Spearman’s ρ for the cross-correlation analysis, autocorrelation coefficients represent Pearson’s r (based on square root transformed data for the counts of impact occurrence). We define strength of correlation as follows: 0–0.1 (no correlation), > 0.1–0.3 (weak), > 0.3–0.6 (moderate), > 0.6–0.9 (strong), and > 0.9 (very strong).

3.2 Random forest modeling

Second, we employed a machine learning approach utilizing an ensemble regression tree approach called “random forest” (Breiman, 2001). Similar to the cross-correlation analysis, the random forest approach also identifies drought indicators best linked to impact occurrence. In addition to extracting predictor importance, the random forest approach is used for obtaining splitting values as estimates of thresholds of impact occurrence, and to model drought impact occurrence.

A “random forest” (Breiman, 2001) is a machine learning algorithm, which constructs a large number of classification or regression trees (CARTs) on bootstrapped subsamples of the data. Non-parametric regression using random forest (RF) consists of the following steps (see Liaw and Wiener, 2002 for details): (1) n_{tree} bootstrap samples are used. The individual cases making up the sample are drawn randomly with replacement from the original data, preserving each month’s pairing of predictand and predictors. The size of each sample is about two-thirds of the size of the total dataset, (2) for each bootstrap sample, an unpruned tree is grown. That is, for each node in turn, a split-in-two of the data is performed for each of m_{try} randomly chosen predictor variables, and the predictor whose split results in the two most homogeneous groups (minimizing the residual sum of squares) of the predictand is chosen as the splitting variable for that node, (3) new data is predicted by averaging predictions over n_{tree} regression trees (Liaw and Wiener, 2002). The user-defined variable n_{tree} was set to 1000. The model parameter m_{try} (number of predictors randomly sampled as

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

candidates at each split) was left as default: one third of the total number of predictors (Liaw and Wiener, 2002). For all other parameters the default was kept as well. The model error is determined by predicting the excluded data (“out-of-bag” data according to Breiman, 2001) at each bootstrap iteration using the tree grown with the bootstrap sample and averaging all errors (Liaw and Wiener, 2002).

For our analysis we applied the R package “randomForest” developed by Liaw and Wiener (2002). The RF predictors for each NUTS1 region included the same indicators as used in the correlation analysis. The response variable is the square root transformed monthly counts of impact data per NUTS1 region. The transformation yielded a near normal distribution of the non-zero data in many regions. Some British NUTS1 regions, however, showed a bi-modal distribution of N_i (NEE, NEW, YHU, and SEE with varying extent), and in some German states the distribution of N_i remained positively skewed after the square root transformation. Results for a log-transform were similar. We then ran models for the same subsets of impacts as in the correlation analysis if there were at least 10 months with impact occurrence: total impacts (N_i), hydrological drought impacts (N_{ih}), non-hydrological drought impacts (N_{io}), and impacts per impact category.

To identify the drought indicators best linked to impact occurrence we used the “variable importance” feature of the RF algorithm. For each predictor it is assessed by how much the prediction error increases when the “out-of-bag” data for that predictor are permuted while all others are left unchanged (Liaw and Wiener, 2002). We use the ranks of percent decrease in accuracy as variable importance measure (e.g. Strobl et al., 2009). Another output from the RF analysis are the splitting values for each predictor. The construction of each regression tree is based on recursively splitting the data into more homogenous groups (nodes). At each node, the best splitting variable and splitting value are determined, with multiple splits possible for the same variable (Strobl et al., 2009). For our analysis we extracted the splitting values corresponding to each predictor, considering all trees and nodes, and visualized their distribution

as boxplot. We regard these splitting values as estimates of thresholds of impact occurrence.

The predictive potential of the random forest models was assessed in two ways. First, the overall model performance was evaluated based on a 10-fold cross-validation. The goal of this assessment (hereafter “RF Predictions”) is to test the performance of RF models as a potential tool for predictive purposes, and to learn about the indicator–impact relationship. The data for cross-validation is the censored time series for each NUTS1 region, i.e. the time series based on the sub-selection of years with drought impact occurrence within 1970–2012. For each of the ten model runs the censored time series was split into 90 % for training and 10 % for prediction; impact occurrence of the left-out 10 % is predicted with a random forest model constructed on the training data. The cross-validation procedure allows evaluation of the predictive performance for “unseen” data excluded from model fitting. As model performance metrics we computed mean absolute error (MAE), root mean squared error (RMSE), and error components according to the Kling–Gupta efficiency (Gupta et al., 2009) modified by Gudmundsson (2012): relative difference in mean ($\Delta\mu$), relative difference in standard deviation ($\Delta\sigma$), and strength of correlation between observed vs. modeled number of impacts (r). Zero is the optimal value of $\Delta\mu$ and $\Delta\sigma$; negative and positive values indicate under- and over-prediction, respectively (Gudmundsson et al., 2012).

The second assessment (hereafter “RF Backwards Learning”) is the application of the RF models that were fitted to the censored time series to predict N_1 per NUTS1 region to those years that had been excluded, i.e. the years within 1970–2012 that have zero impact occurrence. The purpose of this second assessment is to scrutinize the impact data in the EDII database to backwards learn where a year without impacts may either be due to no impacts or due to the lack of reporting or finding reports. As the observations themselves are examined no model performance metrics are presented.

HESSD

12, 9437–9488, 2015

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



A split by season (Fig. 4) also shows distinct differences, yet could not be determined for all regions given limited impact data if partitioned seasonally. In winter and spring the general trend stays the same as for the full series (higher ρ for longer accumulation periods of SPI and SPEI, 12 and 24 months dominating). In the fall, in contrast, there is a notable shift towards intermediate SPI and SPEI time scales as best indicators in many regions. In summer, the correlation pattern is more diverse, but in general is dominated by strong correlations across the majority of indicators and accumulation periods. Also, the correlation with streamflow percentiles in summer is higher and more often significant when compared with year-round data.

In Germany, the overall strength of correlation between times series of N_I and different indicators is in a similar range as in the UK (-0.62 to 0.74). Contrary to the UK, shorter to intermediate accumulation periods of SPI and SPEI best correlate with impact occurrence (Fig. 5). Eleven of the 13 analyzed regions show the highest ρ for SPEI-2 to SPEI-4; for SPI-24 and SPEI-24 a non-significant correlation in inverse direction is found. The difference between SPI and SPEI is slightly more pronounced in Germany, with SPEI performing somewhat better (absolute difference in ρ up to 0.13). Q performs similar to SPI in many cases. Groundwater level percentiles show no or non-significant weak correlation with N_I . In contrast, the sine expression of the month shows a higher and often significant ρ , especially in the northern NUTS1 regions. Similar to the UK, there is no or only a weak correlation with Y . As in the UK, there are regional differences, yet mostly regarding the strength of correlation. Most regions in the north and northeast of Germany display a noticeably lower strength of correlation (mostly weak ρ) than the central and southern regions.

Similar to the UK, a split into I_h and I_o reveals differences in correlation patterns compared with I , yet the picture for I_h and I_o is the opposite: while the correlation pattern for N_{I_o} is rather similar to N_I , there is a noticeable shift towards higher correlation with longer SPI and SPEI time scales for I_h . N_{I_o} dominates over N_{I_h} in some German regions, in contrast to the situation in the UK (Fig. 2). For several states the correlation between Q and N_{I_h} is higher than between Q and N_I . A further split by impact category

uncovered the following: agricultural impacts show highest ρ for SPI and SPEI time scales of 1–4 months, yet most correlations are weak and not significant; there is a shift towards higher correlation with longer SPI and SPEI timescales for impacts on waterborne transportation in some NUTS1 regions; for all other impact categories correlations could only be determined for one or two regions (BV or BW or RP) due to too little impact data. A seasonal split was also not possible to assess due to too few months with I in spring, fall, and winter; the majority of impacts in Germany occurred in summer (Fig. 2).

4.2 Indicator importance in random forest models

For the UK, the general picture from the random forest approach is very similar to the findings from the correlation analysis, both regarding I and different impact subsets (I_h , I_o , and I per impact category) (Fig. 6). Long accumulation periods of SPI and SPEI (12 and 24 months) appear as the highest ranking predictors for most regions, except the more northern regions NEE, NEW and YHU. Q does not show up as important predictor. Distinct differences compared with the correlation analysis include the following: (1) Y plays an important role for I and most impact subsets, (2) for I_o , the RF predictor importance shows a shift to intermediate accumulation periods of SPI and SPEI (7/8 months). This shift is not as clearly discernible in the correlation patterns. The same holds true for the agricultural impacts.

In contrast to the UK, where the RF predictor importance plots look very similar to the correlation analysis plots, there is more variation for Germany (Fig. 7). The RF predictor importance patterns are spottier than the correlation analysis patterns with less smooth transitions between adjacent indicators. Nevertheless, the general tendency of best predictors is confirmed. Short to intermediate accumulation periods of SPI and SPEI are highest ranking predictors; the sine expression of the month is top-scoring in the northern states for I , I_o , and agricultural impacts. Also, there is a shift towards higher correlation with longer SPEI accumulation periods (7/8 months) for I_h and impacts

HESSD

12, 9437–9488, 2015

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



on waterborne transportation. Y shows lower importance than for RF models of UK regions.

4.3 Indicator thresholds in random forest models

While splitting values of all indicators for all impact subsets (I , I_h , I_o , different impact categories) were extracted, we only show the threshold distribution, i.e. splitting value distribution, for SPI and SPEI time scales of 8, 12, and 24 months (best performance for different regions and/or impact subsets) and streamflow percentiles (Figs. 8 and 9). For the UK, the threshold distribution for both meteorological indicators generally shows a considerable range, which decreases with increasing accumulation period (roughly +2 to -2.5/-3.5 for SPI-8, +1.5 to -2.5/-3 for SPI-12, and +1 to -2.5 for SPI-24). For the same accumulation periods of SPEI the range extends to less negative values. Apart from this, the differences between SPI and SPEI are negligible with interquartile ranges (IQR) predominantly between 0 and -2. When only focusing on the median of the distribution, SPI-8 and SPEI-8 values scatter around -1 for most NUTS1 regions. For SPI and SPEI of 12 and especially 24 months duration the scatter around -1 is slightly more variable, and differences among NUTS1 regions are somewhat stronger. Regarding streamflow percentiles the splitting values cover almost the entire range, the IQR is distinctly larger than for SPI and SPEI, and the median ranges between 0.2 and 0.37. The split by impact category results in slightly narrower ranges of threshold distributions for many impact categories, and often a more negative median (not shown). This is not the case for I_h and water supply impacts; yet, there is much variation among indicators. All indicators show regional differences, however without systematic patterns.

For Germany, SPI and SPEI of 3 and 6 months accumulation period are generally well linked to N_I or impact numbers in different impact subsets. Figure 9 shows that the splitting values in the different federal states range from roughly +1.5 to -2/-3 for both SPI and SPEI. Absolute values of the IQR of German regions are similar to the UK. Contrary to the UK, a regional pattern exists regarding the median of the SPI

HESSD

12, 9437–9488, 2015

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



an $R^2 > 0.6$. Overall, the lower agreement between observations and predictions than in the UK concurs with the shorter time series of indicator and impact time series in Germany, and a smaller percentage of months with $N_I > 0$ (Table 1).

The generally small difference in the mean ($\Delta\mu$) of observed vs. modeled impacts for both the UK and Germany (Table 2) suggests that the central tendency is well modeled. However, a closer look at the time series of observed and modeled number of impact occurrences (Fig. 12, time series with gray background (RF Predictions)) reveals that small values are generally over-predicted and large values often under-predicted. Notable under-predictions of peak values include, for example, events in 2003 in Germany, and in 2011/12 for many regions in the UK. The under-prediction of N_I causes lower standard deviations for the modelled values than for the observed ($\Delta\sigma$ between -0.22 and -0.52 , see Table 2).

Furthermore, Fig. 12 shows that predictions and observations in the UK and Germany generally agree well both regarding initiation of impact occurrence and its subsequent temporal evolution. This is also reflected by a moderate to strong correlation between predictions and observations (Table 2). The blue line in Fig. 12 represents an impact threshold of one, as guidance for interpretation: modeled impacts smaller than one may be regarded as an absence of impacts. Taking this into account the temporal dynamics agree even better, especially regarding impact onset. An obvious disagreement between dynamics of observations and predictions is found in many regions in the UK in 1991/92, where modeled N_I is more dynamic than the observed static “block” of N_I following an impact peak. The block-shaped data represent impacts due to a contraction of the stream network in large parts of the south and east of the UK during these years. In Germany, states with larger amplitude of N_I (BV, BW, RP, and NW) tend to have a better agreement of temporal dynamics, especially when only focusing on values above the one-impact-threshold line. For states with low amplitude of N_I , which often concurs with less negative splitting values (see Sect. 4.3), the temporal dynamics are less well modeled (not shown).

HESSD

12, 9437–9488, 2015

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



vs. SPEI vs. Q (vs. G in Germany)) uncovered that in the UK SPI and SPEI perform similarly to each other, and Q performs less well. In Germany SPEI often performed slightly better than SPI, the linkage with Q is better than in the UK, and there is low agreement between G and impact occurrence. (3) The largely congruent findings from the two different statistical approaches independently validate the results.

While much can be speculated about the drivers of region-, impact type-, and season-specific variability, it is nonetheless necessary to explore the underlying mechanisms for the observed differences to rule out purely data-driven, yet physically meaningless, indicator–impact relationships. Regional differences can result from both (1) “real” physical, spatial differences in geographic properties (e.g. climate, geology, soil, land use), vulnerability towards drought, and hazard characteristics, triggering impacts differing in type and response time, and (2) differences due to inherent spatial and temporal biases in the impact data (see Bachmair et al., 2015) on potential EDII error sources).

In the UK we found differences in best SPI and SPEI accumulation periods between most southern/central regions (long periods) vs. more northern regions (shorter periods). This corresponds well to known differences in the nature of the drought hazard, and impacts. Strong regional contrasts in drought occurrence across the UK have been noted previously, with a particular contrast between the upland northern and western UK, which is susceptible to short-term (6 month) summer half-year droughts, and the lowlands of the south-eastern UK that are susceptible to longer-term (18 month or greater) multi-annual droughts (Jones and Lister, 1998; Marsh et al., 2007; Parry et al., 2011). These findings reflect both the climatological rainfall gradient across the UK and the predominance of groundwater dominated catchments in the south-east (Folland et al., 2015). While we also found regional differences in indicator–impact-linkage patterns in Germany, they mostly relate to differences in strength of correlation (weaker correlation in northern/northeastern states). The smaller amplitude of impact time series in these states may explain weaker correlation in contrast to southern/central states with predominantly larger amplitude, i.e. pronounced impact

peaks, as hypothesized by Bachmair et al. (2015). In contrast to the UK, which has seen a limited number of multi-annual droughts, most droughts in Germany have been of shorter duration, although such short (typically summer) droughts are fairly frequent (e.g. Bradford, 2000).

The differences in indicator–impact-relationships between the UK and Germany, and some of the within-country variability, are also very likely a result of the regional composition of drought impact types. It is common knowledge that impacts caused by different types of drought have different response times due to propagation through the hydrological cycle (e.g. Mishra and Singh, 2011; National Drought Mitigation Center, 2015; Wilhite and Glantz, 1985). Some impacts develop quickly (e.g. agricultural impacts) during a precipitation shortfall or heatwave and thus show shorter response times than impacts triggered by more slowly evolving streamflow or groundwater drought. In the UK impacts associated with drought conditions of surface waters and groundwater (I_h) clearly dominate (see Fig. 2). This agrees well with longer SPI and SPEI accumulation periods as best predictors in the UK compared with Germany. While hydrological drought impacts still make up the larger part of impacts in Germany, the fraction of non-hydrological drought impacts (I_o) is distinctly larger than in the UK. Agricultural and forestry impacts in Germany account for roughly 20–70% of impacts depending on the region, and this may explain why short to intermediate SPI and SPEI accumulation periods are the best predictors. In British regions these two impact categories sum up to a maximum of 20%. A subdivision of I_h reveals that in the UK impacts on water supply and freshwater ecosystems are most prominent, whereas in Germany impacts on waterborne transportation and water quality dominate in most regions.

The identification of best-performing indicators for specific impact types is a key outcome of this study. While the absolute values of best SPI and SPEI accumulation periods were not identical in both countries, we found commonalities in the relative shift from total impacts to different impact types. For instance, agricultural and hydrological drought impacts were generally best linked to shorter and longer SPI and SPEI time

HESSD

12, 9437–9488, 2015

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



HESSD

12, 9437–9488, 2015

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



scales, respectively. Here, “shorter” and “longer” refers to different absolute values: 1–4 (DE) and 7–8 months (UK) for agriculture, and 7/8 (DE) and 12/24 months (UK) for I_h . Perhaps unsurprisingly, a universal recommendation about best indicators hence cannot be inferred. However, the similar relative shift in best SPI and SPEI time scales suggests that there are likely to be typical patterns of response for given impact types, but that these are mediated by regional cause-effect-mechanisms. This is in line with two studies introduced earlier, which investigated likelihood of impact occurrence specific to particular impact types across different European countries (Blauhut et al., 2015; Stagge et al., 2014). Seasonal variation in linkage patterns as observed in our study for the UK further complicates recommendations regarding a single best drought indicator. Part of the variation across the seasons is likely to reflect differences in impact type distribution between the seasons (see Fig. 2).

A surprising result is that streamflow did not appear as an important drought indicator in the UK, even after a separation of hydrological drought impacts. In Germany, groundwater level percentiles played only a minor role. There are several possible reasons for these discrepancies. For groundwater level percentiles the mismatch is likely attributable to a lagged groundwater response (Bachmair et al., 2015); cross-correlation for different time lags would be a way to assess if and which delay period is linked to impacts. One probable reason for the lack of relationship between I and Q is the nature of the spatially aggregated streamflow data, which represents different catchments varying in size and characteristics (including degree of human influence), lumped over a large administrative area, which does not coincide with catchment boundaries. A further reason may be the nature of the EDII data, especially regarding the subdivisions of I_h . While in Germany the fraction of instream impacts of I_h is larger (e.g. impaired navigability of streams, water quality, and reduced power plant production due to a lack of cooling water), water supply impacts dominate I_h in the UK. For groundwater or reservoir-fed water supply systems these impacts are, to a certain extent, disconnected from river flows (the purpose of reservoirs being to smooth out variations in instream water availability).

HESSD

12, 9437–9488, 2015

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Overall, despite a rather complex picture in terms of best drought indicator for impact occurrence, the empirical evaluation of drought indicators with text-based impact information proved to be a feasible approach. Given the minor differences in the outcomes of the correlation and the random forest analysis for the UK, both methods appear recommendable. Generally, the strength of the random forest algorithm is that it can handle interactions and nonlinearities among variables, and thus identify non-intuitive relationships (Evans et al., 2011; Hastie et al., 2009). Furthermore, random forests are robust to noise (Breiman, 2001; Hastie et al., 2009), and the bootstrap sampling provides a way to account for the uncertainty of the impact data. Nevertheless, the “black-box” nature of the RF model (Breiman, 2001) may not be as useful when an intuitive method for the choice of best drought indicator is needed (e.g. when working with a wide range of stakeholders from different backgrounds). For Germany, systematic differences in indicator–impact-linkage patterns were easier to perceive in the correlation plots than in the RF predictor importance plots. For large data sets the RF algorithm has the potential to detect relatively complex structures; for small data sets, however, this is unlikely to be the case (Maindonald and Braun, 2006). The generally shorter time series for German regions and stronger zero-inflation of the data may therefore explain the “spottier” pattern of RF predictor importance. The correlation analysis thus yielded more powerful results for Germany. However, this method does not provide further information such as on thresholds of impact occurrence, in contrast to the RF algorithm (see Sect. 4.3). Both approaches therefore complemented each other in our study.

5.2 Indicator thresholds for impact occurrence

The analysis of splitting values used in the random forest construction highlighted a large spread. Yet, when focusing on the median there are differences between the countries and among the regions (medians around -1 for SPI and SPEI of different accumulation periods in the UK, and in DE between ca. 0 and -1 (north/northeast) and -1 and -1.5 (southern/central states)), and, to some extent, impact categories.

occurrence. To our knowledge, there are hardly any publicized studies systematically evaluating the delineation rules of different drought severity classes by using drought impacts (e.g. Sepulcre-Canto et al., 2012) or by stakeholders' experience (e.g. Steinemann and Cavalcanti, 2006). Our analysis and previous findings on indicator thresholds for impact onset in Germany (Bachmair et al., 2015) demonstrate the potential of text-based impact data to do this.

5.3 Lessons learned from random forest predictions

The two parts of the random forest modeling exercise exposed that: (1) there are differences among regions in terms of predictive power, with RF models for regions with better impact data (longer censored time series, a higher percentage of non-zero data, and larger amplitude of the impact time series) showing good agreement between observations and predictions. (2) While the temporal dynamics of impact occurrence were reasonably reproduced, over- and under-prediction of small and large values, respectively, are an issue. (3) Backwards Learning about impact occurrence for years with no observations (through RF models trained on drought years) provided valuable insights into time periods which potentially lack impact data in the EDII.

Overall, the analysis revealed that RF models generally represent a suitable tool for drought M&EW, yet further model tuning is possible (e.g. reduction of predictors, grouping several regions for increasing the number of observations, and impact category specific models). The finding that there is good agreement between observed and predicted number of drought impact occurrences for regions with good data availability is promising. It also underlines the benefit of spending time and resources on impact data collection. Currently the process of impact data collection is not automated but labor intensive. Good model performance for South West England, for example, which is characterized by better data availability than in other regions, makes a strong argument for the value of impact monitoring. The necessity of expanding impact data collection and its benefit for drought M&EW has been reported by others (Lackstrom et al., 2013; Stahl et al., 2015b; Wilhite et al., 2007).

HESSD

12, 9437–9488, 2015

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



(Cole and Marsh, 2006), in part due to a wet summer in 1973. For the 1995–1997 drought, only impacts from the hot, dry summer of 1995 are captured in the EDII, as the summer drought was very extensively reported due to water supply failures and government responses. However, a protracted groundwater drought, with water restrictions in some areas, extended into 1997 (Cole and Marsh, 2006). However no “formal” drought report was written so these later impacts have not been captured by the EDII. Such discrepancies support our choice of time series censoring via drought impact occurrence. Altogether, false-positive impacts provide guidance on which time periods to focus on when searching for additional impact information. This may, in turn, result in more reliable predictions or impact thresholds based on more drought events.

6 Conclusion

The broad goal of our analysis was two-fold: to learn about the relationship between drought indicators and text-based impact information to advance drought monitoring and early warning, and to test methodologies that can be extended to other locations in a next step. We found that drought indicators best linked to impact occurrence are region and impact category specific. In the UK they are additionally season specific to some degree. However, we identified several common traits, allowing the potential grouping of regions and/or impact categories according to their indicator-impact-response. This calls for evaluating continental drought M&EW systems at smaller spatial scales. Also, our analysis provided empirical evidence that impacts associated with different types of drought (e.g. agricultural vs. hydrological drought) have different response times, as reflected by distinct differences in indicator–impact-linkage patterns for each impact category. Regarding methodologies, a random forest machine learning approach proved to be a suitable tool for objectively identifying indicator thresholds for impact occurrence, and to predict the number of drought impact occurrences for regions with sufficient data. We therefore suggest validating any chosen triggers in drought M&EW with impact data as a complementary approach

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

(LUNG), Landesamt für Umwelt, Wasserwirtschaft und Gewerbeaufsicht Rheinland-Pfalz (LUWG), Landesanstalt für Umwelt, Gesundheit und Verbraucherschutz Brandenburg (LUGV, Regionalabteilungen Ost, Süd, West), Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg (LUBW), Landesbetrieb für Hochwasserschutz und Wasserwirtschaft Sachsen-Anhalt (LHW), Landesbetrieb für Küstenschutz, Nationalpark und Meeresschutz Schleswig-Holstein (LKNM), Niedersächsischer Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz (NLWK), Ruhrverband, Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie (LfULG), Staatliches Amt für Landwirtschaft und Umwelt Vorpommern (StALU-VP), Thüringer Landesamt für Umwelt und Geologie (TLUG), Landesamt für Landwirtschaft, Umwelt und ländliche Räume (LLUR), Wasser- und Schifffahrtsverwaltung des Bundes (WSV).

References

- Bachmair, S., Kohn, I., and Stahl, K.: Exploring the link between drought indicators and impacts, *Nat. Hazards Earth Syst. Sci.*, 15, 1381–1397, doi:10.5194/nhess-15-1381-2015, 2015.
- Blauhut, V., Gudmundsson, L., and Stahl, K.: Towards pan-European drought risk maps: quantifying the link between drought indices and reported drought impacts, *Environ. Res. Lett.*, 10, 014008, doi:10.1088/1748-9326/10/1/014008, 2015.
- Bradford, R. B.: Drought events in Europe, in: *Drought and Drought Mitigation in Europe*, vol. 14, edited by: Vogt, J. V. and Somma, F., Springer Netherlands, Dordrecht, 7–20, 2000.
- Breiman, L.: Random forests, *Mach. Learn.*, 45, 5–32, doi:10.1023/A:1010933404324, 2001.
- Bundesamt für Gewässerkunde: *Hydrologischer Atlas von Deutschland*, Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit, Berlin, 2003.
- Bundesanstalt für Geowissenschaften und Rohstoffe: *Bodenarten in Oberböden Deutschlands 1 : 1 000 000*, Hannover, 2007.
- Cole, G. A. and Marsh, T. J.: *Major Droughts in England and Wales from 1800 and Evidence of Impact*, Environment Agency, Bristol, 2006.
- Dieker, E., van Lanen, H. A. J., and Svoboda, M.: Comparison of three drought monitoring tools in the USA, *WATCH Technical Report No. 25*, available at: <http://www.eu-watch.org/publications/technical-reports/3>, last access: 11 September 2015, 2010.

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Dutra, E., Pozzi, W., Wetterhall, F., Di Giuseppe, F., Magnusson, L., Naumann, G., Barbosa, P., Vogt, J., and Pappenberger, F.: Global meteorological drought – Part 2: Seasonal forecasts, *Hydrol. Earth Syst. Sci.*, 18, 2669–2678, doi:10.5194/hess-18-2669-2014, 2014.

European Drought Observatory: Product Fact Sheet: Combined Drought Indicator – EUROPE, available at: <http://edo.jrc.ec.europa.eu/edov2/php/index.php?id=1101> (last access: 20 August 2015), 2013.

Evans, J. S., Murphy, M. A., Holden, Z. A., and Cushman, S. A.: Modeling species distribution and change using random forest, in: *Predictive Species and Habitat Modeling in Landscape Ecology*, edited by: Drew, C. A. Wiersma, Y. F., and Huettmann, F., Springer New York, New York, NY, p. 313, 2011.

Folland, C. K., Hannaford, J., Bloomfield, J. P., Kendon, M., Svensson, C., Marchant, B. P., Prior, J., and Wallace, E.: Multi-annual droughts in the English Lowlands: a review of their characteristics and climate drivers in the winter half-year, *Hydrol. Earth Syst. Sci.*, 19, 2353–2375, doi:10.5194/hess-19-2353-2015, 2015.

Gillette, H.: A creeping drought under way, *Water Sewage Works*, 97, 104–105, 1950.

Gudmundsson, L., Wagener, T., Tallaksen, L. M., and Engeland, K.: Evaluation of nine large-scale hydrological models with respect to the seasonal runoff climatology in Europe, *Water Resour. Res.*, 48, W11504, doi:10.1029/2011WR010911, 2012.

Gudmundsson, L., Rego, F. C., Rocha, M., and Seneviratne, S. I.: Predicting above normal wildfire activity in southern Europe as a function of meteorological drought, *Environ. Res. Lett.*, 9, 084008, doi:10.1088/1748-9326/9/8/084008, 2014.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling, *J. Hydrol.*, 377, 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009.

Hao, Z. and Singh, V. P.: Drought characterization from a multivariate perspective: a review, *J. Hydrol.*, 527, 668–678, doi:10.1016/j.jhydrol.2015.05.031, 2015.

Hargreaves, G. H.: Defining and Using Reference Evapotranspiration, *J. Irrig. Drain. E.-ASCE*, 120, 1132–1139, doi:10.1061/(ASCE)0733-9437(1994)120:6(1132), 1994.

Haslinger, K., Koffler, D., Schöner, W., and Laaha, G.: Exploring the link between meteorological drought and streamflow: effects of climate-catchment interaction, *Water Resour. Res.*, 50, 2468–2487, doi:10.1002/2013WR015051, 2014.

Hastie, T., Tibshirani, R., and Friedman, J.: *The Elements of Statistical Learning*, Springer New York, New York, NY, 2009.

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, *J. Geophys. Res.*, 113, D20119, doi:10.1029/2008JD010201, 2008.
- Hlavinka, P., Trnka, M., Semerádová, D., Dubrovský, M., Žalud, Z., and Možný, M.: Effect of drought on yield variability of key crops in Czech Republic, *Agr. Forest Meteorol.*, 149, 431–442, doi:10.1016/j.agrformet.2008.09.004, 2009.
- Hurlbert, S. H.: Pseudoreplication and the design of ecological field experiments, *Ecol. Monogr.*, 54, 187, doi:10.2307/1942661, 1984.
- Jenkins, G.: The influence of climate on the fishery recruitment of a temperate, seagrass-associated fish, the King George whiting *Sillaginodes punctata*, *Mar. Ecol.-Prog. Ser.*, 288, 263–271, 2005.
- Ji, L. and Peters, A. J.: Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices, *Remote Sens. Environ.*, 87, 85–98, 2003.
- Jones, P. D. and Lister, D. H.: Riverflow reconstructions for 15 catchments over England and Wales and an assessment of hydrologic drought since 1865, *Int. J. Climatol.*, 18, 999–1013, doi:10.1002/(SICI)1097-0088(199807)18:9<999::AID-JOC300>3.0.CO;2-8, 1998.
- Kohn, I., Rosin, K., Freudiger, D., Belz, J. U., Stahl, K., and Weiler, M.: Niedrigwasser in Deutschland 2011, *Hydrol. Wasserbewirts.*, 58, 4–17, doi:10.5675/HyWa_2014,1_1, 2014.
- Kruse, S., Seidl, I., and Staehli, M.: Informationsbedarf zur Früherkennung von Trockenheit in der Schweiz, *Wasser Energie Luft*, 102, 4 pp., 2010.
- Lackstrom, K., Brennan, A., Ferguson, D., Crimmins, M., Darby, L., Dow, K., Ingram, K., Meadow, A., Reges, H., Shafer, M., and Smith, K.: The Missing Piece: Drought Impacts Monitoring, Workshop report produced by the Carolinas Integrated Sciences & Assessments program and the Climate Assessment for the Southwest, 5–6 March 2013, Tucson, AZ, 1–23, 2013.
- Liaw, A. and Wiener, M.: Classification and Regression by Random Forest, *R News*, available at: ftp://131.252.97.79/Transfer/Treg/WFRE_Articles/Liaw_02_ClassificationandregressionbyrandomForest.pdf (last access: 20 August 2015), 2002.
- Loayza, N. V., Olaberría, E., Rigolini, J., and Christiaensen, L.: Natural disasters and growth: going beyond the averages, *World Dev.*, 40, 1317–1336, doi:10.1016/j.worlddev.2012.03.002, 2012.

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Logar, I. and van den Bergh, J. C. J. M.: Methods to assess costs of drought damages and policies for drought mitigation and adaptation: review and recommendations, *Water Resour. Manag.*, 27, 1707–1720, doi:10.1007/s11269-012-0119-9, 2013.
- Maindonald, J. and Braun, J.: *Data Analysis and Graphics Using R: An Example-based Approach*, 3rd edn., Cambridge University Press, New York, USA, 2010.
- Marsh, T., Cole, G., and Wilby, R.: Major droughts in England and Wales, 1800–2006, *Weather*, 62, 87–93, doi:10.1002/wea.67, 2007.
- Martínez-Fernández, J., González-Zamora, A., Sánchez, N., and Gumuzzio, A.: A soil water based index as a suitable agricultural drought indicator, *J. Hydrol.*, 522, 265–273, doi:10.1016/j.jhydrol.2014.12.051, 2015.
- Mavromatis, T.: Drought index evaluation for assessing future wheat production in Greece, *Int. J. Climatol.*, 27, 911–924, 2007.
- McKee, T. B., Doesken, N. J., and Kleist, J.: The Relationship of Drought Frequency and Duration to Time Scales, Preprints, 8th Conference on Applied Climatology, 17–22 January 1993, Anaheim, California, 179–184, 1993.
- Mehta, V. M., Wang, H., Mendoza, K., and Rosenberg, N. J.: Predictability and prediction of decadal hydrologic cycles: a case study in Southern Africa, *Weather and Climate Extremes*, 3, 47–53, doi:10.1016/j.wace.2014.04.002, 2014.
- Mishra, A. K. and Singh, V. P.: Drought modeling – a review, *J. Hydrol.*, 403, 157–175, 2011.
- National Drought Mitigation Center: Types of Drought, available at: <http://drought.unl.edu/DroughtBasics/TypesofDrought.aspx> (last access: 20 August 2015), 2015.
- Ordoyne, C. and Friedl, M.: Using MODIS data to characterize seasonal inundation patterns in the Florida Everglades, *Remote Sens. Environ.*, 112, 4107–4119, doi:10.1016/j.rse.2007.08.027, 2008.
- Parry, S., Hannaford, J., Prudhomme, C., Lloyd-Hughes, B., and Williamson, J.: Objective drought and high flow catalogues for Europe, WATCH Technical Report No. 33, available at: <http://www.eu-watch.org/publications/technical-reports>, last access: 11 September 2015, 2011.
- Parry, S., Prudhomme, C., Wilby, R. L., and Wood, P. J.: Drought termination: concept and characterisation, *Prog. Phys. Geogr.*, in review, 2015.
- Pyper, B. J. and Peterman, R. M.: Comparison of methods to account for autocorrelation in correlation analyses of fish data, *Can. J. Fish. Aquat. Sci.*, 55, 2127–2140, 1998.

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Quiring, S. M. and Papakryiakou, T. N.: An evaluation of agricultural drought indices for the Canadian prairies, *Agr. Forest Meteorol.*, 118, 49–62, 2003.
- Sepulcre-Canto, G., Horion, S., Singleton, A., Carrao, H., and Vogt, J.: Development of a Combined Drought Indicator to detect agricultural drought in Europe, *Nat. Hazards Earth Syst. Sci.*, 12, 3519–3531, doi:10.5194/nhess-12-3519-2012, 2012.
- 5 Shukla, S., Steinemann, A. C., and Lettenmaier, D. P.: Drought monitoring for Washington State: indicators and applications, *J. Hydrometeorol.*, 12, 66–83, doi:10.1175/2010JHM1307.1, 2011.
- Stagge, J. H., Kohn, I., Tallaksen, L. M., and Stahl, K.: Modeling drought impact occurrence based on climatological drought indices for Europe, *J. Hydrol.*, doi:Please provide volume with page range or with article number and DOI. 2014.
- 10 Stagge, J. H., Tallaksen, L. M., Gudmundsson, L., Van Loon, A. F., and Stahl, K.: Candidate distributions for climatological drought indices (SPI and SPEI), *Int. J. Climatol.*, doi:10.1002/joc.4267, online first, 2015.
- 15 Stahl, K., Kohn, I., Blauhut, V., Urquijo, J., De Stefano, L., Acacio, V., Dias, S., Stagge, J. H., Tallaksen, L. M., Kampragou, E., Van Loon, A. F., Barker, L. J., Melsen, L. A., Bifulco, C., Musolino, D., de Carli, A., Massarutto, A., Assimacopoulos, D., and Van Lanen, H. A. J.: Impacts of European drought events: insights from an international database of text-based reports, *Nat. Hazards Earth Syst. Sci. Discuss.*, 3, 5453–5492, doi:10.5194/nhessd-3-5453-2015, 2015a.
- 20 Stahl, K., Kohn, I., De Stefano, L., Tallaksen, L., Rego, F. C., Seneviratne, S. I., Andreu, J., and Van Lanen, H. A.: An impact perspective on pan-European drought sensitivity, in: *Drought: Research and Science-Policy Interfacing*, edited by: Andreu, J., Solera, A., Paredes-Arquiola, J., Haro-Monteagudo, D., and van Lanen, H., CRC Press, London, 329–334, 2015b.
- 25 Steinemann, A.: Drought information for improving preparedness in the Western States, *B. Am. Meteorol. Soc.*, 95, 843–847, doi:10.1175/bams-d-13-00067.1, 2014.
- Steinemann, A. C. and Cavalcanti, L. F. N.: Developing multiple indicators and triggers for drought plans, *J. Water Res. P.-ASCE*, 132, 164–174, 2006.
- 30 Stölzle, M. and Stahl, K.: Wassernutzung und Trockenheitsindikatoren in Baden-Württemberg, *STANDORT – Zeitschrift für Angewandte Geographie*, 35, 94–101, 2011.

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Strobl, C., Malley, J., and Tutz, G.: An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests, *Psychol. Methods*, 14, 323–348, 2009.

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., and Stooksbury, D.: The drought monitor, *B. Am. Meteorol. Soc.*, 83, 1181–1190, 2002.

Trambauer, P., Werner, M., Winsemius, H. C., Maskey, S., Dutra, E., and Uhlenbrook, S.: Hydrological drought forecasting and skill assessment for the Limpopo River basin, southern Africa, *Hydrol. Earth Syst. Sci.*, 19, 1695–1711, doi:10.5194/hess-19-1695-2015, 2015.

Vicente-Serrano, S. M. and López-Moreno, J. I.: Hydrological response to different time scales of climatological drought: an evaluation of the Standardized Precipitation Index in a mountainous Mediterranean basin, *Hydrol. Earth Syst. Sci.*, 9, 523–533, doi:10.5194/hess-9-523-2005, 2005.

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I.: A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index, *J. Climate*, 23, 1696–1718, 2010.

Vicente-Serrano, S. M., Beguería, S., Lorenzo-Lacruz, J., Camarero, J. J., López-Moreno, J. I., Azorin-Molina, C., Revuelto, J., Morán-Tejeda, E., and Sanchez-Lorenzo, A.: Performance of drought indices for ecological, agricultural, and hydrological applications, *Earth Interact.*, 16, 1–27, 2012.

Wetterhall, F., Winsemius, H. C., Dutra, E., Werner, M., and Pappenberger, E.: Seasonal predictions of agro-meteorological drought indicators for the Limpopo basin, *Hydrol. Earth Syst. Sci.*, 19, 2577–2586, doi:10.5194/hess-19-2577-2015, 2015.

Wilhite, D. and Glantz, M.: Understanding: the drought phenomenon: the role of definitions, *Water Int.*, 10, 111–120, doi:10.1080/02508068508686328, 1985.

Wilhite, D. and Svoboda, M.: Drought early warning systems in the context of drought preparedness and mitigation, in: *Early Warning Systems for Drought Preparedness and Drought Management*, edited by: Wilhite, D. A., Sivakumar, M. V. K., and Wood, D. A., World Meteorological Organization, Geneva, 1–21, 2000.

Wilhite, D. A., Svoboda, M. D., and Hayes, M. J.: Understanding the complex impacts of drought: a key to enhancing drought mitigation and preparedness, *Water Resour. Manag.*, 21, 763–774, doi:10.1007/s11269-006-9076-5, 2007.

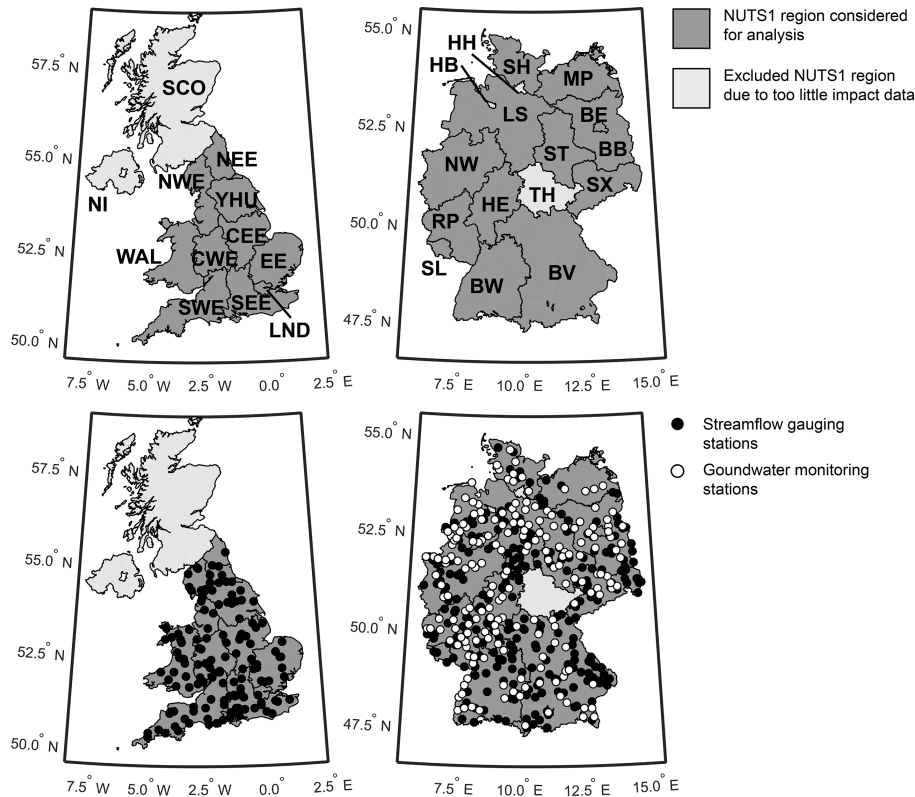


Figure 1. Maps displaying NUTS1 regions in the UK (left) and Germany (right), and the location of streamflow gauging and groundwater monitoring stations. See Table 1 for NUTS1 region abbreviations.

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[⏴](#)

[⏵](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

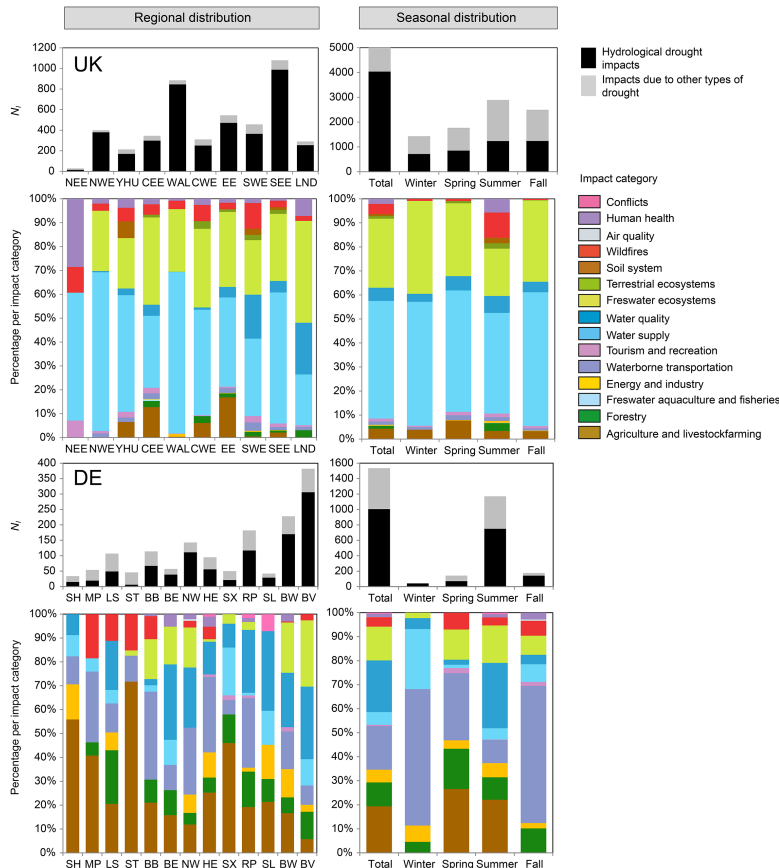


Figure 2. Number of impact occurrences and distribution of impacts per impact category per NUTS1 region and season for the UK (top four plots) and Germany (bottom four plots).

[Title Page](#)

[Abstract](#) [Introduction](#)

[Conclusions](#) [References](#)

[Tables](#) [Figures](#)

[⏪](#) [⏩](#)

[⏴](#) [⏵](#)

[Back](#) [Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

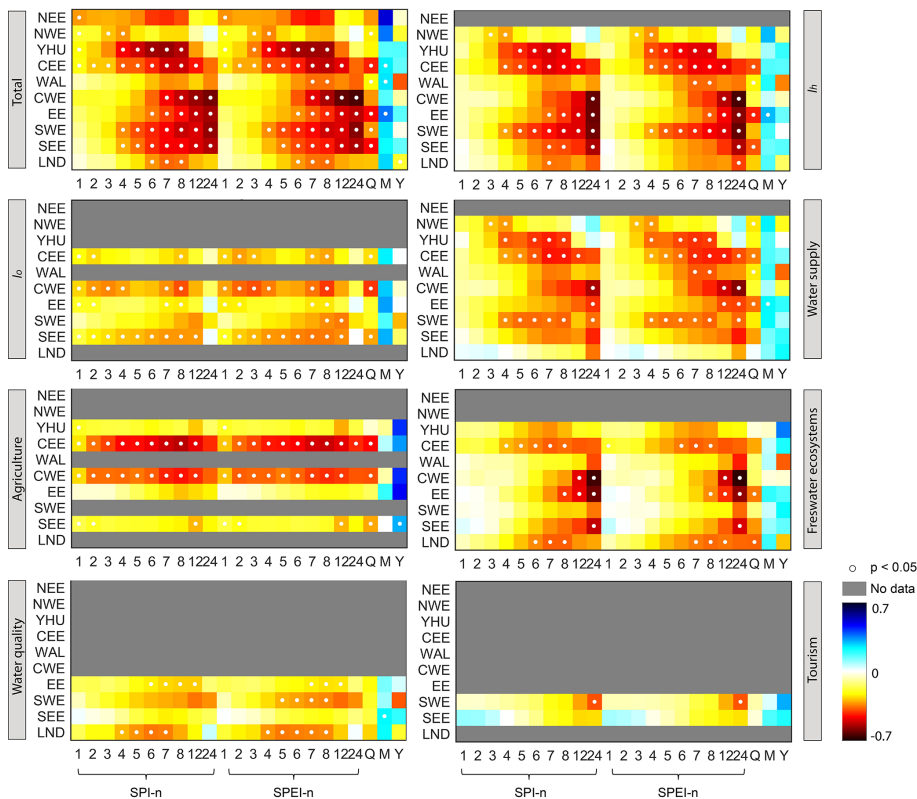


Figure 3. UK: rank correlation coefficients (ρ) between drought indicators and number of impact occurrences for total impacts, hydrological drought impacts (I_h), impacts due to other types of drought (I_o), and selected impact categories per NUTS1 region.

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



HESSD

12, 9437–9488, 2015

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

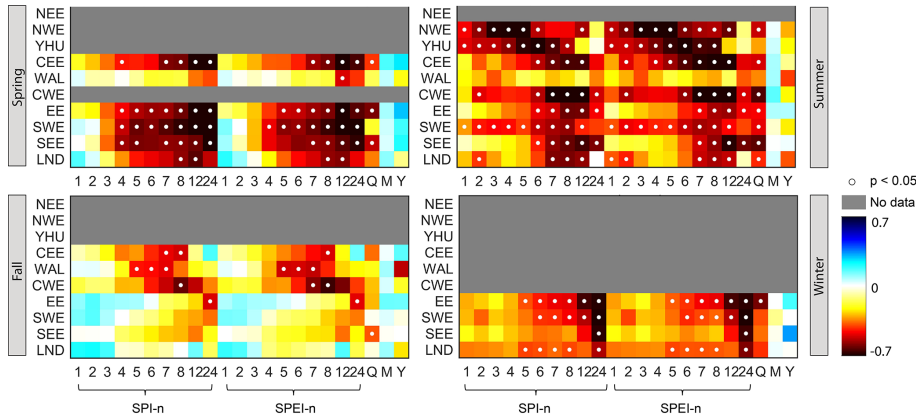


Figure 4. UK: rank correlation coefficients (ρ) between drought indicators and number of impact occurrences per NUTS1 region and season.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

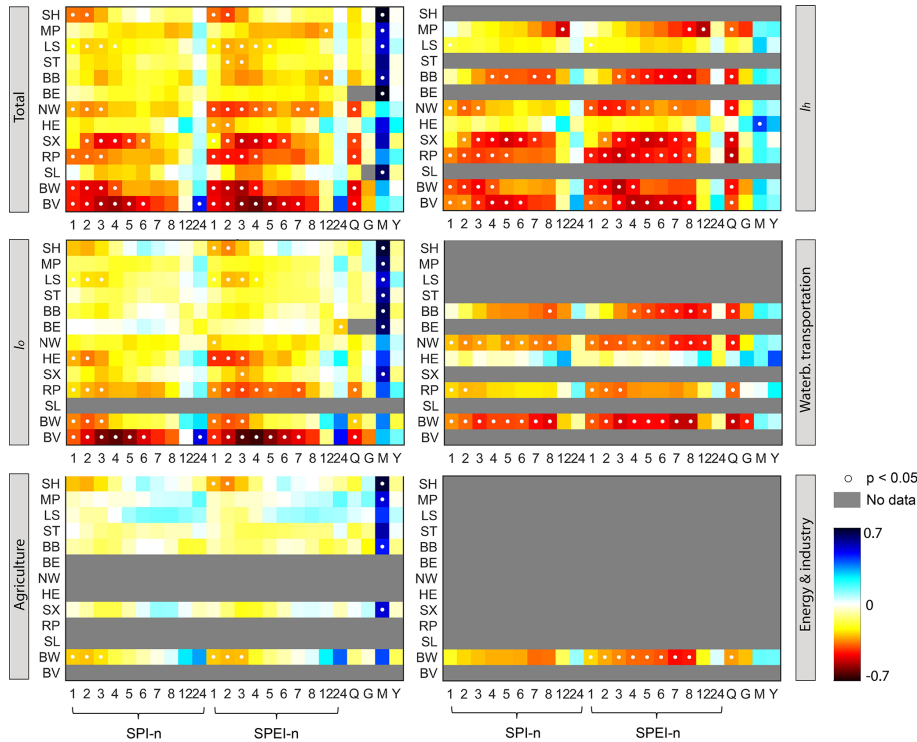


Figure 5. Germany: rank correlation coefficients (ρ) between drought indicators and number of impact occurrences for total impacts, hydrological drought impacts (I_h), impacts due to other types of drought (I_o), and selected impact categories per NUTS1 region.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



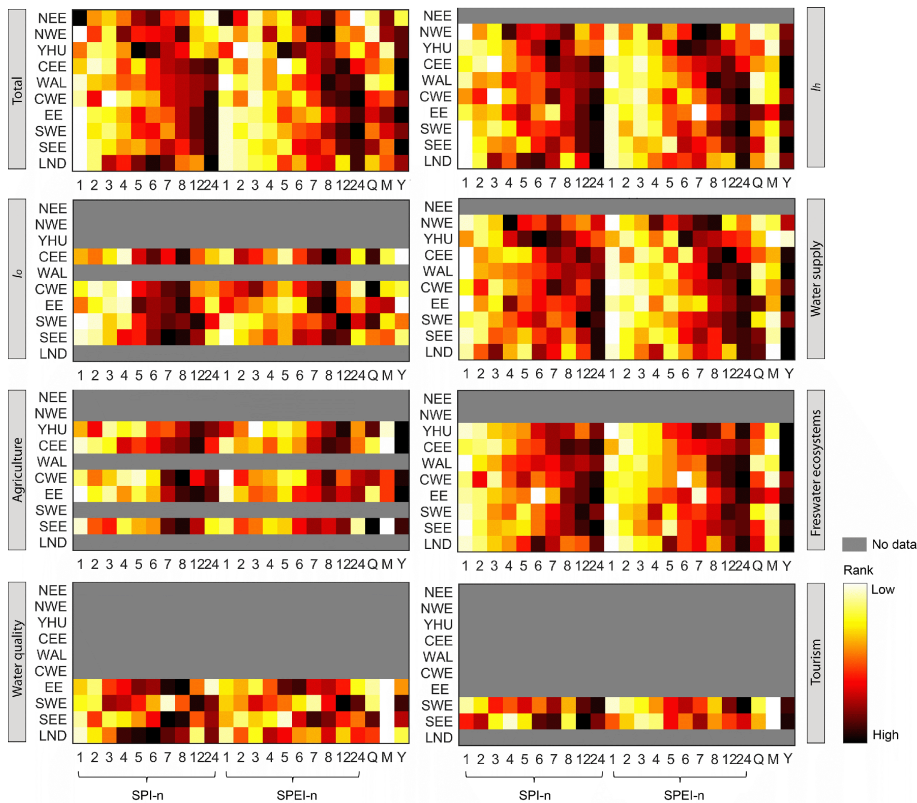


Figure 6. UK: ranks of predictor importance during random forest construction for total impacts, hydrological drought impacts (I_h), impacts due to other types of drought (I_o), and selected impact categories per NUTS1 region.

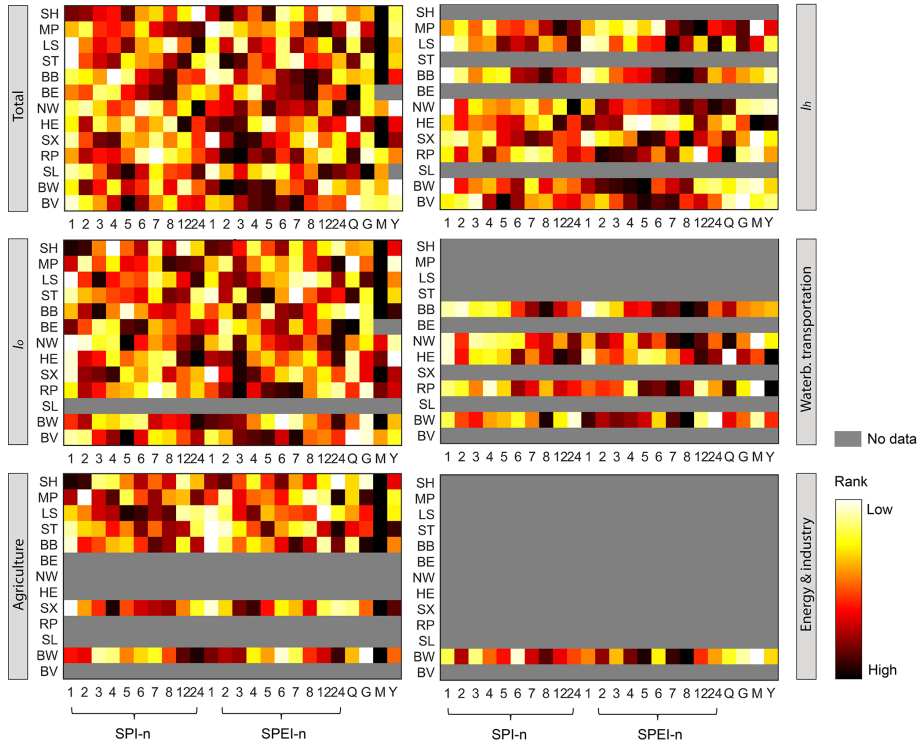


Figure 7. Germany: ranks of predictor importance during random forest construction for total impacts, hydrological drought impacts (I_h), impacts due to other types of drought (I_o), and selected impact categories per NUTS1 region.

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

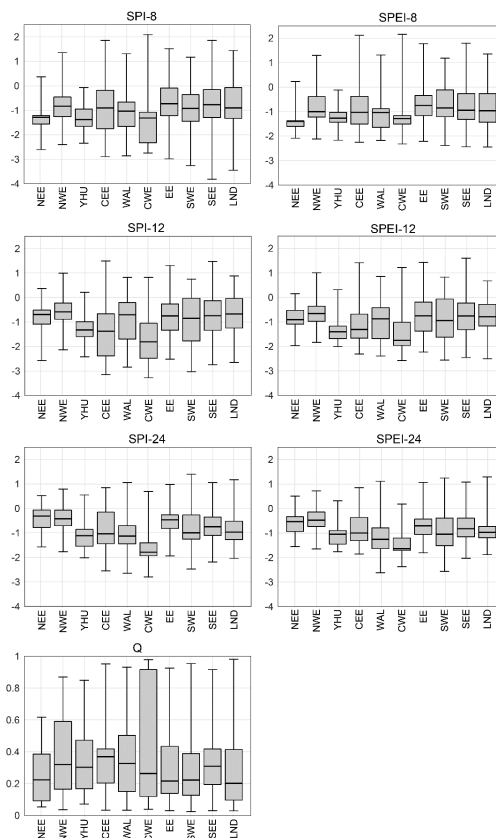


Figure 8. UK: distribution of splitting values during random forest construction (i.e. thresholds of impact occurrence) for selected drought indicator variables for each NUTS1 region. The boxplot whiskers extend to the minimum and the maximum of the distribution, the box encompasses the interquartile range, and the line inside the box displays the median.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

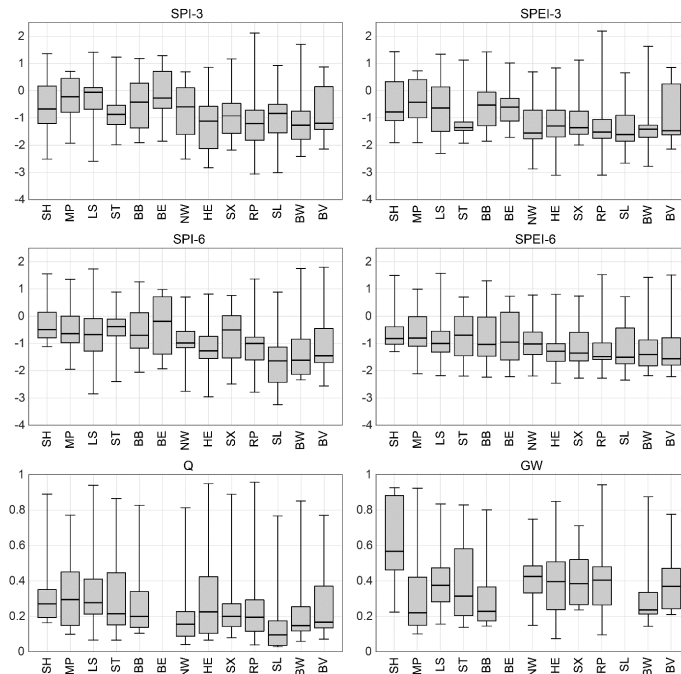


Figure 9. Germany: distribution of splitting values during random forest construction (i.e. thresholds of selected drought indicator variables for each NUTS1 region). Boxplots as Fig. 8.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

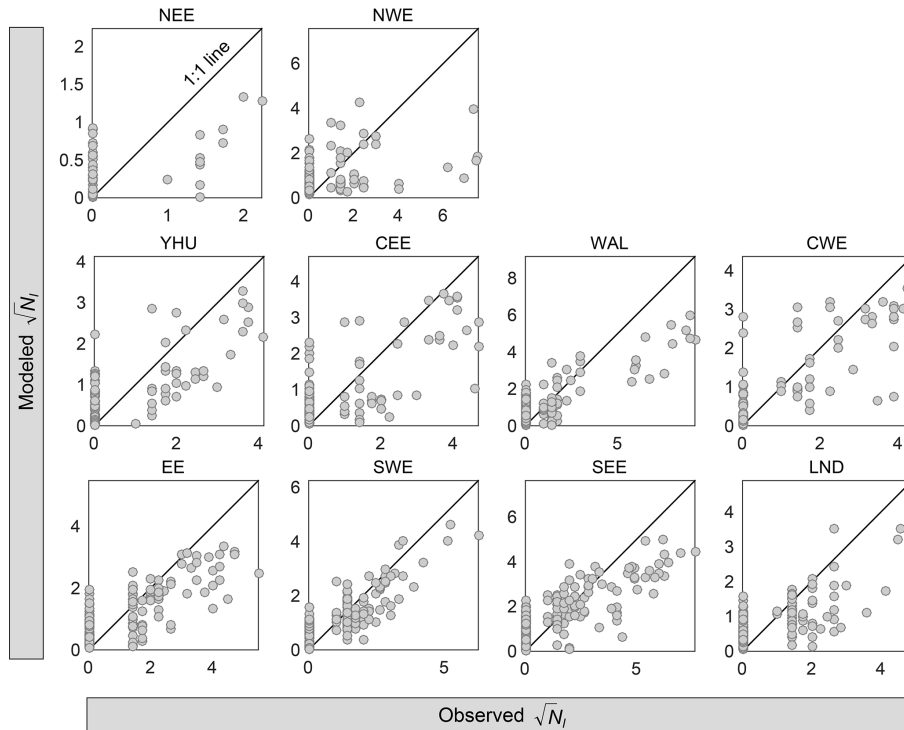


Figure 10. RF Predictions for different regions in the UK (transformed variables).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



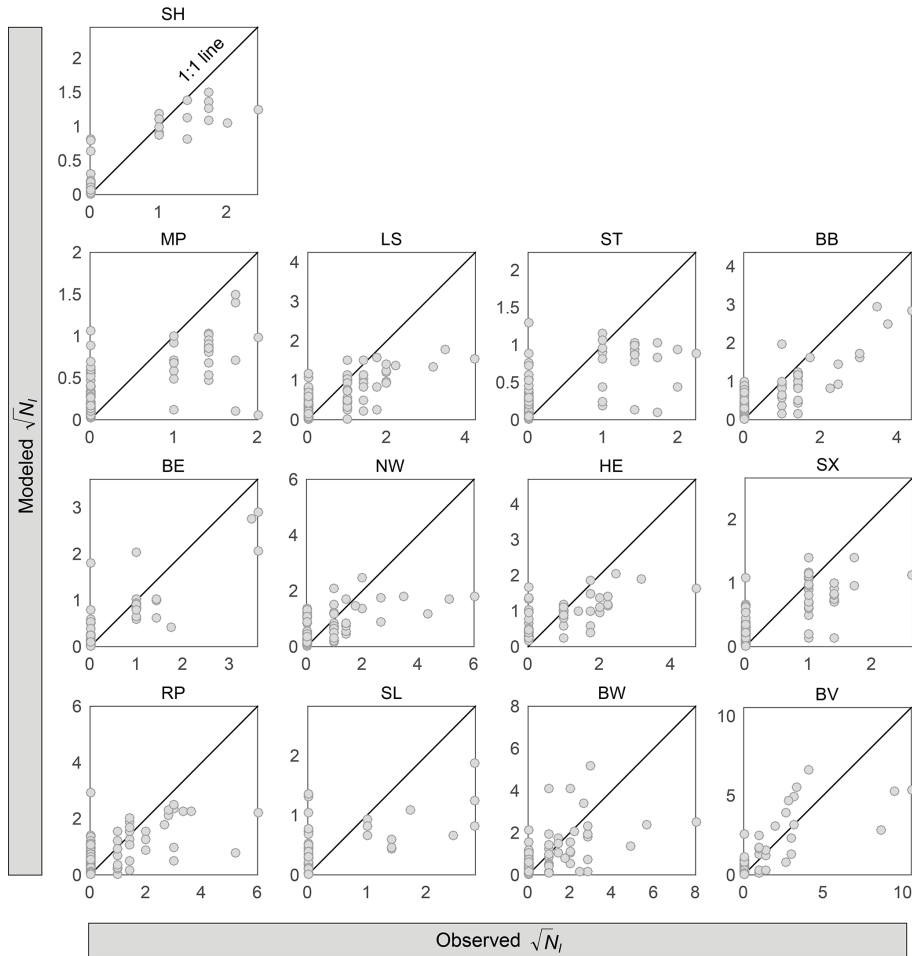


Figure 11. RF Predictions for different regions in Germany (transformed variables).

Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

[Title Page](#)

[Abstract](#) [Introduction](#)

[Conclusions](#) [References](#)

[Tables](#) [Figures](#)

[◀](#) [▶](#)

[◀](#) [▶](#)

[Back](#) [Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Appraising drought indicators and modelling drought impacts

S. Bachmair et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

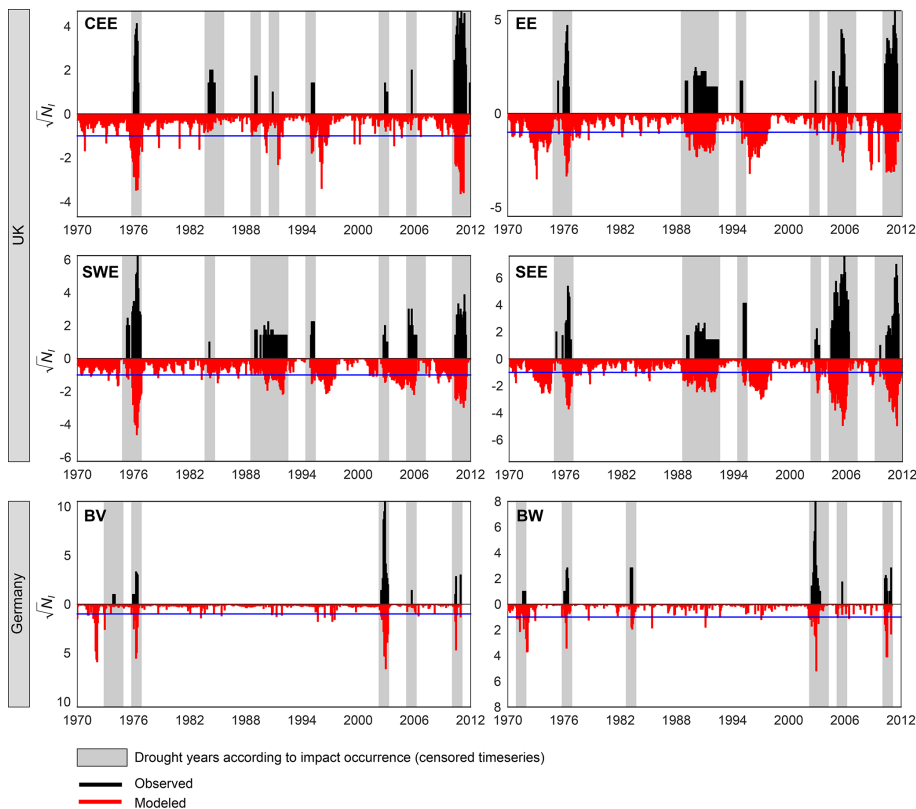


Figure 12. Time series of observed and modeled number of impact occurrences for a selection of NUTS1 regions in the UK and Germany (transformed variables). Grey background: RF Predictions, white background: RF Backwards Learning. The blue line indicates an impact threshold of one: modeled impacts smaller than one should be regarded as absent impact.