The authors' response to the reviewers' comments on "Regional co-variability of spatial and temporal soil moisture - precipitation coupling in North Africa: an observational perspective" *by Irina Y. Petrova et al.*

The authors thank the two reviewers for their time and valuable comments. We highly appreciate the consent of Benoit Guillod to review our paper again, and we value the editor's work and time. The comments of every reviewer are addressed in sequence. The latest implemented corrections/ changes are highlighted in the attached pdf file using *latexhdiff* tracking tool.

The author's response to the Referee #1: Benoit Guillod

The revised manuscript has improved most aspects that were previously raised by the referees. In my view the paper will be suitable for publication after a few minor corrections listed below. It is a good contribution to the topic of soil moisture - precipitation feedback, whereby the more regional focus can also benefit to global studies.

Minor/technical comments:
- Figure 6(c): It is not clear what the "distribution of soil moisture gradient extremes" are. I presume that what is shown (colorbar) is their mean value? Please specify in the caption.
- Page 17, line 4: "The identified in the study regions of the strong SMPC...". This does not make sense and needs rephrasing.
- A few typos and grammatical errors remain and I recommend the authors to carefully read through a last time. Alternatively these typos might be identified by the Copernicus typesetting/proofreading.

We thank Benoit for the careful assessment of the previously implemented corrections, and for his positive evaluation of our work. All the new comments/suggestions were taken into account and implemented into the manuscript text.

The author's response to the Referee #2:

We thank the anonymous reviewer for his/her appreciation of our work, careful evaluation of the manuscript and valuable suggestions. Please, find our response to every item further below.

General comments
This study uses both spatial gradients and temporal anomalies to investigate the relationship between soil moisture and precipitation over the Sahel. It is very well-written and well-presented, and I appreciate its focus on a single region and the thorough meteorological and geographical understanding brought to bear. My primary concern is about the statistical rigor of the significance testing: "Significance is represented by a percentile of the [difference in mean metrics] in a bootstrapped sample." Presumably this is done for each grid point individually, but these grid points are not statistically independent and one is almost guaranteed to have Type I errors (erroneous rejections and overstated results). This statistical dependence between spatial and temporal metrics is discussed on page 13, but not acknowledged within the individual metrics. The more rigorous manner of constructing Figures 4 , 5, or 9 would be to use field significance as in Wilks 2006 JAMC "On Field Significance" and the False Discovery Rate. This would better validate the stated

“strong preference for convective rainfall over spatially drier soils” or that “temporally negative SMPC dominates”. However, it would require substantial reworking of the manuscript in its current form, so this matter is left to the discretion of the authors and editor. Otherwise I have clarifying questions and suggestions:

We thank the reviewer for being critical. However, the issue raised above was not fully clear to us. The reviewer mentions that the estimation of percentile and significance is done for each grid point. In the paper, however, these values are estimated always for an aggregate of rain events in either 5x5 , 2.5x2.5 or 1x1 degree box. Also, the minimum allowed number of events in every box is considered. We hope that this clarification from our side would resolve referees’ concern. Otherwise, we would please ask to send us more detailed formulation of the problem.

Specific comments

We thank the reviewer for pointing out all of the following issues, and we reply to every of them in sequence:

I understand that the statistical framework being used is described in Taylor et al. 2012, but more clarification would still be helpful in reading through the methods:
Page 7, Line 3 – Why is the accumulated precipitation threshold prior to the rain event lowered to 0? This seems overly stringent to me.

In general, the application of zero threshold is more logical, since we want to make sure that no rain occurs in preceding hours. Original threshold of 1 mm among other things considers the typical error range of precipitation detection in current products. In this way this more relaxed threshold may maximize sample size of identified events if a precipitation product tends to overestimate the drizzle, for example. In case of TMPA data set, setting the threshold to zero does not affect the results.

Page 7, Lines 12-13 – The climatological mean is subtracted from S’ prior to Lmax-<Lmin> gradient calculation. The entire current year is excluded so that the rain event does not impact this climatological mean, but are other years also excluded if a rain event occurred then?

We are not sure, we understand the question correctly. In general, the calculation of soil moisture anomaly is done for every single event separate. Hence, the calculation of the climatology of every event excludes the data of the corresponding year in which the event occurred.

Page 7, Lines 15-17 – It is confusing to use Ye as both the spatial gradient and temporal anomaly. The S’e (Lmax) is calculated in this temporal anomaly exactly as within the spatial gradient, right? A sentence stating this explicitly might also be helpful.

We thank the reviewer for pointing this out. We would be in favour to preserve the Ye notation, since we use it in the following paragraph to explain significance calculation as well as in the schematic Figure 2. Yet, we reformulated some sentences to improve the clarity.

Page 7, Lines 19-20 – Again I am curious whether the calculation of Yc accounts for rain events in years other than the one under consideration. The control sample Yc generally contains climatological days where the rain did not occur.

Page 8, Lines 22 to 24 – Although they do not come from this study, I would have appreciated more explanation of the anomalous positive spatial correlations in Figure 5. The mention of “lower

consistency of PERSIANN precipitation and soil moisture variability in time" is not entirely clear.

The results summarized in Figure 5 were obtained based on the visual analyses of the findings of G15. Hence, a detailed analyzed of all locations with positive SMPC were not feasible. Yet, we explored potential reasons for the positive coupling identified in our set up in Fig 4e, which we discussed in L27-30, p9.

Page 9, Lines 31 to 32 – From Figure 4, I think these values for positive δe are listed backwards. It seems that <3% are positive in the 5° grid, while 6% are positive in the 2.5° grid.

Thank you for the careful assessment. The typo was corrected.

Figure 6a – I do not think this Figure is necessary, as it simply illustrates a basic statistical concept.

This figure was added following previous round of the reviews to make the definition of extremes as clear as possible.

Figure 6b – Could you please add a colorbar or mention what the colors mean within the caption?
The colorbar was added following your suggestion. Thank you.

Page 10, Lines 14-15 – This sentence is a bit vague. Which "initial hypothesis"?
The confusing wording was removed. Thank you for pointing at it.

Page 10, Lines 22-33 – It is not clear to me why this linkage of extreme soil moisture gradients to flooded areas is done only for Lmin locations. Would one not be even more interested to pinpoint the locations of intense rainfall (i.e. Lmax) where the drying rate and high soil moisture have low correlation?

The remark has a good point. In general the test for Lmin locations is done in accordance with the "breeze-like" circulation theory. The latter states that local circulation formed between the wetland and its surrounding will cause triggering or re-intensification of rain systems next to the wetland, and its suppression over the wetland (see e.g. Taylor2010). Since the afternoon rainfall maximum is located over Lmax, the Lmin location theoretically should represent the wetland, and therefore location with small soil moisture variability. On the contrary, in Lmax location soil has a chance to dry out, creating a strong (extreme) negative gradient to a flooded region.

Page 10, Line 33 – In my opinion, Figure 11 could be referenced in the wetland breeze discussions because it is quite useful to understand what is being proposed here, even if comes later in the manuscript.
Thank you for the suggestion. We decided to keep it as it is, as more clarifications would be needed in the text, if the reference would be added.

Figures 7c and 7d – The Hovmöller diagram is of rain rate in [mm h-1] (not accumulated precipitation in [mm]), right? Why are there two diurnal profiles one after another?

Indeed, the diurnal cycle is estimated at the original TMPA resolution, i.e. 3-hourly rain rate (mm/hr). The profiles are identical and are duplicated for clarity of their periodicity.

Page 11, Lines 26-27 – I am surprised by the finding that MCS are "short-lived and smaller" in the Western domain. This is not in agreement with some past literature. For example, Zipser et al. 2006

BAMS find some of the most intense thunderstorms globally in West Africa, and studies like Jackson et al. 2009 MWR focus exclusively on the high frequency of West African MCS. Is there a reason for the discrepancy?
Thank you for pointing it out. In the paragraph, we say that the MCS in West Africa are not short-lived, but they are shorter-lived. This is meant in comparison with those typically occurring in the East. In general , this finding is supported and illustrated in the studies of Mathon&Laurant2001 and Laing&Carbone2008;2009. These studies are also references in our manuscript.

Page 12, Lines 9-14 – This result suggest to me that the "choice of a later accumulation time than in T12" stated in Section 3.1.1 has a large impact on, for example, what is shown in Figures 4 and 7. It would be good to mention the potential impact of adjusting accumulation time in the discussion of these two figures.

Indeed, consideration of the complete afternoon time period includes all effects in Figure4. The result of the influence on the SMPC measure is shown and discussed in the Figure 8. That would be equivalent to your suggestion we believe.

Figure 9 – It might be nice to do one sensitivity test of the robustness of the temporal anomaly to the definition of "pre-event" (i.e. for what period of time preceding the rain event does this mostly negative anomaly field persist?).

In their 2015 paper, Guillod et al state that they tested the sensitivity of the temporal SMPC to the soil moisture conditions of the previous day, and it lead to similar results (see in their supplementary material). We did not look into this aspect further.

Page 15, Lines 11-12 – This explanation is a nice mechanistic way of understanding the results, but should adjust with a different soil drying rate (which varies and sometimes depends on initial moisture from Figure A1). This adjustment could be described briefly.
The final point leads to a more general question I had: it seems to me that the current temporal anomaly definition from a climatological mean may not well account for the fact that certain locations simply have much greater soil moisture variability than others. To me, a kind of "soil moisture z-score" normalized by a variability would make more sense. Could the authors comment on this?

The effect of soil moisture variability on the SMPC significance is accounted by the comparison of the event sample to its climatological sample. Yet, this of course does not allow judging on the absolute values of soil moisture and its gradient in general, and the resulting bowen ratio in particular. This also remains being main limitation of the present methodology. The normalization would allow comparison of the e.g. SM gradient strength between different locations, but will not resolve the issue of a missing bowen ratio.

Boreal spring and autumn are the African rainy season, and I suppose that isolating these months from the others should change the fields. Have the authors investigated seasonality at all? Could some discussion be added, even if new Figures are not?

In the considered semi-arid region, rain season is limited to monsoon JJAS months. The effect of varying monsoon activity stage was briefly considered in work of Petrova2017.
Seasonality per se was tested by Guillod2015 study.

Regional co-variability of spatial and temporal soil moisture - precipitation coupling in North Africa: an observational perspective.

Irina Y. Petrova[1*], Chiel C. van Heerwaarden[2], Cathy Hohenegger[1], and Françoise Guichard[3]

[1]Max Planck Institute for Meteorology, Hamburg, Germany
[*]Current affiliation: Laboratory of Hydrology and Water Management, Ghent University, Ghent, Belgium
[2]Meteorology and Air Quality Group, Wageningen University, Wageningen, The Netherlands
[3]CRNM-GAME, CNRS-Météo France, Toulouse, France

Correspondence to: Irina Y. Petrova (irina.petrova@ugent.be)

Abstract. The magnitude and sign of soil ~~moisture - precipitation~~ moisture–precipitation coupling (SMPC) is investigated using a probability-based approach and 10 years of daily microwave satellite data across North Africa at 1° horizontal scale. Specifically, the co-existence and co-variability of spatial (i.e. using soil moisture gradients) and temporal (i.e. using soil moisture anomaly) soil moisture effects on afternoon rainfall is explored. The analysis shows that in the semi-arid environment of the Sahel, the negative spatial and the negative temporal coupling relationships do not only co-exist, but are also dependent of one another. Hence, if afternoon rain falls over temporally drier soils, it is likely to be surrounded by a wetter environment. Two regions are identified as SMPC "~~hot spots~~hot-spots". These are the south-western part of the domain (~~7 - 15~~7–15° N, 10° ~~W - 7~~W–7° E) with the most robust negative SMPC signal, and the South Sudanian region (~~5 - 13~~5–13° N, ~~24 - 34~~24–34° E). The sign and significance of the coupling in the latter region is found to be largely modulated by the presence of wetlands and is susceptible to the amount of long-lived propagating convective systems. The presence of wetlands and an irrigated land area is found to account for about 30 % of strong and significant spatial SMPC in North African domain. This study provides the first insight into regional variability of SMPC in North Africa, and supports potential relevance of mechanisms associated with enhanced sensible heat flux and ~~meso-scale~~ mesoscale variability in surface soil moisture for deep convection development.

1 Introduction

Soil moisture can affect the state of the lower atmosphere through its impact on evapotranspiration and surface energy flux partitioning (e.g., **??**). Especially in the "~~hot spots~~hot-spots" of soil ~~moisture - precipitation~~ moisture–precipitation coupling (SMPC), like the semi-arid Sahel (**???**), soil moisture exerts strong control on evapotranspiration (e.g., **???**), influencing the development of the daytime planetary boundary layer (BL), and hence convective initiation and precipitation variability. Most of the physical understanding on how soil moisture could alter BL properties and affect development of convection comes from 1-D to 3-D model analyses (**??**). Observational evidence of the SMPC largely relies on the measurements of recent field campaigns (like African HAPEX and AMMA: **??**) and hence, is often limited to a short spatio-temporal scale. Such observational analyses present a unique evidence of environmental conditions preceding convection development (e.g., **?**) and can be further used as a test-ground to evaluate and improve the physical parametrizations of models (e.g., **?**). Both,

observational and modelling studies agree reasonably well on the effect of soil moisture availability and heterogeneity on the lower atmospheric stability (e.g., **?**) and convective initiation at ~~meso-scales~~ mesoscales (e.g., **??**). However, the impact of soil moisture on convective precipitation remains more uncertain. At ~~meso-scale~~mesoscale, there is a disagreement in the sign of the SMPC between models which use parametrizations of deep convection and observations (**???**). Recent satellite-based analysis has demonstrated that the choice of soil moisture parameter itself (temporal anomaly or spatial gradient) and related differences in physical mechanisms have a direct effect on the resulting sign of the coupling. **?** found a positive temporal (i.e. using soil moisture anomaly) but negative spatial (i.e. using spatial soil moisture gradient) SMPC over most of the globe at 5° horizontal scale. Our study addresses the question of co-existence of spatial and temporal SMPC at a finer $1^\circ \times 1^\circ$ horizontal grid. Specifically, we use 10-year satellite records of daily soil moisture from the AMSR-E and 3-hourly TMPA precipitation product to investigate spatio-temporal co-variability of observed coupling relationships in the region of North Africa.

Both modelling and observational studies reported the possibility of negative as well as positive SMPC (**?**). Spatial gradients in soil moisture can affect BL state and convection initiation through thermally-induced ~~meso-scale~~ mesoscale circulations on 10 to 100 km scales (**???**). In association with this mechanism and under favourable thermodynamic conditions, convection is likely to be initiated over spatially drier soils, indicating a negative SMPC (**??**). However, whether the further development and propagation of moist convection will occur over drier or wetter soils remains less clear. The modelling study of **?** suggested that negative SMPC is possible under very weak surface wind conditions, and is associated to stationarity of convective systems once initiated. The opposite sign is expected under a stronger horizontal advection, which will support propagation of the developing moist convection downwind, i.e. from dry to wet soils, and its further amplification over wetter areas. Another important factor is related to the life-cycle of ~~meso-convective~~ mesoscale convective systems (MCS) and thus their organization in space and time (**?**). Small-scale convective systems are expected to be particularly sensitive to surface moisture variability and will propagate preferentially towards spatially drier soil, bounded by a wetter surrounding (**?**). Alternatively, larger organized systems have been found to evolve towards wetter soils - areas of increased latent heat flux, convective available potential energy (CAPE) and moist static energy (MSE) (**??**). Hence, these systems are expected to be more sensitive to soil moisture availability.

The impact of temporal anomalies of soil moisture on the atmospheric BL and moist convection is largely governed by thermodynamical processes, and may likewise result in a coupling of both signs. Wet soils are expected to lead to an increase of boundary layer MSE or similarly equivalent potential temperature, through a decreased boundary layer height (BLH) and subsequent less vigorous entrainment (e.g. **??**). The enhanced MSE over wet soils is favourable to convective rainfall formation. Dry soils, on the contrary, are associated with a reduced MSE and thus provide lower potential for convection development and may even suppress existing MCS (**?**) or deviate its propagation direction (**?**). However, modelling and observational evidence indicate that both dry and wet soils, can favour moist convection, depending on the morning stratification of the lower atmosphere (**??**) into which the daytime convective BL is growing (**???**).

The relevance of ~~meso-scale~~ mesoscale spatial heterogeneity of soil moisture in favouring moist convection over wet and dry temporal soil moisture anomalies was demonstrated by e.g. **?** and **?** respectively. However, until recently no attempts were made to directly compare the temporal and spatial aspects. The first comparison of the spatial and temporal effects of soil moisture on

precipitation was presented by the study of **?** (hereafter, G15) at 5° horizontal scale. Applying the probability-based approach of **?** (hereafter, T12) to 10-years of global satellite-based soil moisture and precipitation data, they demonstrated that a negative spatial (rain over spatially drier soils) and a positive temporal (rain over temporally wetter soils) SMPC dominate over most of the globe and do not exclude one another. G15 suggested that the two effects might be interconnected through spatial coupling mechanisms, in which adjacent precipitation would provide required moisture to enhance convection development over spatially but not temporally drier soil. Using multiple data sets, G15 showed that the signal is robust across different input data sets. However, in a few regions, including the Sahel in Africa, an opposite temporal relationship was revealed: spatially and temporally negative coupling was found to co-exist in opposition to the global relationship.

In this study, we further explore spatial and temporal SMPC as well as their co-existence in North African region using the T12 method at a finer $1° \times 1°$ horizontal grid. Furthermore, we provide insight into regional co-variability of the spatial and temporal effects on afternoon rainfall. The analysis is realized following two main steps:

1. Identification of the factors that influence the magnitude and variability of the spatial SMPC measure. By doing so we address the question: which physical processes likely underlie the observed spatial SMPC relationship if any?

2. Analysis of the link between the spatial and temporal effects of soil moisture on precipitation. This part addresses the two following questions: are spatial and temporal negative coupling relationships independent, and if not, how do they inter-relate?

We reproduce and apply the probability-based approach of T12 to 10-years of daily AMSR-E soil moisture and 3-hourly TMPA precipitation records. In contrast to the previous studies, we estimate the temporal and spatial coupling effects at a finer 1° horizontal scale, which reveals previously hidden smaller-scale effects of land cover features on the SMPC relationship.

The first part of the study is embraced in *section 4* and includes an analysis of the regional variability and robustness of the observed spatial SMPC at varying horizontal scales i.e. from the original 5° to 2.5° and 1°. Identification of the factors relevant for the observed spatial SMPC distribution includes a sensitivity analysis of the spatial coupling measure to the presence of soil moisture parameter extremes (*section 4.3*) and to the MCS life cycle (*section 4.4*). The link between the temporal and spatial SMPC is assessed using correlation analysis in *section 5.1*. Summary and discussion *section 6* is delivered in two parts. *Section 6.1* provides main conclusions on the potential physical mechanisms which likely underlie the observed coupling relationships. *Section 6.2* reviews the reasons behind the opposite sign of the temporal coupling identified in the North African region as compared to the dominantly positive relationship identified in G15 over the globe. The paper concludes with more general discussion of the SMPC "hot-spots" and conclusions in *section 7*.

2 Domain and Data

2.1 Study domain

We focus our analyses on the North African region (~~5 - 20~~5–20° N, 20° ~~W - 40~~W–40° E) (Fig. **??**, dashed rectangle) during ~~summertime~~ summer period (JJAS). This region has been repeatedly pointed out as a "~~hot spot~~hot-spot" of ~~land-atmosphere~~

land–atmosphere interactions (**???**), and one of the most vulnerable regions with respect to climate change (**??**). A major feature affecting the Sahelian climate is the West African Monsoon (**?**), which is associated with a high precipitation variability (**?**). During the monsoon, soil moisture fluctuations are strongly influenced by precipitation at a large range of spatial and temporal scales. Atmospheric and surface fields display strong meridional gradients between 10° N and 20° N (Figure **??**, zonal plot) shaped by the migration of the summer time rain belt, also referred as the Inter Tropical Convergence Zone (ITCZ). Wind convergence at the surface is observed further north, around ~~18 - 20~~18–20° N, along the Inter Tropical Discontinuity (ITD), where the cool and moist monsoon flow meets hot and dry Saharan air. Associated with the meridional heat gradient, the monsoon circulation and related large-scale structures like the African Easterly Jet (AEJ), as well as synoptic disturbances like the African Easterly Waves (AEWs) largely modulate convection activity over the region (**??**). Additionally, evidence supporting a significant role of the surface state in the triggering of deep precipitating convection is steadily growing (**?**).

Middle July to August conditions may be less favourable for a strong surface influence on convection. Compared to the drier early and late monsoon months of June and September, the wetter period - from July through August - is characterized by a typically lower level of free convection (LFC) (**??**) and less pronounced spatial contrast between fluxes due to more dense vegetation (**??**). In our study, the role of the monsoon dynamics is not ~~directly~~ addressed to preserve maximum sample size for the sake of statistical significance.

We intentionally extend our analysis further eastwards. Despite the inherent zonal symmetry of surface and atmospheric parameters (as in precipitation in Fig. **??**), considerable differences exist in rainfall and large-scale circulation regimes between East and West. Distinct orography, intensity of surface and upper level jets and wave disturbances are likely to bring dissimilarities in the sensitivity of convection to the surface state between the two regions. The eastern African domain can also remotely influence convection in the western part of the region via the genesis of westward propagating AEWs (**?**) and long-lived MCSs (**?**). Yet, notably few studies have investigated ~~land-atmosphere~~ land–atmosphere interactions in eastern Sahel.

2.2 AMSR-E soil moisture

Soil moisture (SM) data from the Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E, Jun ~~2002 - Oct~~ 2002–Oct 2011) is analyzed in this study. The AMSR-E unit is carried on board of the polar orbiting AQUA satellite, measuring brightness temperatures in 12 channels, at 6 different frequencies (6.9~~ - 89~~ –89 GHz) (**?**). Soil moisture derived from the lowest C-band frequency of 6.9 GHz is used here, as lower frequencies experience less signal contamination from vegetation and surface roughness, and are able to receive emission information from deeper soil layers (still just a few centimetres, **?**). The AQUA orbit is sun-synchronous with typically one overpass per pixel per day at either 13:30 or 01:30 local solar time (LST). In order to capture the surface moisture state shortly before afternoon convection activity, only data of ascending day orbit, i.e. 13:30 LST is used here. It is important to note, that the day overpass is prone to higher biases compared to the night overpass, because of the greater temperature differences between surface and canopy involved in the physics algorithm (**?**).

We utilize the Level 3 estimates of AMSR-E soil moisture derived with the Land Surface Parameter Model (LPRM, **?**) for JJAS ~~2002 - 2011.~~ 2002–2011. The product is available at 0.25° × 0.25° spatial resolution. The LPRM is not valid for dense vegetation and water bodies. Therefore pixels with an optical depth > 0.8 are excluded. Water body and soil moisture quality

masks were adopted from the materials of T12. Accordingly, pixels containing more than 5 % water are excluded, using water body classification of the 1 km Global Land Cover 2000 data set [Available online at http://forobs.jrc.ec.europa.eu/products/glc2000/products.php]. Application of the soil moisture quality mask, based on the correlation analysis between precipitation and soil moisture data sets, is intended to reduce the number of pixels covered with wetlands (for details see suppl. in T12).

Many days (~~40 - 50~~ 40–50 %) do not contain soil moisture information due to satellite revisit times. Over a given longitude per day the number of overpasses below 40° N do not exceed one with occasionally daily or every third day sampling (see Fig. 1 in **?**). The latter significantly reduces the size of the collected rainfall event sample available for the analyses.

The AMSR-E instrument is chosen because it documents a relatively long period and performs better than ASCAT (**?**) over sparsely vegetated and deserted areas. The AMSR-E product also proved to be accurate at the precipitation event scale in capturing rain-related soil moisture variability and timing, when compared with in situ data in the Sahel (**?**).

2.3 TMPA-v7 precipitation

The Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) represents a partial-global coverage (50° ~~S - 50~~S–50° N) product of combined precipitation estimates. Three-hourly precipitation time-series of the TMPA product ~~(1998 - 2015, **?**)~~ (1998–2015, **?**) at 0.25° × 0.25° horizontal resolution are used to estimate locations of afternoon convective precipitation events in the study.

The TMPA algorithm (**?**) involves the following steps: (1) - merging multiple independent passive microwave sensors, (2) - their inter-calibration to the TRMM Combined Instrument (TCI) precipitation estimates, (3) - further blending with preliminary calibrated infra-red products from geostationary satellites, and finally (4) scaling of the estimates to match monthly accumulated Global Precipitation Climatology Center (GPCC) rain gauge data.

In this study we utilize the product version 7 (TRMM-3B42), which includes several modifications to the algorithm and additional satellite data (**?**). Consistent with the soil moisture record, only 10 years (~~2002 - 2011~~2002–2011) of JJAS precipitation data is used. To ensure similar solar forcing on the surface and boundary layer, the 3-h precipitation time-series for the present application are adjusted to LST (based on longitude) by taking the closest 3-h UTC time step. It is important to note that any 3-h TMPA value is not referred directly to its nominal hour, but represents the average of the "best" overpass data within a 3 hourly window, centered around the nominal hour, i.e. +/- 90 min range. Variable time of the TMPA "best" data average is not expected to significantly affect our SMPC results.

3 Methods

3.1 Description of statistical framework after T12

The SMPC in this study is referred to as the relationship between the afternoon convective rainfall and the antecedent soil moisture conditions~~in the few preceding hours~~. Using the method of T12, we examine first whether afternoon rainfall is more likely over soils that are untypically (relative to ~~a control sample~~its climatology) drier or wetter than their surrounding. Next, following the definition of G15, we assess whether afternoon rainfall is more likely on days when soils are untypically drier or

wetter than their temporal mean. Subsequently, the higher than expected probability of convective rainfall events to occur over spatially drier or wetter soils is referred to as *spatial SMPC*, while the higher than expected likelihood of convective rain events to occur over temporally wetter or drier soils quantifies *temporal SMPC*. The following paragraph describes criteria which are used to define a convective precipitation event, and evaluate soil moisture statistics antecedent to every event. The framework algorithm implemented in this study largely follows the method of T12 and is summarized in Fig. **??**.

3.1.1 Definition of convective rainfall event (Fig. ??C)

We define a convective event location, L_{max}, as the location where accumulated afternoon precipitation between 15:~~00 - 21~~00–21:00 LST exceeds a threshold of 6 mm. Then, locations of afternoon accumulated precipitation minima, L_{min}, are identified within a 5 × 5 pixel box (1.25° × 1.25°)[1] centered at L_{max} (Fig. **??**C). The choice of a later accumulation time than in T12 (i.e. 15:~~00 - 21~~00–21:00 LST instead of 12:~~00 - 21~~00–21:00 LST) ensures that the soil moisture measurement at 13:30 LST precedes precipitation without introducing additional filters. The twice larger afternoon accumulated rainfall threshold than in T12 yields qualitatively similar results, though leads to a slightly higher mean SMPC significance over the domain. According to additional sensitivity tests, the choice of higher threshold values in the method mostly influences the amount of significant grid boxes linked to a reduction in the event sample size, yet does not qualitatively affect the dominant preference of the afternoon rainfall over specific soil moisture conditions (**?**).

The following set of assumptions is used to improve the accuracy of the convective event sample. If one of the conditions is not fulfilled, an event is excluded from further calculation:

(1) - accumulated precipitation in the preceding hours (06:~~00 - 15~~00–15:00 LST) in the entire 1.25° × 1.25° box must be zero;

(2) - elevation height difference within the event box must not exceed 300 m. The latter is done to minimize the effect of orographic uplifting on the rainfall variability. The resulting distribution of the orography mask is shown in Fig. **??**a.

(3) - number of identified L_{min} locations within one box must be 3 or more (for averaging reasons). In that case, all L_{min} locations will have the same afternoon accumulated precipitation value, which will most often be zero.

(4) - if boxes overlap, the event with larger afternoon accumulated precipitation value is retained.

3.1.2 Soil moisture statistics in event locations (Fig. ??D - E)

Once events are identified, soil moisture anomaly S' measured prior to the precipitation event (at 13:30 LST) at L_{max}, $\overline{L_{min}}$ or any combination of the two ~~can be~~ is stored and analyzed. $S'^{\overline{L_{min}}}$ represents an average value of S' measured in every identified L_{min} location within a 1.25° × 1.25° event box. S' has its climatological mean subtracted, calculated as ~~a departure from the ± 10 day mean over 10 years~~linear averaging within a 21-day moving window centered on the particular day and across the entire multi-year record. To exclude contribution of a rain event onto the anomaly values, the year of the event is excluded from the climatological mean calculation. ~~In order to investigate whether it rains over spatially wetter or drier soils,~~

[1]Following T12, a box size of 1.25° × 1.25° is selected as minimum possible size to resolve soil moisture variability around the center of the box, taking into account the 50 km footprint of the AMSR-E soil moisture.

To estimate, if soils were anomalously dry or wet in the location of maximum rain (L_{max}) compared to the expected range (climatology) for that location, we store pre-event soil moisture anomaly at L_{max}, i.e. Y_e: ${S'_e}^{L_{max}}$ (Fig. **??**D). To investigate if soils in the location of maximum rain (L_{max}) were drier than in the neighbour region(s) where it rained less or did not rain (L_{min}), we calculate the pre-event soil moisture gradient between L_{max} and $\overline{L_{min}}$ scaled per 100 m, i.e. Y_e = : $\Delta({S'_e}^{L_{max}})$ = ${S'_e}^{L_{max}}$ - ${S'_e}^{\overline{L_{min}}}$ with the dimension of m^3 m^{-3} 100 m^{-1}, where Δ - stands for gradient (Fig. **??**D). To estimate, whether it rains over temporally wetter or drier soils we store pre-event soil moisture anomaly at L_{max} location, i.e. Y_e = ${S'_e}^{L_{max}}$.

For every two each Y_e parameters parameter we define the control sample Y_c, represented by an array of corresponding Y values measured in the same L_{max} and L_{min} event locations in the same calender month, but on the non-event years. The measure of coupling is then quantified by the magnitude of a difference between mean statistics of the event and control samples, $\delta_e = \text{mean}(Y_e) - \text{mean}(Y_c)$, and the measure of δ_e significance (Fig. **??**E). Significance is represented by a percentile, P_e, of the observed δ_e in a bootstrapped sample of δ values that is observed by chance. For that Y_e and Y_c are pooled together and re-sampled without replacement 5000 times.

3.1.3 Definition of temporal and spatial SMPC (Fig. ??F)

Parameters of δ_e and P_e calculated for the soil moisture gradients $\Delta({S'_e}^{L_{max}})$ prior to the event quantify preference of rain to occur over soils drier ($\delta_e < 0$, $P_e \leq 10$ %) or wetter ($\delta_e > 0$, $P_e \geq 90$ %) than its 1.25° × 1.25° environment, and are referred to as negative or positive *spatial SMPC* respectively. The same parameters estimated for the temporal soil moisture anomaly ${S'_e}^{L_{max}}$ instead specify expressed preference of rain to occur over soils drier or wetter than its temporal mean, i.e. negative or positive *temporal SMPC* accordingly.

In this study, estimation of δ_e and its significance P_e for the spatial and temporal coupling is realized over 5° × 5°, 2.5° × 2.5° and 1° × 1° boxes. Aggregation of event statistics at a higher resolution than used in the global studies of T12 and G15 results in a smaller event sample size per grid box, yet allows a reduction of the potential influence of meridional or zonal gradient in the parameter statistics, i.e. makes the spread in underlying surface and atmospheric moisture conditions across the box latitudes smaller (Section **??**). The latter is valuable for the interpretation of obtained statistics in terms of land cover and atmospheric state. Hence, most of the study focuses on the smallest 1° × 1° spatial grid.

3.2 Statistics of convective events

Application of the algorithm to the 10 years of JJAS AMSR-E soil moisture and TMPA precipitation time-series yields 10131 afternoon rainfall events. The distribution of identified events over the domain at 5° × 5° and 1° × 1° grid is shown in Figure **??**b and **??**c respectively. The signature of orography and large-scale dynamic effects on event occurrence becomes more evident at the finer event-aggregation scale, thus giving an advantage to the highest horizontal resolution. Figure **??**c shows that most events occur between 10° N and 18° N, and the occurrence maxima are zonally aligned. Two maxima are found over the central Sahel, covering the area between 10° W - 15W–15° E - aligned with the mean position of the AEJ core (Figure **??**b). Another two maxima are evident at about 22° E and 30° E. Overall, the obtained distribution of identified rain events at 1° × 1° grid resolution is consistent with the observed distribution of intense MCS over the region (**?**).

4 Results of spatial SMPC analysis

4.1 SMPC at 5° horizontal scale. Consistency to previous studies

We start our assessment by investigating the spatial soil ~~moisture - precipitation~~ moisture–precipitation coupling relationship. In agreement with the global-scale studies of T12 and G15, we find a dominantly negative spatial SMPC in the North African domain at the 5° × 5° grid, i.e. a strong preference for convective rainfall events to occur over spatially drier soils (Fig. **??**a). The majority of the 5° × 5° boxes (72 %) have percentile values P_e lower than 10 %, implying a significant negative difference in the mean magnitude of soil moisture gradients $\Delta(S_e'^{L_{max}})$ prior to the events relative to their typical (non-event) state. No significant positive difference between event and non-event conditions is found at the 5° scale (Table **??**).

Figure **??** further compares the percentage of the domain area with significant negative and positive coupling identified in our study, T12 and G15 (see also Table A1). The differences arise due to disparities in the data sets and methodologies. The weakest negative and the strongest positive coupling signal corresponds to the estimates based on PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks, **?**) data set from G15. This is possibly linked to the lower consistency between the PERSIANN precipitation and soil moisture variability in time (T12). On average, results based on different data set combinations summarized in Fig. **??** agree that afternoon precipitation occurs more often than expected over spatially drier soils in 42 % of the 5° × 5° boxes, against only 4 % with a preference over spatially wetter soils (red and blue lines in Fig. **??** respectively).

The variability of spatial SMPC patterns among different data set combinations has shown to be quite strong over the globe and was not analyzed further in G15. We find, however, that in the North African domain, areas of significant negative spatial coupling are fairly consistent. One of the most robust negative spatial SMPC signals is found in the ~~south-western~~ south–western part of the domain (Fig. **??**a,b). Fourteen out of 18 data set combinations summarized in Fig. **??**, including this study, locate the cluster of the lowest percentiles roughly between ~~5 - 15~~5–15° N and 10° ~~W - 10~~W–10° E (Fig. **??**, rhombs). This region occupies a relatively vast and flat area, associated with a reduced orographic forcing on convection development compared to the East, and a regional minimum in cold cloud occurrence (**?**). The effect of large-scale structures like AEWs and AEJ on convection, on the contrary, is expected to be stronger in the western Sahel than further East. However, this does not exclude but may even favour higher sensitivity of convection triggering to soil moisture heterogeneities (**??**). Overall, the identified negative spatial SMPC relationship in the region is consistent with the recent observational- (**???**) and model-based (e.g. **????**) studies in the Western Sahel.

Another cluster of the lowest percentiles and the largest differences in soil moisture state between event and non-event days δ_e is identified in the ~~south-east~~ south–east of the domain (Fig. **??**a,b). The proximity to the Ethiopian Highlands and the presence of extensive seasonally flooded regions in this area makes it generally difficult to isolate effect of surface state on convection. This also possibly led to less agreement in the spatial SMPC estimates identified in our study, G15, and T12 analyses (not shown). Unlike in the western Sahel, no accurate estimates of the SMPC exist in this eastern region.

4.2 Robustness of the negative SMPC at finer 2.5° and 1° scale

To identify factors and potential physical mechanisms that influence the magnitude and variability of the SMPC we reduce the event-aggregation scale to the finer 2.5° × 2.5° and 1° × 1° horizontal grid[2]. In particular, aggregation of the convective rainfall events and corresponding soil moisture statistics over the smallest 1° × 1° grid boxes reveal more details on the effects of land surface conditions on the SMPC.

The percentile maps obtained for the finer scales are presented in Figures **??**c and **??**e. Despite the reduction in the amount of significant δ_e values, largely due to the decreased number of events in every box, negative spatial SMPC relationships remain dominant at the finer scales, and exhibit a similar spatial pattern as at the 5° × 5° grid. The featured regions of significant negative coupling now scale down to the territories of Burkina Faso, Benin, parts of Ivory Coast, Ghana and Mali (~~7 - 15~~7–15° N, 10° ~~W - 7~~W–7° E) in the West, as well as South Sudan (~~5 - 13~~5–13° N, ~~24 - 34~~24–34° E) in the East (Fig. **??**e). In total, 42 % (21 %) of the boxes reveal significant negative difference δ_e for the 2.5° (1°) scale, versus initial 72 % at the 5° × 5° grid (Table **??**).

The overall distribution of the difference δ_e does not change at the finer scales (Fig. **??**d,f). However, multiple pixels with a positive δ_e emerge. For example, a small region enclosed between the Cameroon mountains and Jos Plateu (7° N; 8° E, Fig. **??**f) now indicates a higher likelihood of rainfall to occur over spatially wetter soils. The relationship, though non-significant, is plausible. This area includes part of the Niger river valley and represents a prominent location of intense convection and a local maximum of the cold cloud occurrence, linked to the initiation of convection at the lee side of the high terrain (**?**). The potential link between the land surface characteristics and SMPC parameter is explored in more detail in the following section. In total, 14 % of the 1° × 1° boxes reveal a positive δ_e shift, against less than ~~3 % and 6~~ 6 % and 3 % for coarser 2.5° × 2.5° and 5° × 5° grids respectively.

4.3 Evidence for "~~wetland-breeze~~wetland–breeze" mechanism in the SMPC statistics

In areas like the Cameroon Mountains where orography or floodplains have an effect on deep convection development, persistent wet and dry surface moisture patterns may pre-exist or develop, and therefore, lead to the occurrence of stronger than usual spatial soil moisture gradients. In the SMPC statistics, such gradients occur as extremes in a given distribution of soil moisture gradients $\Delta({S'_e}^{L_{max}})$ within a 1° × 1° box. Here, we define a gradient $\Delta({S'_e}^{L_{max}})$ to be an extreme if it lies outside the following range: (Q_{25}-1.5×IQR) to (Q_{75}+1.5×IQR), where Q_{75} and Q_{25} are the third and first quartiles respectively, and the interquartile range (IQR) is the difference between them (Fig. **??**a).

The distribution and magnitude of the extreme soil moisture gradients identified in the domain are shown in Figure **??**c. We find that 28 % of all valid 1° × 1° boxes contain extreme $\Delta({S'_e}^{L_{max}})$. A large part of the extreme soil moisture gradients are located in the regions of significant negative coupling in the West and East. In these areas, extreme $\Delta({S'_e}^{L_{max}})$ lead to an overestimation of the $\Delta({S'_e}^{L_{max}})$ sample mean and therefore, δ_e magnitude. As a result, extreme gradients appear to predefine

[2]The SMPC statistic for the finer scales is calculated if at least 8 events in a box are present.

SMPC significance. Removal of extremes leads to a decrease in the number of boxes with significant negative spatial coupling by 30 %. However, in most cases the sign of the coupling remains unchanged (not shown).

Further analysis shows that extremes tend to cluster around major rivers and wetland areas in the East and West (Fig. **??**b,c)~~, which supports our initial hypothesis.~~ . Regional distribution of extreme $\Delta(S_e'^{L_{max}})$ is consistent with the distribution of natural wetland fraction presented earlier by **?** and shown in Fig. **??**b, and more recent estimates of inundation regions identified by **?**. Strong positive soil moisture gradients are found around the Senegal river close to the coast, and on the lee side of the Cameroon Mountains. Strong negative soil moisture gradients are more numerous and seen all along the western flow of the Niger river, downwind of the permanent wetlands of Ez Zeraf Game Reserve and irrigated lands of the Gezira Scheme in Sudan. The scatter of the extremes in the East is likely related to the recurrent floods of the White Nile river. The reason for the absence of extremes around the Logone floodplains is obscure.

To verify the link between the location of extreme soil moisture gradients and flooded areas, we additionally identify $1^\circ \times 1^\circ$ boxes with the events, which L_{min} ~~locations are~~ pixels were likely covered with a wetland or a floodplain, and hence ~~are~~ were wetter than the neighbouring L_{max} location (Fig. **??**c~~and appendix Fig. A1~~, crosses). This is done by calculating a linear regression function between climatology of one-day soil moisture drying rate (i.e. difference in soil moisture on the event day 0 and on day -1) and climatology of initial soil moisture value (i.e. soil moisture on day -1)[3] ~~. If~~ (appendix Fig. A1). We consider the L_{min} location to be likely flooded, if (a) all initial soil moisture values are high, and (b) climatological values of drying rate are always small and do not vary much with the initial soil moisture (regression slope is close to zero)~~, while the initial soil moisture values are on the contrary high, then we consider this L_{min} location to be likely flooded. For detailed explanations~~ . For the details the reader is referred to the Appendix A and Figure A1. From the Figure **??**c it is seen that the distribution of grid boxes with the potentially flooded L_{min} locations conforms with the distribution of natural wetland fraction in Fig. **??**b and includes the majority of extreme $\Delta(S_e'^{L_{max}})$ locations. This result ~~further suggests~~ supports the presence of the link between the afternoon rainfall maximum and ~~potentially flooded areas, and supports the potential relevance of thermally-driven "wetland-breeze" circulations on convection triggering.~~ flooded areas.

The identified sensitivity of the afternoon rainfall to the strong negative soil moisture gradients around water bodies is in agreement with the results of the observational-based study of **?**. Analyzing 24 years of Meteosat brightness temperatures over the Niger Inland Delta, he found that convection was initiated more often over and to the east of the wetland in the morning hours. However, later in the day ~~meso-scale~~ mesoscale convective systems tended to develop and propagate away from the wet areas towards drier soils, suggesting formation of deep convection and afternoon precipitation over negative soil moisture gradients. Similarly, observed by **?**, enhancement of rain to the East of irrigated land at 14° N, 33° E and its suppression over the Gezira Scheme itself is consistent with the location of negative (positive) extreme soil moisture gradients to the West (East) of the irrigated region (Fig. **??**c). All the above supports the consistency of the observed afternoon rainfall intensification over the drier soils adjacent to flooded areas with "~~wetland-breeze~~wetland–breeze" mechanism. The ability of the method to capture these effects is also noteworthy.

[3]The climatologies are calculated for every L_{min} location, for the same month as the event but for the non-event years.

4.4 Effect of propagation of deep convective events on the SMPC statistics in eastern and western domains.

Another physical effect that may influence the SMPC relationship is related to the propagation and evolution of ~~meso-scale~~ mesoscale convective systems (not accounted for by the current algorithm). Previous studies indicate that an opposite SMPC relationship might be expected at early versus late stages of MCS development (**??????**). In this respect, a distinct strength or even sign of the spatial SMPC measure may result from separation of the rainfall events into those formed by a weaker and smaller MCSs - mostly found in the early afternoon - or by long-lived and organized MCSs - dominant during late afternoon hours (**?**). Differences in SMPC response to MCS life cycle are also expected to exist between the two regions of significant negative coupling, in the East and West. To characterize these differences, we analyze precipitation diurnal cycles averaged over event days in the East and West first (Fig. **??**), and then estimate sensitivity of the spatial SMPC to varying rainfall accumulation times (Fig. **??**).

The Hovmöller diagram of rainfall averaged over 1000 event days in the Western domain (black rectangle in Fig. **??**) shows that intensification of the moist convection in the region is generally concentrated around main orographical features (Fig. **??**a,c). The peak in precipitation occurs at similar times across the domain, and thereby does not reveal expressed signature of the system propagation. Most of the MCS are therefore expected to be shorter-lived and smaller, ~~suggesting that~~ as their dissipation locations would be found close to their initiation (**?**).

In the East, on the contrary, the strong ~~south-western~~ south–western propagation component of moist convection dominates the zonal progression of the most intense rainfall during diurnal cycle averaged over 754 event days (Fig. **??**d). A large number of MCS initiate at the lee side of the Ethiopian Highlands and propagate westward undergoing cycles of regeneration and growing into a mature and organized MCS (**??**). The emergence of an absolute rain rate maximum downwind of the permanent wetlands of the Ez Zeraf Game Reserve (at 30° E) during afternoon hours supports an influence of the flooded areas on moist convection intensification in the region (Fig. **??**d,f). Consistent with the results of **?** obtained for the Niger Inland Delta, the presence of wetlands in the Eastern domain is expected to increase the number of organized and long-lived propagating MCS in the late afternoon, originating from either locally triggered MCS, i.e. formed at the dry ~~land-wetland~~ land–wetland boundary, or from re-intensified pre-existing westward propagating systems. We may therefore expect a greater sample of long-lived and organized propagating MCS to be found in the late afternoon hours in the Eastern than in the Western domain. Accordingly, the response of the SMPC statistics to propagating MCS is expected to be stronger in the East compared to the West. Figure **??**, which shows the change of the SMPC parameter between different rainfall accumulation time periods confirms this hypothesis. For this assessment an additional area in the North is considered (dashed rectangle in Fig. **??**), as representative of a region where large-scale atmospheric and surface conditions differ from those of the East and West domains.

From Figure **??** it is seen that the earlier rainfall accumulation time periods, i.e. 12:~~00 - 18~~00–18:00 UTC in the East and 15:~~00 - 21~~00–21:00 UTC in the West and North, result in the strongest negative δ_e difference, and hence spatial SMPC relationship in all three domains. No positive δ_e values are found for these time periods, and the fraction of negative soil moisture gradients preceding rainfall events ~~are:~~ is relatively high, i.e. 62 %, 57 % and 55 % for East, West and North accordingly. Later accumulation times lead to a decrease in the magnitude and significance of the coupling parameter δ_e, and an increase in its

spatial variability across the domains. These changes are associated with an increase in the amount of the positive soil moisture gradients in the regions (not shown).

Despite the similarities, differences in the SMPC response exist between the domains. In the East, the spatial SMPC shows the strongest sensitivity to the rainfall accumulation time and switches the sign to a positive one for the 18:~~00 - 24~~00–24:00 UTC period. In accordance with Fig. **??**d, the 18:~~00 - 24~~00–24:00 UTC period reflects the afternoon progression of the mature MCS formed during early afternoon hours at the Ethiopian Highlands and around wetlands. ~~The~~ According to **?** and **?**, large and organized ~~MCS have shown~~ MCSs are expected to be more efficient in developing over wetter soils, associated to a well expressed BL moisture anomaly and higher MSE and CAPE ~~(??)~~ and, at the same time, might get suppressed over drier surfaces (**?**). These observations are consistent with the identified here increase in fraction of positive $\Delta({S'_e}^{L_{max}})$ in all the domains towards late afternoon hours, and the strongest SMPC response in the East.

Additional analysis reveals that the majority of large and negative soil moisture gradients in all domains are linked to the rainfall events that are identified during the first afternoon time step (i.e 12:00 UTC and 15:00 UTC for the East and West respectively) (not shown), and are therefore likely linked to weaker MCSs at the early stage of their development. The smaller and less organized MCSs ~~are known~~ have shown to be more sensitive to the thermally-induced surface convergence zones and are likely to develop over spatially drier soils, adjacent to the strong gradients (e.g, **?**). This knowledge is consistent with the strongest negative δ_e difference identified here and hence with the prominent negative SMPC relationship observed during early afternoon times in all three domains.

5 Results of temporal SMPC analysis

5.1 Co-variability of the spatial and temporal SMPC

To further explore the consistency of the identified negative spatial SMPC with the physical effects we analyze the temporal SMPC relationship. By analogy to the spatial SMPC, we compute the soil moisture anomaly ${S'_e}^{L_{max}}$ prior to the event and its difference δ_e to the typical state. Analysis of ${S'_e}^{L_{max}}$ and its δ_e indicates a strong preference for rainfall events to occur over soils that are drier than their temporal mean (Fig. **??**a) and drier than usual (Fig. **??**b). The percentile values P_e lower than 10% are found in 67 % of the studied $1^\circ \times 1^\circ$ boxes (Table 1). The latter implies that a temporally negative SMPC dominates over the domain, which reaffirms the co-existence of the negative spatial and temporal coupling identified by G15, but at a finer 1° horizontal scale.

The question remains whether the two coupling relationships are independent of one another. To answer this question we calculate the Spearman rank correlation coefficient[4] event-wise between the soil moisture anomaly ${S'_e}^{L_{max}}$ and soil moisture gradients $\Delta({S'_e}^{L_{max}})$ in every $1^\circ \times 1^\circ$ box. The correlation map in Figure **??**c shows that a high and significant correlation exists between ${S'_e}^{L_{max}}$ and $\Delta({S'_e}^{L_{max}})$ anywhere in the domain. The mean correlation of 0.47 over the domain supports the existence of relatively strong and positive monotonic relationship between the magnitude of spatial soil moisture gradient and

[4]Spearman correlation is a measure of monotonic relationship. Therefore, zero or low correlation value does not imply zero relationship between two variables.

soil moisture anomaly measured in L_{max}. For comparison, the mean correlation estimated between soil moisture gradients and mean soil moisture anomaly over the 1.25° event box is small (0.13). All the above suggests that in the North African region the spatial and temporal SMPC relationships, as defined by the current framework, are not independent of each other.

The strong and positive correlation ($in\ time$) identified between the soil moisture anomalies and gradients also yields a regional co-variability of the SMPC patterns. The *spatial* correlation between the two coupling distributions is high (0.64). The largest magnitudes of both ${S'_e}^{L_{max}}$ and $\Delta({S'_e}^{L_{max}})$ parameters and their corresponding δ_e measures are found in the southern part of the domain. These regions are generally characterized as the areas of higher BL moisture and rainfall frequency, and therefore higher variability of soil moisture in time and space.

Mechanistically, the presence of the temporally negative SMPC in the areas of the highest BL moisture in the domain (or lowest lifting condensation level (LCL), Fig. **??**a), is consistent with a higher relevance of mechanisms associated with the BL growth for convection initialization in regions of higher CAPE and lower convective inhibition (CIN) (**???**). In this way, larger negative deviations of the soil moisture amount from its climatological mean, i.e. δ_e, would lead to stronger than usual thermals, which can easier overcome CIN and release CAPE (**?**).

Moreover, in combination with a strong negative spatial gradients, these strong thermals can initiate breeze-like circulations, creating more favourable conditions for bringing BL up to the LFC, especially over the southern regions, where BL moisture is in abundance. The relevance of drier surface conditions for moist convection development on event days over the wetter latitudes is supported by the observed significant increase of the LCL height (decrease in pressure) in the South on event days compared to the typical state (Fig. **??**b, red shading). The higher LCL is associated with a decrease of BL relative and specific humidity (not shown) and supports the relevance of drier surface conditions for convection intensification as opposed to variations in BL water vapour amount prior to the events.

A different picture is observed over the drier latitudes of Northern Sahel at the Sahara margin. The estimated difference in LCL prior to the events relative to the typical state indicates that a significantly lower than usual LCL (higher pressure), associated to a significantly higher amount of water vapour in the BL (not shown), is present on the event days over the dry regions (Fig. **??**b, blue shading). This result is consistent with the previously reported decisive role of low-level moisture on MCS evolution in the drier Sahelian regions (**?**). At these latitudes the northward excursion of moist monsoon air has been shown to favour convective activity (**??**).

Considering also the relatively large number of dry days (10 days on average) preceding rain events in the North, it is less likely that underlying surface heterogeneity caused by a previous rainfall could have an influence on convection development on the event day. In the case study of **?** MCS was initiated due to the arrival of the cold pool and convergence zone emanated by a remote convective system hundreds kilometers away. Similar mechanisms may play a role in moist convection development in Northern Sahel.

6 Summary and discussion

6.1 Role of thermally-induced circulations in moist convection development

The dominant negative spatial SMPC relationship observed over the North African region agrees on the sign of the SMPC suggested by previous case- and modelling studies (e.g. **???**), and reconciles a number of physical effects.

Following the sensitivity analyses, the two main factors are identified, which directly influence the magnitude and significance of the negative spatial SMPC relationship. These are the time of the afternoon rainfall accumulation and the extreme soil moisture gradients. Our results show that the observed relationship between the spatial SMPC measure and the time of rainfall accumulation conforms well with the varying sensitivity of rainfall to the underlying soil moisture conditions for different stages of MCS development (**?????**). This further suggests potentially higher relevance of drier soils and soil moisture heterogeneity on rainfall ~~development~~ when the rain systems are smaller and at early stage of their development, consistently to results of e.g. **?**. ~~The results once again emphasize~~ All the above also emphasizes the importance of ~~consideration of~~ considering the MCS properties as ~~another limiting~~ an additional factor in the ~~probability-type methods when SMPC relationship is analyzed~~ analysis of the SMPC (**?**).

Extreme soil moisture gradients are found to predefine the significance of the negative spatial coupling in 30 % of the domain grid boxes. Concurrently, the extremes tend to cluster in the direct vicinity of major flooded areas and the irrigated land. The identified sensitivity of the afternoon rainfall to the strong ~~gradients in soil moisture~~ soil moisture gradients adjacent to wetland areas shows consistency with the role of ~~wetland-breeze~~ wetland–breeze mechanism in convection intensification over spatially drier soils. If this is true, our results demonstrate for the first time that the ~~wetland-breeze mechanism may well~~ wetland–breeze mechanism could be a systematic feature in the North African region with further implications for the rainfall predictability.

~~The~~ This observed sensitivity of the SMPC measure to the flooded areas and to the MCS life cycle ~~comply~~ complies well with a potential role of ~~soil moisture gradients in afternoon moist convection development via formation of~~ thermally-induced ~~circulations~~circulations in afternoon rainfall development over strong negative soil moisture gradients. Moreover, the identified preference of rainfall to occur simultaneously over temporally drier soils (negative temporal SMPC) and strong negative soil moisture gradients (negative spatial SMPC) might be considered as the most effective combination to initiate thermal circulations~~, which in turn~~ . This would imply a presence of a higher buoyancy and moisture flux at the event location, and hence ~~lead to~~ a higher probability of convection development. Schematic representation of the deep convection initiation by the breeze-like circulations formed under the co-existence of the two SMPC effects is illustrated in Figure **??**.

6.2 Role of rainfall persistence

In the context of this study, the drying of the soil prior to the rainfall events might be considered as the primary process that underlies the magnitude of both SMPC relationships, and helps to explain the opposite sign of the temporal coupling identified in the North African region as compared to the temperate latitudes and wet climates (G15).

Consistently to the observed 2 to 4 day periodicity of rainfall in Western Africa (**??**), 2 to 3 dry days (rain <1 mm) on average are found to precede each convective event day over southern latitudes, suggesting a strong drying of the upper soil

layer in the event locations prior to the rain. The number of dry days reaches 10 over the dry and deserted regions in the North. Following the analysis of **?** an almost complete recovery of the pre-rainfall surface moisture conditions may be expected in ~~2–3~~ 2–3 days following the rainfall. Schematically, this typical variability of rainfall and soil moisture might be illustrated as a sequence of daily rain events separated by the periods of drying (Fig. **??**a). From the Figure it is seen that prior to the rain events the soil dries out, and soil moisture reaches certain minimum value S_{min}. The climatology value S_{clim} of soil moisture in the same location, however, is expected to be higher than any S_{min} in most of the cases, as it includes all, dry and wet event days. Hence, when subtracted from the climatological value, a soil moisture measured prior to the event will very likely yield a negative anomaly - $S_e'^{L_{max}}$, especially when averaged over many events. Therefore, a negative correlation between soil moisture anomaly and rainfall might be expected. Though discussed in the framework of North Africa, similar behaviour might be expected in other water-limited regions of the world.

A different situation might occur in the wet temperate latitudes, where the variability of rainfall is to a large extent linked to fluctuations between passage of a cyclone and a blocking situation (**?**). Such a behavior might be illustrated as a multi-day sequence of rain events, associated with precipitation persistence as defined by the persistence in the weather regimes (Fig. **??**b, see also Fig. 2 in **?**). During these periods soil moisture increases and remains relatively high. Hence, a higher fraction of events might be expected to occur over soils that are wetter than usual, resulting in a positive soil moisture anomaly $S_e'^{L_{max}}$ prior to the event. The above relationship is consistent with the positive temporal SMPC, identified in G15. It is important to note, however, that the rainfall persistence may not be solely atmospherically driven, but may also reflect effects of the land surface (**???**).

The modulation of the SMPC sign depending on the large-scale weather regime was studied e.g. by **?** over France. The analysis showed that the synoptic blocking situations generally associated with drier conditions lead to a negative SMPC, while positive correlation of rainfall to drier soil conditions was observed in wet weather regime. Similarly, most pronounced effect of negative soil moisture gradients on convection initiation over Europe and a higher correlation of the gradients to land surface temperatures was observed for the period with less antecedent rainfall (**?**).

7 Conclusions

In this study, we revisit the negative spatial and negative temporal SMPC relationships identified earlier by T12 and G15 at the 5° × 5° horizontal grid. We use the probability-based approach of T12 and 10 years of satellite-based soil moisture and precipitation data (i) - to identify the potential link of the observed statistical relationships to the physical mechanisms and (ii) - to study the regional co-variability of the SMPC effects.

We find that the negative spatial coupling dominates over the region of North Africa. The result is independent of the choice of the observational data sets and is robust with respect to the event aggregation (spatial) scale. Compared to the previously used coarser 5° × 5° grid, the interest of the finest considered 1° × 1° event-aggregation scale is that it reveals links to the wetland areas and rivers which can not be captured at the coarser scale.

The co-variability analysis of the two SMPC relationships indicates that spatial and temporal effects of soil moisture on afternoon precipitation in the North African region do not only co-exist but are also dependent of one another. The latter suggests that if rain falls over temporally drier soils, it is likely to be surrounded by a wetter environment. This combination is consistent with the relevance of processes associated with the dominance of sensible heat flux and boundary layer growth on convection initiation, and supports the role of ~~meso-scale~~ mesoscale variability in surface soil moisture for deep convection development.

The identified negative sign of the temporal coupling in the semi-arid conditions of the Sahelian environment is not unexpected. The drying of the soil for several days prior to the rainfall events is likely to underlay the preference of rain to occur over temporally drier soils. This additionally may play a role in the opposite sign of the temporal coupling in the North African region as compared to the positive relationship identified in wetter climates by G15. For the same reason, the predictability potential of the temporal effect on rainfall in the North African region is expected to be lower compared to the spatial effect.

Analysis of the spatial SMPC measure and factors which can ~~affect~~ influence its magnitude and variability in particular reveals two "~~hot spot~~hot-spot" regions, where predictability skill of spatial soil moisture variability on rainfall might be higher. These are the Western African domain (~~7 - 15~~7–15° N, 10° ~~W - 7~~W–7° E), and South Sudan in the East (~~5 - 13~~5–13° N, ~~24 - 34~~24–34° E). In the Western domain, the negative spatial SMPC signal is indicated to be more robust. In the East, the spatial coupling is found to be largely modulated by the presence of wetlands and is susceptible to the amount of longer-lived propagating MCS. The analysis of the BL moisture conditions (here, LCL) preceding the rainfall events further supports potential relevance of spatially and temporally drier soils for convection development in the South of the domain, where BL moisture is in abundance. In the drier northern latitudes the variability of BL moisture, associated to intrusions of moisture from the south, seems to be more decisive.

Following our analysis, a number of potential improvements to the framework might be summarized. Apparent non-local effects of water bodies, that are originally excluded by the method, hints on the potential gaps in the filtering procedure and emphasizes potential role of moist convection evolution and propagation that are neglected by the method. The presence of wetland regions itself, as we have shown, complicates interpretation of the SMPC relationships. The uncertainty estimates of the soil moisture parameter derived over the recursively flooded regions are still missing. In the future, ~~application of~~ dynamical wetland products like the Global Inundation Extent from Multi-Satellites (GIEMS, **??**) may be used to better isolate the effect of water bodies on moist convection development.

Notwithstanding these limitations, the present study demonstrates the ability of probability-based methods to identify characteristic features of physical effects. Considering continuous increase in availability, time span and quality of satellite data, ~~the~~ development of similar statistical methods should be valuable. The ~~identified in the study regions~~ region of the strong ~~SMPC would also~~ negative SMPC identified in the East would highly benefit from more modelling and observational ~~studies especially in eastern North Africa~~analysis in the eastern Sahel. The knowledge on the regional variability of the SMPC presented here can be further explored in drought and climate change research, observational campaigns and for validation of GCMs.

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

Acknowledgements. The authors would like to thank Max Planck Institute for Meteorology (MPI-M) and International Max Planck ~~Reasearch~~ Research School (IMPRS) for providing facilities, material and scientific support which made publication of this paper possible. The authors would like to acknowledge Christopher Taylor, Benoit Guillod and Alexander Mahura for helpful comments on the study, and Stephan Kern for data support. We also thank George Huffman and Robert Parinussa for their clarifications related to TMPA and AMSR-E data products respectively.

Table 1. General statistics of the average event number, percentile P_e and δ_e difference estimated at various scales for three soil moisture parameters: soil moisture gradient $\Delta(S_e'^{L_{max}})$, temporal soil moisture anomaly $S_e'^{L_{max}}$, and (not presented in the methodology) soil moisture variance over the 1.25° box, $\sigma S_e^{1.25}$. Percentiles $P_e < 10$ % (> 90 %) indicate significant negative (positive) δ_e difference, and hence negative (positive) SMPC relationship.

$Parameter$	$Scale$	$\overline{Num_e}$	$P < 10$, [%]	$P > 90$, [%]	$\delta_{ev} < 0$, [%]	$\delta_{ev} > 0$, [%]
$\Delta(S_e'^{L_{max}})$:	$5 \times 5°$	309	72	0	92	2.8
	$2.5 \times 2.5°$	84	42	1.4	73	5.6
	$1 \times 1°$	17	21	1.3	43	14
$S_e'^{L_{max}}$:	$1 \times 1°$	-	67	0.8	92	8
$\sigma S_e^{1.25}$:	$1 \times 1°$	-	33	3	78	22

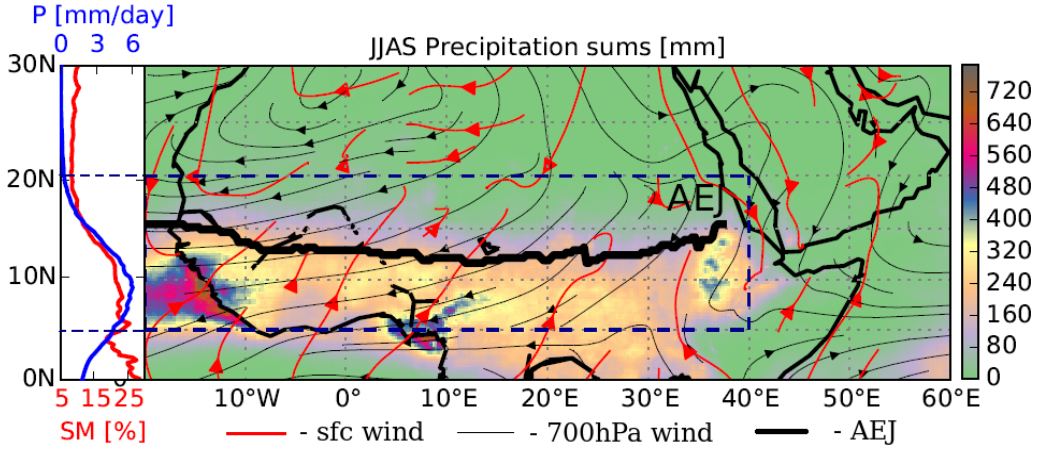

Figure 1. JJAS TMPA precipitation (shading), mean surface (red streamline) and 700 hPa (black streamline) ERA-Interim (1979 - present, **?**) wind climatology averaged over 2002 - 2011 period. The black thick line shows mean location of the African Easterly Jet (AEJ). Inset plot to the left shows zonal means of daily AMSR-E soil moisture (red) and TMPA precipitation (blue) climatology calculated over a study domain (dashed rectangle, 5° - 20°N; 20°W - 40°E).

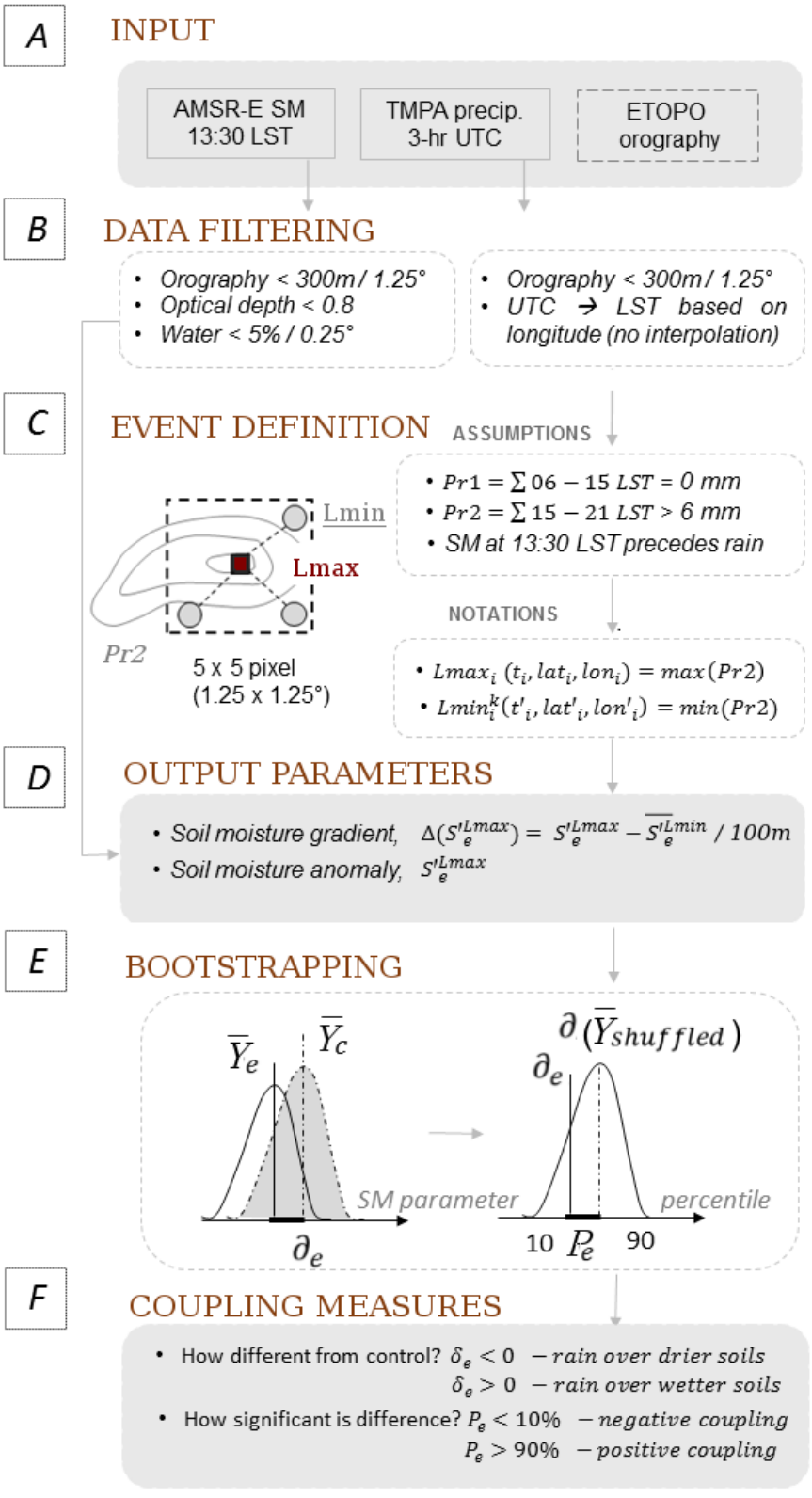

Figure 2. Schematic of data post-processing and statistical framework protocol implemented in the study.

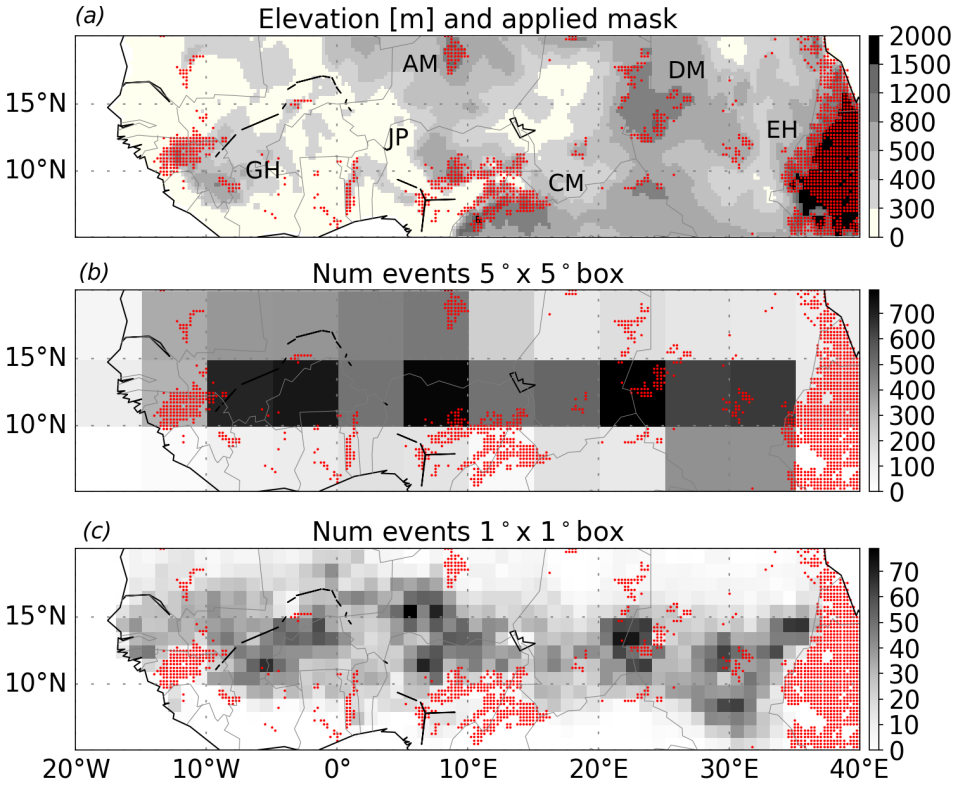

Figure 3. (a) - Elevation map based on 1 Arc-Minute Global Relief Model data ETOPO1 (**?**) (grey shading) and orography mask used in the study (red dots). Main orographic features of the region are: AM - Air Mountains, DM - Darfur Mountains, EH - Ethiopian Highlands, CM - Cameroon Mountains, JF - Jos Plateau, GH - Guinea Highlands. (b-c) - Number of events in every (b) $5^\circ \times 5^\circ$ and (c) $1^\circ \times 1^\circ$ box (gray shading) and applied orography mask at the 0.25° horizontal resolution (red dots).

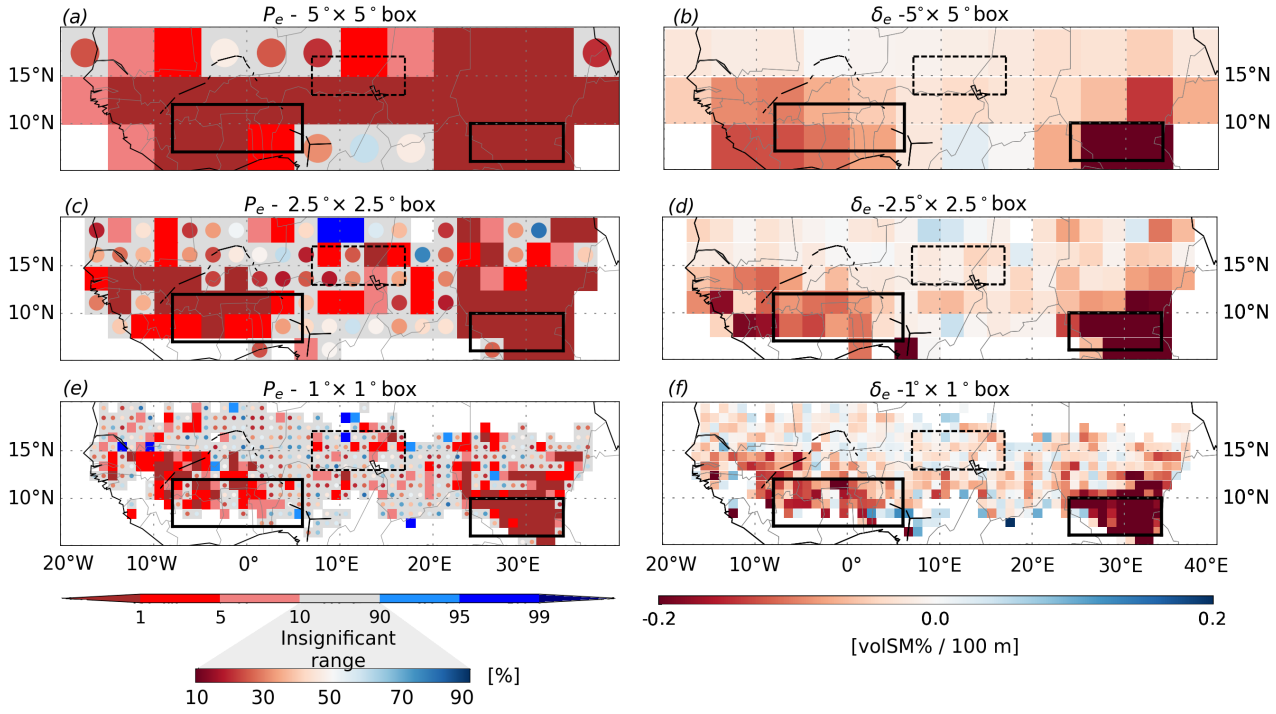

Figure 4. Distribution of percentiles P_e (*left*) of the observed δ_e difference (*right*), estimated over (a-b) $5^\circ \times 5^\circ$, (c-d) $2.5^\circ \times 2.5^\circ$ and (e-f) $1^\circ \times 1^\circ$ boxes. Percentiles $<10\%$ indicate significant negative coupling, i.e rain over spatially drier soils, and percentiles >90 % - significant positive coupling, i.e. rain over spatially wetter soils. The percentile values lying outside the significance range (i.e. in-between 10 % and 90 % percentiles) are illustrated by circles. Black rectangles on the maps indicate featured domains selected for an in-depth analysis.

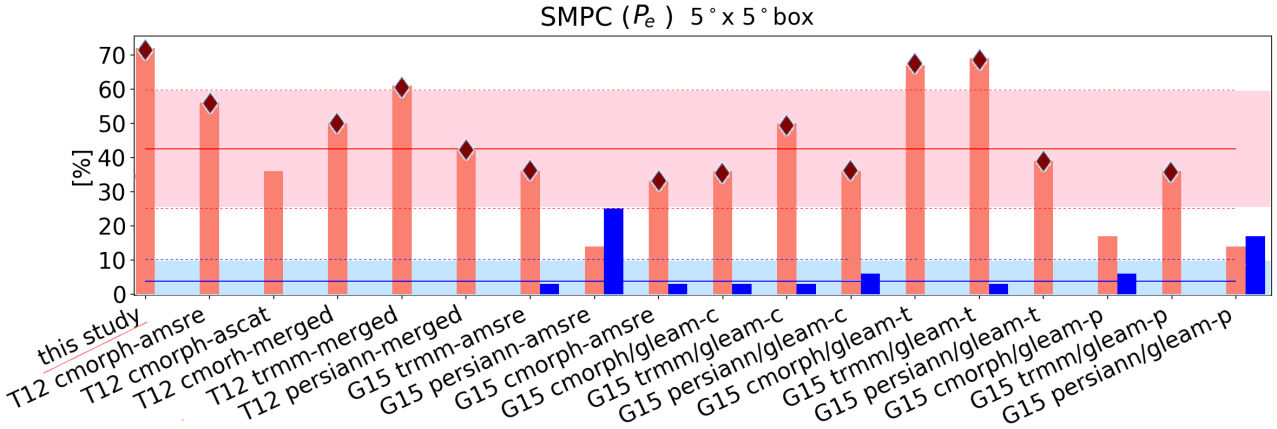

Figure 5. Percentage of $5^\circ \times 5^\circ$ grid boxes with significantly negative ($P_e < 10$ %, in red) and positive ($P_e > 90$ %, in blue) spatial SMPC over the North African domain in this study and previous studies of T12 and G15. Different data set combinations used in T12 and G15 are listed. Mean and st.dev. of the negative (positive) SMPC fractions across the experiments are shown as a red (blue) solid line and shading accordingly. Following visual inspection, the experiments, in which significant negative SMPC relationship exists in the western region of the North African domain are marked with a $rhomb$.

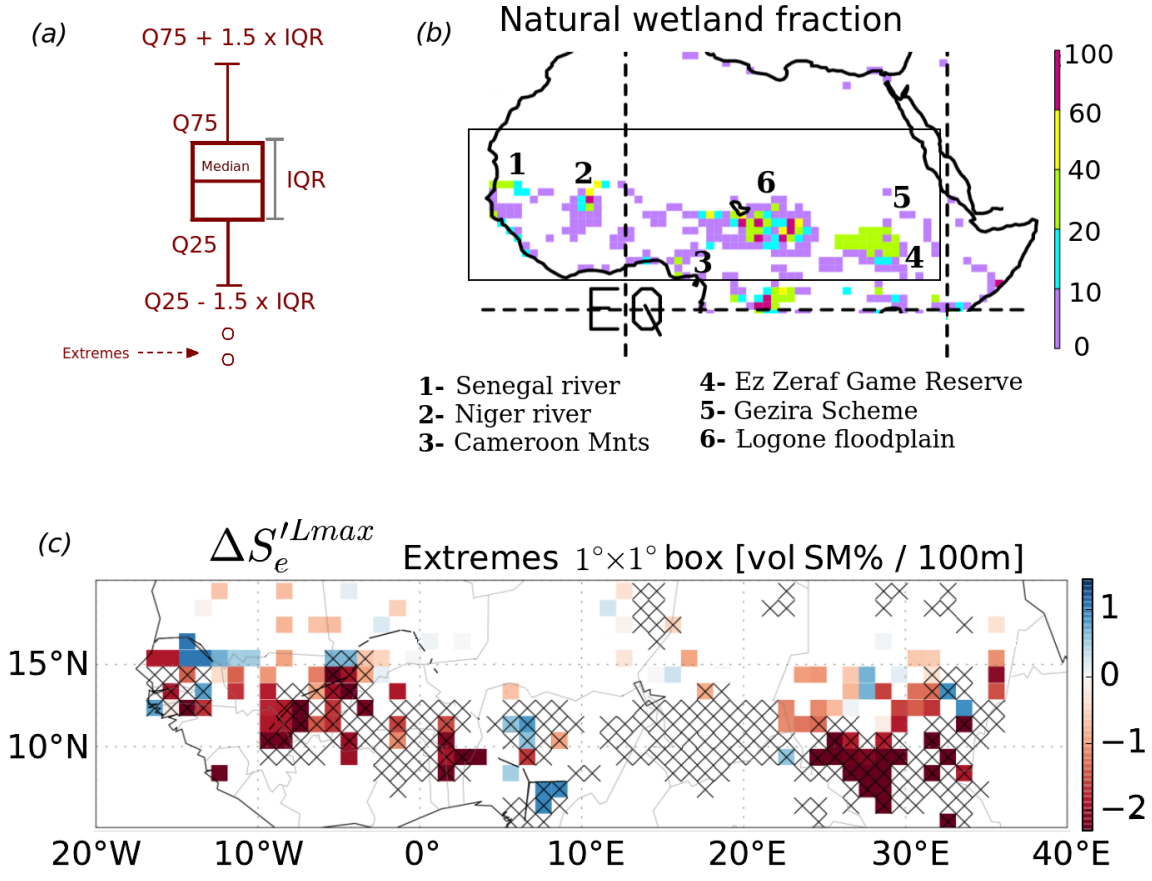

Figure 6. (a) - Schematic box plot illustrating the quantile range (Q_{25}-1.5 $\times IQR$, Q_{75}+1.5 $\times IQR$) ~~range~~ used to identify extreme $\Delta(S_e'^{Lmax})$ values. Here, Q_{75} and Q_{25} are the third and first quartiles respectively, and the interquartile range (IQR) is the difference between them. (b) - Natural wetland fraction from **?** on a 1° × 1° grid (adopted from Fig.3 in **?**). Names of major wetland and river locations are marked with a number. (c) - Distribution of soil moisture gradient $\Delta(S_e'^{Lmax})$ extremes in the corresponding event sample of a 1° × 1° box (color). $\Delta(S_e'^{Lmax})$ is considered to be an extreme if it lies outside the (Q_{25}-1.5 $\times IQR$, Q_{75}+1.5 $\times IQR$) range. If more than one extreme is found in one grid box, then its average value is shown. Black crosses indicate boxes containing $Lmin$ locations, in which climatology of daily soil moisture drying rates does not vary much with different absolute soil moisture values. These relationship is equivalent to low soil moisture variability in time, and is representative for a wet (flooded) locations. For detailed algorithm the reader is referred to the Appendix Figure A1. The distribution of identified potentially flooded locations is consistent with the natural wetland fraction (b).

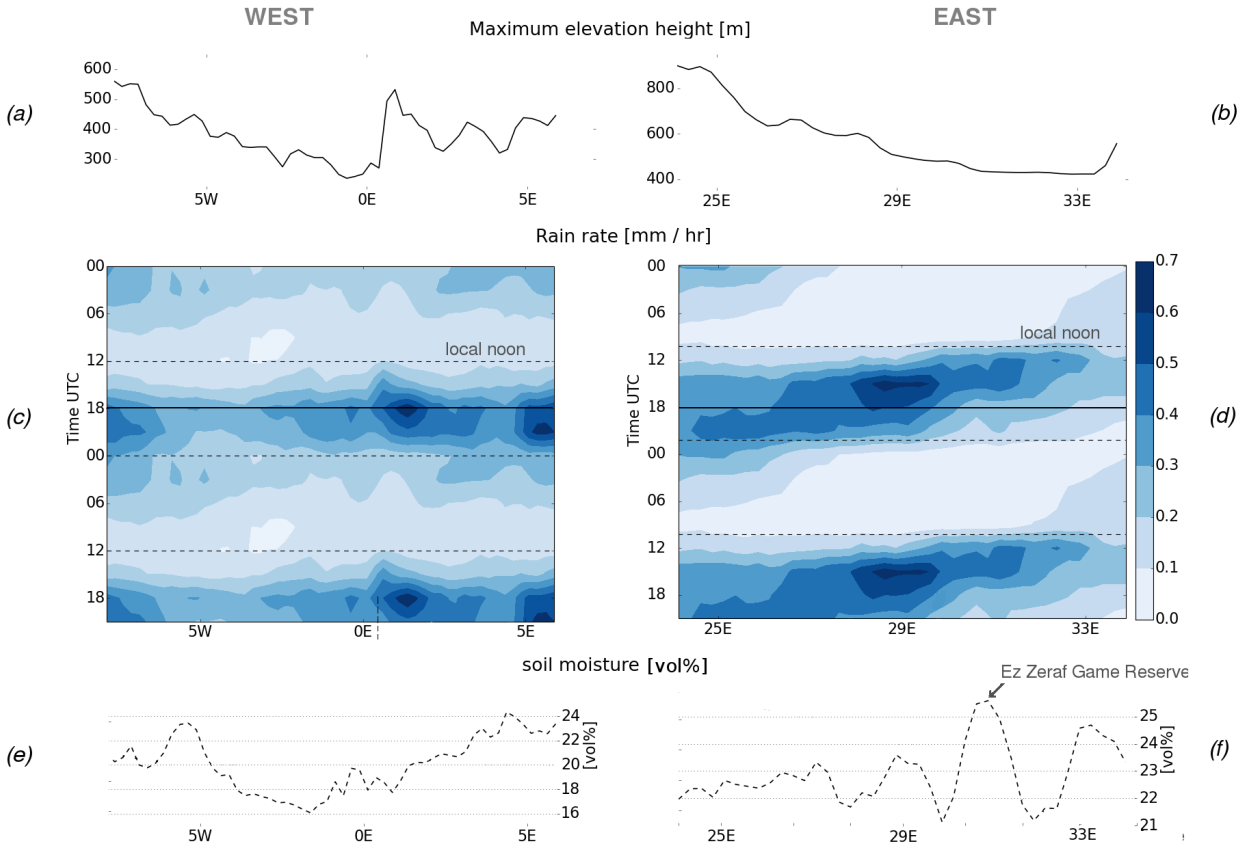

Figure 7. (a),(b) - Longitudinal cross-sections of maximum elevation height in the Western and Eastern domains respectively, (c),(d) - diurnal cycles of the rain rate averaged over event days and domain latitudes, and (e),(f) - Longitudinal cross section of soil moisture averaged over domain latitudes. Location of the Ez Zeraf Game Reserve permanent wetlands is marked by an arrow. All the times are given in UTC. Note, the UTC+2 hour difference to LST in the East.

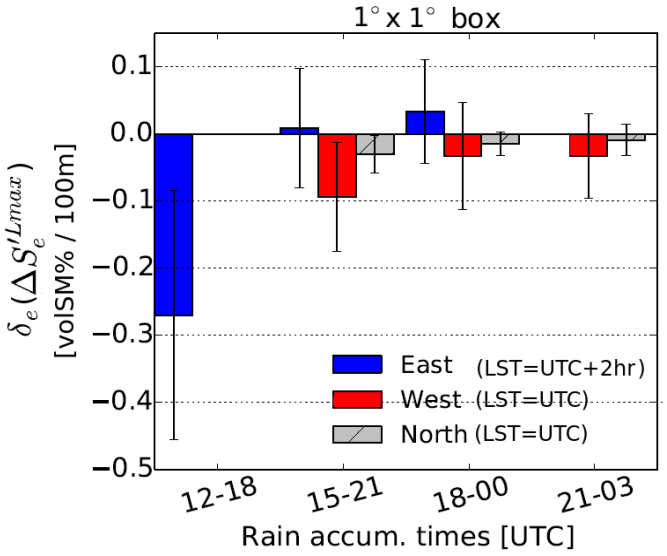

Figure 8. Value of the spatial coupling measure δ_e calculated for various afternoon rainfall accumulation times, and averaged over selected domains, i.e East (6 - 10° N, 24 - 34° E), West (7 - 12° N, 8° W - 6° E) and North (14 - 17° N, 7 - 14° E). Locations of the domains are shown in Fig. 4. Error bars indicate one std.dev. of δ_e values in every domain. Note, that all times are indicated in UTC.

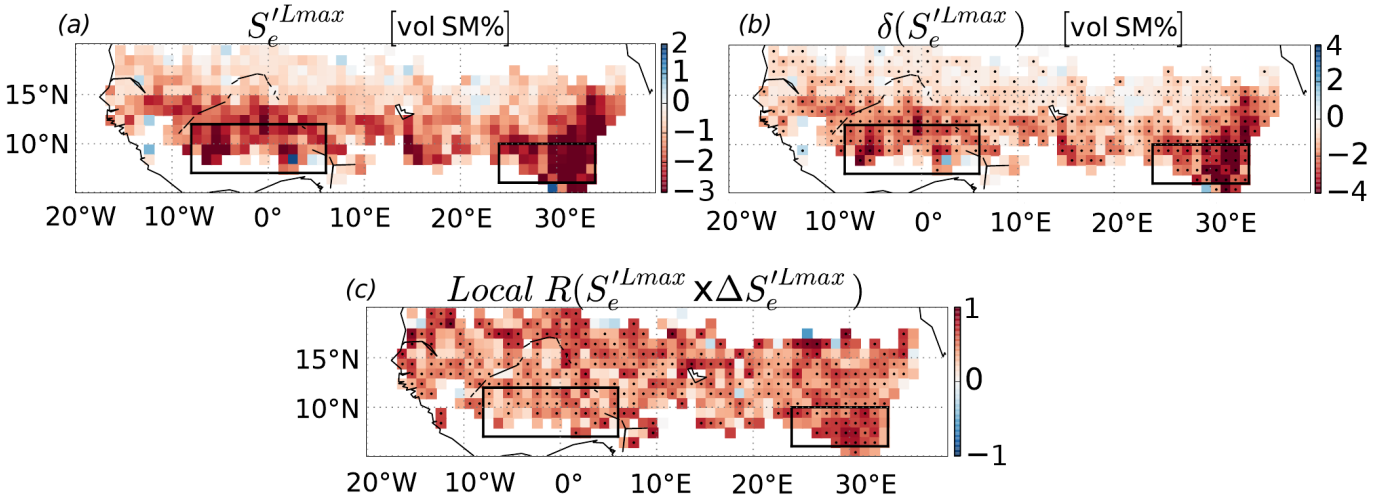

Figure 9. Distribution of the (a) pre-event soil moisture anomaly value $S_e'^{L_{max}}$ in event locations and (b) its departure from the typical non-event conditions, δ_e, averaged over $1^\circ \times 1^\circ$ boxes; and (c) *Local* Spearman rank correlation coefficient calculated event-wise between soil moisture anomaly $S_e'^{L_{max}}$ and spatial soil moisture gradients $\Delta S_e'^{L_{max}}$ in every 1° box. Significant δ_e values with percentiles P_e below 10 % (above 90 %) and correlation coefficients with p-values lower than 0.05 are indicated by black dots.

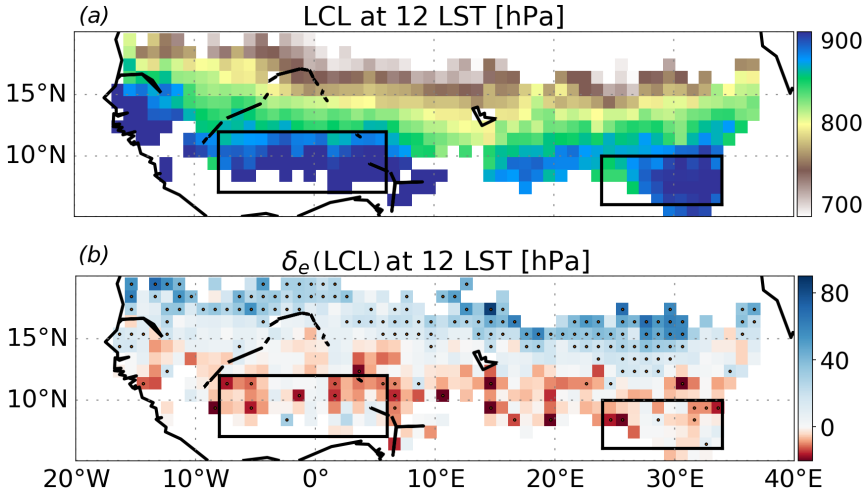

Figure 10. (a) - Lifting condensation level (LCL) value derived from the 6-hourly ERA-Interim temperature, specific humidity profile and surface pressure data on event days at 12:00 LST and averaged over the $1^\circ \times 1^\circ$ box. (b) - Corresponding δ_e difference of the mean LCL prior to the events relative to their typical state (climatology). The dot indicates significant δ_e values with percentiles P_e below 10 % (above 90 %). The positive (negative) δ_e values indicate lower (higher) than usual LCL (relative to the ground level).

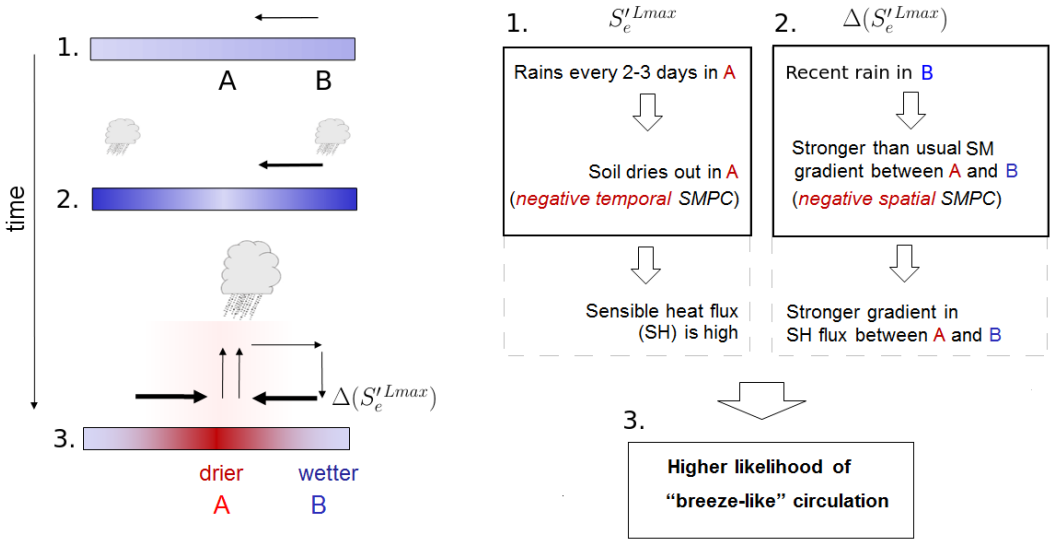

Figure 11. Conceptual diagram, illustrating intensification of moist convection by the initiated "breeze-like" circulations under favourable conditions of co-existing negative spatial and negative temporal SMPC effects. On the one hand, typically observed 2 to 4 day periodicity of rainfall in Western Africa leads to a strong drying of the upper soil layer in the location A prior to the rain, and therefore increases sensible heat and buoyancy flux locally. Simultaneously, recent rainfall in B produces wet soils - a potential moisture supply area for the location A. Strong spatial gradients in soil moisture between locations A and B together with a relatively strong buoyancy flux in A can favour formation of thermally-induced circulations under benign wind conditions. Considering a relatively weak mean surface wind of 2-3 m/s observed prior to the rainfall events over Southern latitudes, the ~~meso-scale~~ mesoscale circulations are likely to be initiated.

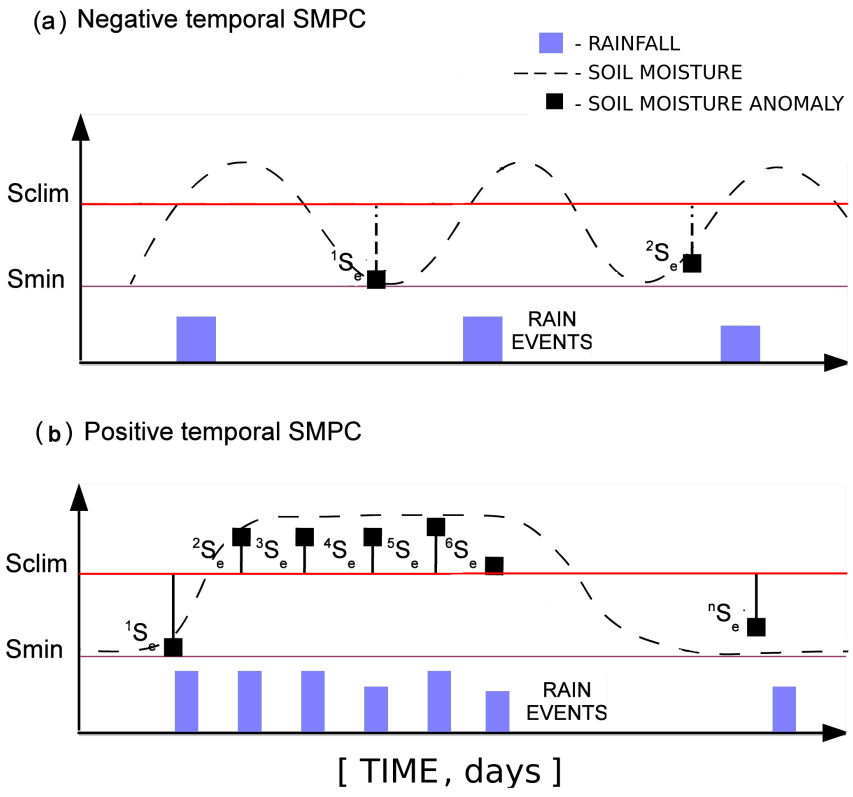

Figure 12. Conceptual diagrams of the relationship between daily rainfall occurrence and associated to it surface moisture variability in time as representative for (a) West Africa and temporally negative SMPC, and (b) Central and Northern Europe and positive temporal SMPC.

Appendix A: : Additional material to the Section 4.3

Method used for identification of potentially flooded locations

The following paragraph describes methodology used to identify events, which L_{min} locations are likely flooded. Following this methodology, a climatology of soil moisture drying rates is computed in every L_{min} location first. Climatology is calculated from values measured in the same L_{min} location, in the same month as the event but during the non-event years. Soil moisture drying rate is computed as the difference between soil moisture at 13:30 LST on the event day 0 and the previous day -1. Days with non-zero precipitation between two soil moisture measurements are excluded.

As a next step, we compute a climatology of soil moisture values on day -1 to estimate a potential range of soil moisture conditions in every L_{min} location. Finally, for every L_{min} location we identify a linear regression function which fits best into a scatter plot relationship between drying rates and initial soil moisture values. Based on the slope of the linear regression and a climatological range of drying rates we stratify L_{min} locations as being potentially "always dry", "normal" or "always wet" (Fig. A1). Locations, which are potentially "always dry" may be representative for rocky sand areas, and hence will most often show low soil moisture values and drying rates close to zero. The "always wet" cases should indicate mostly high soil moisture values, but at the same time small drying rates under condition that water supply is present. The "normal" case is expected to show a clear relationship between the drying rate and initial soil moisture content. The higher the soil moisture is the larger the drying rate is expected to be.

Because the identification procedure requires selection of thresholds, the distribution of the potentially flooded locations will slightly vary depending on the selected threshold of the regression slope, minimum drying rate value or minimum absolute soil moisture. For the calculation of the potential wetland locations presented in Fig. **??**c we chose two thresholds: the slope had to be larger than -0.15 and the soil moisture values - larger than 20 %.

Table A1. Percentage of $5^\circ \times 5^\circ$ grid boxes with significantly negative ($P_e < 10$ %) and positive ($P_e > 90$ %) spatial SMPC over the North African domain in this study and previous studies of T12 and G15. Different data set combinations used in T12 and G15 are listed. $Merged$ represents an integrated product composed of grid boxes in which either AMSR-E and ASCAT soil moisture data set was identified to be "best" following a quality-control check. Following visual inspection, the experiments, in which significant negative SMPC relationship exists in the western region of the Sahelian domain are indicated with a $*$ - sign. For more details on various data set combinations the reader is referred to the original papers.

#	$Experiment$	$Frac.\ P < 10\%$, [%]	$Frac.\ P > 90\%$, [%]
1*	This study (trmm-amsre)	72	0
2*	T12 cmorph-amsre	56	0
3	T12 cmorph-ascat	36	0
4*	T12 cmorph-merged	50	0
5*	T12 trmm-merged	61	0
6*	T12 persiann-merged	42	0
7*	G15 trmm-amsre	36	3
8	G15 persiann-amsre	14	25
9*	G15 cmorph-amsre	33	3
10*	G15 cmorph-gleam(c)	36	3
11*	G15 trmm-gleam(c)	50	3
12*	G15 persiann-gleam(c)	36	6
13*	G15 cmorph-gleam(t)	67	0
14*	G15 trmm - gleam(t)	69	3
15*	G15 persiann-gleam(t)	39	0
16	G15 cmorph-gleam(p)	17	6
17*	G15 trmm-gleam(p)	36	0
18	G15 persiann-gleam(p)	14	17
	$mean$	42	4
	$std.dev$	10	1

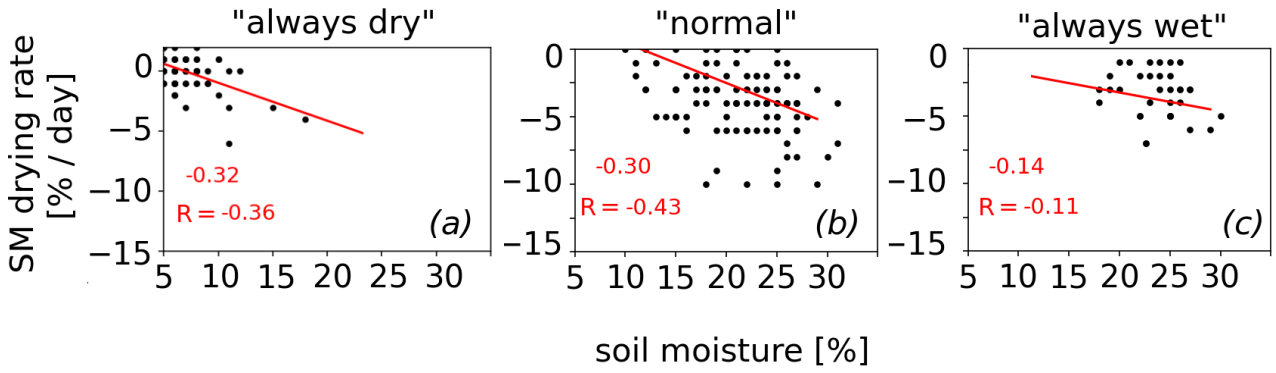

Figure A1. Examples of three possible relationships between climatology of soil moisture drying rate and absolute soil momisture in L_{min} locations. The red numbers indicate values of a slope and Spearmann correlation respectively.