We thank both reviewers for their time and their helpful comments. A point-to-point response follows.

Reviewer 1

This is an interesting manuscript, and outlines a novel application of multivariate return periods. In my opinion it deserves publication: it is clear, well written, and the mathematical tools used are well referenced.

Thank you for this positive evaluation.

I only have two minor remarks:
1- There are two typos on page 5, line 10: it is Pr(X<x), not F(X<x), and similarly for Y.

Thanks, we have corrected this in the revised version.

2- On page 3, lines 9-13, the Authors should add that the concept of "critical regions" has recently been well mathematically formalized in the paper mentioned below, via the notions of Hazard Scenarios and Upper Sets (which provide a consistent mathematical framework): such a piece of information may be useful for the interested readers.

G. Salvadori, F. Durante, C. De Michele, M. Bernardi, and L. Petrella. A multivariate Copula-based framework for dealing with Hazard Scenarios and Failure Probabilities. Water Resources Research, 53:3701–3721, 2016. doi: 10.1002/2015WR017225.

Thank you for this suggestion. We have used this reference to better motivate the use of "critical regions". "The concept of critical regions has recently been well mathematically formalized via the notions of Hazard Scenarios and Upper Sets, providing a consistent mathematical framework. The interested reader is referred to Salvadori et al. (2016).

Reviewer 2

This study seeks the linkage between multivariate climate conditions and crop yields. This paper is one of the first which employs bivariate return periods of temperature and precipitation as the indicator of climate variability to explain crop yield variability. It is clearly written and obtains the interesting finding that the combination of temperature and precipitation can explain more crop yield variability than models relying directly on temperature and precipitation as predictors on average in Europe. The result also reveals different sensitivities of crops to climate conditions. A need to incorporate the nonlinear impacts into the climate-crop yield assessment is highlighted. For all these reasons, I recommend publication after addressing a few comments regarding the statistical examinations.

Thank you for this positive evaluation.

One comment is related to the calculation of return periods. As explained by the authors, a return period of RP, also known as a recurrence interval is an estimate of the likelihood of an event to occur, i.e., an occurrence probability of 1/RP every year. If the event of interest happens every year or the annual maximum/minimum data is in use, the mean interarrival time is 1 yr (i.e., the numerator in Eq. 1). If I understand correctly, the authors applied copulas to the seasonal and 2-month averaged climate variables for return period calculation. For such case, why the mean interarrival time is 1 yr? Instead, the mean interarrival time should be calculated as the length of data in years (L)/ the number of occurred events in the length of data (n), i.e., L/n.

We used values averaged over two month periods but looked at only one value per year (e.g. only spring temperature and precipitation). Hence the interarrival time is 1 yr. We have added a note to avoid confusion: "Note that in our application the interarrival time \mu=1 year because we consider yearly values of climate variables averaged over a specific time period (section 2.2)."

The authors examined 6 types of copulas in order to represent different combined effects of temperature and precipitation, i.e. dry and hot, dry and cold, wet and hot, wet and cold. My understanding is that the impact of these combinations can be due to a single variable or both variables being in an extreme state. The examined copulas, however, do not include extreme-value copulas, which are usually considered more appropriate for reproducing the interrelationship/interdependence structure between rare events. Did the authors compare the result using extreme-value copulas?

The t-copula and some of the Archimedean copulas allow tail dependence, i.e. dependence in the extremes, see e.g. Schoelzel and Friederichs (2008) Sections 4.1.2 and 4.2. The t-copula is symmetric such that if the upper tails are dependent, the lower tails are as well. In contrast, the Gumbel and Joe copulas can model only upper tail dependence whereas the Clayton copula can model lower tail dependence (all three are asymmetric). Note that in this application we aim for modeling the whole distribution to also capture events with high return frequency and not only the extremes. Hence, only using extreme value copulas might not be appropriate. We have added a note in Section 2.4 to highlight that the copulas that we use are able to capture dependence in the extremes.
"The t copula and some of the Archimedean copulas allow tail dependence, that is, dependence in the extremes, (see e.g. Schoelzel and Friederichs, (2008)). The t copula is symmetric such that if the upper tails are dependent, the lower tails are as well. In contrast, the Gumbel and Joe copulas can model only upper tail dependence, whereas the Clayton copula can model lower tail dependence (all three are asymmetric). Note that in this application, we aim for modeling the whole distribution to also capture events with low return periods and not only the extremes. Hence, only using extreme value copulas might not be appropriate."

Here the authors applied copulas to the seasonal and 2-month averaged climate variables. One of the prerequisites to apply copula is the assumption of temporal independence of variables, e.g. by examining the autocorrelation. Did the input variables meet this requirement?

Thank you for this comment. We have now tested for autocorrelation in the yearly time series of temperature and precipitation averaged over the different time scales use in the study (66 years). 8.8% of the temperature time series and 4.4% of the precipitation time series are significantly autocorrelated for lag 1 at the 5% level. Since we look at bivariate distributions, we conclude that autocorrelation should not have a significant effect on the conclusions of the paper. We have add this information in the revised manuscript in Section 2.2: "We tested for autocorrelation in the yearly time series of temperature and precipitation averaged over these different time scales. 8.8% of the temperature time series and 4.4% of the precipitation time series are significantly autocorrelated for lag 1 at the 5% level. Because our focus is on bivariate distributions, we conclude that autocorrelation does not have a significant effect on our conclusions."

Bivariate return periods of temperature and precipitation explain a large fraction of European crop yields

Jakob Zscheischler[1], Rene Orth[1], and Sonia I. Seneviratne[1]

[1]Institute for Atmospheric and Climate Science, ETH Zurich, Universitätstrasse 16, 8092 Zurich, Switzerland

Correspondence to: Jakob Zscheischler (jakob.zscheischler@env.ethz.ch)

Abstract. Crops are vital for human society. Crop yields vary with climate and it is important to understand how climate and crop yields are linked to ensure future food security. Temperature and precipitation are among the key driving factors of crop yield variability. Previous studies have investigated mostly linear relationships between temperature and precipitation, and crop yields variability. Other research has highlighted the adverse impacts of climate extremes such as drought and heat waves on crop yields. Impacts are, however, often non-linearly related to multivariate climate conditions. Here we derive bivariate return periods of climate conditions as indicators for climate variability along different temperature-precipitation gradients. We show that in Europe, linear models based on bivariate return periods of specific climate conditions explain on average significantly more crop yield variability (42%) than models relying directly on temperature and precipitation as predictors (36%). Our results demonstrate that most often crop yields increase along a gradient from hot and dry, to cold and wet conditions with lower yields associated with hot and dry periods. The majority of crops are most sensitive to climate conditions in summer and to maximum temperatures. The use of bivariate return periods allows the integration of nonlinear impacts into climate-crop yield analysis. This offers new avenues to study the link between climate and crop yield variability and suggests that they are possibly more strongly related than what is inferred from conventional linear models.

1 Introduction

Agriculture is essential to the well-being of humans and is directly affected by changes in climate (Lobell et al., 2011; Porter et al., 2014). Continuing climate change will likely increase the pressure on agriculture in the future with adverse impacts on food security (Porter et al., 2014; Wheeler and Von Braun, 2013; Rosenzweig et al., 2014). Understanding how crop yields

5 vary with climate may help to ensure future food security through increased predictability of yields and adequate adaptation measures. Trying to keep variability of yields low is a key objective as lower yield variability leads to more stable farmer incomes (Reidsma et al., 2010) and food supply (Slingo et al., 2005). Biological crop models exist at least since the 1970s (Loomis et al., 1979). However, such models are limited in their ability to quantify the impact of climate variability on crop yields over larger scales. For this purpose statistical models are applied. Globally, variations in precipitation and temperature

10 account for about a third of the observed yield variability (Lobell and Field, 2007; Ray et al., 2015). Due to their sensitivity to climate, crop yields are also strongly susceptible to climate extremes. In particular, disastrous droughts and heat waves often severely impact crop yields, reducing national cereal production by up to 9-10% (Lesk et al., 2016). How climate extremes impact crop production, however, varies across locations. Crops in the northern latitudes, for instance, might thrive under very warm conditions. In addition, impacts of climate extremes on vegetation strongly depends on when the climate extreme occurs

15 and in which developing stage the underlying vegetation is (De Boeck et al., 2011; Frank et al., 2015; Sippel et al., 2016). Because crops are often strongly adapted to specific climate conditions and covary with local climate (Osborne et al., 2009), it can be expected that the timing of climate anomalies is even more crucial for crops.

Crop yields can respond nonlinearly to changes in climate conditions and extremes (Porter and Semenov, 2005). Nonlinear impacts of extremely dry and hot conditions have been shown previously on terrestrial carbon fluxes. At the global scale,

20 the impacts of concurrent dry and hot periods on carbon fluxes exceed the sum of impacts from either hot or dry conditions (Zscheischler et al., 2014a). Studies investigating the variability of crop yields with climate usually rely on linear models (Osborne and Wheeler, 2013; Ray et al., 2015) though sometimes nonlinear transformations of climate variables are used as predictors (Ray et al., 2015). Building more complex nonlinear models is often not feasible because crop yield time series are too short (mostly not more then 60 years). However, climate extremes might have disproportionately large impacts on crops

25 (Lesk et al., 2016). Due to the sensitivity of crop yields to both temperature and precipitation (Ray et al., 2015) and extremes therein (Lesk et al., 2016), a variable incorporating the degree of extremeness of both variables at the same time might be more robustly related to crop yields. In the univariate setting, the magnitude of an extreme event is sometimes approximated by its return period. Here we use this analogy and derive bivariate return periods of temperature and precipitation which we then relate to crop yields. We explore how well bivariate return periods of temperature and precipitation are linked to crop yields.

30 Multivariate return periods have been studied increasingly in the recent past, mostly in the field of hydrology, for instance to study floods (Grimaldi and Serinaldi, 2006; Salvadori et al., 2011, 2013) and droughts (Shiau and Modarres, 2009; Wong et al., 2009; De Michele et al., 2013). The introduction of the concept of copulas into meteorology and climate research has boosted the usage and modeling of multivariate distributions (Schoelzel and Friederichs, 2008). A copula is a multivariate probability distributions for which the marginal probability distribution of each variable is uniform. Copulas can be used to describe the

dependence between random variables. They can greatly simplify calculations involving multivariate distributions and have lead to suite of definitions of multivariate return periods (Salvadori and De Michele, 2004; Serinaldi, 2015). Most studies so far have focused on return periods of droughts and floods in a multivariate setting. AghaKouchak et al. (2014) were probably the first to compute return periods of events involving both precipitation and temperature.

5 Generally, there is no unique way to define multivariate return periods (Serinaldi, 2015). As there is no natural ordering in higher dimensional spaces, one has to decide first in what direction one wants to "look". For instance in our case, we can compute return periods of hot and dry events, hot and wet events, cold and dry events, or cold and wet events. Furthermore, we have to decide how to set a threshold to compute threshold exceedances which will ultimately result in return periods of specific events. Various suggestions have been proposed in the recent years (Serinaldi, 2015). Here we use the definition motivated by

10 the notion of "critical regions" (Salvadori et al., 2011) whose return periods can be computed using so-called survival functions and survival copulas (Salvadori et al., 2013). The concept of critical regions has recently been well mathematically formalized via the notions of Hazard Scenarios and Upper Sets, providing a consistent mathematical framework. The interested reader is referred to Salvadori et al. (2016). A survival function is the complementary cumulative distribution function, i.e., for a distribution function $F(x) = \Pr(X \le x)$ the survival function, denoted by $\bar{F}$, is given by $\bar{F}(x) = \Pr(X > x)$. Survival copulas

15 are copulas based on survival functions.

Critical regions are separated from noncritical regions by a so-called critical layer $\{(x,y) \in R^2 : \bar{F}(x,y) = \Pr(X > x, Y > y) = t\}$, where all points $(x,y) \in R^2$ have the same probability $\bar{F}(x,y) = t$ in the joint survival distribution $\bar{F}$. The critical region is then defined as $\bar{R}_t^< = \{(x,y) \in R^2 : \bar{F}(x,y) < t\}$ (De Michele et al., 2013). A return period RP of an event can be defined as the inverse of the probability of falling in the critical region,

$$RP = \frac{\mu}{\Pr((X,Y) \in \bar{R}_t^<)} \tag{1}$$

where μ is the average interarrival time of events (1 year in our case). As we will see, from this definition it follows that events in which only one variable is extreme (e.g., either extreme heat but no drought, or extreme drought but no heat) and events in which both variables are moderately extreme (e.g., moderate heat and drought) may have similar return periods. Crop yield vary with climate, most importantly temperature and precipitation (Ray et al., 2015). If we assume that temperature and

25 precipitation at potentially different time periods are of similar relevance for crop yields, a relationship between bivariate return periods of precipitation and temperature can be expected.

AghaKouchak et al. (2014) applied the concept of survival copulas to calculate return periods of critical regions (Salvadori et al., 2013) with precipitation and temperature data from California and estimated a return period of 200 years for the very hot and dry year 2014. Here we apply this approach to climate data in Europe and investigate relationships between bivariate

30 return periods of combinations of temperature and precipitation with crop yield variability. By "looking in all directions" in the temperature-precipitation space we assess whether an intensification along a certain direction (dry and hot, dry and wet etc.) leads to an increase or decrease of crop yields. By computing return periods for different combinations of months (e.g. temperature in spring, precipitation in June and July) we estimate which combinations of months and climate conditions mostly affect crop yields. We compare this analysis with linear models fitted to the two predictors precipitation and temperature.

2 Methods

2.1 Data

All analyses are focused on Europe. We obtained annual EUROSTAT national crop yield statistics for 19 European countries for the period 1955-2015 from the European Commission (http://epp.eurostat.ec.europa.eu). Due to their importance for the European crop production (Leff et al., 2004) and good spatiotemporal coverage, we selected the following crops (following the official EUROSTATS terminology): Cereals (excluding rice), Wheat (including spelt), Maize (grain maize and corn-cob mix), Potatoes (including early potatoes and seed potatoes), and Sugar beet (excluding seed). Wheat does not include winter and spring varieties. Sugar beet is sown in spring and harvested in autumn in Europe. In the EUROSTAT database, yield is reported as the amount of dry matter suitable for consumption (Moors et al., 2010). Not every country grows every crop every year, accordingly some time series contain missing values. In total we collected 3552 crop yield years distributed over 19 countries and 5 crops. Most data was available for cereals (781 crop yield years) and potatoes (766 crop yield years) followed by wheat (734), sugar beet (672), and maize (599 crop yield years).

We further used daily minimum (T_{min}), mean (T) and maximum temperature (T_{max}), and precipitation (P) data from E-OBS (Haylock et al., 2008) at 0.25° spatial resolution and computed spatial averages for the 19 countries for which we have collected crop yields. Because crop area potentially changed over time, for each country we averaged climate data over the entire country. In this way we investigate how climate variability at country scale is related to annual crop yields.

2.2 Data preprocessing

Long-term trends in crop yields are mostly caused by technical progress such as breeding, and changes in land use policy and management practices. Various corrections are typically used to adjust for this effect. Often the difference in values from one year to the next (Lobell and Field, 2007), or linear or quadratic detrending (Ray et al., 2015) are applied to obtain anomalies of crop yields independent of long-term trends which are more closely related to climate variability. Here we rely on a different, more adaptive approach (Wu et al., 2014). We fitted cubic smoothing splines to each crop yield time series to capture the long-term trend. The original time series was then divided by this trend to obtain a dimensionless index. Values greater than 1 indicate positive yield anomalies whereas values below 1 indicate negative yield anomalies. The cubic smoothing splines were fitted with a frequency cutoff of 50% at 2/3 of the time series length to remove low frequency variability. This procedure is analogous to standard dendrochronological procedures, where the aim is to filter out low-frequency trends related to tree geometry and not to climate (Babst et al., 2012; Cook and Peters, 1981). We used the R package *dplR* (Bunn, 2008) for these computations and denote the resulting yield anomalies by Y. We used the same normalization to subtract long-term trends in climate data (T_{min}, T, T_{max}, and P) for each country, similar to (Wu et al., 2014). The resulting indices preserve high-frequency (i.e., interannual) variability, but not the trends or low-frequency signals (Babst et al., 2012; Cook and Peters, 1997).

We averaged normalized climate variables over different time periods to assess the sensitivity of crops towards different climate variables and different time periods. In particular, we computed seasonal and 2-month averages during the growing

season and averaged climate variables (T_{min}, T, T_{max}, and P) over the months 3-8, 3-5, 6-8, 3-4, 4-5, 5-6, 6-7, 7-8, and 8-9. We selected these periods to cover seasonal and bimonthly climate variations over the time period relevant for the crops. We tested for autocorrelation in the yearly time series of temperature and precipitation averaged over these different time scales. 8.8% of the temperature time series and 4.4% of the precipitation time series are significantly autocorrelated for lag 1 at the 5% level. Because our focus is on bivariate distributions, we conclude that autocorrelation does not have a significant effect on our conclusions. Subsequently, we computed bivariate return periods for P and all variants of T for all combinations of averaging periods as described below (section 2.3). All our analyses were done in the R programming language version 3.1.2 (R Core Team, 2014).

2.3 Bivariate return periods

The computation of bivariate return periods was done based on copulas similar to (AghaKouchak et al., 2014). The joint distribution of two variables X (e.g., temperature) and Y (e.g., precipitation) can be written as (Sklar, 1959; Salvadori and De Michele, 2004)

$$F(x,y) = \Pr(X \leq x, Y \leq y) = C(F_X(x), F_Y(y)) \tag{2}$$

with the marginal cumulative distribution functions $F_X(x) = F(X \leq x), F_Y(y) = F(Y \leq y)$ $F_X(x) = \Pr(X \leq x), F_Y(y) = \Pr(Y \leq y)$, and a copula C. Copulas are multivariate probability distributions with uniform marginal distributions designed to model the dependence between multiple variables (Nelsen, 1999). As we are interested in exceedance probabilities, we rely on survival functions. The so-called joint survival distribution $\bar{F}(x,y) = \Pr(X > x, Y > y)$ can be obtained using the concept of survival copulas (Salvadori et al., 2011, 2013)

$$\bar{F}(x,y) = \hat{C}(\bar{F}_X(x), \bar{F}_Y(y)) \tag{3}$$

with marginal survival functions $\bar{F}_X = 1 - F_X$ and $\bar{F}_Y = 1 - F_Y$, and $\hat{C}$ the survival copula. For any given $(x,y) \in R^2$, there exists a unique survival critical layer on which the set of realizations of X and Y share the same probability $t \in (0,1)$ (De Michele et al., 2013; AghaKouchak et al., 2014): $\bar{L}_t^{\bar{F}} = \{(x,y) \in R^2 : \bar{F}(x,y) = t\}$. The survival critical layer partitions R^2 into a "safe" and a "critical" region $\bar{R}_t^< = \{(x,y) \in R^2 : \bar{F}(x,y) < t\}$. A bivariate return period RP can now be defined as the probability of falling inside the critical region $\bar{R}_t^<$ (Salvadori et al., 2013; De Michele et al., 2013) through

$$RP = \frac{\mu}{1 - \bar{K}(t)}. \tag{4}$$

where $\mu = 1$ year and $\bar{K}$ is the Kendall's survival function associated with $\bar{F}$, defined as

$$\bar{K}(t) = \Pr(\bar{F}(X,Y) \geq t) = \Pr(\hat{C}(\bar{F}_X(x), \bar{F}_Y(y)) \geq t). \tag{5}$$

There exist analytical expressions for Kendall's survival function if the copula is Archimedean (Salvadori and De Michele, 2004). Otherwise it can be obtained through simulations (Salvadori et al., 2011). Note that in our application the interarrival

time $\mu = 1$ year because we consider yearly values of climate variables averaged over a specific time period (section 2.2). For a given temperature-precipitation couple (x, y) we can now (i) determine the survival critical layer $\bar{L}_t^{\bar{F}}$ associated with $t = \bar{F}(x, y) = \hat{C}(\bar{F}_X(x), \bar{F}_Y(y))$, (ii) compute the probability of an event being at least as extreme as (x, y) (i.e., falling inside the critical region defined by t) by $p = 1 - \bar{K}(t)$, and finally (iii) compute return periods via (4).

2.4 Evaluation

For each combination (choice of T, averaging period of T, averaging period of P, choice of bivariate direction: dry and hot, dry and cold, wet and hot, wet and cold) we fitted 6 copulas: the normal copula, the t copula (Schoelzel and Friederichs, 2008), and the four Archimedean copulas Clayton, Frank, Gumbel, and Joe (Embrechts et al., 2001). The t copula and some of the Archimedean copulas allow tail dependence, that is, dependence in the extremes, (see e.g. Schoelzel and Friederichs, 2008). The t copula is symmetric such that if the upper tails are dependent, the lower tails are as well. In contrast, the Gumbel and Joe copulas can model only upper tail dependence, whereas the Clayton copula can model lower tail dependence (all three are asymmetric). Note that in this application, we aim for modeling the whole distribution to also capture events with low return periods and not only the extremes. Hence, only using extreme value copulas might not be appropriate. For further analysis we only kept the copula with the best fit for each combination using the Bayesian Information Criterion implemented in the R package *VineCopula* (Schepsmeier et al., 2015). Goodness of fit was tested based on the ~~Cramer-von-Mises~~ Cramér-von Mises functional (Genest et al., 2009) implemented in the R package *copula* (Hofert et al., 2015). 3 out of 91 copulas had a p-value below 0.05, which is in the expected range. We subsequently computed return periods with (4). We then fitted linear models to explain crop yield anomalies as a function of bivariate return periods using the natural logarithm of the bivariate return periods as a predictor (i.e., $Y_c = a + b \ln(RP)$) and only kept the model with the highest fraction of explained variance. This model was denoted by M_{rp}, using rp as abbreviation for return period. Furthermore, we fitted ordinary linear models to each combination of climate and time period using P and the different variants of T as predictors (i.e., crop yields are modeled as $Y_c = a + bT + cP$ with $T \in \{T_{min}, T, T_{max}\}$) and also only kept the model with the highest fraction of explained variance. We denoted this model by M_{TP} to refer to the two predictors T and P which are used in this linear model. We then compare M_{rp} and M_{TP} with respect to the fraction of explained variance using the adjusted R^2 to account for the different number of predictors (one in M_{rp}, two in M_{TP}). We use the non-parametric Mann-Whitney U test to test whether the models M_{rp} achieve an overall significantly higher R^2 than the models M_{TP}.

3 Results

We first present three examples for particular country-crop combinations illustrating our approach, and the comparison between M_{rp} and M_{TP} (section 3.1). The examples are chosen to illustrate (i) the various sensitivities of different crops to different temperature-precipitation gradients and (ii) differences in the selected time periods and the selected temperature variable between the models M_{rp} and M_{TP}. Subsequently summary statistics of all countries and crops are presented (section 3.2).

3.1 Examples

Yields often vary along a specific precipitation-temperature gradient. In figure 1 we present an example of our approach for maize yield anomalies in Germany. Maize yield anomalies are high when maximum temperatures in July and August are low and precipitation in June and July is high (lower left of figure 1a). Yields decrease with increasing temperatures and decreasing precipitation. Return periods of hot and dry events increase along the same gradient, as denoted by the curves of equal return periods. Accordingly, bivariate return periods along this gradient capture the variability of maize yield anomalies well, explaining 72% of their variance (figure 1b). This variability is less well captured if a linear model is directly fitted to the same predictors (M_{TP}, $R^2 = 0.62$, figure 1c).

In a similar fashion, figure 2 presents an example of cereal yields in Lithuania. In this case, yields are high when maximum temperatures are low in June and July, and when precipitation is low in August and September (lower left of figure 2a). They decrease with increasing temperatures and, in contrast to our first example, increasing precipitation most likely because Lithuania's climate is already wet (it has humid continental climate with an average annual rainfall of around 620 mm). Return periods of hot and wet conditions increase along the same gradient and thus they capture the variability of cereal yield anomalies well ($R^2 = 0.56$, figure 2b). Also in this example, the variability in yield anomalies is less well captured if a linear model is fitted to the same predictors (M_{TP}, $R^2 = 0.47$, figure 2c).

Finally, figure 3 presents an example of yields of sugar beet in Slovenia. Here, yields are high when maximum temperatures are low in April and May, and when precipitation is high in July and August (upper right in figure 3a). Hence, yields increase with decreasing temperatures and increasing precipitation. Return periods of cold and wet events increase along this gradient and capture most of the variability of yield anomalies in cereals ($R^2 = 0.68$, figure 3b). Note that here, because of their direction, the bivariate return periods are positively related to yield anomalies. The variability in yield anomalies is slightly less well captured if a linear model is fitted to the maximum temperature anomalies averaged over April and May, and precipitation anomalies averaged over July and August (M_{TP}, $R^2 = 0.64$, figure 3c and d). This is also suggested when one compares how well the curves of equal return periods (figure 3a) and the levelsets of the best linear model (M_{TP}, figure 3c) capture yield anomalies that lie in a similar range.

3.2 Summary statistics

The models M_{rp} perform significantly better than M_{TP} ($p < 0.05$, figure 4). This is also evident for the individual crop-country combinations (figure 5). On average the fraction of explained variance is 42% for M_{rp} (colored bars) versus 36% for M_{TP} (white bars). This corresponds to an increase in predictability of 17%. Crop yield anomalies mostly vary strongest along a hot-dry to cold-wet gradient with low yields during hot and dry conditions (41 out of 91, 45%, red bars). Crop-country combinations where yield anomalies increase along the hot-wet to cold-dry gradient (green bars) represent 29 of all 91 cases (about 32%). The other two gradients appear less often, cold-wet to hot-dry 10 times (blue bars) and cold-dry to hot-wet 11 times (purple) out of 91.
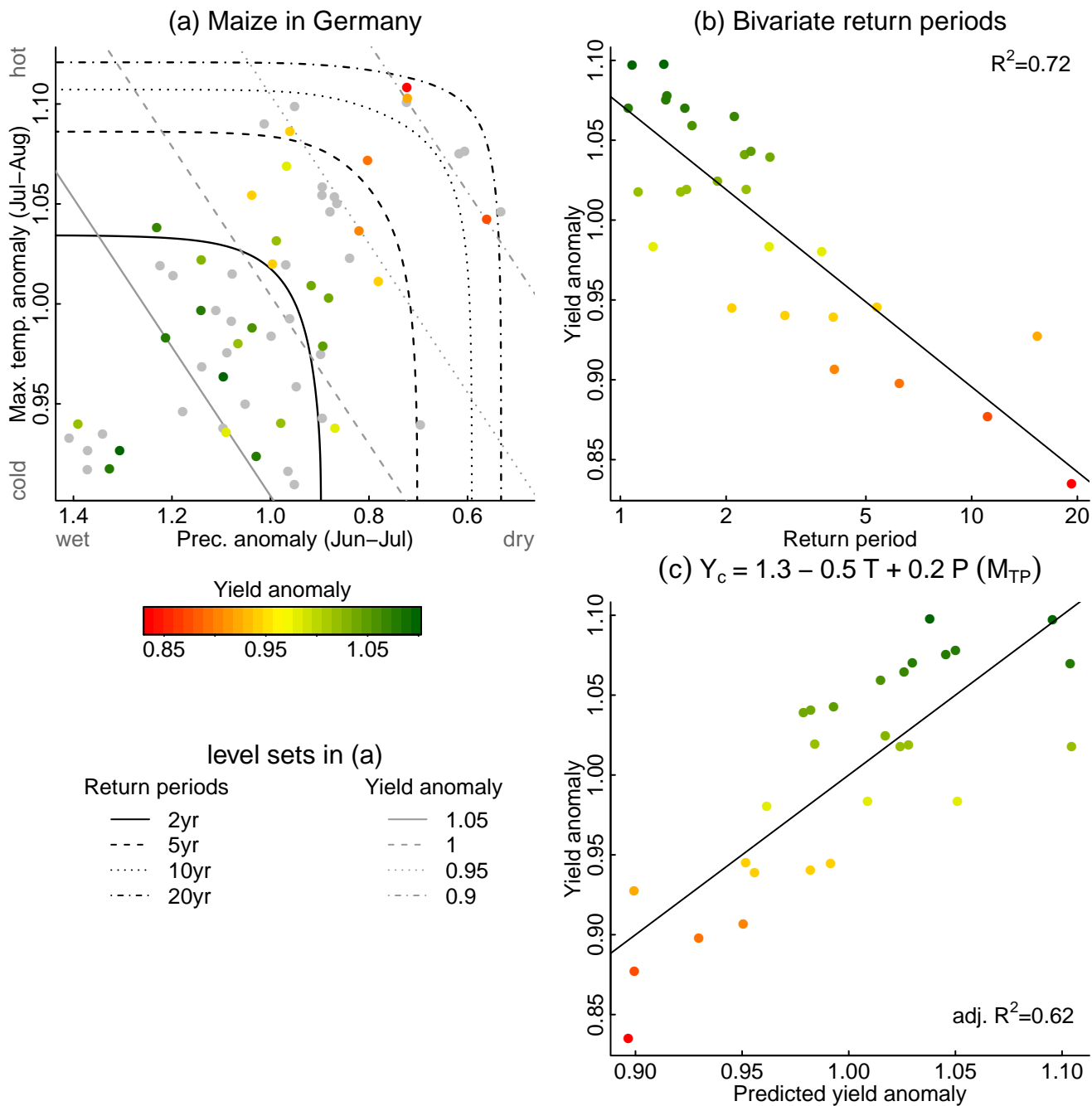
Figure 1. Maize in Germany. (a) Maximum temperature anomalies averaged over July and August versus precipitation anomalies averaged over June and July. Dots are colored according to maize yield anomalies if yield data was available and gray otherwise. Cold and wet conditions are on the lower left. Bivariate return periods increase from the lower left to the upper right as indicated by the curves. Straight lines indicate levels of equal yield anomalies as predicted by the linear model M_{TP}. (b) Maize yield anomaly versus bivariate return periods. The line indicates the best linear fit with $R^2 = 0.72$. (c) Linear model using predictors from (a) (M_{TP}). The adjusted R^2 is 0.62.
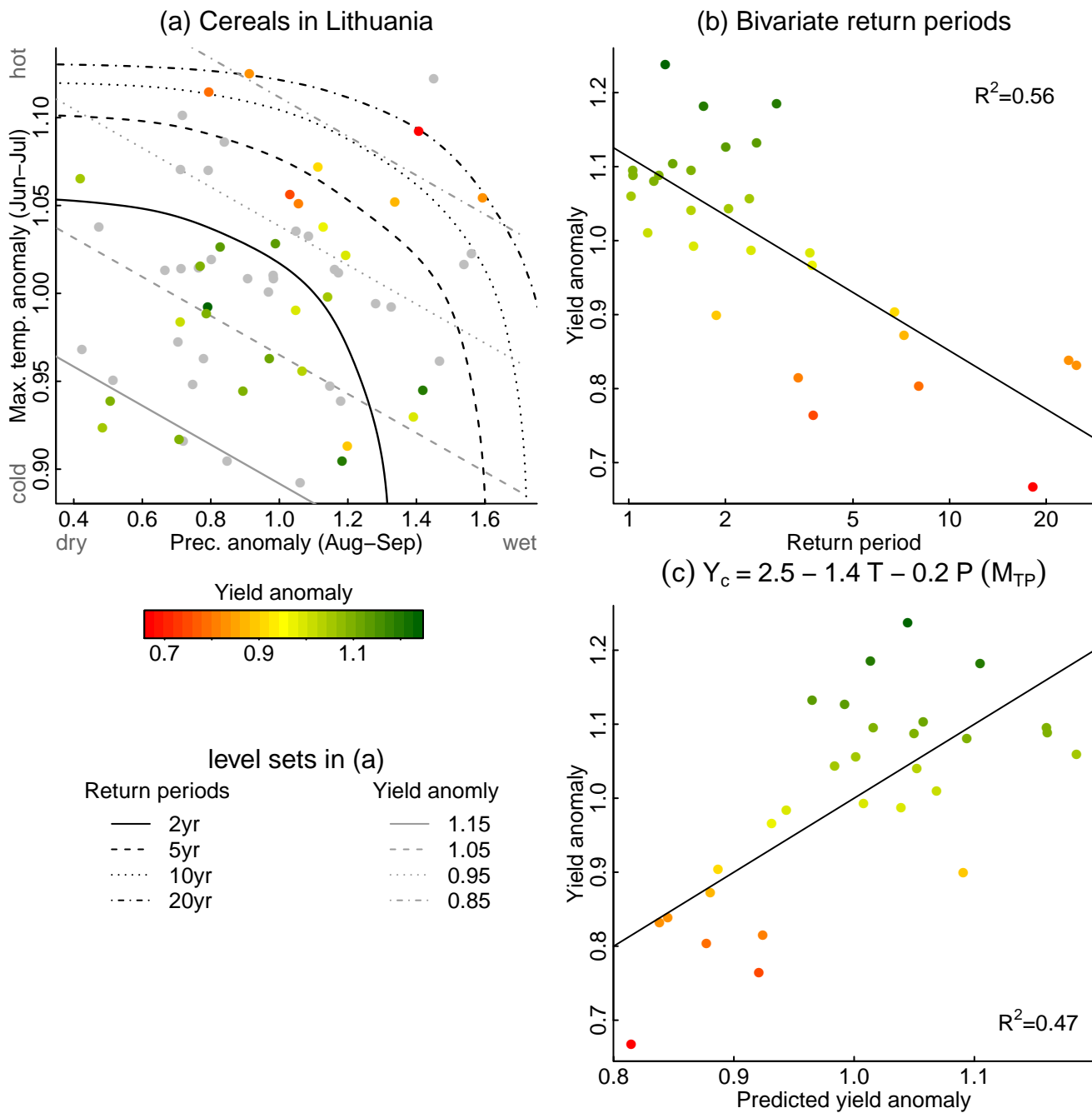
Figure 2. Cereals in Lithuania, similar to figure 1. (a) Maximum temperature anomalies averaged over June and July versus precipitation anomalies averaged over (August and September). Dots are colored according to cereal yield anomalies if yield data was available and gray otherwise. Dry and cold conditions are on the lower left. (b) Cereals yield anomaly versus bivariate return periods. The line indicates the best linear fit with $R^2 = 0.56$. (d) Linear model using predictors from (a) (M_{TP}). The adjusted R^2 is 0.47.
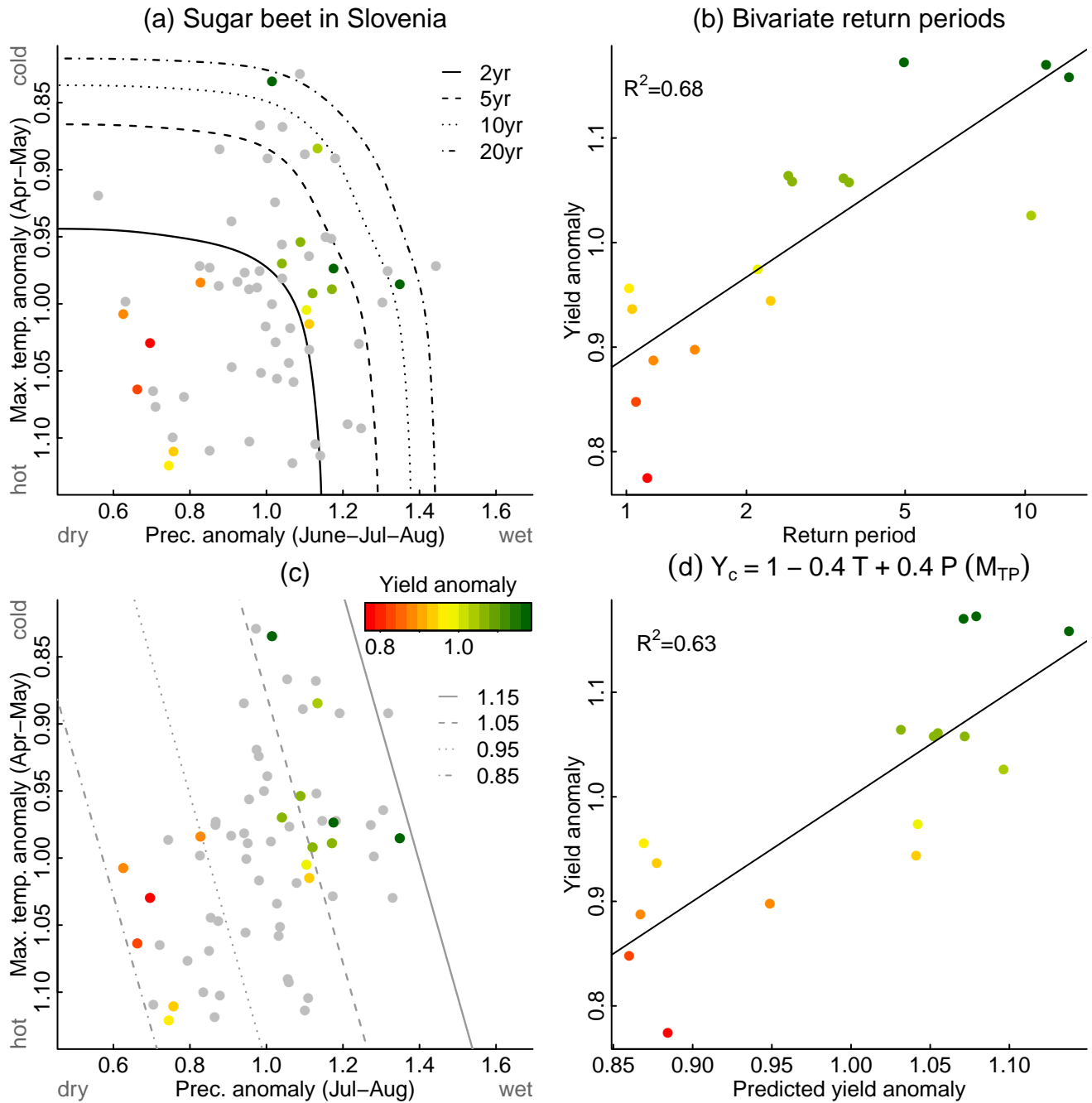
Figure 3. Sugar beet in Slovakia, similar to figure 1. (a) Maximum temperature anomalies averaged over April and May versus precipitation anomalies averaged over summer (June, July and August). Dots are colored according to sugar beet yield anomalies if yield data was available and gray otherwise. Dry and hot conditions are on the lower left. (b) Yield anomalies of sugar beet versus bivariate return periods. The line indicates the best linear fit with $R^2 = 0.68$. (c) As in (a) but with precipitation averaged over July and August. Lines indicate levels of equal yield anomalies as predicted by the linear model M_{TP}. (d) Linear model using predictors from (c) (M_{TP}). The adjusted R^2 is 0.64.
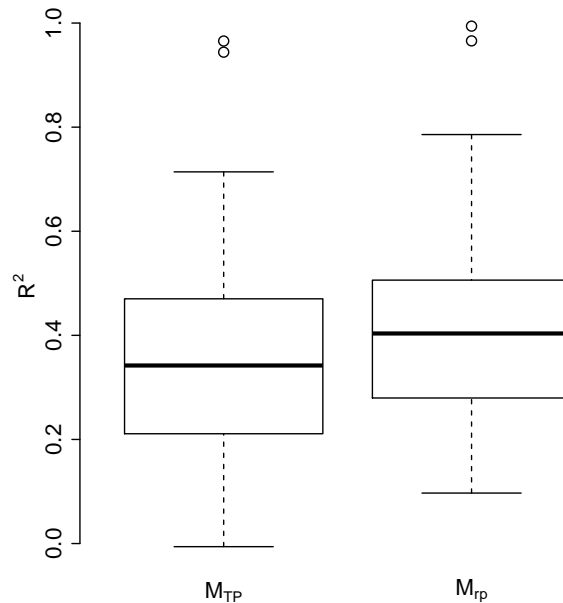
Figure 4. R^2 from all countries and crops (see figure 5) summarized in boxplots. Medians are 0.34 for M_{TP} and 0.4 for M_{rp}.

Most often, T_{max} is selected as temperature variable which is most relevant to determine crop yield variability. For the M_{rp} model, T_{max} is chosen 46 out of 91 times (51%). For the M_{TP} model , it is chosen 38 times (42%). T_{min} is chosen 29 (32%) and 33 times (36%), respectively.

For the M_{rp} model there is a large variety in the time periods to which crop yields are most sensitivity towards (figure 6).

5 Notably, specific precipitation sums over longer time periods (mostly 3 months and more) seems to be most relevant for crops. In contrast, crop yields are sensitive to temperature conditions at shorter time scales (generally less than 3 months). Precipitation averaged over the whole spring and summer (13 incidences) or over one season (28 incidences) was selected in nearly 41% of the cases while averaging periods for temperatures were only 2 months in nearly 86% of the cases (78 incidences out of 91). In total, for precipitation 2.5 times more months are selected than for temperature (524 versus 211). Overall, climate conditions

10 in summer (JJA) seem to be more relevant for crops and climate during other periods. For both temperature and precipitation, 63% of the selected months fall in the summer.

4 Discussion and Conclusions

Linear models based on return periods of specific climate conditions (M_{rp}) can better explain variability in annual crop yield anomalies than linear models based directly on precipitation and temperature (M_{TP}, 42% versus 36% of explained variance in

15 crop yield anomalies on average, $p < 0.05$). The reason for this might be that yields react nonlinearly to more extreme climate
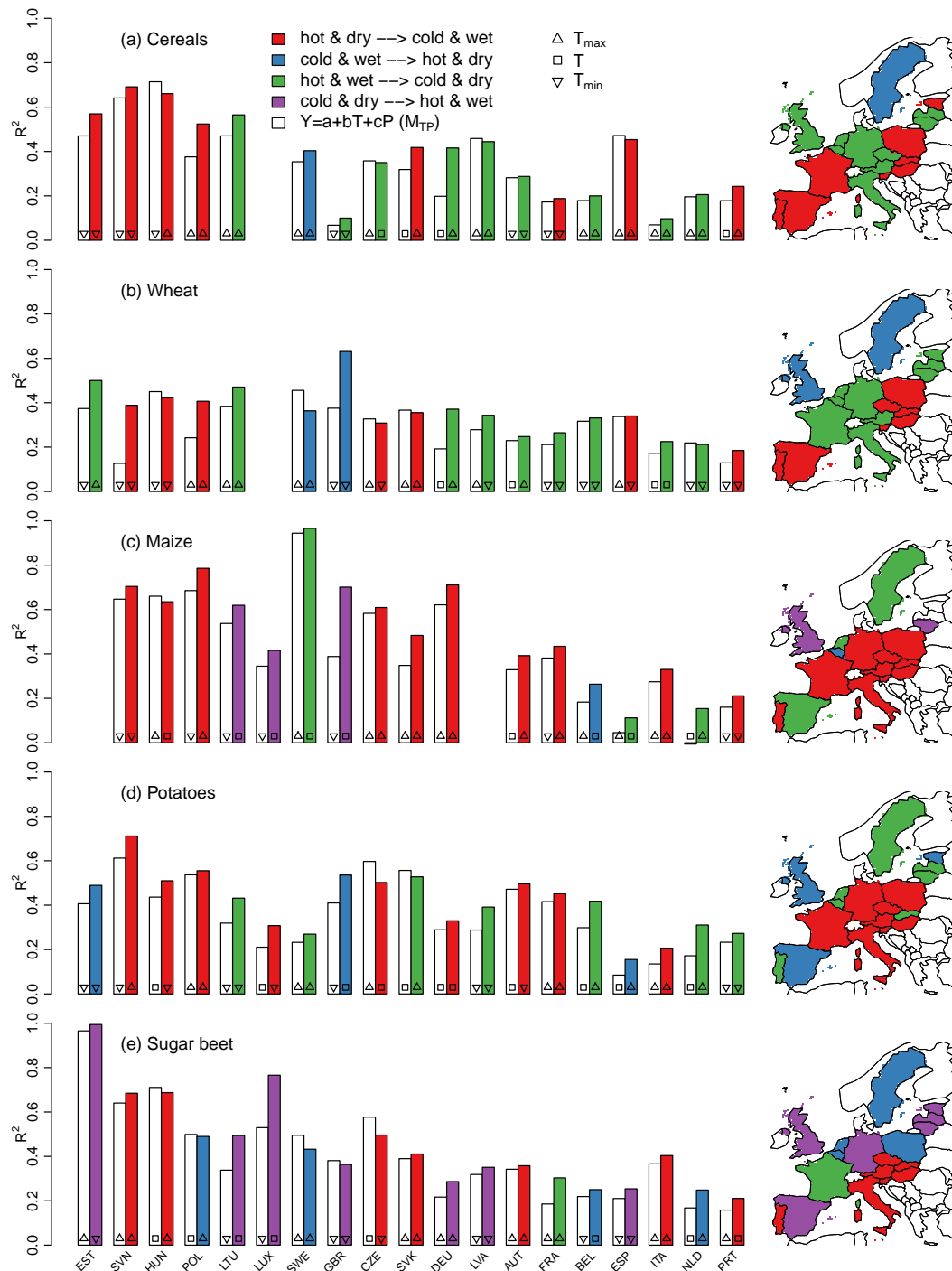
Figure 5. Summary of all R^2 derived from linear models using either return periods (M_{rp}, colored bars) or temperature and precipitation anomalies (M_{TP}, white bars) as predictors for yield anomalies. The temperature variable which was used (T_{max}, T, T_{min}) is denoted with symbols at the bottom of each bar. Yield anomalies co-vary positively with the climate gradients along the four directions denoted by the legend. Countries are sorted by the average performance of the models M_{rp}.
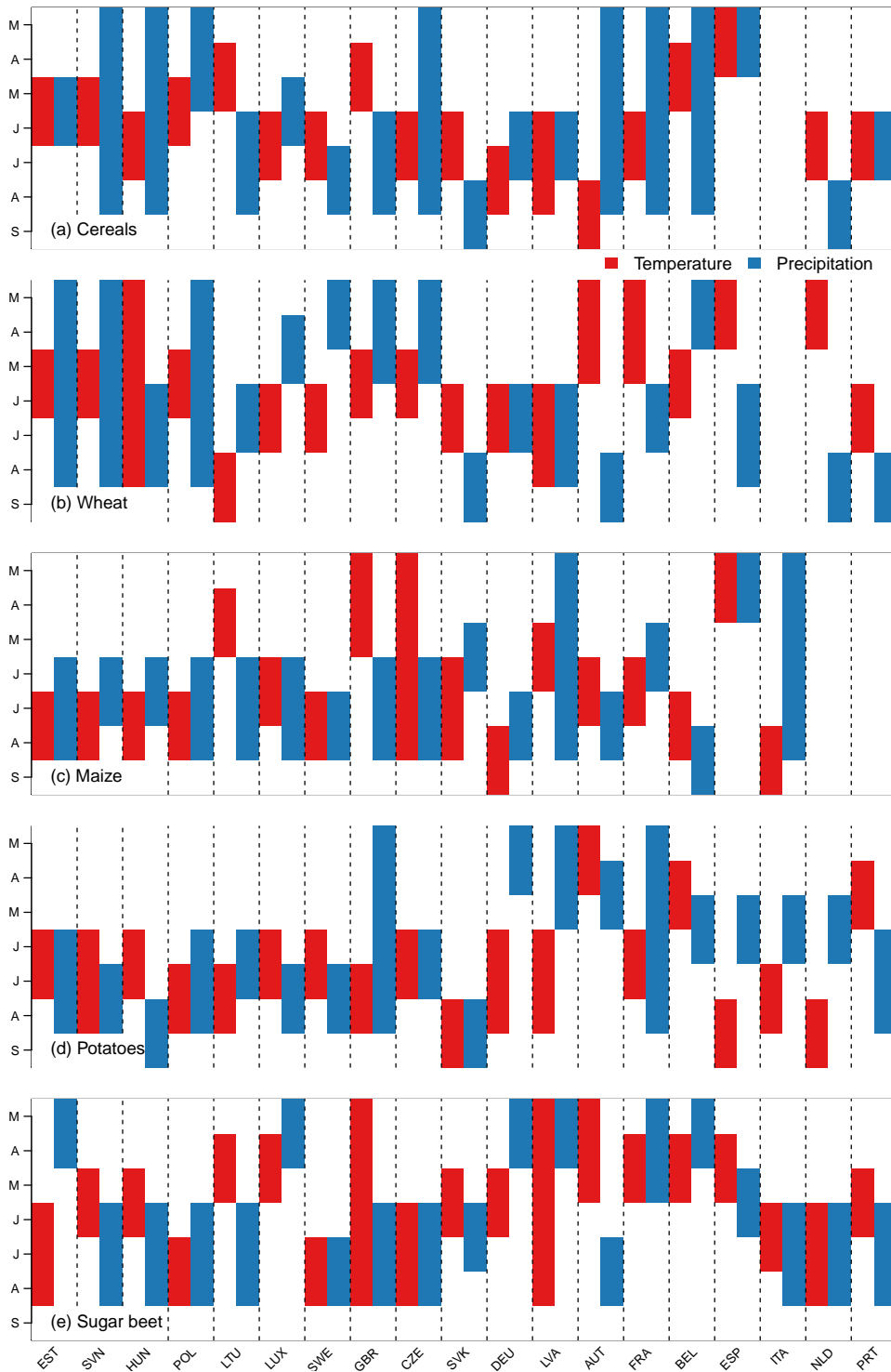
Figure 6. Summary of which months are selected to obtain the model with the best performance based on bivariate return periods of temperature and precipitation (M_{rp}). Months range from March till September. Red bars denote temperature, blue bars precipitation.

conditions. Including more variables such as radiation, humidity, and soil moisture might further improve the prediction of crop yield anomalies. However, due to the shortness of yield time series (more than 50% of the yield series contain less than 30 years of data), robust models with many predictors are difficult to build. Our analysis demonstrates the different sensitivities of crops to climate conditions. Figures 5 and 6 can be used in combination to disentangle the sensitivities of crop yields to climate

5   conditions at certain time periods. For example, in most countries Maize yields decrease for hot and dry conditions. However, in countries such as Lithuania, Luxembourg and Great Britain, maize yields increase under hot and wet conditions (figure 5). For all three countries, maize yields are most sensitive to precipitation in summer (figure 6). Overall, crops are most sensitive to climate variations along the hot-dry to cold-wet gradient. This in line with what can be expected from univariate assessments (Lesk et al., 2016), showing that crop yields can decrease substantially under heat or drought. Further, crops are most sensitive

10   to maximum temperatures (in contrast to mean and minimum temperatures) whereas the most important time period largely depends on the country and the crop. Because crops can exhibit a thresholds behavior at high temperatures resulting in crop damage (Porter and Semenov, 2005), maximum temperatures might be the most crucial variable for yield variability in most countries and crops.

4.1   Uncertainties in the estimation procedure

15   Estimating return periods is not free of uncertainties. Bivariate copulas are defined on the uniform marginal distributions, which is achieved by ranking of the original marginal distributions. Copulas are thus related to the original bivariate distribution by (3). Naturally, the accuracy of this transformation improves with increasing sample size. After the data have been transformed to the unit square, copulas have to be fitted. The fitted copulas we used passed standard goodness-of-fit tests though longer climate time series certainly help to obtain more robust fits (E-OBS data currently covers 1950-2015). Uncertainties in the

20   fitting procedure may be particularly large for extreme return periods (Serinaldi, 2015).

4.2   Return periods of climate events and their relation to impact-related variables

Despite the above-mentioned uncertainties, which are particularly relevant for extreme return periods, we argue that bivariate return periods of temperature and precipitation conditions can be used as a robust estimate of the climate variation along a specific temperature-precipitation gradient, outperforming conventional linear models. In particular if both temperature and

25   precipitation are strongly affecting crop yields, bivariate return periods will be related to yield variation. This assumption holds for many crops and regions in Europe (Peltonen-Sainio et al., 2010) but might be different for other regions in the world (Ray et al., 2015). The gradient which is most strongly related to crop yield variability or other impact variables can be chosen by testing different directions in the temperature-precipitation space. In addition, due to the nonlinear nature of return periods, extreme conditions might be captured more appropriately. In particular, extreme heat and drought can have extreme impacts

30   on crops which might exceed the impacts predicted by linear models (Porter and Semenov, 2005; Zscheischler et al., 2014a; Lesk et al., 2016).

For most country-crop combinations, crops are sensitive to climate along the hot-dry to cold-wet gradient, with hotter and drier conditions leading to lower yields, though there is large variability among countries (figure 5). This is in line with previous

research on the relationship between climate and other ecosystem variables. For instance, photosynthetic carbon uptake also generally increases along the hot-dry to cold wet gradient (Zscheischler et al., 2014a; Ahlström et al., 2015). Furthermore, most negative impacts on terrestrial carbon uptake are associated with hot and dry conditions (Zscheischler et al., 2014b).

We have demonstrated how bivariate return periods of climate conditions as an indicator for climate variability along a certain climate gradient can be used to predict crop yield anomalies. Our proposed approach explains significantly more yield variability than linear models based on temperature and precipitation. However, climate time series should be long enough to robustly estimate bivariate return periods (preferably at least 50 years). Crops largely rely on precipitation and temperature at the right time periods, depending on the specific crop and the location. Our analysis illustrated that crops are often most sensitive to summer climate, maximum temperatures, and specific precipitation conditions over longer time periods (at least 3 months). Other ecological variables sensitive to impacts of climate variability such as tree rings and terrestrial carbon uptake and release may be analyzed in a similar fashion. If other driver variables despite precipitation and temperature are crucially relevant for an impact variable (such as for instance radiation, humidity, and soil moisture) extensions to higher-dimensional return periods are possible, though the choice of the direction of a gradient becomes more difficult in higher dimensions. In conclusion, bivariate return periods allow to summarize bivariate climate conditions in a novel univariate measure which then can be used more easily to investigate associated impacts.

Competing interests. The authors declare that they have no conflict of interest.

References

AghaKouchak, A., Cheng, L., Mazdiyasni, O., and Farahmand, A.: Global warming and changes in risk of concurrent climate extremes: Insights from the 2014 California drought, Geophysical Research Letters, 41, 8847–8852, 2014.

Ahlström, A., Raupach, M. R., Schurgers, G., Smith, B., Arneth, A., Jung, M., Reichstein, M., Canadell, J. G., Friedlingstein, P., Jain, A. K.,
5   Kato, E., Poulter, B., Sitch, S., Stocker, B. D., Viovy, N., Wang, Y. P., Wiltshire, A., Zaehle, S., and Zeng, N.: The dominant role of semi-arid ecosystems in the trend and variability of the land CO2 sink, Science, 348, 895–899, doi:10.1126/science.aaa1668, 2015.

Babst, F., Carrer, M., Poulter, B., Urbinati, C., Neuwirth, B., and Frank, D.: 500 years of regional forest growth variability and links to climatic extreme events in Europe, Environmental Research Letters, 7, 045 705, 2012.

Bunn, A. G.: A dendrochronology program library in R (dplR), Dendrochronologia, 26, 115–124, 2008.

10 Cook, E. R. and Peters, K.: The smoothing spline: a new approach to standardizing forest interior tree-ring width series for dendroclimatic studies, Tree-ring bulletin, 1981.

Cook, E. R. and Peters, K.: Calculating unbiased tree-ring indices for the study of climatic and environmental change, The Holocene, 7, 361–370, 1997.

De Boeck, H. J., Dreesen, F. E., Janssens, I. A., and Nijs, I.: Whole-system responses of experimental plant communities to climate extremes
15   imposed in different seasons, New Phytologist, 189, 806–817, 2011.

De Michele, C., Salvadori, G., Vezzoli, R., and Pecora, S.: Multivariate assessment of droughts: frequency analysis and dynamic return period, Water Resources Research, 49, 6985–6994, 2013.

Embrechts, P., Lindskog, F., and McNeil, A.: Modelling Dependence with Ccopulas and Applications to Risk Management, in: Handbook of Heavy Tailed Distributions in Finance, edited by Rachev, S., pp. 329–384, Elsevier, 2001.

20 Frank, D., Reichstein, M., Bahn, M., Frank, D., Mahecha, M. D., Smith, P., Thonicke, K., van der Velde, M., Vicca, S., Babst, Flurin, Beer, C., Buchmann, N., Canadell, J. G., Ciais, P., Cramer, W., Ibrom, A., Miglietta, F., Poulter, B., Rammig, A., Seneviratne, S. I., Walz, A., Wattenbach, M., Zavala, M. A., and Zscheischler, J.: Effects of climate extremes on the terrestrial carbon cycle: concepts, processes and potential future impacts, Global Change Biology, doi:10.1111/gcb.12916, 2015.

Genest, C., Rémillard, B., and Beaudoin, D.: Goodness-of-fit tests for copulas: A review and a power study, Insurance: Mathematics and
25   economics, 44, 199–213, 2009.

Grimaldi, S. and Serinaldi, F.: Asymmetric copula in multivariate flood frequency analysis, Advances in Water Resources, 29, 1155–1167, 2006.

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, Journal of Geophysical Research: Atmospheres, 113, doi:10.1029/2008JD010201,
30   2008.

Hofert, M., Kojadinovic, I., Maechler, M., and Yan, J.: copula: Multivariate dependence with copulas, R package version 0.999-13, 2015.

Leff, B., Ramankutty, N., and Foley, J. A.: Geographic distribution of major crops across the world, Global Biogeochemical Cycles, 18, 2004.

Lesk, C., Rowhani, P., and Ramankutty, N.: Influence of extreme weather disasters on global crop production, Nature, 529, 84–87, 2016.

35 Lobell, D. B. and Field, C. B.: Global scale climate–crop yield relationships and the impacts of recent warming, Environmental research letters, 2, 014 002, 2007.

Lobell, D. B., Schlenker, W., and Costa-Roberts, J.: Climate trends and global crop production since 1980, Science, 333, 616–620, 2011.

Loomis, R. S., Rabbinge, R., and Ng, E.: Explanatory models in crop physiology, Annual Review of Plant Physiology, 30, 339–367, 1979.

Moors, E. J., Jacobs, C., Jans, W., Supit, I., Kutsch, W. L., Bernhofer, C., Béziat, P., Buchmann, N., Carrara, A., Ceschia, E., Elbers, J., Eugster, W., Kruijt, B., Loubet, B., Magliulo, E., Moureaux, C., Olioso, A., Saunders, M., and Soegaard, H.: Variability in carbon exchange of European croplands, Agriculture, ecosystems & environment, 139, 325–335, 2010.

5 Nelsen, R. B.: An introduction to copulas, Springer, New York, 1999.

Osborne, T., Slingo, J., Lawrence, D., and Wheeler, T.: Examining the interaction of growing crops with local climate using a coupled crop-climate model, Journal of Climate, 22, 1393–1411, 2009.

Osborne, T. M. and Wheeler, T. R.: Evidence for a climate signal in trends of global crop yield variability over the past 50 years, Environmental Research Letters, 8, 024 001, 2013.

10 Peltonen-Sainio, P., Jauhiainen, L., Trnka, M., Olesen, J. E., Calanca, P., Eckersten, H., Eitzinger, J., Gobin, A., Kersebaum, K. C., Kozyra, J., et al.: Coincidence of variation in yield and climate in Europe, Agriculture, ecosystems & environment, 139, 483–489, 2010.

Porter, J., Xie, L., Challinor, A., Cochrane, K., Howden, S., Iqbal, M., Lobell, D., and Travasso, M.: Food security and food production systems, pp. 485–533, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014.

Porter, J. R. and Semenov, M. A.: Crop responses to climatic variation, Philosophical Transactions of the Royal Society of London B: Biological Sciences, 360, 2021–2035, 2005.

15 R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/, 2014.

Ray, D. K., Gerber, J. S., MacDonald, G. K., and West, P. C.: Climate variation explains a third of global crop yield variability, Nature communications, 6, 2015.

20 Reidsma, P., Ewert, F., Lansink, A. O., and Leemans, R.: Adaptation to climate change and climate variability in European agriculture: the importance of farm level responses, European journal of agronomy, 32, 91–102, 2010.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A. C., Müller, C., Arneth, A., Boote, K. J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Piontek, F., Pugh, T. A. M., Schmid, E., Stehfest, E., Yang, H., and Jones, J. W.: Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison, Proceedings of the National Academy of Sciences, 111, 3268–3273, 2014.

25 Salvadori, G. and De Michele, C.: Frequency analysis via copulas: Theoretical aspects and applications to hydrological events, Water Resources Research, 40, 2004.

Salvadori, G., Michele, C. D., and Durante, F.: On the return period and design in a multivariate framework, Hydrology and Earth System Sciences, 15, 3293–3305, 2011.

Salvadori, G., Durante, F., and Michele, C.: Multivariate return period calculation via survival functions, Water Resources Research, 49, 30 2308–2311, 2013.

Salvadori, G., Durante, F., De Michele, C., Bernardi, M., and Petrella, L.: A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities, Water Resources Research, 52, 3701–3721, doi:10.1002/2015WR017225, 2016.

Schepsmeier, U., Stoeber, J., Brechmann, E. C., Graeler, B., Nagler, T., and Erhardt, T.: VineCopula: Statistical Inference of Vine Copulas, http://CRAN.R-project.org/package=VineCopula, R package version 1.6-1, 2015.

35 Schoelzel, C. and Friederichs, P.: Multivariate non-normally distributed random variables in climate research–introduction to the copula approach, Nonlin. Processes Geophys., 15, 761–772, 2008.

Serinaldi, F.: Can we tell more than we can know? The limits of bivariate drought analyses in the United States, Stochastic Environmental Research and Risk Assessment, pp. 1–14, 2015.

Shiau, J.-T. and Modarres, R.: Copula-based drought severity-duration-frequency analysis in Iran, Meteorological Applications, 16, 481–489, 2009.

Sippel, S., Zscheischler, J., and Reichstein, M.: Ecosystem impacts of climate extremes crucially depend on the timing, Proceedings of the National Academy of Sciences, 113, 5768–5770, 2016.

5 Sklar, M.: Fonctions de répartition à n dimensions et leurs marges, Université Paris 8, 1959.

Slingo, J. M., Challinor, A. J., Hoskins, B. J., and Wheeler, T. R.: Introduction: food crops in a changing climate, Philosophical Transactions of the Royal Society of London B: Biological Sciences, 360, 1983–1989, 2005.

Wheeler, T. and Von Braun, J.: Climate change impacts on global food security, Science, 341, 508–513, 2013.

Wong, G., Lambert, M., Leonard, M., and Metcalfe, A.: Drought analysis using trivariate copulas conditional on climatic states, Journal of
10 Hydrologic Engineering, 15, 129–141, 2009.

Wu, X., Babst, F., Ciais, P., Frank, D., Reichstein, M., Wattenbach, M., Zang, C., and Mahecha, M. D.: Climate-mediated spatiotemporal variability in the terrestrial productivity across Europe, Biogeosciences, 11, 3057–3068, 2014.

Zscheischler, J., Michalak, A. M., Schwalm, Christopher Mahecha, M. D., Huntzinger, D. N., Reichstein, M., Berthier, G., Ciais, P., Cook, R. B., El-Masri, B., Huang, M., Ito, A., Jain, A., King, A., Lei, H., Lu, C., Mao, J., Peng, S., Poulter, B., Ricciuto, D., Shi, Xiaoying Tao,
15 B., Tian, H., Viovy, N., Wang, W., Wei, Y., Yang, J., and Zeng, N.: Impact of Large-Scale Climate Extremes on Biospheric Carbon Fluxes: An Intercomparison Based on MsTMIP Data, Global Biogeochemical Cycles, 28, 585–600, 2014a.

Zscheischler, J., Reichstein, M., von Buttlar, J., Mu, M., Randerson, J. T., and Mahecha, M. D.: Carbon cycle extremes during the 21st century in CMIP5 models: Future evolution and attribution to climatic drivers, Geophysical Research Letters, 41, 8853–8861, 2014b.