Improving the quantification of climate change hazards by hydrological models: A simple ensemble approach for considering the uncertain effect of vegetation response to climate change on potential evapotranspiration

Thedini Asali Peiris[1] and Petra Döll[1,2]

[1]Institute of Physical Geography, Goethe University Frankfurt, Frankfurt am Main, Germany
[2]Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Frankfurt am Main, Germany
Correspondence: Thedini Asali Peiris <ThediniAsaliPeiris@em.uni-frankfurt.de>

Abstract.

Almost no hydrological model takes into account that changes in evapotranspiration are affected by how the vegetation responds to changing CO_2 and climate. This severely limits their ability to quantify the impact of climate change on evapotranspiration and thus water resources. As the simulation of vegetation responses is both complex and very uncertain. We recommend a simple approach for considering, in climate change impact studies with hydrological models, the uncertainty that the vegetation response causes for the estimation of future potential evapotranspiration (PET). To quantify this uncertainty in a simple manner, we propose to run the hydrological model in two variants, with its standard PET approach and with a modified approach for computing PET. In the case of PET equations containing stomatal conductance, the modified approach can be implemented by adjusting the conductance. We introduce a modified approach for hydrological models that computes PET as a function of net radiation and temperature only, i.e., with the Priestley-Taylor (PT) equation. The new PT-MA approach is based on the work of Milly and Dunne (2016) (MD), who compared the change of non-water-stressed actual evapotranspiration (NWSAET) as computed by an ensemble of global climate models (GCMs), which simulate vegetation response as well as interactions between the atmosphere and the land surface, with various methods for computing PET change. Based on this comparison, MD proposed to estimate the impact of climate change on PET as a function of only the change in net energy input at the land surface. PT-MA retains the impact of temperature on daily to interannual as well as spatial PET variations but removes the impact of the long-term temperature trend on PET such that long-term changes in future PET are driven by changes in net radiation only. We implemented PT-MA in the global hydrological model WaterGAP 2.2d and computed daily time series of PET between 1901 and 2099 using the bias-adjusted output of four GCMs. Increases of GCM-derived NWSAET between the end of the 20th and the end of the 21st century for RCP8.5 are simulated well by WaterGAP if PT-MA is applied but are severely overestimated with the standard PT. Application of PT-MA in WaterGAP results in smaller future decreases or larger future increases in renewable water resources (RWR) as compared to standard PT, except in a small number of grid cells where increased inflow from upstream due to increased upstream runoff leads to enhanced evapotranspiration from surface water bodies or irrigated fields. On about 20% of the global land area, PT-MA leads to an increase of RWR that is more than

20% higher than in the case of standard PT, while on more than 10% of the global land area, the projected RWR decrease is
25   reduced by more than 20%. While the modified approach for computing PET is likely to avoid the overestimation of future
drying in many if not most regions, the vegetation response in other regions may be such that application of the standard PET
leads to more likely changes in PET. As these regions cannot be identified with certainty, the proposed ensemble approach with
two hydrological model variants serves to represent the uncertainty of hydrological changes due to the vegetation response to
climate change that is not represented in the model.

30   *Copyright statement.*

TEXT

Abbreviations

Abbreviation	Definition
AET	Actual Evapotranspiration
CMIP5	Coupled Model Intercomparison Project 5
CO_2	Carbon-dioxide
DC	Relative difference of change
DGVMs	Dynamic Global Vegetation Models
GCM	Global Climate Model
MD	Milly and Dunne (2016)
NWSAET	Non-water-stressed- actual evapotranspiration
PET	Potential Evapotranspiration
PET-EO	Potential Evapotranspiration derived from Energy only method
PET-PM	Potential Evapotranspiration derived from Penman-Monteith method
PET-PT-MA	Potential Evapotranspiration derived from PT-modified approach (the new approach)
PM	Penman-Monteith
PT	Priestley-Taylor Equation
PT-MA	PT-modified approach (the new approach)
RCP	Representative Concentration Pathway
RWR	Renewable Water Resources
WGHM	WaterGAP Global Hydrological Model

1 Introduction

Appropriate estimation of evapotranspiration is essential for assessing water flows and storages on the continents, including renewable water resources, groundwater recharge and streamflow, and how they develop under climate change (Vörösmarty et al., 1998; Milly and Dunne, 2017). On average, about two-thirds of the precipitation over the continents (excluding Antarctica and Greenland) evapotranspirate (Müller Schmied et al., 2021), ranging from about 50% in very humid areas to more than 90% in arid areas (Zhao et al., 2013). Thus, small relative changes in evapotranspiration cause large relative changes in renewable water resources, particularly in the dry regions of the globe. The rate of evapotranspiration that occurs when there is an unlimited water supply is called potential evapotranspiration (PET), while actual evapotranspiration (AET) is often limited by available soil moisture. Hydrological models generally compute AET as a function of PET and soil moisture (Telteu et al., 2021).

PET is a variable that cannot be easily measured and has a high estimation uncertainty. According to Lu et al. (2005), there are about 50 different PET estimation techniques provided in the literature. They can be categorized into three groups, 1) temperature-based methods (Thornthwaite, Hamon, Hargreaves-Samani, Linacre, ...), 2) radiation-based methods (Makkink, Priestley-Taylor, Turc, ...), where PET is a function of temperature and radiation, and 3) combination methods (Penman-Monteith. ...), where PET is a function of radiation, temperature, wind speed and humidity (Zhao et al., 2013). PET values as calculated by different PET methods may differ significantly (Zhao et al., 2013; Lu et al., 2005; Weiß and Menzel, 2008; Kingston et al., 2009; Vörösmarty et al., 1998), and the same is true for computed impacts of climate change on PET (Kingston et al., 2009).

PET over land (i.e., not over open water surfaces) integrates both transpiration by plants and evaporation from canopy and soil. Therefore, PET depends on vegetation characteristics and processes that may change with anthropogenic climate change. PET is affected by three types of vegetation response to changing atmospheric CO_2 concentrations: the physiological effect, the structural effect (also called fertilization effect) and biome shifts (Gerten, D., Betts, R., Döll, P., 2014). For photosynthesis, plants take up CO_2 and release water through the leaf's stomata, tiny pores on the leaf surface that regulate the exchange between the plant and the atmosphere. With higher atmospheric CO_2 concentrations, the stomata close such that transpiration is reduced (physiological effect) (Purcell et al., 2018). At the same time, higher CO_2 concentrations may stimulate photosynthesis and thus biomass production and leaf area of C3 plants, thus increasing transpiration and evaporation from the canopy (structural effect) (Atwell, B. J., Kriedemann, P. E., & Turnbull, C. G., 1999; Berg and Sheffield, 2019). Climatic changes affect plant growth and plant type distribution and may lead to biome shifts, affecting evapotranspiration (Davie et al., 2013; Gerten, D., Betts, R., Döll, P., 2014; Berg and Sheffield, 2019).

Vegetation response to changing atmospheric CO_2 concentrations and climate is simulated by dynamic global vegetation models (DGVMs). DGVMs simulate physiological processes, such as photosynthesis and respiration, and biogeochemical cycles and include the effects of fire, atmospheric CO_2 concentration and competition between plant life forms for light, water and nutrients on vegetation dynamics but still neglect other relevant vegetation responses (Cramer et al., 2001; Thonicke et al., 2001). Quantification of the overall effect of the vegetation response on changes of evapotranspiration and other hydrological

variables is still uncertain. Depending on the region and the model, the overall effect can be an increase or a decrease of evapotranspiration, but there is a strong tendency towards a decrease of evapotranspiration as compared to assuming no response (Davie et al., 2013; Gerten, D., Betts, R., Döll, P., 2014; Milly and Dunne, 2016; Reinecke et al., 2021). However, responses

70 of different DGVM for a specific region may differ strongly and even in the sign of change (Davie et al., 2013; Reinecke et al., 2021). Analyzing the impact of future climate change on groundwater recharge based on four models that simulate the impact of increasing atmospheric CO_2 concentrations and climate on vegetation and four models that do not do this, Reinecke et al. (2021) found that the former models simulate a lower increase in AET than the latter in 19 out of 24 world regions. Exceptions are five of the regions with projected decreases in precipitation. In particular, for these regions, the range of computed changes

75 of groundwater recharge is much smaller for the models that do not simulate vegetation processes, underlining the uncertainty of simulating the effect of vegetation processes on AET.

Typical hydrological models, however, do not take into account the vegetation response to changing atmospheric CO_2 concentrations and climate when computing PET as affected by climate change. The impact of climate change on water resources, i.e., streamflow, groundwater recharge, or other hydrological variables, is almost exclusively estimated by hydrological models

80 that do not compute dynamic vegetation processes. While these hydrological models may be able to simulate historic streamflow dynamics well, they lack the capacity to simulate the change in evapotranspiration and thus streamflow due to changing vegetation processes such that computed climate change impacts on hydrological variables are likely biased. This is true for hydrological models at any scale.

Global climate models (GCMs) simulate atmospheric, vegetation and soil processes as well as their interactions owing

85 to the fact that DGVMs are integrated into their land surface models. GCMs typically compute AET from land based on a system of process-based equations, distinguishing canopy evaporation, transpiration, and evaporation from the soil. Thus, the representation of processes affecting AET is more comprehensive in GCMs than in hydrological models, but the uncertainty of computed AET and AET changes remains high (Sepulchre et al., 2020; Milly and Dunne, 2016; Jones et al., 2011; Watanabe et al., 2010; Randall et al., 2007; Cramer et al., 2001). When computing AET, GCMs do not use the PET concept; together

90 with the need for bias-correction of the GCM output for hydrological modeling studies, this prevents those hydrological models can benefit from the complex simulation of vegetation processes and land surface interactions performed by GCM (Milly and Dunne, 2017).

The GCM output that can be utilized for informing on how PET might change under climate change, which is the AET computed by GCMs in locations and times without water stress. This is equivalent to the PET. Milly and Dunne (2016) (MD)

95 compared the future change of non-water-stressed AET (NWSAET) as simulated by GCMs with PET calculated by different PET methods using the climate variables of the GCMs. They analyzed NWSAET changes between the reference period 1981-2000 and the period 2081-2100 using the output of 16 CMIP5 climate models under the Representative Concentration Pathway 8.5 (RCP8.5), considering the mean changes of all GCM-specific grid cells and months without water stress in the reference period. They found that, for all GCMs, the PET changes calculated with two Penman-Monteith (PET-PM) variants are about

100 twice as large as the changes of NWSAET. When calculating PET-PM with changing surface resistance, MD concluded that the most significant contribution to overestimating PET originates from the negligence of the physiological effect but that

feedbacks between vegetation and atmosphere, which cannot be simulated by any hydrological or land surface model that is not coupled to an atmospheric model also important. MD showed that long-term changes in NWSAET were best approximated by simply assuming that PET change equals 80% of the change in net radiation (PET-EO, Eq. 8 in MD). Regarding the ensemble

105 mean of the 16 GCMs, the change in PET-EO was equal to the change in NWSAET, while the differences between PET-EO and NWSAET change for the individual GCMs was much smaller than the difference between PET-PM and NWSAET.

Yang et al. (2019) analyzed the same set of CMIP5 climate models under RCP8.5 as MD and found that the long-term changes in annual mean surface resistance of non-water-stressed grid cells and months increased linearly with atmospheric CO_2 during 1861-2100, with model-specific sensitivities between $0.05\%/\text{ppm}$ and $0.15\%/\text{ppm}$. They stated that an increase

110 in evapotranspiration caused by a warming-induced vapor pressure deficit increase is almost entirely offset by a decrease in evapotranspiration caused by increased surface resistance (i.e. decreased stomatal conductance) as driven by rising CO_2. They proposed that those hydrological models that use PET-PM adjust the PET equation such that surface resistance is expressed as a function of atmospheric CO_2, using the ensemble mean sensitivity of $0.09\%/\text{ppm}$.

Calibrating a very simple hydrological model of annual AET as a function of PET and precipitation to the AET as computed

115 by 24 GCMs for RCP8.5, Milly and Dunne (2017) found that averaged over the global land surface and for five out of nine large river basins, the humid Columbia, Mississippi, Amazon, Congo and Danube basins, PET-PM as well as PET computed by the Priestley-Taylor approach (PET-PT) strongly overestimate the future increase in AET, while PET-EO leads to a good fit to the GCM ensemble mean. However, for two snow-dominated basins (Mackenzie and Ob) and two semi-arid basins (Colorado and Yellow River), AET computed with PET-PM and PET-PT fit better to the GCM AET than AET computed with PET-EO,

120 which leads to an underestimation of the future AET increase.

With this paper, we propose the following approach for climate change impact studies done with hydrological models that do not simulate vegetation processes. To approximately represent the uncertainty of future hydrological changes caused by the uncertainty of the vegetation response to future climate change (including increasing CO_2), hydrological models should be run in two variants. In variant A, their standard (net radiation-based) PET approach is used for estimating conditions under

125 future climate change, while in variant B, PET changes in the future are assumed to occur due to changes in net radiation only, according to the studies of Milly and Dunne (2016), Milly and Dunne (2017) and Yang et al. (2019). Considering the studies of Milly and Dunne (2017) and Reinecke et al. (2021), variant B is expected to lead to more reliable hydrological changes in most regions, but it is not yet clear in which regions. We suggest that these two variants approximately represent the uncertainty bounds related to the vegetation response. They thus serve to generate an improved ensemble of future hydrological changes,

130 which, as a standard, includes model runs driven by the (bias-adjusted) output of multiple GCM or the output of multiple hydrological models e.g.,Davie et al. (2013); Reinecke et al. (2021).

While hydrological models using PET-PM can apply the approach of Yang et al. (2019) to implement variant B, this paper presents an approach that is suitable for hydrological models that do not compute PET as a function of surface resistance, such as Priestley-Taylor (PT). The approach is applicable for estimating the change of hydrological variables between a reference

135 period and a period in the future. The proposed approach aims at leading to a similar effect on PET and runoff as the complex GCMs (with DGVMs) show (at least on average). The new "modified approach" (hereafter PT-MA) to compute PET, removes

the long-term temperature trend in the PET computation such that, following MD, PET changes due to climate change occur only due to changes in net radiation. The effect of short-term and spatial temperature variations on PET is still considered. PT-MA enables estimation of spatially variable daily PET time series as a function of net radiation and temperature while approximately considering the net effect of vegetation response to changing CO_2 and climate on PET. It is validated by implementing PT-MA in the global hydrological model WaterGAP 2.2d using the bias-adjusted output of four GCMs available on the ISIMIP data portal (Frieler et al. 2017) and comparing PET changes simulated by WaterGAP to NWSAET changes of three GCMs included in MD.

The following section describes the new PT-MA approach, its integration into WaterGAP, and the model experiments performed for this study. Section 3 first presents the validation results and then compares future changes in global-scale PET and renewable water resources (RWR) as computed by the standard (PT) and the new (PT-MA) method, using climate scenarios derived by four GCMs. In addition, the effect of applying PT-MA on other hydrological variables and under four different emissions scenarios is presented. Section 4 compares the PET uncertainty due to the PET approach to the uncertainty due to the GCMs and presents the caveats of the proposed approach. Conclusions are drawn in section 5.

2 Methods and Data

2.1 The global hydrological model WaterGAP 2

With a spatial resolution of $0.5° \times 0.5°$ grid, the global hydrological model WaterGAP 2 computes human water use from either groundwater or surface water (via the GWSWUSE submodel of WaterGAP) and takes these into account when, using the submodel WGHM, daily water fluxes (e.g., AET and streamflow) and storages (e.g., in groundwater and surface water bodies) are calculated (Müller Schmied et al., 2021). WGHM is driven by daily inputs of temperature, precipitation, downward shortwave radiation, and downward longwave radiation as well as by net abstractions from groundwater and surface water.

In WGHM, water flows between the water storage compartments canopy, snow, soil, groundwater, surface water bodies (this includes Wetlands, Lakes and Reservoirs) and river are simulated (Müller Schmied et al., 2021). Total AET is the sum of canopy evaporation, snow sublimation, evapotranspiration from the soil and evaporation from surface water bodies. Canopy evaporation is calculated as a function of PET and leaf area index. AET from the snow (i.e., sublimation) is determined as the fraction of PET that remains after canopy evaporation. AET from soil is a function of soil PET (calculated as the difference between total PET, snow sublimation and canopy evaporation) and soil water saturation. When computing soil AET, transpiration of plants is not distinguished from evaporation from the soil, and like typical hydrological models, vegetation responses to changing atmospheric CO_2 and climate that affects transpiration such as stomatal closure and changing leaf area are not simulated. AET of open water bodies is equal to PET. Per default, PET $[\mathrm{mm/day}]$ is computed in WGHM according to the Priestley-Taylor (PT) equation following Shuttleworth (1993) as

$$PET = \alpha \frac{\delta R_n}{\delta + \gamma} \tag{1}$$

where α is an empirical constant accounting for the effect of the vapor pressure deficit not taken into account directly in PT [-] (for humid areas $\alpha = 1.26$ and for arid/semi-arid $\alpha = 1.74$), R_n is net radiation [mm/day], γ is the psychometric constant [kPa/°C] and δ is the slope of the saturation vapor pressure-temperature relationship [kPa/°C], with

$$\delta = \frac{4098(0.6108e^{\frac{17.27T}{T+237.3}})}{(T+237.3)^2} \tag{2}$$

where T is daily temperature [°C]

R_n is calculated using the climate input data downward shortwave radiation and downward longwave radiation as well as upward shortwave radiation and upward longwave radiation, both of which are computed in WGHM (Müller Schmied et al., 2016). Upward shortwave radiation is computed as a function of land cover-specific albedo, while upward longwave radiation is computed as a function of land cover-specific emissivity and temperature (Müller Schmied et al., 2016). If snow storage exceeds 3 mm in a 0.5° grid cell, a land cover-specific snow albedo is applied, while R_n of surface water bodies is set at 0.08 (Müller Schmied et al., 2021).

In this paper, WaterGAP 2.2d as described in Müller Schmied et al. (2021) is applied for two purposes, 1) to validate the PT-MA method against changes in NWSAET computed by three GCMs as analyzed by MD and 2) to investigate the impact of the modified approach when computing PET and respective impact on renewable water resources (RWR) and other hydrological variables. WGHM outputs used in this study are R_n, total PET, total AET and streamflow that is derived from bias-adjusted outputs of GCMs. RWR of each grid cell is calculated as the difference between the streamflow leaving the cell and the streamflow entering it.

WGHM has been calibrated against observed mean annual streamflow at 1319 gauging stations Müller Schmied et al. (2014) using the EWEMBI (E2OBS, WFDEI, and ERA-Interim data set merged and bias-adjusted for ISIMIP) (Frieler et al., 2017; Lange, 2016) climate data set. The new PET calculation method (PT-MA) is implemented in WGHM as an alternative PET scheme to be used specifically for climate change studies.

2.2 PT-MA approach for adjusting PET computed according to Priestley-Taylor

In the case of the PT method, PET increases with temperature due to the temperature dependence of δ (Eq. 2). In the case of $\alpha = 1.26$ and $T = 16$ °C, for example, PET = $0.80R_n$, while for $T = 18$ °C, PET = $0.84R_n$. Thus, PT-derived PET, therefore, increases with global warming. According to both MD and Yang et al. (2019), the impact of the temperature increase on PET is approximately canceled by the impact of changes of other processes that are taken into account by GCMs but not by typical hydrological models. MD proposed that the long-term change of NWSAET and thus PET is best approximated by the change of R_n multiplied by 0.80, which they called PET-EO (energy-only). Therefore, in the proposed PT-MA approach, the daily temperature values obtained from GCM-derived climate scenarios that are used as input to hydrological models are modified such that the long-term temperature trend of the future time period is removed (hereafter modified temperature).

We chose the period 1981-2000 as the reference period for the implementation of PT-MA in WGHM, and trend removal started in 2001. Selecting this reference period enabled a direct comparison between the results of our implementation with MD, but the PT-MA approach can be implemented with other reference periods, too. To compute modified daily temperature

($T_{modified}$ [ºC]), a grid-cell specific temperature reduction factor T_{diff} [ºC] is calculated for each year, with

$$T_{diff,i} = T_{mean,(i-10)-(i+9)} - T_{mean,1981-2000} \qquad \text{for i = 2001-2099 (if i>2091, replace (i+9) by 2099)} \qquad (3)$$

where $T_{mean,(i-10)-(i+9)}$ is the annual mean temperature of the 20-year period around year i (i.e., if i = 2001, it is the annual mean temperature of 1991-2010). T_{diff} removes the long-term temperature trend from the daily temperature time series in the future period. For a given day in year i,

$$T_{modified,day,i} = T_{day,i} - T_{diff,i} \qquad (4)$$

Use of $T_{modified}$ in the calculation of δ to determine PET with PT-MA keeps the 20-year-mean temperature at the level of the reference period while it still varies at the daily to inter-annual scales (Eq. 2).

When the PT-MA option is selected in WGHM, $T_{modified}$ is used, starting in 2001, for computing δ (Eq. 2) but not when computing evaporation from open water bodies, as the temperature effect on PET of open water bodies is not reduced by the closure of any stomata. Computation of upward longwave radiation is always done using T, in accordance with PET-EO of MD.

2.3 Data and modelling experiments

To assess the proposed approach, a series of GCM-driven WGHM simulations were conducted. Bias-adjusted GCM-derived climate data (daily data for temperature, precipitation as well as shortwave down and longwave down radiation) that are available on the ISIMIP2b data portal (Frieler et al., 2017) were used as input data for the pre-calibrated WGHM. Thirty-two model runs were conducted that combined four GCMs (GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR, and MIROC5), four RCPs (RCP2.6, RCP4.5, RCP6.0, and RCP8.5), and two PET schemes (PT and PT-MA). Each simulation was done for the period 1901-2099.

For validating the PT-MA method, changes in PET between the reference period (1981-2000) and 2080-2099 as computed by WGHM were compared with the GCM-derived NWSAET changes and PET-EO changes of MD. Monthly values corresponding to non-water-stressed months and grid cells identified by MD were compared. By definition, PET computed for grid cells and months without water stress should be equal to the AET (i.e., NWSAET) computed by the GCMs for the same grid cells/months (Milly and Dunne, 2016; Yang et al., 2019). The authors of the Milly and Dunne (2016) study provided us with non-water-stressed grid cells/months for GCMs GFDL-ESM2M, HadGEM2-ES, and IPSL-CM5A-LR. MIROC5 is not included in the MD study; hence MIROC5 derived outputs are not included in the validation analysis of this study. Grid cells/months where the reference level air temperature is less than 10°C were removed in MD to avoid frozen water. Non-water-stressed cells of the three selected GCMs are concentrated in Southeast Asia and South America (Fig. 1).
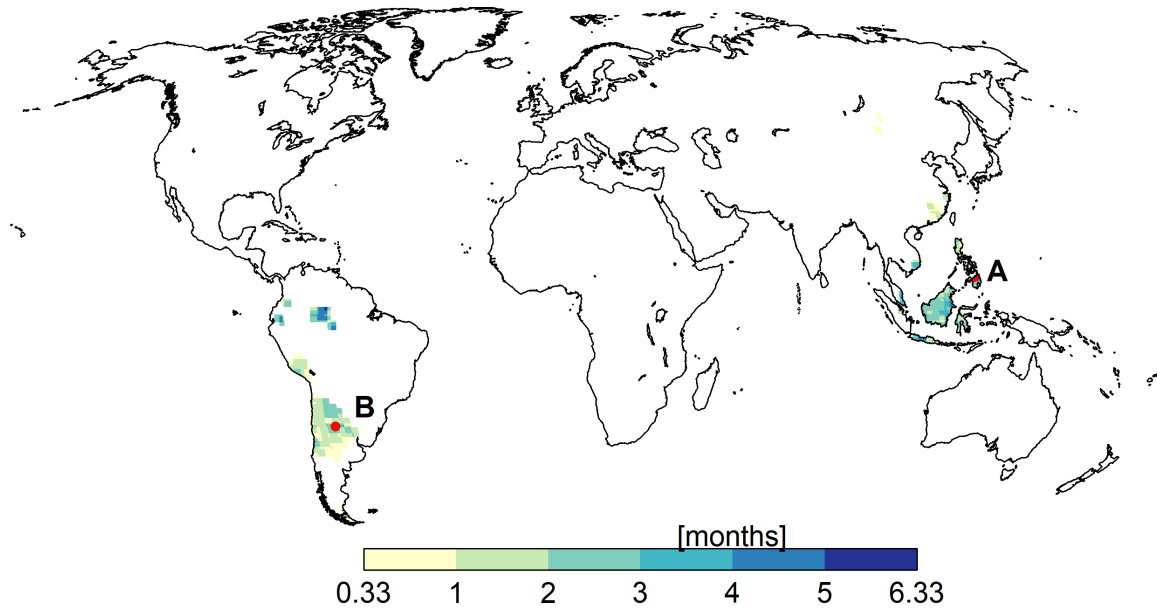
Figure 1. Mean number of calendar months per year when evapotranspiration is non-water-stressed over the reference period (1981-2000). The values shown here were obtained by averaging the outputs of three GCMs.

To understand the behavior of the PT-MA method at the grid cell level, two grid cells were selected. Cell A (cell center at 8.75°latitude, 124.75° longitude) is located in the Philippines, cell B (cell center at -29.75° latitude, -64.25° longitude) in Argentina (Fig. 1). According to GFDL-ESM2M, grid cells A and B are non-water-stressed for one and two months per year, respectively, during the reference period. In the case of IPSL-CM5A-LR, the corresponding values are four and zero months, respectively, and in the case of HadGEM2-ES, both cells are under non-water-stressed conditions for four months.

2.4 A metric to quantify the impact of the PT-MA approach.

We calculated a metric for quantifying the magnitude of the impact of the PT-MA approach on the change of PET and RWR with respect to the standard PT approach. The metric is a signal-to-change ratio, the relative difference of change (DC) [%], with

$$DC_{var} = \frac{(dvar_{PT-MA} - dvar_{PT})}{dvar_{PT}} * 100 \tag{5}$$

where var is the variable (i.e., PET or RWR), $dvar_{PT-MA}$ is the change (between future and reference periods) of the variable computed according to PT-MA approach and $dvar_{PT}$ is the change of the variable computed according to PT approach.

The DC metric can be interpreted as follow. If DC is less than -100%, the changes as computed by the two approaches do not agree in the sign; otherwise, both approaches show either increases or decreases. PET is expected to generally increase in the future, and the increase in the case of PT-MA should be smaller than in the case of the standard PT such that negative

DC values should prevail. The DC value indicates how many percent smaller the PET increase is with PT-MA as compared to the standard PT. RWR, however, may increase or decrease in the future. As future RWR should generally be higher with PT-MA as compared to the standard PT, due to the smaller PET change in case of PT-MA, positive DC values should generally occur where RWR increases in the future and negative values where RWR decreases. If DC is positive, the RWR increase with PT-MA is DC% larger than with PT; if DC is negative, the RWR decrease in the case of PT-MA is DC% smaller. For example, if RWR is projected to decrease by 20 $\mathrm{mm/year}$ in the case of the standard PT but by only 10 $\mathrm{mm/year}$ by PT-MA, DC = -50%.

3   Results

3.1   Validation of the PT-MA approach

Performance of the PT-MA method is analyzed based on the area-weighted average changes of PET and R_n over non-water-stressed grid cells and months (Figure 1), considering the changes between the reference period 1981-2000 and the future period 2080-2099 for RCP8.5 as only this RCP was considered in MD (Table 1). Table 1 also presents the respective values for the reference period. PET-PT is the PET as calculated by the standard WGHM, while PET-PT-MA is the result of the PT-MA method presented in section 2.2. For both variants, R_n is computed based on the bias-adjusted output of the three GCMs (section 2.1). PET-EO and NWSAET values for three GCMs were extracted from the MD study and are therefore not affected by any bias adjustment, as they had been derived by MD using the original GCM output.

The results indicate that averaged over the three GCMs, the ensemble GCM mean PET-PT-MA change is, with 0.31 $\mathrm{mm/day}$, only about half of the PET change computed with the standard PT (0.57 $\mathrm{mm/day}$). This is similar to the reduction of the ensemble mean PET change if NWSAET or PET-EO was used instead of PET-PM for the 16 GCMs in MD. The PET-PT-MA change is much closer to the NWSAET change of 0.19 $\mathrm{mm/day}$ and the PET-EO change of 0.25 $\mathrm{mm/day}$ than the PET-PT change but still overestimates both.

For each individual GCM, too, PET change in the case of PT-MA is much closer to NWSAET, PET-EO and R_n than in the standard approach but both PET-PT-MA and PET-EO overestimate NWSAET change (Table 1). For two of the GCMs (GFDL-ESM2M and HadGEM2-ES), PET-PT-MA and PET-EO are very similar in terms of change. However, for IPSL-CM5A-LR, PET-PT-MA overestimates the change compared to both NWSAET and PET-EO. For the three GCMs, the PT-MA method reduces the PET change by 0.18-0.31 $\mathrm{mm/day}$ relative to standard PT. This is a significant reduction when compared to the range of NWSAET changes between the three GCMs of 0.10-0.27 $\mathrm{mm/day}$ (Table 1).

Table 1. Comparison of PET and PET changes as computed by WGHM using the standard PT and the newly developed PT-MA approach to the actual evapotranspiration computed by global climate models under non-water-stressed conditions (NWSAET) and the PET-EO approach of MD (their Fig. 1 and S2). Area-weighted averages over all non-water-stressed grid cells/months of these variables as well as of WGHM net radiation (R_n), are shown for the reference period 1981-2000. PET-PT, PET-PT-MA and R_n are computed by WGHM using the bias-adjusted output of the listed GCMs. "Change" refers to the change between the reference period and 2080-2099 in the case of emissions scenario RCP8.5. All values are in $\mathrm{mm/day}$.

	GFDL-ESM2M		HadGEM2-ES		IPSL-CM5A-LR		Ensemble Mean	
	1981-2000	Change	1981-2000	Change	1981-2000	Change	1981-2000	Change
PET-PT	3.78	0.32	3.06	0.62	4.28	0.78	3.70	0.57
PET-PT-MA	3.78	0.14	3.06	0.32	4.28	0.47	3.70	0.31
Rn	4.09	0.14	2.91	0.36	4.58	0.48	3.86	0.33
PET-EO*	3.50	0.13	3.60	0.32	4.30	0.31	3.80	0.25
NWSAET*	3.90	0.10	3.80	0.20	3.70	0.27	3.80	0.19

* Milly and Dunne (2016) considered 2081-2100 as future period.

Differences in the PET values for the reference period, as computed with the different approaches, do not help to understand the differences in the PET changes. In the case of HadGEM2-ES, for example, where the PET-PT-MA change is larger than the NWSAET change, PET-PT-MA is much smaller than NWSAET in the reference period, while in the case of IPSL-SC5A-LR, both PET-PT-MA during the reference period and PET-PT-MA change are larger than NWSAET and its change, respectively.

275   The fact that changes of PET-PT-MA are very similar to changes of R_n indicates that PT-MA successfully implements the MD proposal that in climate change impact studies, PET should change with R_n only (Table 1). We cannot expect a perfect agreement between the changes of PET-PT-MA computed by WGHM and PET-EO computed in MD, one reason being that the spatial and temporal resolutions are different. Possibly more important are differences in the R_n computation. R_n of WGHM depends not only on the downward shortwave and longwave radiation provided by the GCMs but also on the WGHM estimation

280   of upward shortwave and longwave radiation. In addition, the climate input for PET-PT-MA is bias-adjusted GCM output, while there was no bias adjustment for PET-EO. When comparing R_n change with PET-EO change, which is computed as the change of $0.8R_n$, it can be concluded that R_n change is likely higher (GFDL-ESM2M and HadGEM2-ES) or lower (IPSL-CM5A-LR) in the original GCMs than in the WGHM, where R_n is computed from bias-adjusted GCM output (Table 1).

The temporal development of the two PET variants and R_n between 1901 and 2099 for the non-water-stressed cells and

285   months, as computed by WGHM, does not show, for all three GCMs, appreciable trends in the 20$^{\text{th}}$ century for both PET and R_n (Figure 2 a, c and e). In the 21$^{\text{st}}$ century, the variables increase strongly. Reflecting the PT-MA method (section 2.2), PET-PT and PET-PT-MA only start to deviate from each other after the end of the selected reference period (here 1981-2000) of the climate change study. PET-PT-MA increases less strongly after 2000 than PET-PT and is very similar to R_n. By the end of the 21$^{\text{st}}$ century, the 9-19% PET increase projected for the NWSAET cells/months by the standard PT method for the three

290   GCMs is reduced to 4-11% if PT-MA is applied (compared to the start of the 21$^{\text{st}}$ century). Based on Koster and Mahanama

(2012), MD proposed, with PET-EO, that climate-change driven PET change is not equal to the R_n change but to only 80% of the R_n change. The slopes of the PET-PT to R_n regression lines are much larger than 0.8, ranging from 1.42 to 1.73, while the values for PET-PT-MA are reduced to 0.9-1.02 (see Figure 2 b, d and f). Thus, according to the PT-MA approach implemented in WaterGAP, a larger fraction of the additional net radiation evaporates under non-water-stressed conditions than assumed by MD. This may be explained by the selection of grid cells with a relatively high mean temperature (compare Figure 1). While in case of $\alpha = 1.26$, a slope of 0.80 results from a temperature value of 16 °C according to Eq. 1, slopes of 0.9 and 1.02 result from temperature values of 22 °C and 33 °C, respectively. The interannual variability of the PET time series closely follows the variability of the R_n time series (Fig. 2 a, c and e).
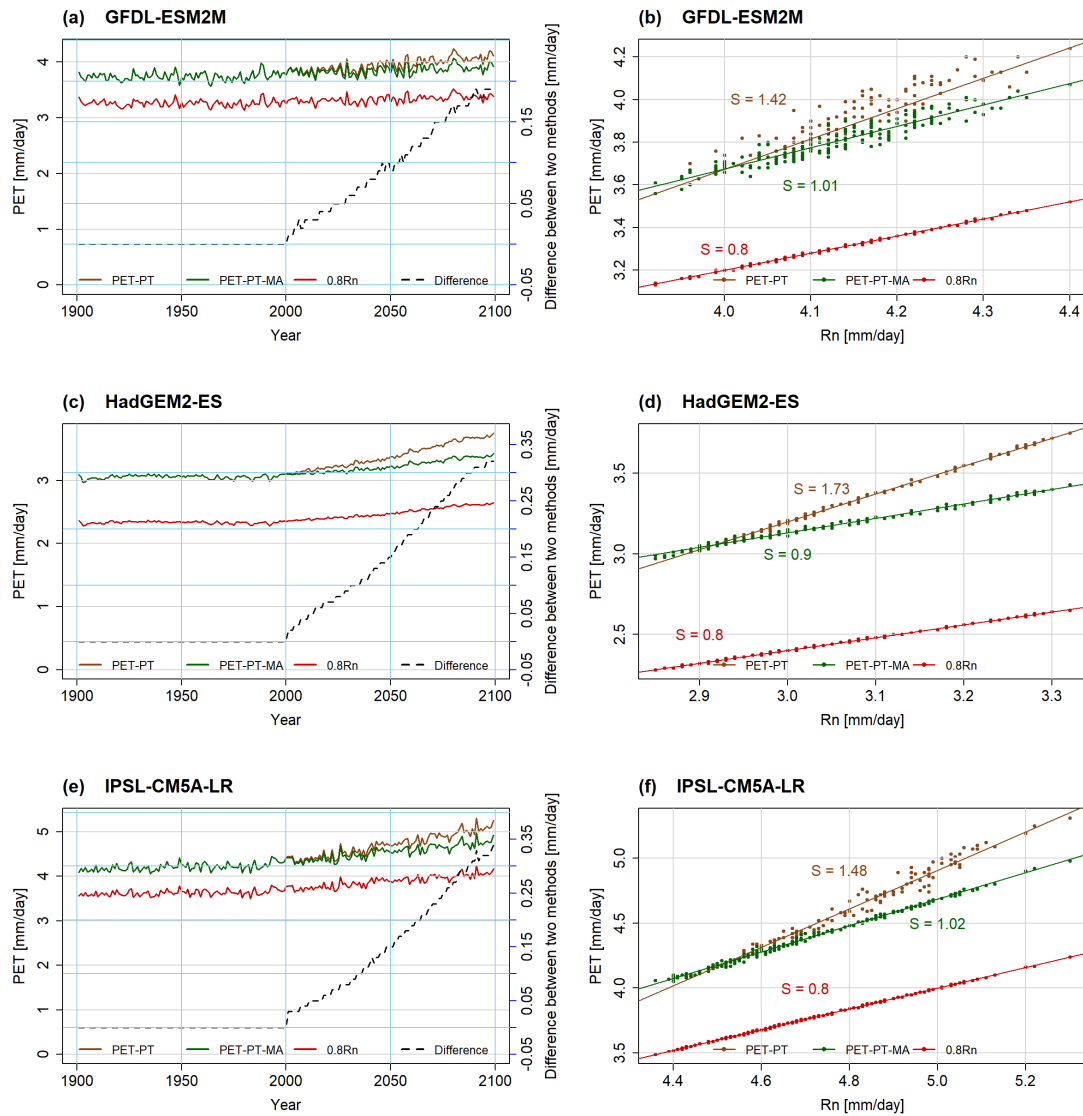
Figure 2. Area-weighted average over non-water-stressed grid-cells/months of PET-PT (brown), PET-PT-MA (green) and $0.8R_n$ (red) for all years of 1901-2099 computed based on daily WGHM output, forced by GCM bias-adjusted climate data under RCP8.5. Time series plots : **(a)**, **(c)** and **(e)**. Scatter plots between R_n and two PET schemes: **(b)**, **(d)** and **(f)**. In the time series plots, the difference between PET-PT and PET-PT-MA (Difference) is given on the secondary y-axis (black dashed line). Please note that for each GCM, different grid cells and months are aggregated. "S" is the slope of the trend line

3.2 Temporal development of PET at two locations

300 At the selected locations A and B (compare Figure 1), PET is projected to increase in the future except for location A in case of the GFDL-ESM2M climate model (Figure 3 and Appendix B1). With GFDL-ESM2M climate input, a slightly decreasing trend for both PET-PT and, more so, PET-PT-MA is observed at location A (as well as in very small areas elsewhere), corresponding to the relatively small temperature increase computed by this GCM for location A and globally (see Figure B1 c).

When comparing the two time series of PET-PT and PET-PT-MA until 2001, there is no difference between the two methods,
305 as intended. From 2001 onwards, the rate of PET increase with the PT-MA method is smaller than with the standard PT method. As a result, the difference between PT and PT-MA is increasing over time and varies among the GCMs and locations (black dashed line in Figure 3 a, b, e and f). Removal of the long-term temperature trend (section 2.2) is successfully done in the PT-MA method (compare Figure S1-S4 in Supplement). Both $T_{modified}$ and PT-MA to R_n ratios do not show a trend in cells A and B in the 21st century but still a GCM-specific inter-annual variability (see Figure 3 c, d, g and h).
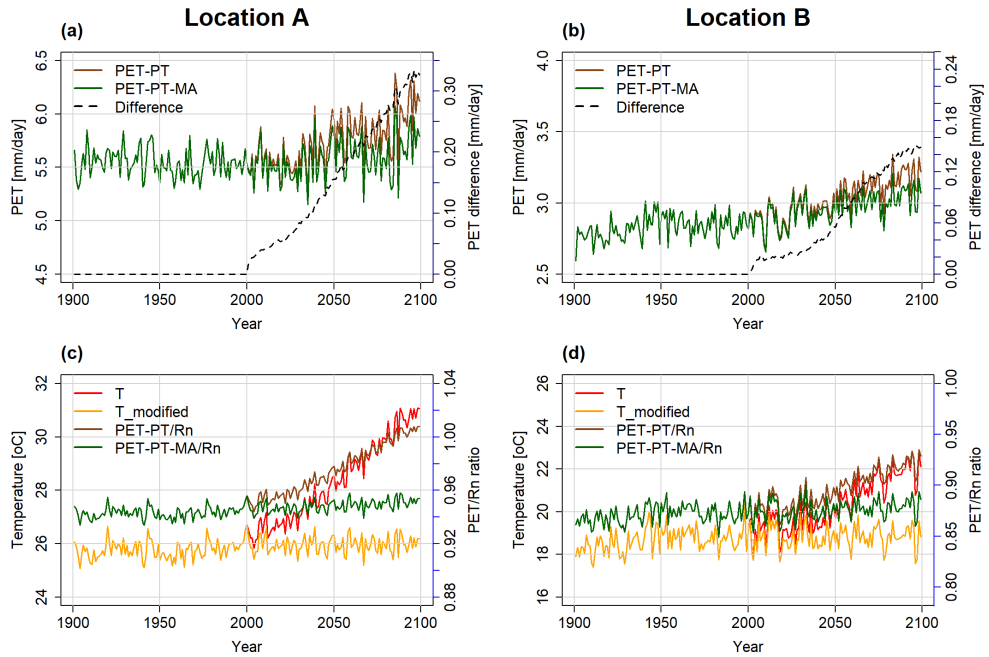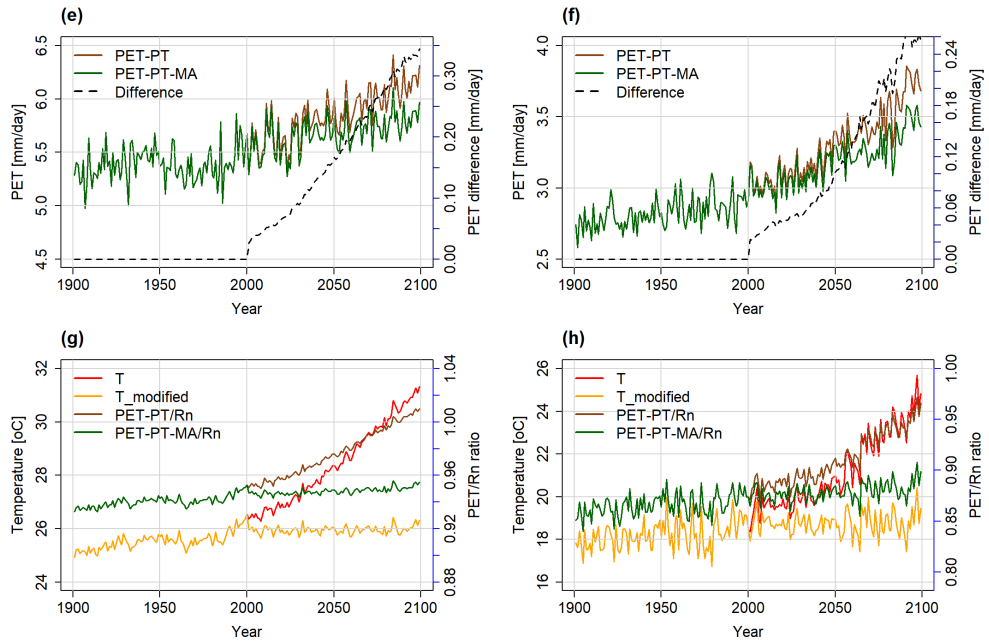
Figure 3. Annual time series of PET and temperature at location A (left) and location B (right), as computed by the WGHM forced by HadGEM2-ES (first two rows) and IPSL-CM5A-LR (last two rows) under RCP8.5 scenario. In (a), (b), (e), and (f), PET with PT (brown) and PET with PT-MA (green) are shown on the primary y-axis. The difference between the two methods (black dashed line) is on the secondary y-axis. In (c), (d), (g), and (h), PET to Rn ratio for PT method (brown) and PT-MA (green) method are on the secondary y-axis, bias-adjusted input temperature (red) and modified temperature (yellow) are on the primary y-axis.

3.3 Spatially heterogeneous future mean PET changes

We compared the projected future absolute changes of mean annual net radiation, standard Priestley-Taylor PET-PT and the modified PET-PT-MA(Figure 4 a - f and appendix B2 a - f). Hereafter in the main text, we present the results corresponding to only two GCMs, while as the results for the other two GCMs are shown in the Appendix.

Change refers to the difference between the period 2080-2099 and the reference period 1981-2000 under RCP8.5. According to the HadGEM2-ES and IPSL-CM5A-LR GCMs, net radiation increases in the future, with the IPSL-CM5A-LR model projecting a stronger increase. The spatial patterns of change in PET-PT-MA (dPET-PT-MA) are very similar to those of net radiation change (dRn), as is intended by the PT-MA approach for computing PET. Increases in PET-PT are much higher than increases in PET-PT-MA. dPET-PT values for the HadGEM2-ES climate model, which has smaller dRn values than the IPSL-CM5A-LR model, are even higher than dPET-PT-MA values for the IPSL-CM5A-LR model.

The Wilcoxon signed-rank test for pair samples was applied to understand the significance of the differences between the time series of annual PET 2080-2099 as computed either by the standard PT or the modified PT-MA approach. The null hypothesis is "The median difference between the values computed by two methods is zero (i.e., no significant difference)"; it is rejected at a 95% level of significance. For all GCM, it was found that there is a significant difference between the PET values computed by the two alternative approaches in all grid cells.
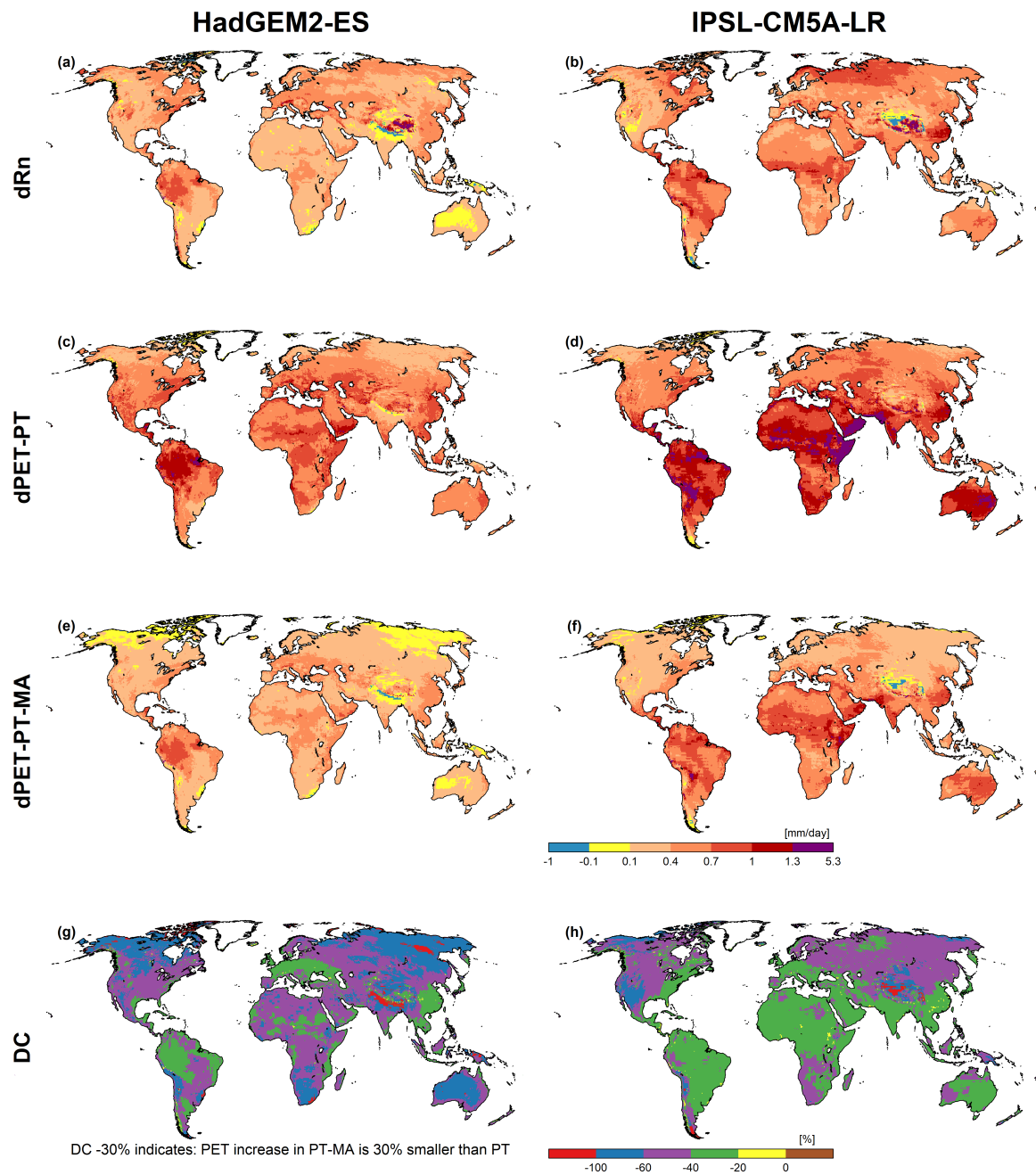
Figure 4. Net radiation (Rn) and PET changes computed by WGHM forced with HadGEM2-ES (left) and IPSL-CM5A-LR (right) climate data between the reference (1981 -2000) and the future (2080 - 2099) period under RCP8.5. (a) and (b) Average Rn change (dRn), (c) and (d) average PET change with PT (dPET-PT), (e) and (f) average PET change with PT-MA (dPET-PT-MA), (g) and (h) percent difference of dPET-PT-MA and dPET-PT (DC). A DC value of, e.g., -50% indicates that PET-PT-MA increases only half as much as PET-PT.

325 The relative difference of change, DC ((dPET-PT-MA-dPET-PT)/dPET-PT) is, as expected, almost everywhere negative, and differs between the GCMs (see g and h of Figure 4 and Figure B2). As explained in section 2.4, the more negative the DC values are, the more the PET increase computed by PT is reduced in case of PT-MA relative to dPET-PT. Relative reductions of PET increase are larger for the HadGEM2-ES GCM than for the IPSL-CM5A-LR GCM, likely due to the lower absolute increases of PET-PT. In some parts of Australia, for example, the new PET approach leads to a reduction of the PET increase

330 by more than 60% and 20-40% in the case of HadGEM2-ES and IPSL-CM5A-LR, respectively. In the few grid cells with red color (DC is less than -100%), PET-PT increases while PET-PT-MA decreases. There are very few grid cells projecting a decrease (brown color areas) in the average PET for both methods (compare g and f of Figure 4 and Figure B2).

3.4 Spatially heterogeneous future mean changes of renewable water resources

While a reduction of projected PET increase due to the new PT-MA method is expected to lead to a higher projected runoff

335 and thus higher renewable water resources (RWR) as compared to the standard approach, the relative increase depends on how much AET is limited by water availability in the soil and open water bodies, and thus depend also on changes in precipitation. If a large part of a grid cell consists of open water bodies, the RWR increase will be small as the PET of open water is not affected by the PT-MA method. In addition, higher runoff and, thus, upstream streamflow may lead to increased AET and, thus, decreased RWR in downstream grid cells due to evaporation from surface water bodies or irrigated fields. Therefore, the

340 effect of the alternative PT-MA approach for computing PET on AET and subsequently RWR is more intricate than its effect solely on PET.

  Changes in precipitation (dP) as projected by GCMs, with different global patterns (Figure 5 a and b), are the main drivers of changes in RWR (dRWR). Increasing P will mostly result in RWR increases unless P increase is small such that increased AET leads to decreased RWR (Figure 5 c-f). For instance, when comparing Figure 5 a and c, precipitation slightly increases in

345 grid cells in the Prairie Provinces of Canada but RWR slightly decreases.

  To assess the significance of the differences between the RWR values computed with the PT and the PT-MA variants at the grid cell level, we employed the Wilcoxon signed-rank test for the time series of annual RWR 2080-2099, similar to the approach used for the PET analysis (section 3.3). For 0.5° grid cells shown in light gray in 5 g and h and B3 e-h, differences resulting from the two alternative variants are insignificant.
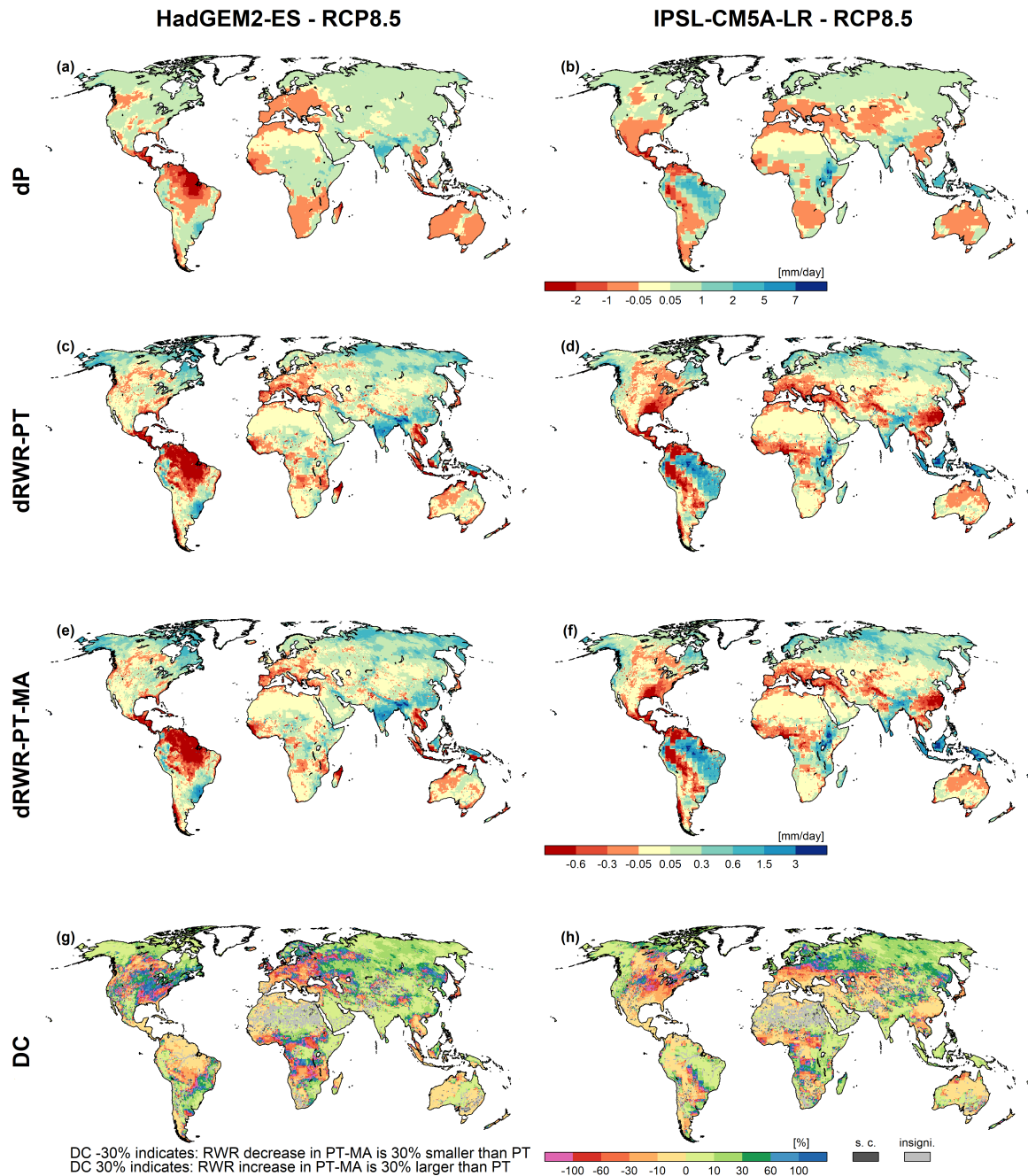
Figure 5. Precipitation (P) and renewable water resources (RWR) changes for HadGEM2-ES (left) and IPSL-CM5A-LR (right) climate data between the reference (1981 -2000) and the future (2080 - 2099) period under RCP8.5. (a) and (b) Absolute change of average precipitation between future and reference periods (dP), (c) and (d) absolute change of average RWR computed with the PT method (dRWR-PT), (e) and (f) absolute change of average RWR computed with the PT-MA method (dRWR-PT-MA), (g) and (h) percent difference of dRWR-PT-MA and dRWR-PT (DC). Cells where dRWR-PT-MA < dRWR-PT are denoted as special cases (s.c.). The cells labeled as "insigni." represents instances where the computed change in RWR using the two methods is not statistically significant As the difference is positive in all other cells, a negative DC value indicates a future decrease of RWR-PT, and a positive DC value a future increase.

350   The spatial patterns of future increases and decreases of dRWR-PT and dRWR-PT-MA appear to be very similar (Figure 5 c-f), and the expected higher future RWR-PT-MA, with dRWR-PT-MA > dRWR-PT, i.e., higher RWR increases and lower RWR decreases, is not well visible. This is due to both the strong impact of dP on dRWR and the strong spatial variability of projected increases and decreases. Stronger RWR decreases by the PT approach as compared to the PT-MA approach are visible only by, e.g., slightly larger red areas in Africa and South America for the IPSL-CM5A-LR GCM. The DC metric (Figure 5

355 g and h) directly quantifies the differences in the future change of RWR between the two PET variants. Negative DC values indicate areas with future decreases of RWR-PT, where the standard PET approach overestimates drying as compared to the modified approach. Positive values indicate areas with future increases in RWR-PT, where the standard method underestimates the increases as compared to the modified approach. For example, the RWR decrease in Central America and the downstream Amazon in the case of HadGEM2-ES is less than 10% smaller in the case of computing PET-PT-MA than in the standard

360 approach (yellow color), while it is up to 30% less in the upstream Amazon, in parts of Central and Western Europe (orange color). Higher values of up to 100% are computed for areas scattered around the world, for all four investigated GCMs (compare g and h of Figure 5 and e and f of Appendix B3). Green to blue indicates areas where the standard approach underestimates the future increase of RWR. Like in the case of decreases, differences of less than 10% dominate but higher discrepancies of up to 30% occur in large areas of Siberia. They can exceed even 100% where projected increases are small. The pink color identifies

365 grid cells where the RWR slightly decreases in case of the standard approach, but slightly increases in case of PT-MA.

    In a small number of grid cells, 3% in case of HadGEM2-ES and 4% in case of IPSL-CM5A-LR, dRWR-PT-MA is smaller than dRWR-PT. These special cases can be explained by the lateral flow processes in combination with lakes and wetland and/or irrigation that are represented in WaterGAP but not in GCMs. RWR in WaterGAP represents the net amount of liquid water that is added or removed per time step by the processes within a grid cell, averaged over the analysis period. It is computed as

370 the difference between the streamflow leaving the cell and the streamflow entering the cell. In most grid cells, outflow is larger than inflow and the positive RWR reflects the part of the precipitation on the land and the surface water bodies of the grid cell that does not evapotranspirate. In the case of lakes and wetlands that are recharged by upstream streamflow as well as in the case of irrigation, outflow from the cell may become smaller than inflow; then RWR is negative. Cells where PT-MA does not lead to higher RWR as compared to PT are mainly cells with lakes and wetlands (including internal sink cells) that receive

375 inflow from upstream cells (e.g., wetlands along the Amazon or the Niger) or cells with a high irrigation water use. In the case of PT-MA, cells with such lakes or wetlands receive a higher inflow from upstream, which leads to an increased surface water area and thus increased evaporation from the surface water body. So even if the grid cell runoff from soil increases in the case of PT-MA, the increased evaporation from the surface water body can dominate and lead to a decreased RWR of the grid cell. The same can happen in grid cells with large irrigation water demand from surface water bodies that can be fulfilled better in

380 case of increased streamflow, leading to increased evapotranspiration in the cell and thus possibly lower RWR. Note that to calculate PET from surface water bodies or irrigation water demand, PET is computed according to the standard PT approach. The PT-MA approach for computing PET change increases total globally averaged RWR more strongly than the standard PT approach, for all four GCMs (Table 2). However, while the increase of RWR is more than 100 % larger for HadGEM2-ES, it is only 50 % larger for GFDL-ESM2M and even less for the other two GCMs.

Table 2. Projected change [%] of renewable water resources between 1981-2000 and 2080-2099 for the RCP 8.5 emissions scenario in the case of computing PET with the modified (PT-MA) and the standard (PT) approach, averaged over the global land area (excluding Antarctica and Greenland).

GCM	PT-MA [%]	Standard PT [%]
GFDL-ESM2M	6.96	4.67
HadGEM2-ES	7.39	3.65
IPSL-CM5A-LR	13.93	10.46
MIROC5	16.61	13.28

3.5   Analyses for other RCPs

MD analyzed only GCM simulations that implemented the high emissions scenario RCP8.5. As for individual GCMs, increases in CO_2 concentrations and temperature roughly correlate (Humlum et al., 2013). We also applied the PT-MA approach for other RCPs. In RCP2.6, projected precipitation changes are smaller than under RCP8.5 (compare Figure 6 a, b with Figure 5 a, b). For each GCM, spatial DC patterns for RWR under RCP2.6 are rather similar to those under RCP8.5 (compare Figure 6 c, d with Figure 5 g, h). Under RCP2.6, the differences between dRWR computed by the two approaches are only slightly smaller than under RCP8.5 (Figure 7b).
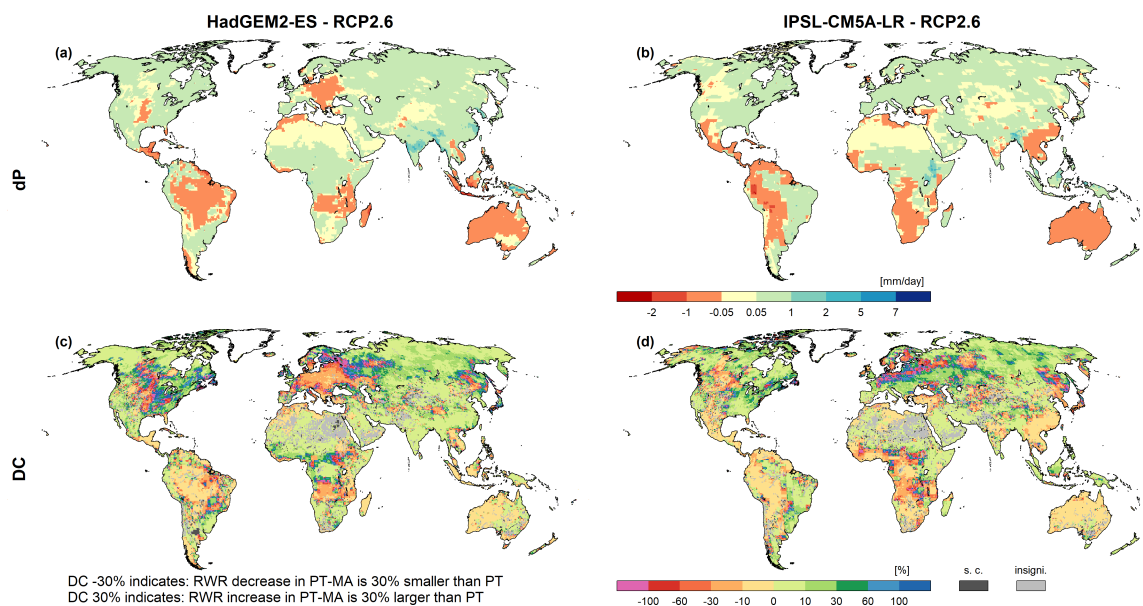
Figure 6. Projected RWR changes for HadGEM2-ES (left) and IPSL-CM5A-LR (right) climate data between the reference (1981 -2000) and the future (2080 - 2099) period under RCP2.6. (a) and (b) Absolute change of average precipitation (dP), (c) and (d) percent difference of dRWR-PT-MA and dRWR-PT (DC). Cells where dRWR-PT-MA < dRWR-PT are denoted as special cases (s.c.). The cells labeled as "insigni." represent instances where the computed change in RWR using the two methods is not statistically significant)
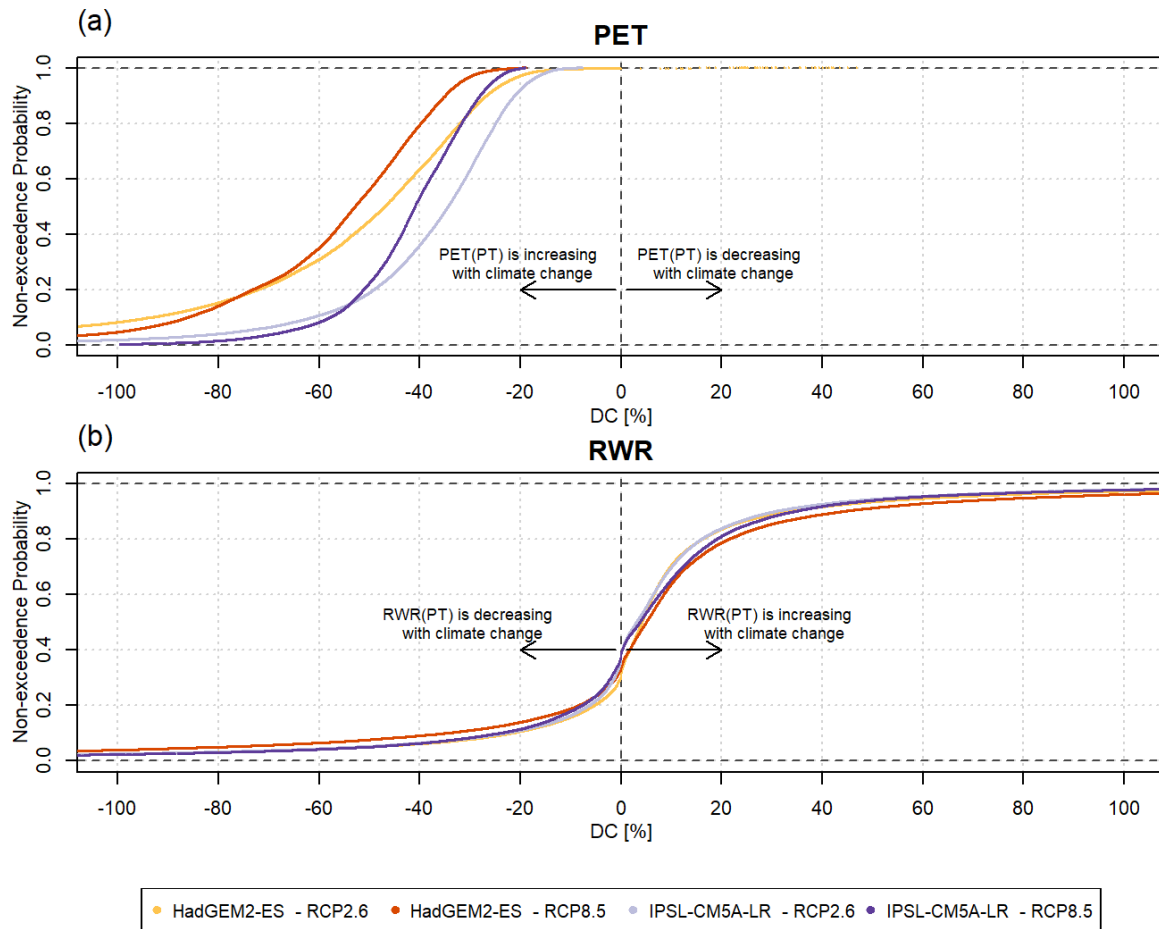
Figure 7. The cumulative probability distributions of DC, i.e. the relative difference between the change computed with the PT-MA approach and the change computed with the standard PT method for (a) PET and (b) RWR. The results are based on WGHM output forced with HadGEM2-ES and IPSL-CM5A-LR climate models under RCP8.5 and RCP2.6. The lighter colors correspond to RCP2.6 and the darker colors to RCP8.5. Antarctica and Greenland are excluded

Global DC-RWR distributions for the two GCMs are similar, while HadGEM2-ES leading to a slightly higher impact of PT-MA than IPSL-CM5A-LR (Figure 7b). Regardless of the RCP or the GCM, most grid cells have DC between -10% and 20%. About 60% of the DC values are positive, indicating that also about 60% of all grid cells indicate an increase in RWR with both PT and PT-MA methods in the future period compared to the historical period.

In contrast to DC-RWR, the cumulative probability distributions of DC-PET differ appreciably between RCP8.5 and RCP2.6 and between the GCMs (Figure 7a). The reduction of PET increase in the PT-MA approach as compared to the standard PT method is smaller in the case of RCP2.6. Reductions are larger for HadGEM2-ES (-60% - -30% under RCP8.5) than for IPSL-CM5A-LR (-50% - -20%) which is likely due to the higher temperature increase in the northern hemisphere and thus the stronger temperature reduction in case of HadGEM2-ES (see Appendix A1).

Averaged globally and over all four GCMs, the three land water balance components precipitation, AET and streamflow into oceans and internal sinks (i.e., RWR) are projected to increase in the future due to climate change, with increases increasing with greenhouse gas emissions (Figure 8). This is also the case for PET. With the proposed PT-MA approach global-scale PET increases of 6.9-18.9% (four RCPs) for the standard approach are strongly decreased to 4.4-10.7%. The effects of PT-MA on global changes of AET and RWR are of similar magnitude. While AET increases of 3.1-5.7% are decreased to 2.4-3.7% (Figure 8c), RWR increases of 1.6-8.0% are increased to 2.8-11.2% (Figure 8c).
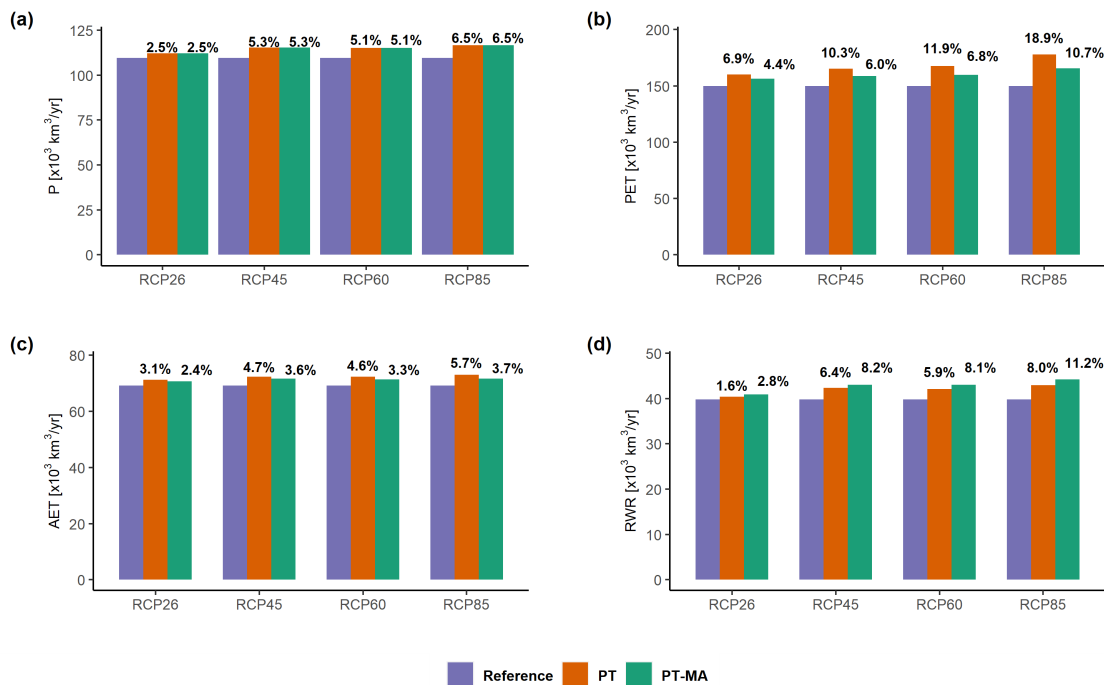
Figure 8. Effect of the PT-MA and the standard PT approach for computing PET on the water balance of the global land area. Ensemble mean across the four GCMs of globally aggregated (a) precipitation (P), (b) potential evapotranspiration (PET), (c) actual evapotranspiration (AET) and (d) streamflow into oceans and inland sinks (renewable water resources (RWR)) according to WGHM for the periods 1981-2000 and 2080-2099. The percentage values at the top of the bars indicate the relative change of the variable compared to the reference period value. Antarctica and Greenland are excluded.

4   Discussion

4.1   Sources of uncertainty of the projected PET change

To understand the magnitude of the uncertainty in PET changes related to the choice of PET approach (PT-MA or PT), we compared it to the magnitude of the uncertainty related to the choice of GCM. The computed absolute change of PET between

the periods 1981-2000 and 2080-2099 computed by WGHM that was forced by four GCMs under the RCP8.5 scenario is used in this uncertainty analysis. The uncertainty range that stems from the four GCMs (GCM_{range}) is computed as

$$GCM_{range} = \frac{\sum_{m=1}^{2} |\Delta PET_{GCM(m, highest\,out\,of\,four\,GCMs)} - \Delta PET_{GCM(m, lowest\,out\,of\,four\,GCMs)}|}{2} \tag{6}$$

where m = 1-2 corresponds to the two PET estimation methods (PT-MA and PT).

The uncertainty that originates from the two different PET computation approaches ($Approach_{range}$) is calculated as

$$Approach_{range} = \frac{\sum_{n=1}^{4} |\Delta PET_{GCM(n, PET-PT-MA)} - \Delta PET_{GCM(n, PET-PT)}|}{4} \tag{7}$$

where n = 1-4 corresponds to the four GCMs applied in this study.

Figure 9 shows the GCM uncertainty and the approach uncertainty as percentages of the total uncertainty, i.e., the sum of GCM_{range} and $Approach_{range}$. In most regions, GCM uncertainty is the dominant source of uncertainty. Globally averaged, the choice of GCM is responsible for 62% of the total uncertainty. PET approach uncertainty is dominant in large parts of North America, Central Asia and Africa. In most regions of the globe, the PET approach causes at least 30% of the total uncertainty. The minimum contribution of GCM uncertainty is 16%, while the minimum contribution of the PET approach is zero, due to grid cells that only consist of open water. This analysis shows that the choice of the PET approach has a significant impact on computed PET change when compared to the impact of the well-known large uncertainty of climate projections that is caused by applying different climate models.

Using an ensemble of model outputs derived from two variants of PET calculation methods is crucial for capturing the full range of uncertainty associated with vegetation response in climate change impact assessment studies.
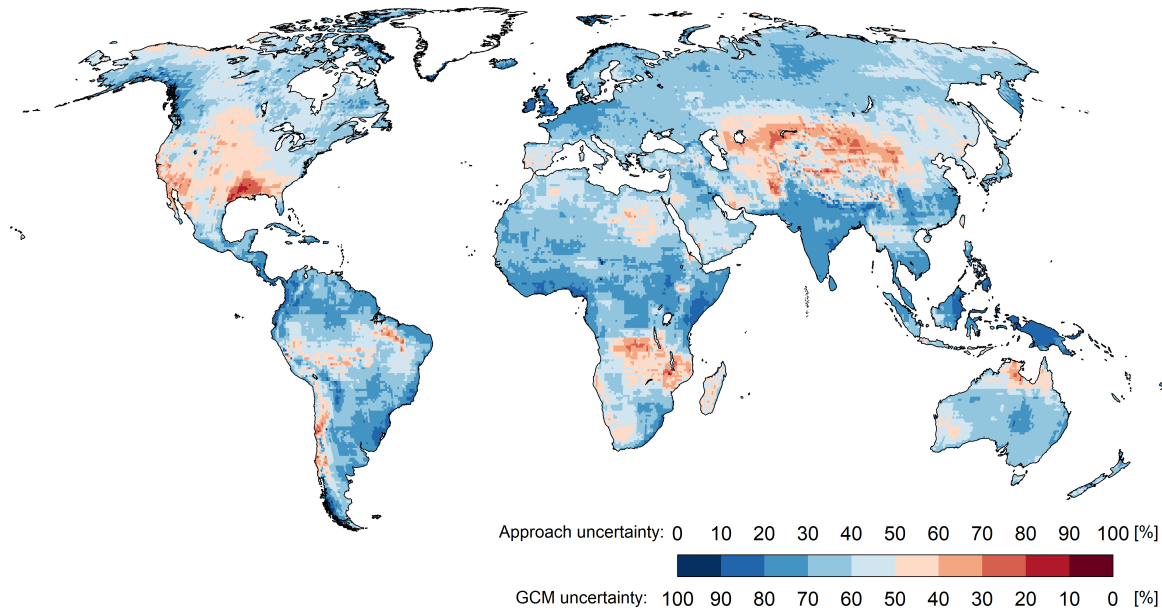
Figure 9. Source of uncertainty of the projected PET change based on WGHM outputs that is derived from two PET calculation approaches and four GCMs under the RCP8.5 scenario.

4.2 Caveats and applicability

This paper presents a method that enables hydrological models that compute PET as a function of net radiation and temperature to simulate future PET in a way that PET changes with net radiation only. The estimation of PET changes by the PT-MA approach is obviously also affected by uncertainties in the estimation of net radiation changes. In our study, the latter were computed from changes in downward shortwave and longwave radiation as provided by GCMs and changes in upward shortwave and longwave radiation as computed by WGHM.

That PET change under climate change can be approximated by the change in net radiation was derived by MD, who analyzed changes in actual evapotranspiration as computed by a number of GCMs for only a small number of grid cells/months in tropical climates that do not experience water stress. This means that our approach assumes that vegetation response to climate change around the globe is similar to that of tropical vegetation under a given change in net radiation and temperature. Such a spatially homogeneous response cannot be expected. In particular, in the PT-MA approach, PET is always less compared to the standard PT approach. While different DGVMs tend to simulate this behavior for many regions, the total balancing of the warming effects seen in the non-water-stressed tropical grid cells of the GCMs might not be true for other biomes. How much if at all, in a certain biome, the physiological effect (closure of stomata) dominates over other effects such as the structural effect (increase in biomass/leaf area) and biome shifts, is, however, now well-known. Different from the PET change computed with PT-MA, DGVMs predict a spatially heterogeneous vegetation response to climate change, which leads to a relative increase of PET in some regions. Unfortunately, there are considerable differences between the simulation results of different DGVMs, and there is no agreement on where the vegetation response leads to increased or decreased PET (Reinecke et al., 2021).

Therefore, we propose that two alternative model variants are routinely used in climate change impact studies with hydrological models that do not simulate vegetation processes. In one variant, the standard PT approach is selected for computing PET of the vegetated land surface, and in the other variant, the PET-MA approach is selected. In this way, the uncertainty of PET due to the vegetation response can be taken into account. The disadvantage is that the number of model runs is doubled, as the two variants need to be run for each alternative GCM and RCP that determines the model input. While application of only the standard PET approach has been shown to overestimate future drying, i.e. overestimate the decrease of water resources, or underestimate the increase of future water resources (Milly and Dunne, 2016; Yang et al., 2019), application of the newly proposed PT-MA might overestimate the impact of stomatal closure due to increasing atmospheric CO_2 concentrations in some regions, and PET estimation with the standard PT approach is more appropriate, as indicated by the studies of Milly and Dunne (2017) and Reinecke et al. (2021). Therefore, without further knowledge on the local vegetation-specific response to climate change, it is best to simulate hydrological changes under climate change with an ensemble approach, i.e. running the model with both PET variants.

The PT-MA approach can be used if PET is calculated as a function of net radiation and temperature only, as in the PT equation. To our knowledge, the PT equation is the sole PET equation that computes PET as a function of temperature and net radiation only. If a hydrological model uses a Penman-Monteith type equation to estimate PET, the approach of Yang et al. (2019), where stomatal conductance is adjusted, is suitable for computing PET changes under climate change. However, the application of the Penman-Monteith type equation demands more input data than the PT equation, requiring data on humidity and wind speed, which is why many local and regional models do not use this type of PET equation (Lu et al., 2005; Koedyk and Kingston, 2016). Furthermore, if hydrological models use a Penman-Monteith type equation, assessment of climate change impacts requires down-scaled humidity and wind speed projections, which are quite uncertain (Randall et al., 2007). This limits the usability of Penman-Monteith type equations.

Therefore, the use of the PT equation is a good option for simulating PET in hydrological models, and these models can then implement the PT-MA approach for computing scenarios of future hydrological hazards due to climate change. Widely used basin-scale hydrological models such as HBV/HBV-Light or SWAT are well suited for applying the PT-MA approach as they use daily PET time series as input that can be computed using the PT equation (Rajib et al., 2018; Koedyk and Kingston, 2016). The user can pre-compute time series of daily PET for the study area using the PT-MA approach and then use them as a model input. Other radiation-based PET equations such as Jensen-Haise, Makkink and Turc that do not consider net radiation but components of net radiation such as shortwave radiation should not be used for climate change studies (if the hydrological model does not simulate the interactions between the atmosphere and land surface), and these methods should not be adapted according to the proposed approach.

5 Conclusions

Application of standard equations for estimating PET in hydrological models, such as Penman-Monteith or Priestley-Taylor type equations, was shown to lead to an overestimation of future PET for many regions of the globe, mainly due to neglecting

the impact of vegetation responses to changing atmospheric CO_2 concentrations and climate. As a result, future decreases

480 in renewable water resources may be overestimated, or future increases underestimated. With the proposed method for PET computation in hydrological models that do not simulate vegetation processes, future PET changes occur only in response to changes in net radiation. This was shown by Milly and Dunne (2016) to be consistent with the changes of non-water-stressed AET computed by a number of GCMs, which simulate the complex interaction between soil, vegetation and the atmosphere, including the vegetation response to changing atmospheric CO_2 concentrations and climate. However, at least in

485 some regions, which however cannot be defined with certainty, the vegetation response is such that standard PET computation leads to more reliable results. To take into account the uncertainty of future PET changes due to the vegetation response, we propose that climate change impact studies by hydrological models should always include runs with two model variants, one with the standard PET approach and the other with a modified approach. The two variants, which approximately represent the uncertainty bounds related to the vegetation response to climate change, improve the estimation of the total uncertainty of

490 future hydrological changes.

The modified approach for computing Priestley-Taylor PET under climate change, PT-MA, is suitable if PET is computed according to the Priestley-Taylor equation, while the approach of Yang et al. (2019) should be applied if one of the more data-intensive Penman-Monteith type equations can be used in the hydrological model. In the latter case, we propose to check whether the thus computed change in PET is approximately equal to the change in net energy input to the land surface.

495 Implementation of the PT-MA approach is very simple, and the reference period for the climate change study can be easily adjusted.

When implementing the PT-MA approach in the global hydrological model WaterGAP 2.2d, the projected increase of global renewable water resources and streamflow into oceans is enhanced. The PT-MA approach leads to reduced drying (smaller decrease of renewable water resources) or increased wetting (stronger increase of renewable water resources) as compared

500 to the standard Priestley-Taylor equation for almost all grid cells, for both RCP2.6 and RCP8.5. Only in a few cells (3-4%), consideration of the vegetation response leads to lower renewable water resources as compared to the standard approach; these are grid cells where increased lateral inflow leads to an increased extent of surface water bodies and thus evapotranspiration or where large irrigation demands can be fulfilled better by increased streamflow.

Regarding future research, we propose to compare, at the global scale, changes of PET and renewable water resources as

505 computed by WaterGAP using the PT-MA approach to changes computed by various DGVMs (or land surface models that simulate the vegetation response), to better understand for which climates and vegetation types the PT-MA approach is too simplistic and cannot capture, for example, that the vegetation response leads to an increase of PET. In addition, we propose to compare PT-MA derived PET projections to those obtained by applying the Yang et al. (2019) approach to Penman-Monteith type equations, to determine whether both approaches lead to similar PET changes for any given GCM climate scenario.

510   *Data availability.*   Gridded WGHM model outputs of monthly PET (with PT and PT-MA methods) and RWR (with PT and PT-MA methods)

for 1981-2099, forced by the bias-adjusted output of four GCMs, each implementing RCP2.6 and RCP8.5 are available at https://zenodo.org/record/659313

(DOI: 10.5281/zenodo.6593136).

Appendix A: Temperature reduction factor
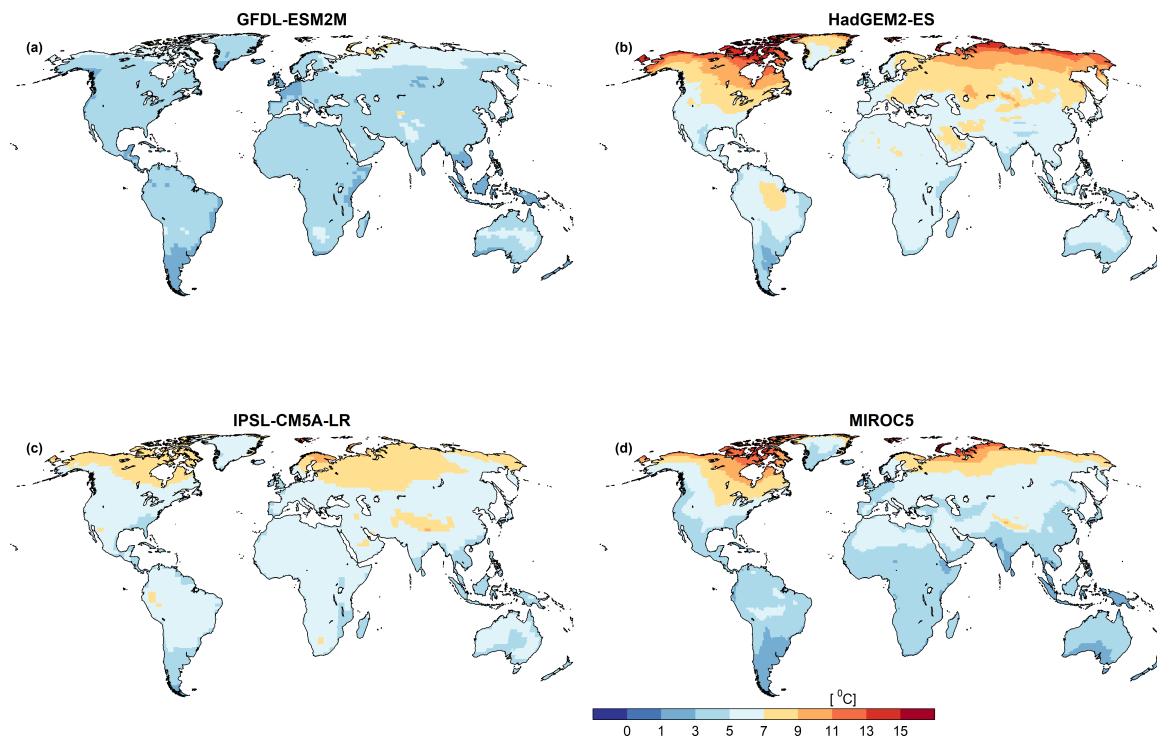
Figure A1. Spatial distribution of the temperature reduction factor for the year 2089

Appendix B: Additional results
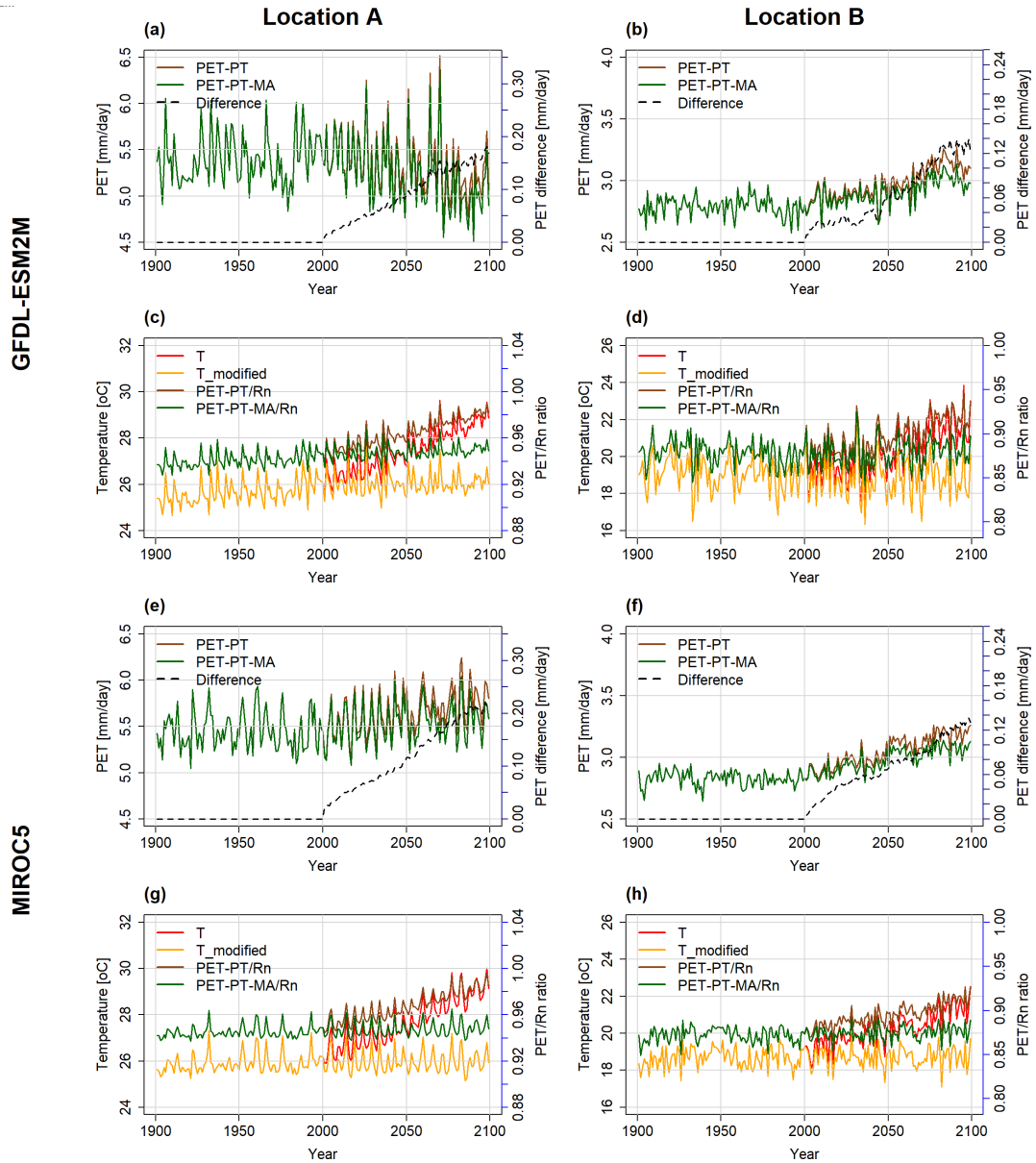
Figure B1. Temporal development of the PET and temperature variables at location A (left) and location B (right), as computed by the WGHM forced by GFDL-ESM2M (first two rows) and MIROC5 (last two rows) under RCP8.5 scenario. In (a), (b), (e), and (f), PET with PT (brown) and PET with PT-MA (green) are on the primary y-axis, the difference between the two methods (black dashed line) is on the secondary y-axis. In (c), (d), (g), and (h), PET to Rn ratio with PT method (brown) and PT-MA (green) method are on the secondary y-axis, bias-adjusted input temperature (red) and modified temperature (yellow) is on the primary y-axis.

GFDL-ESM2M **MIROC5**

DC -30% indicates: PET increase in PT-MA is 30% smaller than PT
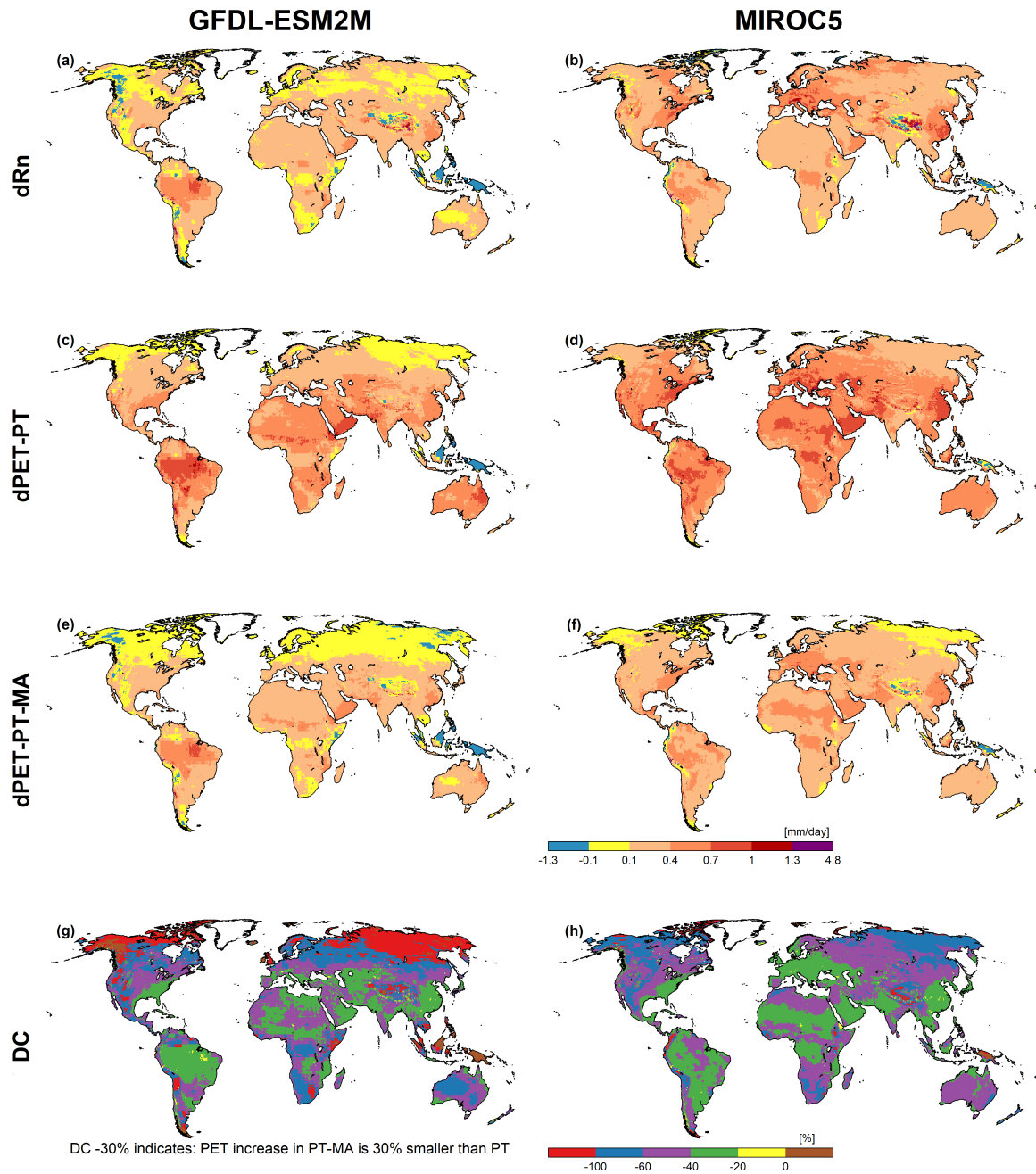
Figure B2. Net radiation (Rn) and PET computed by WGHM forced with GFDL-ESM2M (left) and MIROC5 (right) climate data between the reference (1981-2000) and the future (2080-2099) period under RCP8.5. (a) and (b) Average Rn change (dRn), (c) and (d) average PET change with PT (dPET-PT), (e) and (f) average PET change with PT-MA (dPET-PT-MA), (g) and (h) percent difference of dPET-PT-MA and dPET-PT (DC). A DC value of e.g., -50% indicates that PET-PT-MA increases only half as much as PET-PT.
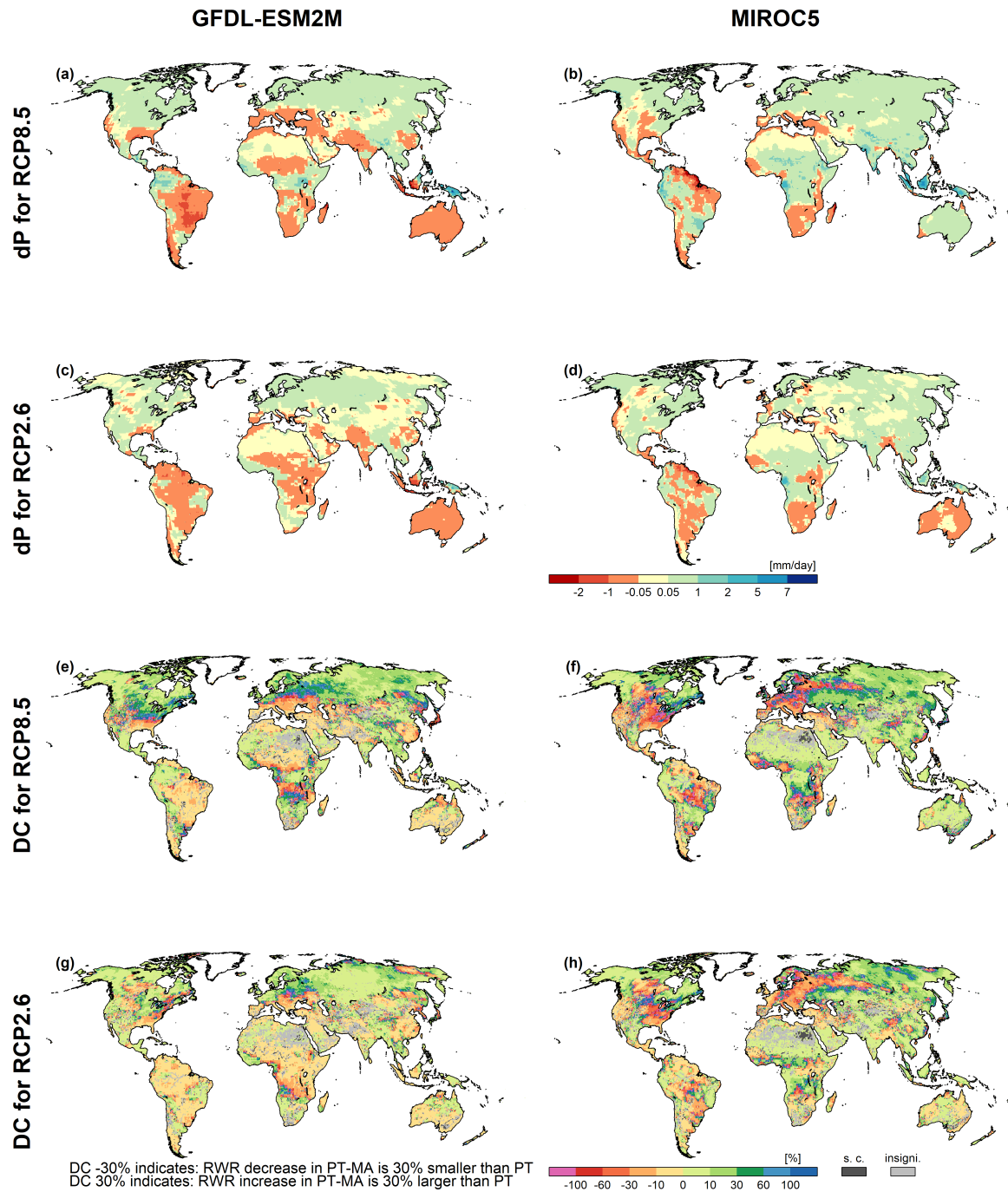
Figure B3. Precipitation (P) and percent difference of dRWR-PT-MA and dRWR-PT (DC) changes for GFDL-ESM2M (left) and MIROC5 (right) climate data between the reference (1981-2000) and the future (2080-2099) periods under RCP8.5 and RCP2.6. Absolute change of average precipitation between future and reference periods (dP) (a), (b) for RCP8.5 and (c), (d) for RCP2.6 and DC for RCP8.5 (e), (f) and for RCP2.6 (g), (h). Cells where dRWR-PT-MA < dRWR-PT are denoted as special cases (s.c.). The cells labeled as "insigni." represent instances where the computed change in RWR using the two methods is not statistically significant.

Author contributions. The concept for the approach and the paper were developed by PD, supported by AP. AP conducted the literature review, implemented the PET-PT-MA method in the WaterGAP model code, conducted the simulations, analyzed the model outputs, prepared the tables and figures, and drafted the initial paper, with substantial revisions by PD.

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

Acknowledgements. We would like to thank the authors of Milly and Dunne (2016) and Yang et al. (2019) for supporting this study by providing the locations of the non-water-stressed grid cells used in their respective studies. We also thank Denise Cáceres from the Institute of Physical Geography, Goethe University Frankfurt for her constructive comments to improve the quality of this paper.

Financial support. This research has been carried out in the framework of Co-development of Methods to utilize uncertain multi-model based Information on freshwater related hazards of Climate Change (CO-MICC) project, funded by the German Federal Ministry of Education and Research (BMBF) (grant no. 690462 — ERA4CS — H2020-SC5-2014-2015/H2020-SC5-2015-one-stage 51) in the framework of European Research Area for Climate Services (ERA4CS).

References

Atwell, B. J., Kriedemann, P. E., & Turnbull, C. G.: Plants in action: adaptation in nature, performance in cultivation: Chapter2, Macmillan Education AU, Australia, https://books.google.de/books?id=chWs4ewSzpEC&pg=PA21-IA7&dq=C3+plants&hl=en&sa=X&ved=2ahUKEwjC5o-Z1onuAhXDKewKHeliDR0Q6AEwAHoECAYQAg#v=onepage&q=C3%20plants&f=false, 1999.

530 Berg, A. and Sheffield, J.: Evapotranspiration Partitioning in CMIP5 Models Uncertainties and Future Projections, Journal of Climate, 32, https://doi.org/10.1175/JCLI-D-18-0583.s1, 2019.

Cramer, W., Bondeau, A., Woodward, F. I., Prentice, I. C., Betts, R. A., Brovkin, V., Cox, P. M., Fisher, V., Foley, J. A., Friend, A. D., et al.: Global response of terrestrial ecosystem structure and function to CO2 and climate change: results from six dynamic global vegetation models, Global change biology, 7, 357–373, 2001.

535 Davie, J., Falloon, P., Kahana, R., Dankers, R., Betts, R., Portmann, F., Wisser, D., Clark, D., Ito, A., Masaki, Y., et al.: Comparing projections of future changes in runoff from hydrological and biome models in ISI-MIP, Earth System Dynamics, 4, 359–374, 2013.

Frieler, K., Lange, S., Piontek, F., Reyer, C. P., Schewe, J., Warszawski, L., Zhao, F., Chini, L., Denvil, S., Emanuel, K., et al.: Assessing the impacts of 1.5 C global warming–simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b), Geoscientific Model Development, 10, 4321–4345, 2017.

540 Gerten, D., Betts, R., Döll, P.: Cross-chapter box on the active role of vegetation in altering water flows under climate change: Climate change, https://doi.org/10.1029/2009JD013408, 2014.

Humlum, O., Stordahl, K., and Solheim, J.-E.: The phase relation between atmospheric carbon dioxide and global temperature, Global and Planetary Change, 100, 51–69, 2013.

Jones, C., Hughes, J., Bellouin, N., Hardiman, S., Jones, G., Knight, J., Liddicoat, S., O'connor, F., Andres, R. J., Bell, C., et al.: The
545 HadGEM2-ES implementation of CMIP5 centennial simulations, Geoscientific Model Development, 4, 543–570, 2011.

Kingston, D. G., Todd, M. C., Taylor, R. G., Thompson, J. R., and Arnell, N. W.: Uncertainty in the estimation of potential evapotranspiration under climate change, Geophysical Research Letters, 36, 2009.

Koedyk, L. and Kingston, D.: Potential evapotranspiration method influence on climate change impacts on river flow: a mid-latitude case study, Hydrology Research, 47, 951–963, 2016.

550 Koster, R. D. and Mahanama, S. P.: Land surface controls on hydroclimatic means and variability, Journal of Hydrometeorology, 13, 1604–1620, 2012.

Lange, S.: EartH2Observe, WFDEI and ERA-Interim data Merged and Bias-corrected for ISIMIP (EWEMBI), GFZ Data Serv, 2016.

Lu, J., Sun, G., McNulty, S. G., and Amatya, D. M.: A Comparison of Six Potential Evapotranspiration Methods for Regional Use in the Southeastern United States 1, JAWRA Journal of the American Water Resources Association, 41, 621–633, 2005.

555 Milly, P. and Dunne, K. A.: A hydrologic drying bias in water-resource impact analyses of anthropogenic climate change, JAWRA Journal of the American Water Resources Association, 53, 822–838, 2017.

Milly, P. C. D. and Dunne, K. A.: Potential evapotranspiration and continental drying, Nature Climate Change, 6, 946–949, https://doi.org/10.1038/nclimate3046, 2016.

Müller Schmied, H., Eisner, S., Franz, D., Wattenbach, M., Portmann, F. T., Flörke, M., and Döll, P.: Sensitivity of simulated global-scale
560 freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration, Hydrology and Earth System Sciences, 18, 3511–3538, 2014.

Müller Schmied, H., Müller, R., Sanchez-Lorenzo, A., Ahrens, B., and Wild, M.: Evaluation of radiation components in a global freshwater model with station-based observations, Water, 8, 450, 2016.

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., Peiris, T. A., Popat, E., Portmann, F. T., Reinecke, R., et al.: The global water resources and use model WaterGAP v2. 2d: Model description and evaluation, Geoscientific Model Development, 14, 1037–1079, 2021.

Purcell, C., Batke, S., Yiotis, C., Caballero, R., Soh, W., Murray, M., and McElwain, J. C.: Increasing stomatal conductance in response to rising atmospheric CO2, Annals of botany, 121, 1137–1149, 2018.

Rajib, A., Merwade, V., and Yu, Z.: Rationale and efficacy of assimilating remotely sensed potential evapotranspiration for reduced uncertainty of hydrologic models, Water Resources Research, 54, 4615–4637, 2018.

Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., et al.: Climate models and their evaluation, in: Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC (FAR), pp. 589–662, Cambridge University Press, 2007.

Reinecke, R., Müller Schmied, H., Trautmann, T., Andersen, L. S., Burek, P., Flörke, M., Gosling, S. N., Grillakis, M., Hanasaki, N., Koutroulis, A., et al.: Uncertainty of simulated groundwater recharge at different global warming levels: A global-scale multi-model ensemble study, Hydrology and Earth System Sciences, 25, 787–810, 2021.

Sepulchre, P., Caubel, A., Ladant, J.-B., Bopp, L., Boucher, O., Braconnot, P., Brockmann, P., Cozic, A., Donnadieu, Y., Dufresne, J.-L., et al.: IPSL-CM5A2–an Earth system model designed for multi-millennial climate simulations, Geoscientific Model Development, 13, 3011–3053, 2020.

Shuttleworth, W.: Evaporation, in: Handbook of Hydrology, edited by Maidment, D., pp. 1–4, McGraw-Hill, New York, 1993.

Telteu, C.-E., Müller Schmied, H., Thiery, W., Leng, G., Burek, P., Liu, X., Boulange, J. E. S., Seaby Andersen, L., Grillakis, M., Gosling, S. N., et al.: Understanding each other's models: a standard representation of global water models to support improvement, intercomparison, and communication, Geoscientific Model Development, 2021.

Thonicke, K., Venevsky, S., Sitch, S., and Cramer, W.: The role of fire disturbance for global vegetation dynamics: coupling fire into a Dynamic Global Vegetation Model, Global Ecology and Biogeography, 10, 661–677, 2001.

Vörösmarty, C. J., Federer, C. A., and Schloss, A. L.: Potential evaporation functions compared on US watersheds: Possible implications for global-scale water balance and terrestrial ecosystem modeling, Journal of Hydrology, 207, 147–169, 1998.

Watanabe, M., Suzuki, T., O'ishi, R., Komuro, Y., Watanabe, S., Emori, S., Takemura, T., Chikira, M., Ogura, T., Sekiguchi, M., et al.: Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity, Journal of Climate, 23, 6312–6335, 2010.

Weiß, M. and Menzel, L.: A global comparison of four potential evapotranspiration equations and their relevance to stream flow modelling in semi-arid environments, Advances in Geosciences, 18, 15–23, 2008.

Yang, Y., Roderick, M. L., Zhang, S., McVicar, T. R., and Donohue, R. J.: Hydrologic implications of vegetation response to elevated CO2 in climate projections, Nature Climate Change, 9, 44–48, https://doi.org/10.1038/s41558-018-0361-0, 2019.

Zhao, L., Xia, J., Xu, C.-y., Wang, Z., Sobkowiak, L., and Long, C.: Evapotranspiration estimation methods in hydrological models, Journal of Geographical Sciences, 23, 359–369, 2013.