


# Empirical Estimate of Forestation-Induced Precipitation Changes in Europe

## Journal Article

### Author(s):

Meier, Ronny ; Schwaab, Jonas; Seneviratne, Sonia I. ; Sprenger, Michael; Lewis, Elizabeth; Davin, Edouard Léopold 

### Publication date:

2021-07

### Permanent link:

<https://doi.org/10.3929/ethz-b-000493339>

### Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

### Originally published in:

Nature Geoscience 14, <https://doi.org/10.1038/s41561-021-00773-6>

### Funding acknowledgement:

172715 - CLimate IMPacts of Utilizing Land in Switzerland and Europe (CLIMPULSE) (SNF)

# Empirical Estimate of Forestation-Induced Precipitation Changes in Europe

Accepted preprint version of <https://doi.org/10.1038/s41561-021-00773-6><sup>1</sup>

Ronny Meier<sup>1</sup>, Jonas Schwaab<sup>1</sup>, Sonia I. Seneviratne<sup>1</sup>, Michael Sprenger<sup>1</sup>, Elizabeth Lewis<sup>2</sup>, and Edouard L. Davin<sup>1</sup>

<sup>1</sup>ETH Zurich, Institute for Atmospheric and Climate Science, Universitaetstrasse 16, 8092 Zurich, Switzerland

<sup>2</sup>Newcastle University, School of Engineering, Newcastle upon Tyne, NE1 7RU, United Kingdom

Land cover changes can affect the climate by altering the water and energy balance of the land surface. Numerous modeling studies have indicated that alterations at the land surface can result in considerable changes in precipitation. Yet, land cover-induced precipitation changes remain largely unconstrained by observations. Here we use an observation-based continental-scale statistical model to show that forestation of rainfed agricultural land in Europe triggers substantial changes in precipitation. Locally, we find an increase in precipitation following forestation, in particular in winter, which is supported by a paired rain gauge analysis. In addition, forests are estimated to increase downwind precipitation in most regions during summer. In contrast, the downwind effect in winter is positive in coastal areas but near-neutral and negative in Continental and Northern Europe, respectively. The combined local and non-local effects of a realistic reforestation scenario, constrained by sustainability safeguards, are estimated to increase summertime precipitation by  $7.6 \pm 6.7\%$  on average over Europe ( $0.13 \pm 0.11$  mm/day), potentially offsetting a substantial part of the projected precipitation decrease from climate change. We therefore conclude that land cover-induced alterations of precipitation should be considered when developing land management strategies for climate change adaptation and mitigation.

Observations show that land cover changes (LCCs) alter local temperatures considerably through biogeophysical processes<sup>2-5</sup>. Such observational constraints have enabled the evaluation of the land cover (LC)-temperature coupling in climate models<sup>6-10</sup>. Alongside temperature, water availability is amongst the most important climate drivers for life on Earth. The high precipitation amounts over tropical rainforests are partly sustained by the presence of the forests themselves<sup>11,12</sup>. On the other hand, precipitation can increase locally after clearing a forest patch in the Amazon<sup>13</sup>. Outside of the tropics, modeling studies suggest that LCCs can trigger substantial changes in precipitation<sup>14,15</sup>. However, such studies lack solid observational constraints at regional to global scales, as little work has been conducted on estimating the LC-precipitation coupling in the mid-latitudes based on observations<sup>16</sup>. This is likely owed to the difficulty of isolating such a signal. The distribution of a certain LC type itself is often influenced by the precipitation climatology, by other factors (e.g., elevation) that in turn affect precipitation, or systematic preferences for certain terrain types from human land use<sup>17</sup>. A further complication is that LC-precipitation couplings do not necessarily occur at the location of the LC, but might also have an influence on precipitation further downwind. Thus, they are difficult to observe especially in regions of heterogeneous LC, as in Europe.

In this study, we estimate the impact of converting rainfed agricultural land ( $AL_r$ ; including pasture) to forest on precipitation over Europe. We focus on this LC conversion since it has been the dominant LCC over the last decades in many European regions<sup>18-21</sup> (Supplementary Fig. 1). An estimated 20 % of the European land area has undergone de-, re-, or afforestation between 1900 and 1990 mainly in exchange for cropland and grassland<sup>20</sup>. In addition, reforestation has been proposed as a tool to mitigate greenhouse gas emissions<sup>22</sup>, while potentially providing additional co-benefits for biodiversity<sup>23</sup>, soil protection<sup>23</sup>, and the local climate<sup>2-4</sup>. For simplicity, we assume here that forestation corresponds only to the conversion of  $AL_r$  to forest. We acknowledge however that other forms of human land use, such as settlements, might also be suitable for forestation. Also, we use forestation as a general term, not distinguishing between afforestation and reforestation. Due to the lack of constraints from previous studies, we employ two independent methodologies to corroborate our results (Extended Data Fig. 1): (1) We search for suitable closely-located rain gauge site pairs in the Global Sub-Daily Rainfall Dataset<sup>24</sup> (GSDR) and Global Historical Climatology Network<sup>25,26</sup> (GHCN), which differ in the  $AL_r$  and forest fractions. Such an approach has previously been applied to assess the effect of LC on temperature<sup>2,27</sup> and the surface energy budget<sup>28-30</sup>. (2) We statistically disaggregate the monthly

precipitation climatology over 1986 to 2015 from MSWEP v2.2<sup>31</sup> (Multi-Source Weighted-Ensemble Precipitation version 2.2; Supplementary Fig. 7) with a Generalized Additive Model (GAM) to infer the contribution of LC on the spatial distribution of precipitation (Extended Data Table 1). Statistical models were successfully used as a tool to assess both local<sup>32</sup> and remote<sup>33</sup> temperature changes induced by LCC. For the GAM presented here, we consider topographic effects, which strongly modulate the spatial distribution of precipitation<sup>34</sup>, by including a number of topographic metrics as predictors (Supplementary Fig. 3). The ERA5-Land 2 m temperature climatology<sup>35</sup> serves as an indicator for climatic conditions, as the amount of precipitation varies considerably across climate zones. Further, the atmospheric circulation is an important driver of the spatial distribution of precipitation<sup>36</sup> and is therefore represented in the GAM by various metrics based on ERA5 air parcel trajectories<sup>37</sup>. Finally, we employ the CORINE LC data<sup>38</sup> as predictors in the GAM to examine the relation between LC and precipitation. Alongside the local LC, we consider the upwind LC fractions, calculated from the backwards ERA5 air parcel trajectories (Extended Data Fig. 2). Therefore, we can not only isolate the local effect of forestation but also its downwind effect.

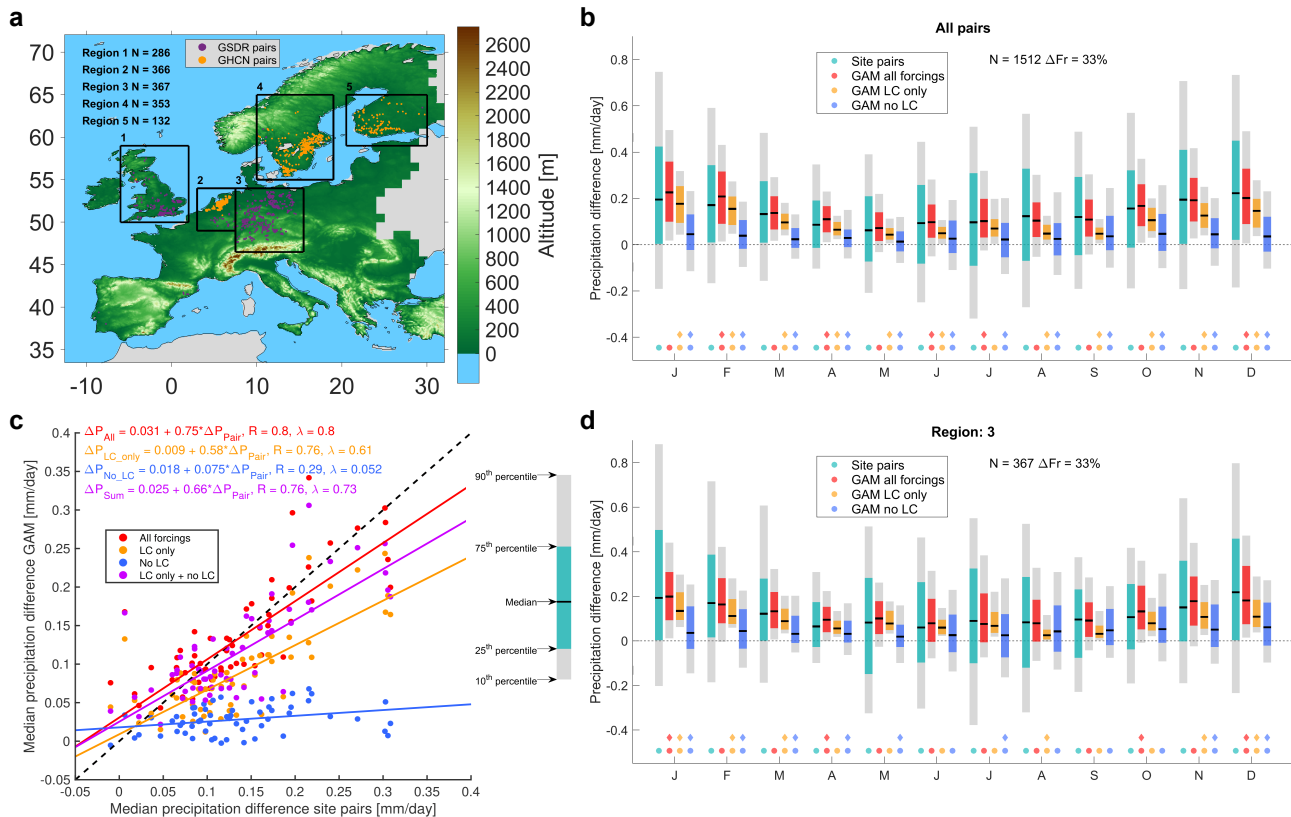
## Local effect of land cover on precipitation in site pairs

Among the 3481 rain gauge stations of GSDR and the 2376 stations of GHCN included in our analysis (Extended Data Fig. 3), we find a total of 1512 station pairs that differ at least by 20 % in the  $AL_r$  and forest LC fractions according to the CORINE LC data at 0.1° resolution and fulfill the additional selection criteria defined to mitigate the influence of potential confounding factors. Here, we focus on five clusters of station pairs, as approximately 100 sites pairs are required for a robust signal (Fig. 1 a). Precipitation is significantly higher at the sites dominated by forests in all regions, despite a considerable spread among individual station pairs, with the exception of Region 5 from April to July (green bars in Figs. 1 b, d and Extended Data Fig. 4). A stronger signal appears during the winter months, while the summer signal tends to be only slightly positive. Further, the precipitation differences between the sites with more forest minus the site with more  $AL_r$  ( $\Delta P_{loc}$ ) is more pronounced under the oceanic climate of Regions 1 and 2 and decreases in strength as the regions become more continental. In relative terms, the median  $\Delta P_{loc}$  correspond to 5-15 % and 0-10 % of the total precipitation during winter and summer, respectively (Extended Data Fig. 5).

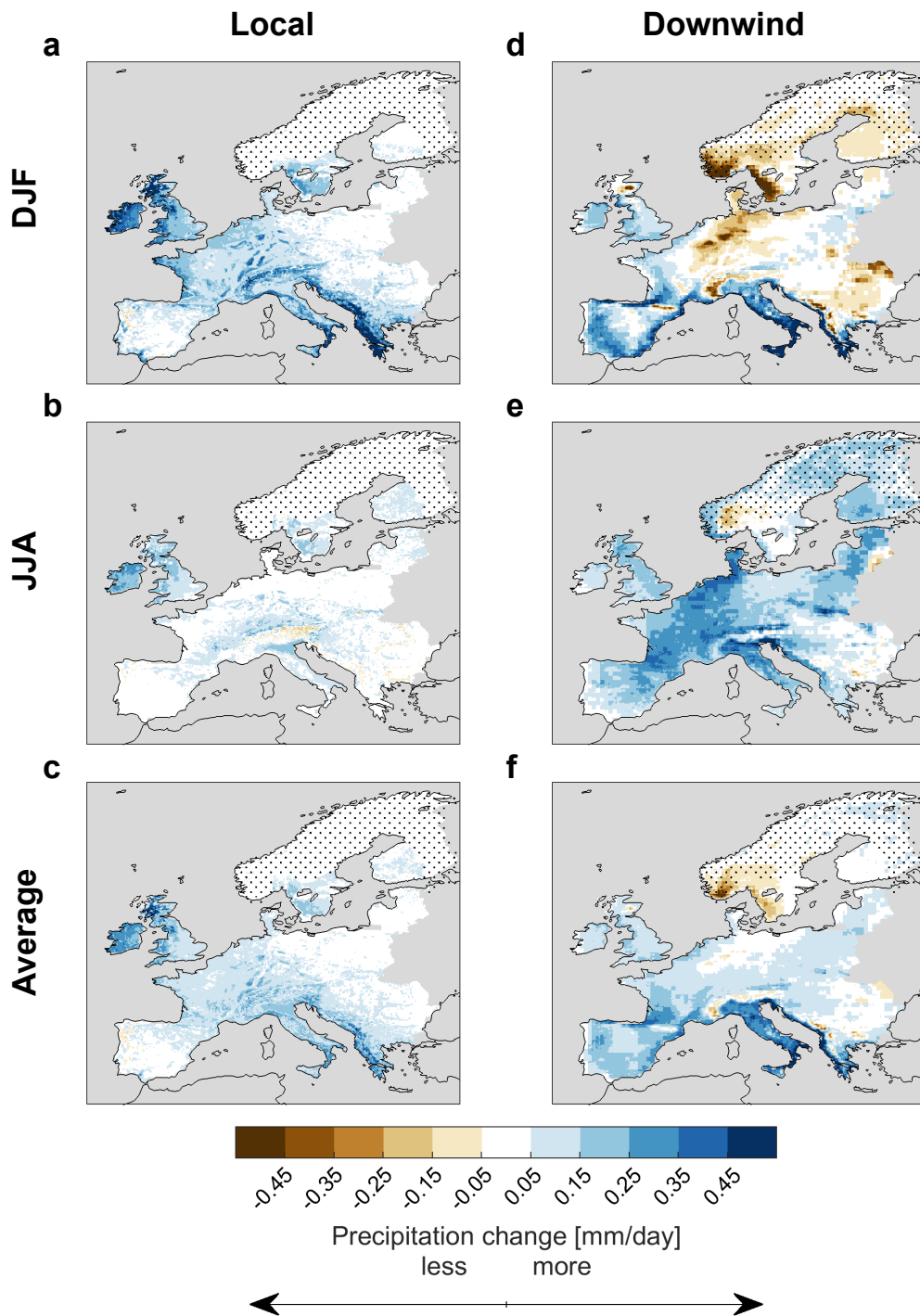
In order to infer the relationship between forestation and precipitation in a spatially-explicit way across Europe, we use a GAM, which we first validate here at the site pairs. The  $\Delta P_{loc}$  predicted by the GAM at the site pairs has a smaller spread than the one of the rain gauge pairs but gives a similar result for the magnitude and seasonality of the median  $\Delta P_{loc}$  over the five regions, supported by an index of agreement (IA)<sup>39</sup> of 0.80 (red/'all forcings' in Fig. 1 b-d and Extended Data Fig. 4). To ensure that the observed  $\Delta P_{loc}$  does not originate from systematic differences in variables other than the local  $AL_r$  and forest LC fractions, we make a second GAM prediction, using the forcing data from the first station for all predictors at both sites except for these two LC fractions (orange/'LC only' in Fig. 1 b-d and Extended Data Fig. 4). While the signal in this GAM prediction is somewhat weaker in all regions, it shows a similar pattern as the paired sites (IA of 0.61), indicating that most of the precipitation difference at the site pairs can be explained by the difference in the  $AL_r$  and forest LC fractions. For a second confirmation, we fitted another GAM without the LC-related predictors to the MSWEP precipitation climatology. Reassuringly, the precipitation difference predicted by this GAM for the site pairs captures only a small fraction of the signal from the site pairs (blue/'no LC' in Fig. 1 b-d and Extended Data Fig. 4, IA of 0.052).

## Theoretical effect of uniform forestation on precipitation

We employ the GAM that was evaluated in the previous section to estimate the effect of a 20 % increase in forest cover uniformly across Europe, with the exception of Northern Scandinavia, where no agricultural land is available for forestation (Extended Data Fig. 6). Locally, such a forestation increases precipitation in almost all locations with a mean of 0.087 mm/day across the area that is reforested (Fig. 2 c).  $\Delta P_{loc}$  is more positive during winter (mean of 0.13 mm/day) and tends to be stronger at locations close to the coast (Fig. 2 a). During the summer months, precipitation locally increases only weakly or even decreases slightly due to forestation (Fig. 2 b, mean of 0.043 mm/day). Besides the overall local increase in precipitation associated with forestation, we find substantial downwind effects from this LCC on precipitation ( $\Delta P_{dw}$ ). Foresting 20 % of the land surface decreases winter downwind precipitation over Northern Europe, exhibits a weak signal in Central and Eastern Europe, and increases precipitation in coastal areas of Western and Southern Europe (Fig. 2 d, mean of 0.043 mm/day). On the other hand, the GAM predicts a downwind increase due to forestation in summer precipitation over most of Europe (Fig. 2 e, mean of 0.17 mm/day). The  $\Delta P_{dw}$  during winter originates mostly from forestation in regions close to the coast, while the relevance of forestation in continental regions for downwind precipitation is confined to the summer season (Supplement A).



**Figure 1. Local precipitation difference between forest and agricultural land at site pairs.** Panel a, locations of the site pairs in GSDR (purple) and GHCN (orange). To the right the median (black line), interquartile range (colored shading), and range between 10<sup>th</sup> and 90<sup>th</sup> percentile (grey shading) of the monthly precipitation difference between the site with more forest coverage minus the site with more rainfed agricultural land ( $\Delta P_{loc}$ ) over all site pairs (b) and Region 3 (d), which exhibits an intermediate signal strength (other regions are shown in Extended Data Fig. 4). Bars correspond to distribution amongst the site pairs in the rain gauge data sets (green), the generalized additive model (GAM) prediction using all the forcings at the individual stations (red), the contribution to  $\Delta P_{loc}$  in the GAM from the differences in the local rainfed agricultural land and forest fractions (orange), and prediction of a GAM that was fitted without the land cover (LC) information (blue). Colored circles indicate samples which are significantly different from zero in a t-test and diamonds GAM prediction samples which are significantly different from the sample of the site pairs in a Welch's t-test both at 5% confidence level. N is number of site pairs and  $\Delta Fr$  median difference in forest/rainfed agricultural land fraction among site pairs used to estimate the contribution from those LCs in the GAM. Panel c shows monthly medians of  $\Delta P_{loc}$  over the five regions in the site pairs against monthly medians of the all-forcing GAM (red), LC only GAM (orange), no LC GAM (blue), and medians of LC only GAM plus medians of no LC GAM (purple). R denotes Pearson correlation coefficient and  $\lambda$  the index of agreement.



**Figure 2. Theoretical precipitation change from foresting 20 % of the land surface across Europe.** Estimated local (a-c) and downwind precipitation change (d-f) after foresting 20 % of the land surface in the entire study domain. Stippling depicts areas that were not forested in the respective scenario. Top row boreal winter (DJF), middle row boreal summer (JJA), and bottom row annual mean.

**Table 1.** Area-weighted average plus/minus area-weighted mean absolute deviation from average of relative precipitation change under RCP4.5 (Fig. 4 a, d), following forestation (Fig. 4 b, e), and sum of the two (Fig. 4 c, f) over biogeographical regions (Fig. 3 k). Columns 2-4 boreal winter (DJF) and columns 5-7 boreal summer (JJA). In bold numbers of forestation that reduce the climate change signal.

$\Delta P \pm MAD$ [%]	DJF			JJA		
Region	RCP 4.5	Forestation	Sum	RCP 4.5	Forestation	Sum
Mediterranean	-0.19 ± 4.9	<b>23 ± 9.7</b>	23 ± 11	-19 ± 3.7	<b>19 ± 8.9</b>	-0.63 ± 9.1
Atlantic	6.3 ± 2.3	16 ± 11	23 ± 11	-8.7 ± 6.0	<b>14 ± 5.6</b>	5.7 ± 7.7
Alpine	7.4 ± 4.0	<b>-0.79 ± 4.2</b>	6.7 ± 5.6	3.2 ± 9.9	1.7 ± 2.5	4.9 ± 11
Continental	7.9 ± 3.2	<b>-1.0 ± 9.3</b>	6.9 ± 10	-6.0 ± 5.2	<b>4.8 ± 3.6</b>	-1.2 ± 6.4
Boreal	11 ± 2.2	<b>-2.1 ± 4.1</b>	8.6 ± 4.5	7.6 ± 3.1	2.9 ± 3.3	11 ± 5.1

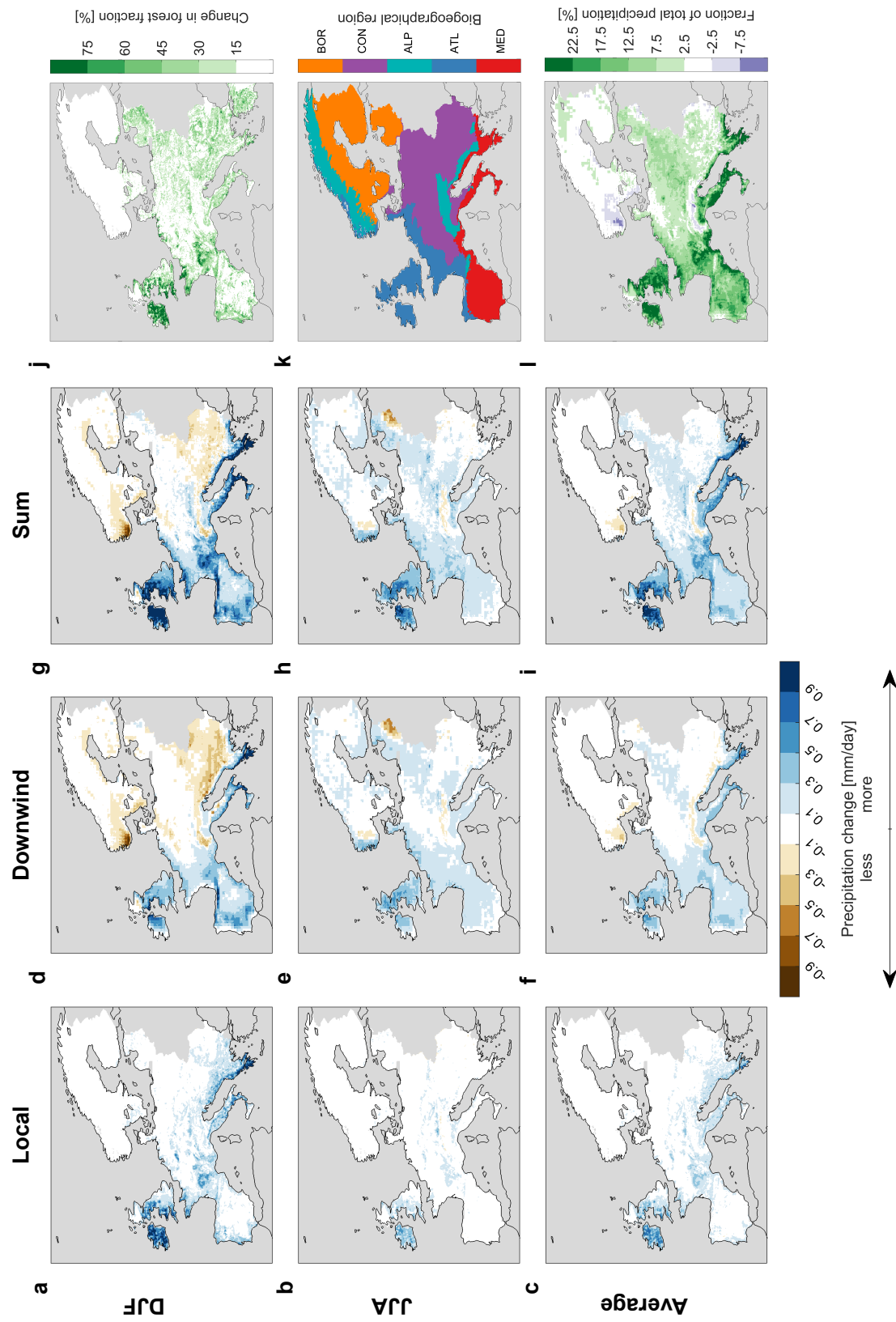
## Implications for a realistic forestation scenario

The forestation scenarios presented so far were stylized and might not be applicable in all regions of Europe. Therefore, we estimate the changes in precipitation if the reforestation potential according to the Global Reforestation Potential Map<sup>22,40</sup> was realized. This map was constrained by safeguards for food and fiber security, biodiversity, and against adverse biogeophysical temperature effects<sup>2-4</sup>. In total, 14.4 % of the land surface in the study domain is suitable for forestation, with concentrations over the British Isles, western and southern France, Portugal, Italy, and Eastern Europe (Fig. 3 j). The  $\Delta P_{loc}$  due to this scenario often exceeds 0.1 mm/day during winter, while the summer signal rarely exceeds this value except for the British Isles. Additionally, this forestation scenario triggers substantial non-local changes in the precipitation field (Fig. 3 d-f). These downwind effects reach regions with no significant forestation potential, such as Scandinavia, where reforestation was prohibited due to the local biogeophysical warming effect of forests in those areas<sup>22,40</sup>. Adding the two contributions up, precipitation is increased by 0.14 mm/day in winter, 0.13 mm/day in summer, and 0.16 mm/day for the annual mean averaged over the continent. The precipitation change amounts to more than 10 % of the annual mean precipitation for 27 % of the European area (Fig. 3 l).

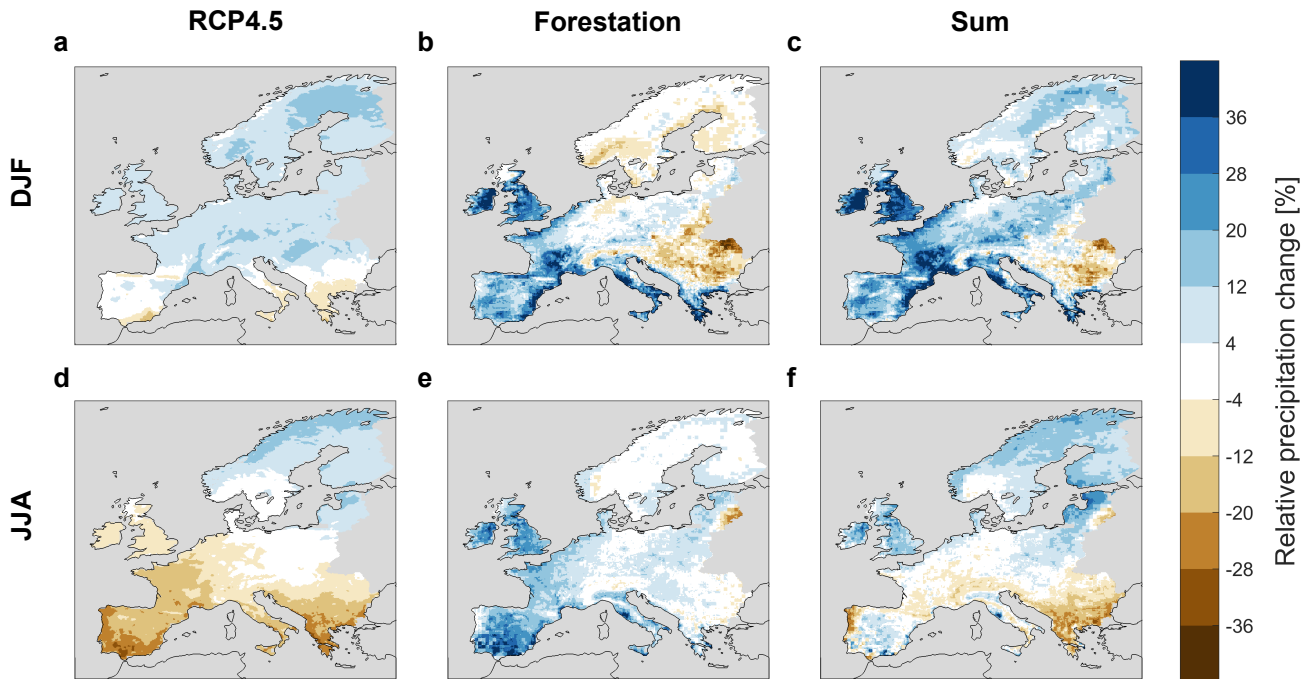
The precipitation changes induced by realistic forestation have the potential to compensate part of the climate change signal, as estimated by ten EUR-11 ensemble members (Fig. 4 and Table 1). During summer, realistic forestation is estimated to alleviate the precipitation decreases under RCP4.5 in the Mediterranean, Atlantic, and the Continental biogeographical regions (Fig. 3 k) and to compensate a considerable fraction of the precipitation signal under RCP8.5 (Supplementary Fig. 4 and Supplementary Table 1). Therefore, forestation could play a crucial role in adapting to the increased risks of summer droughts due to climate change<sup>41</sup>. Further, forestation is estimated to attenuate the wintertime precipitation increase in the Boreal region and even overcompensate the precipitation reduction in the Mediterranean (Fig. 3 k). On the other hand, it could also have potentially adverse effects by further intensifying climate change-induced precipitation increases in the Atlantic region. It needs to be noted that the forestation-induced precipitation changes presented here are determined under recent climatic conditions and may be altered under a warmer climate. Also, the EUR-11 ensemble exhibits substantial (mostly positive) biases in the simulated precipitation distribution of the recent past<sup>42,43</sup>, which could also affect the estimated precipitation changes due to climate change.

## Physical drivers and potential caveats

Here, we provide a first observational estimate of forestation-induced changes in precipitation over Europe. Therefore, it is difficult to assess the robustness of our estimate with previous studies. Forest differs in several biogeophysical characteristics from agricultural land, which can explain the identified precipitation changes. Firstly, forests typically have a higher surface roughness than agricultural land. We hypothesize that this might alter the local precipitation amount by inducing more turbulence and slowing movement of precipitating air masses. The increased surface roughness over urban areas has been identified as an important contributor to the amplification of precipitation over cities and downwind of cities<sup>44</sup>. Similarly, the high surface roughness of forests was an important contributor to the precipitation increase from afforestation in the Sahel region in a model experiment<sup>14</sup>. Further supporting this hypothesis, we find that higher wind speeds, as an anti-correlated proxy for the surface roughness, are associated with lower precipitation in particular during winter (Fig. B1 and Supplement B). The surface roughness has also implications for  $\Delta P_{dw}$ . A decreased number of cyclones was found following large-scale reforestation of Europe in a regional climate model, mainly due to the high surface roughness of forests<sup>15</sup>. Especially forests close to the coast might therefore hinder the propagation of frontal systems into the interior of the continent. As cyclones contribute more to the



**Figure 3. Potential precipitation change from realistic forestry.** Estimated local (a-c), downwind (d-f), and total precipitation change (g-i) associated with forestry in panel j. Top row boreal winter (DJF), middle row boreal summer (JJA), and bottom row annual mean. Panel k shows biogeographical regions over which averages are calculated in Table. 1 and panel l the precipitation change in panel i as a fraction of the annual mean precipitation.



**Figure 4. Comparison of precipitation changes from climate change and forestation.** Change in precipitation in EUR-11 multi model ensemble under RCP4.5 scenario between 1986-2015 and 2071-2100 relative to modeled precipitation climatology of 1986-2015 (a, d), estimated precipitation changes due to realistic forestation scenario relative to precipitation climatology of 1986-2015 in MSWEP (b, e), and sum of the two (c, f). Top row boreal winter (DJF) and bottom row boreal summer (JJA).

total precipitation during winter than during summer<sup>45,46</sup>, this can explain why forestation leads to a downwind increase of precipitation in areas close to the coast during winter, while the signal is neutral in Central and Western Europe and even negative in Northern Europe. Indeed, forestation is estimated to decrease the propagation speed of the calculated wind trajectories further away from the coasts, leading to a reduction of the rainfall amount (Supplement B). Secondly, observations indicate that forests typically sustain higher evapotranspiration than  $AL_r$ , in particular during the summer season<sup>3,4,47</sup>. We hypothesize that this is an important driver behind the downwind summertime precipitation increase from forestation in most locations of Europe, which is supported by applying a moisture source diagnostic to the wind trajectories used in this study (Supplement C). Higher evapotranspiration of forests was also linked to increased precipitation in the tropics<sup>11,12</sup> and the Sahel region<sup>14</sup>. Thirdly, forestation warms the land surface during winter but cools it during summer<sup>3,4,47</sup>. This could explain the seasonal cycle of the local signal we observe, as warmer temperatures at the land surface destabilize the planetary boundary layer, thereby favoring the creation of precipitation. In fact, the urban heat island effect was linked to locally increased precipitation<sup>44</sup>. Overall, future sensitivity experiments with models could provide a valuable contribution to understand the processes underlying the signals observed here.

Despite the fact that our results are consistent with the current understanding of mechanisms at play, a number of potential caveats in the analysis should be highlighted. Our results are based on a spatial rather than a temporal comparison of the precipitation amount over different LCs. In consequence, the observed  $\Delta P_{loc}$  can be the result of a spatial redistribution of precipitation rather than an alteration of the overall amount of precipitation, which would imply that the estimated  $\Delta P_{loc}$  is not directly attributable to the effect of forestation at larger scales. Further, causality is inherently uncertain in such observation-based approaches. By choosing precipitation as a response variable and LC as a predictor, we implicitly assume that the latter influences the former. In reality, the distribution of the different LC types may also be influenced by precipitation amounts. However, the results of the site pair analysis are robust against the stringency of the selection criteria (Supplement D and Fig. D1). As for the GAM analysis, the influence of potential confounding factors is minimized by including them as predictors. While the two estimates of  $\Delta P_{loc}$  based on different data sets and methodologies show good agreement, they are not fully independent, as MSWEP v2.2 incorporates rain gauge measurements besides reanalysis products and remote sensing data sets. Hence, we cannot exclude that systematic biases in the rain gauge measurements translate to the GAM analysis. The undercatch of rain gauges measurements is exacerbated by high wind speeds<sup>48</sup>. As



forests typically have a higher surface roughness than  $AL_r$ , lower surface wind speeds occur at the sites in regions with more forest (Supplement B). This could result in weaker undercatch at those sites compared to the sites with more  $AL_r$  in their proximity. Further, the  $\Delta P_{loc}$  in the GAM could not be evaluated in the southern half of Europe, due to the scarcity of rain gauge data in those areas. Thus, the compensatory effect of forestation on the projected reduction of precipitation due to climate change in Southern Europe is prone to larger uncertainties than the precipitation changes from forestation in other regions (Note however that the GAM exhibits a similar signal as the rain gauge data over the Iberian Peninsula with relaxed selection criteria; Extended Data Fig. 7). Uncertainties are larger for  $\Delta P_{dw}$ , for which we rely solely on the GAM. Overall, our results are consistent with observations in the tropics<sup>11</sup> and modeling studies<sup>12,15</sup>. Nevertheless, the estimated downwind effect of forestation is sensitive to both the construction of the GAM and the starting heights of the wind trajectory calculation, although qualitatively robust (Fig. E2 and Supplement E). We therefore encourage a more thorough analysis of  $\Delta P_{dw}$ .

Overall, our results imply that forestation could trigger substantial changes in precipitation over Europe. Forestation-induced precipitation changes appear to be subject to spatial trade-offs due to downwind effects. While we find a local increase in precipitation due to forestation across Europe, forestation might reduce precipitation further downwind in winter. On the other hand, forestation increases precipitation downwind in summer, likely due to higher moisture supply by forests than by  $AL_r$ . This downwind enhancement of precipitation might come at the cost of a local reduction of runoff<sup>49,50</sup>, adding yet another dimension to the spatial interconnections. Overall, our results highlight that LCCs, such as forestation, can alter precipitation in the mid-latitudes considerably, both locally and further downwind. Hence, the consequences of human land-use for water availability should be considered alongside biogeochemical effects and the biogeophysical alteration of temperatures. As droughts are projected to become more severe with changing climate in Europe<sup>41</sup>, the interplay between LC and water availability deserves more attention.

## References

1. Meier, R. *et al.* Empirical estimated of forestation-induced precipitation changes in Europe. *Nat. Geosci.* **XXX**, YYY–YYY, DOI: [10.1038/s41561-021-00773-6](https://doi.org/10.1038/s41561-021-00773-6) (2021).
2. Lee, X. *et al.* Observed increase in local cooling effect of deforestation at higher latitude. *Nature* **479**, 384–387, DOI: [10.1038/nature10588](https://doi.org/10.1038/nature10588) (2011).
3. Li, Y. *et al.* Local cooling and warming effects of forests based on satellite observations. *Nat. Commun.* **6**, DOI: [10.1038/ncomms7603](https://doi.org/10.1038/ncomms7603) (2015).
4. Duveiller, G., Hooker, J. & Cescatti, A. The mark of vegetation change on earth’s surface energy balance. *Nat. Commun.* **9**, DOI: [10.5194/essd-2018-24](https://doi.org/10.5194/essd-2018-24) (2018).
5. IPCC 2019c: Jia, G. *et al.* Landclimate interactions. In: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D.C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, J. Malley, (eds.)]. in press (2019).
6. Lejeune, Q., Seneviratne, S. I. & Davin, E. L. Historical land-cover change impacts on climate: Comparative assessment of LUCID and CMIP5 multimodel experiments. *J. Clim.* **30**, 1439–1459, DOI: [10.1175/JCLI-D-16-0213.1](https://doi.org/10.1175/JCLI-D-16-0213.1) (2017).
7. Winckler, J., Reick, C. H. & Pongratz, J. Robust identification of local biogeophysical effects of land-cover change in a global climate model. *J. Clim.* **30**, 1159–1176, DOI: [10.1175/JCLI-D-16-0067.1](https://doi.org/10.1175/JCLI-D-16-0067.1) (2017). <https://doi.org/10.1175/JCLI-D-16-0067.1>.
8. Duveiller, G. *et al.* Biophysics and vegetation cover change: a process-based evaluation framework for confronting land surface models with satellite observations. *Earth Syst. Sci. Data Discuss.* 1–24, DOI: [10.5194/essd-2018-24](https://doi.org/10.5194/essd-2018-24) (2018).
9. Meier, R. *et al.* Evaluating and improving the community land model’s sensitivity to land cover. *Biogeosciences* **15**, 4731–4757, DOI: [10.5194/bg-15-4731-2018](https://doi.org/10.5194/bg-15-4731-2018) (2018).
10. Meier, R., Davin, E. L., Swenson, S. C., Lawrence, D. M. & Schwaab, J. Biomass heat storage dampens diurnal temperature variations in forests. *Environ. Res. Lett.* **14**, 084026, DOI: [10.1088/1748-9326/ab2b4e](https://doi.org/10.1088/1748-9326/ab2b4e) (2019).
11. Spracklen, D., Arnold, S. & Taylor, C. Observations of increased tropical rainfall preceded by air passage over forests. *Nature* **489**, 282–285, DOI: [10.1038/nature11390](https://doi.org/10.1038/nature11390) (2012).
12. Lejeune, Q., Davin, E. L., Guillod, B. P. & Seneviratne, S. I. Influence of amazonian deforestation on the future evolution of regional surface fluxes, circulation, surface temperature and precipitation. *Clim. Dynam.* **44**, 2769–2786, DOI: [10.1007/s00382-014-2203-8](https://doi.org/10.1007/s00382-014-2203-8) (2015).

13. Khanna, J., Medvigy, D., Fueglistaler, S. & Walko, R. Regional dry-season climate changes due to three decades of amazonian deforestation. *Nat. Clim. Chang.* **7**, 200–204, DOI: [10.1038/nclimate3226](https://doi.org/10.1038/nclimate3226) (2017).
14. Yosef, G. *et al.* Large-scale semi-arid afforestation can enhance precipitation and carbon sequestration potential. *Sci. Rep.* **8**, DOI: [10.1038/s41598-018-19265-6](https://doi.org/10.1038/s41598-018-19265-6) (2018).
15. Belušić, D., Fuentes-Franco, R., Strandberg, G. & Jukimenko, A. Afforestation reduces cyclone intensity and precipitation extremes over europe. *Environ. Res. Lett.* **14**, 074009, DOI: [10.1088/1748-9326/ab23b2](https://doi.org/10.1088/1748-9326/ab23b2) (2019).
16. Perugini, L. *et al.* Biophysical effects on temperature and precipitation due to land cover change. *Environ. Res. Lett.* **12**, 053002, DOI: [10.1088/1748-9326/aa6b3f](https://doi.org/10.1088/1748-9326/aa6b3f) (2017).
17. Sandel, B. & Svenning, J. Human impacts drive a global topographic signature in tree cover. *Nat. Commun.* **4**, DOI: [10.1038/ncomms3474](https://doi.org/10.1038/ncomms3474) (2013).
18. Fuchs, R., Herold, M., Verburg, P. H. & Clevers, J. G. P. W. A high-resolution and harmonized model approach for reconstructing and analysing historic land changes in europe. *Biogeosciences* **10**, 1543–1559, DOI: [10.5194/bg-10-1543-2013](https://doi.org/10.5194/bg-10-1543-2013) (2013).
19. Hansen, M. C. *et al.* High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853, DOI: [10.1126/science.1244693](https://doi.org/10.1126/science.1244693) (2013).
20. Fuchs, R., Herold, M., Verburg, P. H., Clevers, J. G. & Eberle, J. Gross changes in reconstructions of historic land cover/use for europe between 1900 and 2010. *Glob. Chang. Biol.* **21**, 299–313, DOI: [10.1111/gcb.12714](https://doi.org/10.1111/gcb.12714) (2014).
21. McGrath, M. J. *et al.* Reconstructing european forest management from 1600 to 2010. *Biogeosciences* **12**, 4291–4316, DOI: [10.5194/bg-12-4291-2015](https://doi.org/10.5194/bg-12-4291-2015) (2015).
22. Griscom, B. W. *et al.* Natural climate solutions. *P. Natl. Acad. Sci. USA* **114**, 11645–11650, DOI: [10.1073/pnas.1710465114](https://doi.org/10.1073/pnas.1710465114) (2017). <https://www.pnas.org/content/114/44/11645.full.pdf>.
23. Navarro, L. M. & Pereira, H. M. *Rewilding Abandoned Landscapes in Europe*, 3–23 (Springer International Publishing, Cham, 2015).
24. Lewis, E. *et al.* GSDR: A global sub-daily rainfall dataset. *J. Clim.* **32**, 4715–4729, DOI: [10.1175/JCLI-D-18-0143.1](https://doi.org/10.1175/JCLI-D-18-0143.1) (2019). <https://doi.org/10.1175/JCLI-D-18-0143.1>.
25. Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E. & Houston, T. G. An overview of the global historical climatology network-daily database. *J. Atmospheric Ocean. Technol.* **29**, 897–910, DOI: [10.1175/JTECH-D-11-00103.1](https://doi.org/10.1175/JTECH-D-11-00103.1) (2012). <https://doi.org/10.1175/JTECH-D-11-00103.1>.
26. Menne, M. J. *et al.* Global Historical Climatology Network - Daily (GHCN-Daily), Version 3.20, DOI: [10.7289/V5D21VHZ](https://doi.org/10.7289/V5D21VHZ) (2012).
27. Zhang, M. *et al.* Response of surface air temperature to small-scale land clearing across latitudes. *Environ. Res. Lett.* **9**, DOI: [10.1088/1748-9326/9/3/034002](https://doi.org/10.1088/1748-9326/9/3/034002) (2014).
28. Liu, H., Randerson, J. T., Lindfors, J. & Chapin III, F. S. Changes in the surface energy budget after fire in boreal ecosystems of interior alaska: An annual perspective. *J. Geophys. Res.* **110**, DOI: [10.1029/2004JD005158](https://doi.org/10.1029/2004JD005158) (2005).
29. Juang, J.-Y., Katul, G., Siqueira, M., Stoy, P. & Novick, K. Separating the effects of albedo from eco-physiological changes on surface temperature along a successional chronosequence in the southeastern united states. *Geophys. Res. Lett.* **34**, DOI: [10.1029/2007GL031296](https://doi.org/10.1029/2007GL031296) (2007).
30. Vanden Broucke, S., Luysaert, S., Davin, E. L., Janssens, I. & van Lipzig, N. New insights in the capability of climate models to simulate the impact of luc based on temperature decomposition of paired site observations. *J. Geophys. Res.-Atmos.* **120**, 5417–5436, DOI: [10.1002/2015JD023095](https://doi.org/10.1002/2015JD023095) (2015).
31. Beck, H. E. *et al.* MSWEP V2 global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *B. Am. Meteorol. Soc.* **100**, 473–500, DOI: [10.1175/BAMS-D-17-0138.1](https://doi.org/10.1175/BAMS-D-17-0138.1) (2019).
32. Schwaab, J. *et al.* Increasing the broad-leaved tree fraction in european forests mitigates hot temperature extremes. *Sci. Rep.* **10**, DOI: [10.1038/s41598-020-71055-1](https://doi.org/10.1038/s41598-020-71055-1) (2020).
33. Cohn, A. S. *et al.* Forest loss in brazil increases maximum temperatures within 50 km. *Environ. Res. Lett.* **14**, 084047, DOI: [10.1088/1748-9326/ab31fb](https://doi.org/10.1088/1748-9326/ab31fb) (2019).
34. Houze Jr., R. A. Orographic effects on precipitating clouds. *Rev. Geophys.* **50**, DOI: [10.1029/2011RG000365](https://doi.org/10.1029/2011RG000365) (2012). <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2011RG000365>.

35. Copernicus Climate Change Service (C3S). C3S ERA5-Land reanalysis (2019).
36. Daly, C. *et al.* Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous united states. *Int. J. Clim.* **28**, 2031–2064, DOI: [10.1002/joc.1688](https://doi.org/10.1002/joc.1688) (2008).
37. Sprenger, M. & Wernli, H. The LAGRANTO Lagrangian analysis tool - version 2.0. *Geosci. Model. Dev.* **8**, 2569–2586, DOI: [10.5194/gmd-8-2569-2015](https://doi.org/10.5194/gmd-8-2569-2015) (2015).
38. Kosztra, B., Büttner, G., Hazeu, G. & Arnold, S. Updated CLC illustrated nomenclature guidelines (2019).
39. Duveiller, G., Fasbender, D. & Meroni, M. Revisiting the concept of a symmetric index of agreement for continuous datasets. *Sci. Rep.-UK* **6**, 19401, DOI: [10.1038/srep19401](https://doi.org/10.1038/srep19401) (2016).
40. Griscom, B. W. *et al.* Global reforestation potential map, DOI: [10.5281/zenodo.883444](https://doi.org/10.5281/zenodo.883444) (2017).
41. Sheffield, J. & Wood, E. F. Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations. *Clim. Dynam.* **31**, 79–105, DOI: [10.1007/s00382-007-0340-z](https://doi.org/10.1007/s00382-007-0340-z) (2008).
42. Kotlarski, S. *et al.* Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geosci. Model. Dev.* **7**, 1297–1333, DOI: [10.5194/gmd-7-1297-2014](https://doi.org/10.5194/gmd-7-1297-2014) (2014).
43. Prein, A. F. *et al.* A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.* **53**, 323–361, DOI: [10.1002/2014RG000475](https://doi.org/10.1002/2014RG000475) (2015). <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2014RG000475>.
44. Liu, J. & Niyogi, D. Meta-analysis of urbanization impact on rainfall modification. *Sci. Rep.* **9**, DOI: [10.1038/s41598-019-42494-2](https://doi.org/10.1038/s41598-019-42494-2) (2019).
45. Van der Ent, R. J. & Savenije, H. H. G. Length and time scales of atmospheric moisture recycling. *Atmospheric Chem. Phys.* **11**, 1853–1863, DOI: [10.5194/acp-11-1853-2011](https://doi.org/10.5194/acp-11-1853-2011) (2011).
46. Rüdüsühli, S., Sprenger, M., Leutwyler, D., Schär, C. & Wernli, H. Attribution of precipitation to cyclones and fronts over Europe in a kilometer-scale regional climate simulation. *Weather. Clim. Dyn. Discuss.* **2020**, 1–38, DOI: [10.5194/wcd-2020-18](https://doi.org/10.5194/wcd-2020-18) (2020).
47. Schultz, N. M., Lawrence, P. J. & Lee, X. Global satellite data highlights the diurnal asymmetry of the surface temperature response to deforestation. *J. Geophys. Res.-Biogeo.* **122**, 903–917, DOI: [10.1002/2016JG003653](https://doi.org/10.1002/2016JG003653) (2017).
48. Pollock, M. D. *et al.* Quantifying and mitigating wind-induced undercatch in rainfall measurements. *Water Resour. Res.* **54**, 3863–3875, DOI: [10.1029/2017WR022421](https://doi.org/10.1029/2017WR022421) (2018). <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2017WR022421>.
49. Trabucco, A., Zomer, R. J., Bossio, D. A., Straaten], O. v. & Verchot, L. V. Climate change mitigation through afforestation/reforestation: A global analysis of hydrologic impacts with four case studies. *Agr. Ecosyst. Environ.* **126**, 81–97, DOI: [10.1016/j.agee.2008.01.015](https://doi.org/10.1016/j.agee.2008.01.015) (2008). International Agricultural Research and Climate Change: A Focus on Tropical Systems.
50. Padrón, R. S., Gudmundsson, L., Greve, P. & Seneviratne, S. I. Large-scale controls of the surface water balance over land: Insights from a systematic review and meta-analysis. *Water Resour. Res.* **53**, 9659–9678, DOI: [10.1002/2017WR021215](https://doi.org/10.1002/2017WR021215) (2017). <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2017WR021215>.

## Methods

Extended Data Fig. 1 provides an overview of the overall procedure and how the various datasets (Supplementary Table 2) and analysis steps are connected. Supplementary Table 3 shows all abbreviations used.

### Gridded data

We employ a number of gridded data sets, which are used for two purposes: (1) They serve as criteria to find suitable site pairs in the GSDR and GHCN data. (2) They are used as response variables and predictors for the GAM. All the data are used at a regular 0.1° grid, which is the native resolution of the MSWEP data set. The study domain of the GAM is confined to the coverage of the CORINE LC data. Turkey was removed from the analysis, since it is surrounded by areas not covered by CORINE, which hampers the calculation of the upwind LC fractions.

### **Precipitation fields**

We use the monthly precipitation climatologies from MSWEP v2.2 over the period 1986-2015 as the response variable to train the GAMs. This data set is the latest version of the Multi-Source Weighted-Ensemble Precipitation data set<sup>31</sup>. For simplicity we refer to it as MSWEP. It is available globally at 0.1° spatial and 3-hourly temporal resolution. As indicated by its name, multiple rain gauge data sets, remote sensing products, and reanalysis products were used to create MSWEP<sup>50</sup>. This data set systematically outperforms other gridded precipitation data sets when evaluated with rain gauge data and streamflow measurements<sup>51</sup> as well as radar measurements<sup>52</sup>. We calculate monthly precipitation climatologies from MSWEP over the period 1986-2015 (Supplementary Fig. 2).

### **Topography**

Topography plays a crucial role in the spatial distribution of precipitation<sup>34,36,53</sup>. We use the Digital Elevation Model over Europe version 1.1 (EU-DEM v1.1)<sup>54</sup> to derive a number of relevant metrics. First, we compute the mean elevation at 0.1° from the original 30 m resolution. Besides, we calculate the slope, exposition of the slope, the Topographic Position Index (TPI), and the Terrain Ruggedness Index (TRI) using the gdaldem tool of the Geospatial Data Abstraction software Library 2019 (GDAL, Supplementary Fig. 3). The TPI describes the elevation of a point in comparison to the surrounding areas. It is defined as the elevation difference between the focal grid cell and the mean of the eight surrounding grid cells. The TRI is a measure of terrain variability and calculated as the mean of the absolute elevation differences between the focal grid cell and the eight adjacent grid cells. The exposition of the slope and the TPI are calculated from the 0.1° elevation data, whereas the slope and TRI are computed at the original resolution of EU-DEM and then averaged over the 0.1° grid.

### **Land cover information**

For the LC information, we regrid the CORINE Land Cover (CLC) data from the year 2000 at 100 m resolution to 0.1° resolution<sup>38</sup>. The original 44 LC classes are aggregated to the classes artificial surfaces, agricultural land, forest, shrubland, natural low vegetation, bare soil, wetland, and open water according to Supplementary Table 4. Further, we differentiate irrigated agricultural land using the Global Map of Irrigation Areas version 5 (GMIA5)<sup>55</sup>, because irrigation itself can trigger substantial modifications of the precipitation distribution<sup>56,57</sup>. The fraction of irrigated agricultural land in a grid cell from GMIA5 is used, unless it exceeds the fraction of agricultural land in CLC. In this case, it is set to the fraction of agricultural land in CLC, to assure the fraction of irrigated agricultural land does not exceed the total fraction of agricultural land in CLC (Supplementary Fig. 5). The remainder of the agricultural land in CLC is then attributed to  $AL_r$ .

### **Trajectory-based fields**

The spatial distribution of precipitation is not only affected by the properties of the land surface at a given location, but also the atmospheric motion prior to reaching this location. In particular the distance to the coast along the path of atmospheric motion is an important driver of how much moisture is available for precipitation<sup>36</sup>, as the oceans are the most important moisture source. To account for the relationship between the atmospheric circulation and precipitation, we use the Lagrangian analysis tool (LAGRANTO) based on hourly ERA5 reanalysis data<sup>37,58</sup>. Every six hours between January 1986 and December 2015, we calculate two kinematic trajectories five days backward in time for all land grid cells where the precipitation exceeded 1 mm/hr. We choose the two pressure levels that have produced most precipitation in the ERA5 data as starting heights, where the precipitation production at a specific level was calculated from the volumetric rain and snow water content of this level minus the volumetric rain and snow water content of the level above (Supplementary Fig. 6). The calculations are done in time steps of 10 min using hourly ERA5 wind field data at 0.5° resolution. Lagranto enables to track the position of the air parcel, the specific humidity, the land-sea mask, the elevation of the surface, and precipitation at the surface of ERA5, as well as the LC fractions from CLC. In the post-processing, we follow the trajectories backwards in time until the coast to determine the distance of the final precipitation event to the coast along the trajectory ( $uw\_cd$ ), the time it took to reach the coast ( $uw\_ct$ ), and the elevation difference between the highest topographic point the air mass surmounted on its path from the coast and the end point ( $uw\_hd$ ; Extended Data Fig. 2 d and e). In addition, we compute the average fraction of each LC class the air mass overpassed during the previous five days, which are called upwind LC fractions here ( $uw\_pct\_LC_i$ ; Extended Data Fig. 2 a-c). In the calculation of the upwind LC fractions we stop tracking the trajectory if the specific humidity of the parcel falls below 0.05 g/kg or if a land point is reached according to the ERA5 land-sea mask that is outside of the CLC domain. Collecting all trajectories for a given location and month, we then compute the median value for each of those metrics to retrieve the predictor fields for the GAMs. Using a similar approach, we calculate the median of the elevation difference between highest topographic point the air mass surmounted during the day following the precipitation event and the point where the precipitation occurred ( $dw\_hd$ ; Extended Data Fig. 2 f).

### **2m Temperature**

Finally, we use the ERA5-Land monthly 2 m temperature climatology from 1986-2015 as a predictor in the GAM, which is available at the required 0.1° resolution<sup>35</sup>.

## Site pair analysis

For the paired rain gauge station analysis we employ the GSDR data set<sup>24</sup> of the following countries: Belgium, Spain (Catalonia), Finland, Germany, Ireland, Italy, Norway, Portugal, Switzerland, and the UK. In addition, we include the data from the Integrated Surface Database<sup>59</sup>, which is not country specific. GSDR has been extensively quality controlled for common mechanical, recording, and storage errors in rainfall data as well as being checked against neighboring gauges<sup>60</sup>. The station data are aggregated from the original hourly to a three-hourly resolution to match with the temporal resolution of MSWEP. We exclude stations with more than 50 % missing data arriving at a total of 3481 stations in the study domain. Besides GSDR, we also employ the GHCN-Daily v3.20<sup>25,26</sup> rain gauge collection, which is available at daily resolution and has 2376 stations with less than 50 % missing days available in the study domain (Extended Data Fig. 3). Given the differing temporal resolution and quality checks of the two data sets, we only look for site pairs within one data set. Further, we neglect any measured data prior to 1986. For all of these stations we extract the values of the gridded data sets at the respective location to check for the following criteria to find station pairs: (1) The first site has at least 20 % more forest coverage and at least 20 % less coverage by  $AL_r$  than the second site according to the CLC data. Note that the site comparison does not correspond to a pure local comparison, but rather to a comparison of having 20 % more/less of  $AL_r$ /forest within a grid cell of roughly 11 km by 7 km. The isolated signal does therefore not only include the effect of foresting one specific location but also the effect of foresting areas in close proximity. (2) The sum of the  $AL_r$  and forest fractions is at least 50 % at each site. (3) The great-circle distance between the two sites is less than  $0.75^\circ$  corresponding to 83.5 km. (4) They have at least three overlapping years of measurements. (5) They differ less than 25 m in altitude,  $5^\circ$  in slope, 3 m in TRI, 20 % in the open water fraction, 20 % in the upwind open water fraction, and  $0.25^\circ$  in the annual mean  $uw\_cd$ . For each station in a pair, we then calculate the monthly rainfall climatology considering only times during which both stations have collected data. As an additional quality check, we discard station pairs that had a smaller sample size than 400 and 50 for any month in GSDR and GHCN, respectively (corresponding to 50 measurement days). An analysis of the sensitivity of the site pair analysis on the selection criteria is presented in Supplement D.

## GAM construction

GAMs express the expected value,  $\mathbb{E}$ , of a response variable,  $Y_i$ , as the sum of a number of smooth functions<sup>61</sup>:

$$g(\mathbb{E}(Y_i)) = f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + \dots \quad (1)$$

, where  $g$  is a link function,  $f_j$  are the smooth functions, and  $x_k$  are different predictor variables. GAMs possess a number of features that are advantageous for our analysis: (1) The smooth functions allow to empirically capture non-linear relations between the predictor and the response variables without a prior assumption regarding the shape of this relation. (2) The smooth functions can depend on several predictor variables. (3) The additive structure of GAMs makes them readily interpretable. (4) The smooth functions are penalized for their 'wiggleness', which prevents overfitting of the model. In this case, the response variable is the monthly precipitation climatology of MSWEP, which is assumed to follow a gamma distribution,  $g$  is the natural logarithm, and the  $f_j$  terms are summarized in Extended Data Table 1. The logarithmic link function was chosen to avoid harming the assumption of constant variance of residuals. Precipitation changes from forestation are estimated in a spatially-explicit way with our GAM, by including interaction terms between the LC fractions and other spatially-varying variables ('prox\_LC<sub>i</sub>, t2m, alt' and 'uw\_pct\_LC<sub>i</sub>, t2m, uw\_ct' smooths) and by directly including interaction terms of the LC fractions with the spatial dimensions ('prox\_LC<sub>i</sub>, lat, lon' and 'uw\_pct\_LC<sub>i</sub>, lat, lon' smooths).

We employ the extensive R package Mixed GAM Computation Vehicle with Automatic Smoothness Estimation ('mgcv')<sup>61,62</sup> to construct a GAM for each month of the MSWEP precipitation climatology. To shorten computation times, the model is fitted with the 'bam' function, which is numerically optimized for large datasets, using the fast restricted maximum likelihood method (fREML) to estimate the smoothing parameters (if not specified) and discretization of covariate values, which increases the computational efficiency of the fREML method<sup>63,64</sup>. We use thin plate regression splines ('s' smooths) for the one-dimensional terms, which are invariant in space and therefore called global terms here, and tensor product smooths ('ti' smooths) for the multi-dimensional terms, which are suitable if other terms using the respective predictors are used (Supplementary Fig. 7). To avoid unrealistically noisy results, the smoothing parameters were chosen manually for the multi-dimensional terms. For the local effect of LC, we also account for proximity effects by computing proximity LC fractions giving a third weight to the central grid cell and a sixth weight to each of the four grid cells that share an edge with the central grid cell (prox\_LC<sub>i</sub>).

## Performance assessment of GAM

Across all months, the GAM achieves an adjusted  $R^2$  of 0.93, a RMSE of 0.30 mm/day, and an index of agreement<sup>39</sup> of 0.96. Overall, the residuals still show some spatial correlation up to a distance of approximately  $1.5^\circ$ , which we accept in order to not have to over-fit the precipitation field with the purely spatial term (Supplementary Figs. 8 and 9). The GAM slightly deteriorates

the agreement with the precipitation climatologies at the rain gauges compared to the original MSWEP data (Supplementary Fig. 10). However, considerable differences are already present between the rain gauges and MSWEP, likely owed to the different spatial and temporal scales of these data sets. These differences are not systematic however. The significance of the individual smooths was tested in a type III ANOVA, using the 'mgcv' function 'anova'. Overall, all the smooths remain highly significant throughout the year, with a few exceptions mostly for interactions terms of some variables with latitude and longitude (Extended Data Table 1). Note however, that the significance of the terms is likely to be slightly overestimated due to the spatial auto-correlation of the residuals<sup>65</sup>.

To compare the results of the GAM with the site pairs, we make a GAM prediction extracting the forcing data from each individual site in a site pair. Due to the differing spatial scales of the rain gauge data (point scale) and the MSWEP data (0.1° grid), we use the local CLC fractions at the sites (within 0.1° grid cell) as the proximity when conducting the GAM prediction. Besides this, we make a second GAM prediction, using only the local CLC AL<sub>r</sub> and forest LC information from the individual sites, but setting all other predictor variables to the values at the location of the site with more forest coverage (Site 1) at both sites. The AL<sub>r</sub> and forest LC fractions at the second site are computed as follows to ensure the LC fractions still add up to 100 %:

$$\Delta Fr = \min(pct_{for1}, 100 - pct_{agr1}, \text{mean}(pct_{for1} - pct_{for2}, pct_{agr2} - pct_{agr1})),$$

$$pct_{agr1}^* = pct_{agr1} + \Delta Fr, pct_{for1}^* = pct_{for1} - \Delta Fr \quad (2)$$

, where  $pct_{for1}$  and  $pct_{agr1}$  are the fraction forest and fraction AL<sub>r</sub> at the site with more forest,  $pct_{for2}$  and  $pct_{agr2}$  the respective fractions at the site with less forest,  $\Delta Fr$  the imposed LCC, and  $pct_{agr1}^*$  and  $pct_{for1}^*$  the new fractions of AL<sub>r</sub> and forest to compute the contribution from the differences in those two LCs. A GAM prediction at Site 1 with these updated LC fractions is then subtracted from the original prediction at Site 1 to isolate the actual contribution from the differences in the AL<sub>r</sub> and forest fractions, while excluding contributions from other factors, such as elevation. To further corroborate our results, we fit a GAM without the local LC fractions as predictors (blue terms in Extended Data Table 1). Note that the linear fits in Fig. 1, Extended Data Fig. 7, and Supplementary Fig. 10 are produced with the Matlab function 'fitlm' using the option 'RobustOpts' to be less affected by outliers when fitting. A discussion on the sensitivity of the GAM results on its structure is provided in Supplement E.

### Estimating the theoretical effect of forestation on precipitation

When estimating the forestation-induced precipitation change, we confine ourselves to foresting 20 % of the land surface, to (1) ensure the imposed LCC is of realistic magnitude and (2) prevent moving outside of the boundaries of the predictor space for the AL<sub>r</sub> and forest LC fractions. Further, the GAM estimate of  $\Delta P_{loc}$  was evaluated for foresting 20 % of the land surface with the site pair analysis. We determine the local effect of foresting 20 % of the land using one LC map with 10 % more forest and 10 % less AL<sub>r</sub> and a second LC map with 10 % less forest and 10 % more AL<sub>r</sub> relative to CLC. When creating these maps, we prevent LC fractions from falling below 0 % or above 100 %. In case it is impossible to convert 10 % of the land surface in one direction, we allow for more than 10 % conversion in the other direction up to a total conversion of 20 %:

$$\Delta LC_{tot} = \min(20\%, prox_{agr} + prox_{for}), \Delta LC_1 = \min(prox_{agr}, \max(10\%, \Delta LC_{tot} - prox_{for})),$$

$$\text{and } \Delta LC_2 = \min(prox_{for}, \max(10\%, \Delta LC_{tot} - prox_{agr})) \quad (3)$$

, where  $prox_{agr}$  and  $prox_{for}$  are the proximities AL<sub>r</sub> and forest in the CLC,  $\Delta LC_{tot}$  is the total LCC between the map with more forest cover and the map with more AL<sub>r</sub>, and  $\Delta LC_1$  and  $\Delta LC_2$  are the imposed changes in the AL<sub>r</sub> and forest fraction for the forested map and the deforested map, respectively.  $\Delta LC_1$  is then added to the forest fraction and subtracted from the fraction of AL<sub>r</sub> to create the forested LC map ( $LC_1$ ), while the opposite is done with  $\Delta LC_2$  to create the deforested map ( $LC_2$ ). To avoid the estimation of the forestation effect in regions with no AL<sub>r</sub> available for forestation in Scandinavia, we prevent forestation at grid cells where the average fraction of AL<sub>r</sub> within a moving window of 3.1° is below 8 %. With this procedure, it is possible to achieve 20 % forestation almost everywhere in Europe (Extended Data Fig. 6 a). Then, we predict precipitation with each of the two new LC maps ( $P_{pred}(LC_1)$  and  $P_{pred}(LC_2)$ ) and calculate the difference of the two predictions to estimate the local change in precipitation due to foresting 20 % of the land surface:

$$\Delta P_{loc} = P_{pred}(LC_1) - P_{pred}(LC_2) \quad (4)$$

Using the same procedure, we calculate the upwind LC fractions if the land was transformed according to the mentioned rules when processing the wind trajectories (Extended Data Fig. 6 b and c;  $uw\_pct\_LC_i(LC_1)$  and  $uw\_pct\_LC_i(LC_2)$ ). Again, the two GAM predictions for each of the resulting upwind LC fractions are subtracted to determine  $\Delta P_{dw}$ :

$$\Delta P_{dw} = P_{pred}(uw\_pct\_LC_i(LC_1)) - P_{pred}(uw\_pct\_LC_i(LC_2)) \quad (5)$$

### Realistic forestation scenario

Finally, we estimate the precipitation change of reforestation Europe according to the World Reforestation Potential Map<sup>40</sup> (WRPM), which estimates the reforestation potential relative to the LC state from 2000-2009. Starting from the original binary map indicating whether a location is suitable for reforestation at  $0.0083^\circ$  resolution, we calculate the area fraction which is suitable for reforestation at  $0.1^\circ$  resolution. This fraction we then add to the forest and subtract from the  $AL_r$  fractions of CLC, again making sure the  $AL_r$  cover fraction does not fall below 0 % and the forest fraction does not exceed 100 %. The estimate of the precipitation change due to this reforestation scenario then amounts to the difference of a GAM prediction with the forested local and upwind LC fractions minus a GAM prediction with local and upwind LC fractions according to CLC.

### Comparison to climate change signal

The EURO-CORDEX EUR-11 regional climate simulations, used for comparison to the climate change signal, stem from the CH2018 collection<sup>66</sup>. EURO-CORDEX simulations at  $0.44^\circ$  resolution are excluded from the analysis, since they generally represent present day precipitation fields worse than the ensemble at  $0.11^\circ$  resolution<sup>67</sup>. Compared to the original EUR-11 ensemble<sup>68</sup>, some simulations were removed in CH2018 following quality checks. We use data from nine different RCM/GCM combinations with a total of ten available ensemble members at  $0.11^\circ$  resolution for the RCP4.5 and RCP8.5 scenarios to calculate multi-model mean precipitation climatologies for the periods 1986-2015 and 2071-2100 (Supplementary Table 5, note that MPI-CSC-REMO2009/MPI-M-MPI-ESM-LR provides two ensemble members). The original data on a rotated grid were interpolated to a regular  $0.1^\circ$  grid using distance-weighted remapping of CDO. Then, we compute the area-weighted change in precipitation between 2071-2100 and 1986-2015 over five aggregated biogeographical regions compared to the climatology of 1986-2015. For the realistic forestation scenario, we compute the precipitation changes from this scenario compared to the MSWEP precipitation climatology of 1986-2015.

## References

50. Beck, H. E. *et al.* MSWEP: 3-hourly  $0.25^\circ$  global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrol. Earth Syst. Sc.* **21**, 589–615, DOI: [10.5194/hess-21-589-2017](https://doi.org/10.5194/hess-21-589-2017) (2017).
51. Beck, H. E. *et al.* Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrol. Earth Syst. Sci.* **21**, 6201–6217, DOI: [10.5194/hess-21-6201-2017](https://doi.org/10.5194/hess-21-6201-2017) (2017).
52. Beck, H. E. *et al.* Daily evaluation of 26 precipitation datasets using stage-IV gauge-radar data for the CONUS. *Hydrol. Earth Syst. Sc.* **23**, 207–224, DOI: [10.5194/hess-23-207-2019](https://doi.org/10.5194/hess-23-207-2019) (2019).
53. Lu, N. Scale effects of topographic ruggedness on precipitation over Qinghai-Tibet Plateau. *Atmos. Sci. Lett.* **20**, e904, DOI: [10.1002/asl.904](https://doi.org/10.1002/asl.904) (2019). <https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/asl.904>.
54. EEA. EU-DEM statistical validation (2014).
55. Siebert, S., Henrich, V., Frenken, K. & Burke, J. *Global Map of Irrigation Areas version 5*. Rheinische Friedrich-Wilhelms-University, Bonn, Germany / Food and Agriculture Organization of the United Nations, Rome, Italy (2013).
56. DeAngelis, A. *et al.* Evidence of enhanced precipitation due to irrigation over the great plains of the united states. *J. Geophys. Res.-Atmos.* **115**, DOI: [10.1029/2010JD013892](https://doi.org/10.1029/2010JD013892) (2010). <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2010JD013892>.
57. Thiery, W. *et al.* Present-day irrigation mitigates heat extremes. *J. Geophys. Res.* **122**, 1403–1422, DOI: [10.1002/2016JD025740](https://doi.org/10.1002/2016JD025740) (2017).
58. Wernli, B. H. & Davies, H. C. A lagrangian-based analysis of extratropical cyclones. I: The method and some applications. *Q. J. Roy. Meteor. Soc.* **123**, 467–489, DOI: [10.1002/qj.49712353811](https://doi.org/10.1002/qj.49712353811) (1997). <https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.49712353811>.
59. Smith, A., Lott, N. & Vose, R. The integrated surface database: Recent developments and partnerships. *B. Am. Meteorol. Soc.* **92**, 704–708, DOI: [10.1175/2011BAMS3015.1](https://doi.org/10.1175/2011BAMS3015.1) (2011). [https://journals.ametsoc.org/bams/article-pdf/92/6/704/3738664/2011bams3015\\_1.pdf](https://journals.ametsoc.org/bams/article-pdf/92/6/704/3738664/2011bams3015_1.pdf).

60. Blenkinsop, S., Lewis, E., Chan, S. C. & Fowler, H. J. Quality-control of an hourly rainfall dataset and climatology of extremes for the UK. *Int. J. Climatol.* **37**, 722–740, DOI: [10.1002/joc.4735](https://doi.org/10.1002/joc.4735) (2017).
61. Wood, S. N. *Generalized additive models : an introduction with R*. Chapman & Hall/CRC texts in statistical science (CRC Press/Taylor & Francis Group, Boca Raton, 2017), second edition edn.
62. Wood, S. N. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc.* **73**, 3–36, DOI: [10.1111/j.1467-9868.2010.00749.x](https://doi.org/10.1111/j.1467-9868.2010.00749.x) (2011). <https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9868.2010.00749.x>.
63. Wood, S. N., Li, Z., Shaddick, G. & Augustin, N. H. Generalized additive models for gigadata: Modeling the U.K. black smoke network daily data. *J. Am. Stat. Assoc.* **112**, 1199–1210, DOI: [10.1080/01621459.2016.1195744](https://doi.org/10.1080/01621459.2016.1195744) (2017). <https://doi.org/10.1080/01621459.2016.1195744>.
64. Li, Z. & Wood, S. N. Faster model matrix crossproducts for large generalized linear models with discretized covariates. *Stat. Comput.* **30**, 19–25, DOI: [10.1007/s11222-019-09864-2](https://doi.org/10.1007/s11222-019-09864-2) (2020).
65. Dormann, C. *et al.* Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography* **30**, 609–628, DOI: [10.1111/j.2007.0906-7590.05171.x](https://doi.org/10.1111/j.2007.0906-7590.05171.x) (2007).
66. CH2018. 2018 Climate scenarios for Switzerland. Tech. Rep., National Centre for Climate Services (2018).
67. Prein, A. F., Gobiet, A., Truhetz, H. *et al.* Precipitation in the euro-cordex 0.11° and 0.44° simulations: high resolution, high benefits? *Clim. Dynam.* **46**, 383–412, DOI: [10.1007/s00382-015-2589-y](https://doi.org/10.1007/s00382-015-2589-y) (2016).
68. Jacob, D. *et al.* EURO-CORDEX: new high-resolution climate change projections for European impact research. *Reg. Environ. Chang.* **14**, 563–578, DOI: [10.1007/s10113-013-0499-2](https://doi.org/10.1007/s10113-013-0499-2) (2014).
69. Defense Mapping Agency (DMA) & U.S. Geological Survey (USGS). 30" DEM from Digital Chart of the World. *in USGS, ed., 1997* (1996).

## Acknowledgements

We are thankful for the help and advice of Dr. Andreas Papritz, Dr. Franziska Scholder-Aemisegger, Michael Windisch, and Dr. Anina Gilgen. This work is part of the CLIMPULSE project, which was funded by the Swiss National Science Foundations (SNSF; <http://p3.snf.ch/Project-172715>; grant no. 200021\_172715) and the Swiss Federal Office for the Environment (FOEN).

The GSDR data analyzed in this study contain measurements of the following Institutes: Service Puplic de Wallonia, Finnish Meteorological Institute, Météo-France, Deutscher Wetterdienst, Met Éireann, Meteo Trentino, Agrometeorologico Siciliano, Autonome Provinz Bozen–Südtirol, Norwegian Meteorological Institute, Sistema Nacional de Informação de Recursos Hídricos, Instituto Português do Mar e da Atmosfera, Servei Meteorologic de Catalunya, Meteo Schweiz, UK Met Office (Met Office (2006): MIDAS UK Hourly Rainfall Data. NCAS British Atmospheric Data Centre, May 5th 2020. <https://catalogue.ceda.ac.uk/uuid/bbd6916225e7475514e17fdbf11141c1>), Environment Agency UK, the Scottish Environment Protection Agency (Contains public sector information licensed under the Open Government Licence v3.0), and Natural Resources Wales (Contains Natural Resources Wales information © Natural Resources Wales and database right. All rights reserved). The post-processed model data of CORDEX are provided by the Center for Climate Systems Modeling (C2SM), ETH Zurich (Jan Rajczak, Silje Soerland, Urs Beyerle, Curdin Spirig, and Elias Zubler).

## Author contributions statement

RM and ELD conceptualized the study with inputs from JS and SIS. RM and JS prepared the gridded data and conducted the GAM analysis. RM and MS calculated the wind trajectories with the Lagranto tool. RM and EL prepared and analyzed the GSDR station data. RM conducted the analysis of the GHCN data. RM drafted the manuscript with the help of ELD, JS, and SIS. All authors contributed to the interpretation of results and preparation of the text.

## Competing of interests

The authors declare no competing interests.

## Data availability

MSWEP v2.2 can be retrieved from [https://platform.princetonclimate.com/PCA\\_Platform/](https://platform.princetonclimate.com/PCA_Platform/). EU-DEM v1.1 can be downloaded from <https://land.copernicus.eu/imagery-in-situ/eu-dem/eu-dem-v1.1>,

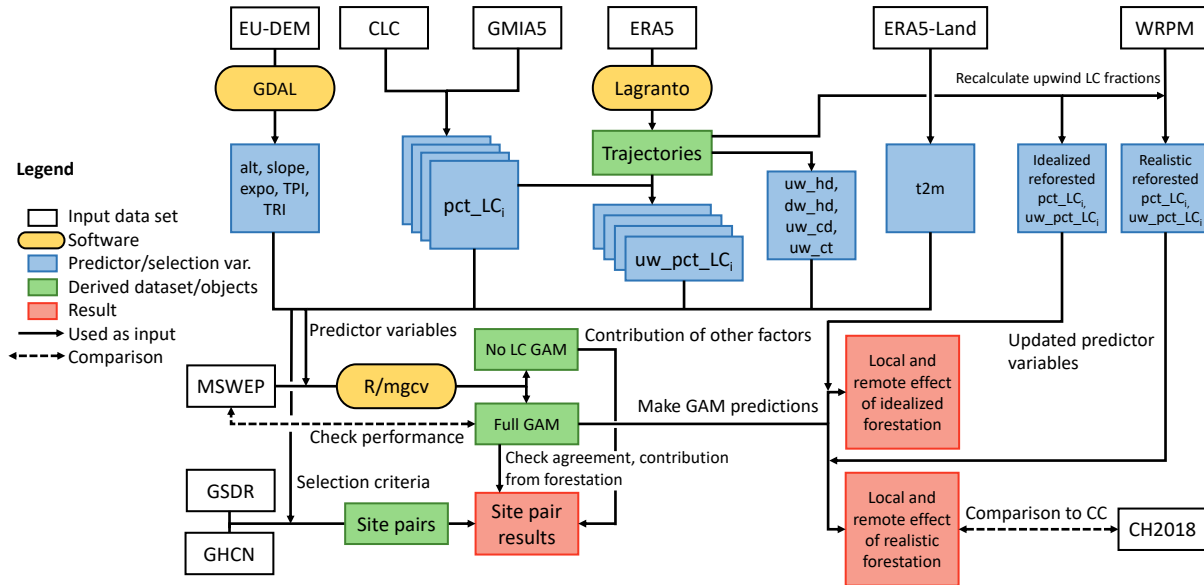


CORINE LC 2000 from <https://land.copernicus.eu/pan-european/corine-land-cover/clc-2000>, the Global Map of Irrigation Areas version 5 from <http://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas>, and the ERA5(-Land) reanalysis data from <https://cds.climate.copernicus.eu>. The GSDR station data are available upon request from Elizabeth Lewis ([Elizabeth.Lewis2@newcastle.ac.uk](mailto:Elizabeth.Lewis2@newcastle.ac.uk)), but require a permission to use the respective data from several institutions<sup>24</sup>. GHCN\_Daily v3.x can be retrieved from <ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily/>. The World Reforestation Potential map can be downloaded from <https://zenodo.org/record/883444>. To access the CH2018 collection follow the instructions at <https://doi.org/10.18751/climate/scenarios/ch2018/1.0>, where you also find a contact form to access specialized data. The base maps of all figures displaying Europe were generated with the standard 'landareas' file of Matlab, which is based on the Digital Chart of the World<sup>69</sup>. The source data files are available at the ETH Research Collection under doi:[10.3929/ethz-b-000448232](https://doi.org/10.3929/ethz-b-000448232) (not active yet).

## Code availability

The R library mgvc can be found at <https://cran.r-project.org/web/packages/mgcv/index.html> and the GDAL library at <https://cran.r-project.org/web/packages/rgdal/index.html>.

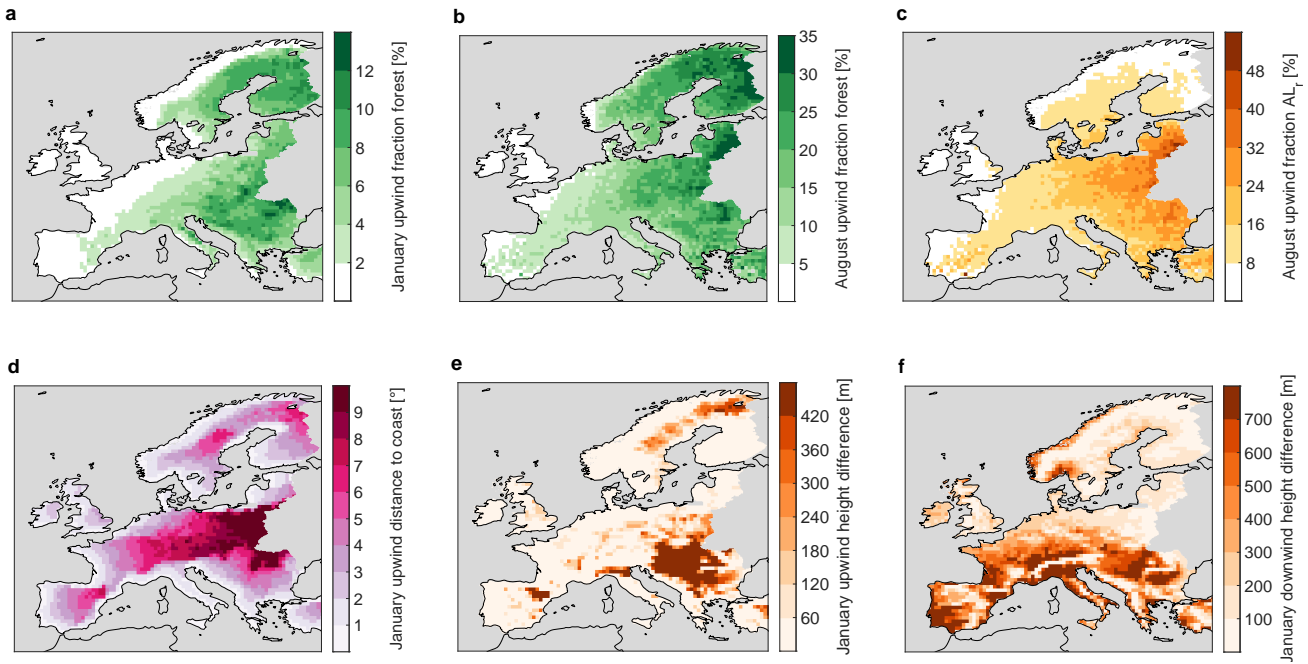
## Extended Data Figures and Tables



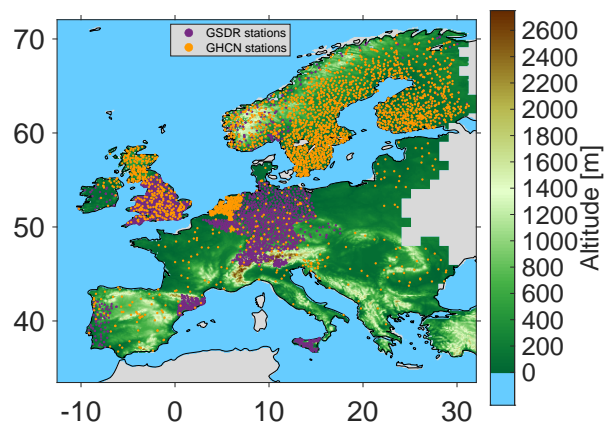
**Extended Data Figure 1. Overview of the steps, softwares, and input data sets described in the Methods and how they are interconnected.** White boxes, input data sets and yellow boxes softwares used in the study (Supplementary Table 2). The blue boxes show variables that were used as predictor variables in the GAMs and for the selection criteria in the site pair analysis, the green boxes derived data sets and objects, and red boxes the final results.

**Extended Data Table 1. Smooths used to construct the GAM.** First column, the predictor variables used in the respective smooth. Second column, names the type of smooth used. Third column, shows the maximum number of nodes of the smooth (higher k more flexibility) and column four the smoothing parameter applied (higher sp results in a smoother function). In blue smooths that were removed when fitting the 'no LC' GAM. Last column shows the maximum p-value across the GAMs for the individual months of the respective smooth and in brackets number of months for which this smooth had a p-value larger than 0.01. For the LC-related smooths values are given in order of forest, AL<sub>r</sub>, shrubland, natural low vegetation, artificial surfaces, irrigated agricultural land, natural bare land, open water, and wetland.

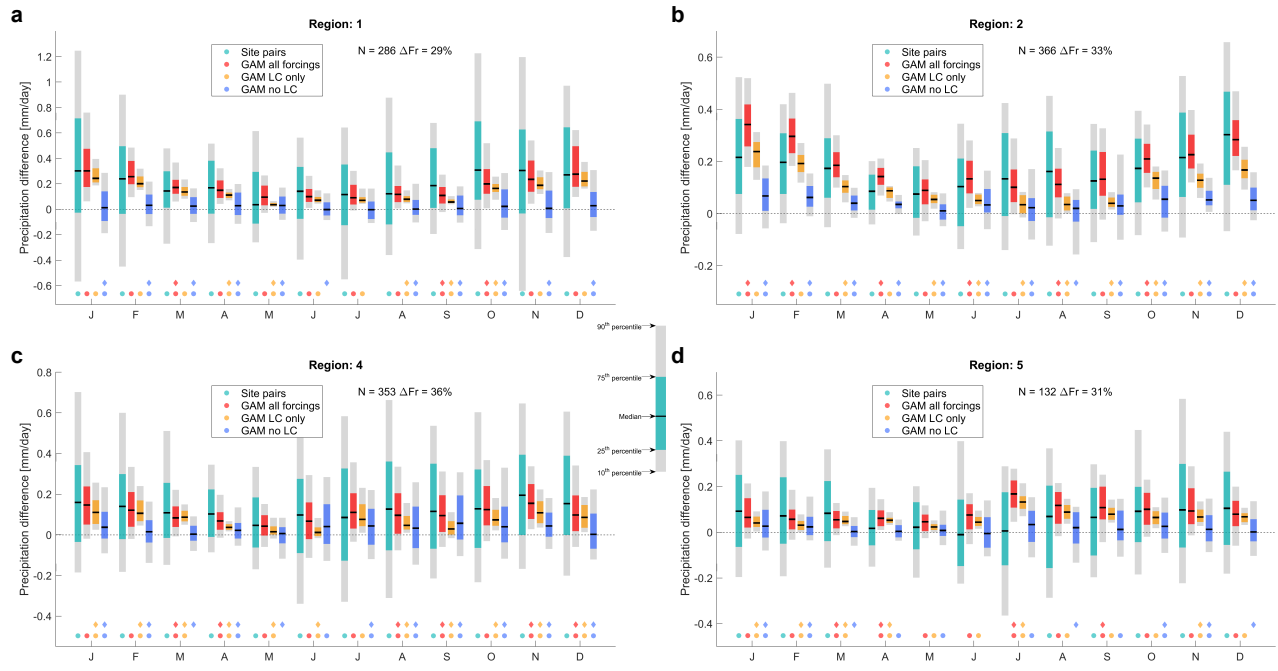
Variable(s)	Type	k	sp	Description	Max p-value (number of p-values above 0.01)
alt	s	15	estimated	global altitude term	3.1e-29 (0)
alt, lat, lon	ti	5	5	spatially varying altitude term	7.1e-19 (0)
TPI	s	10	estimated	global TPI term	7.4e-42 (0)
expo	s	10	estimated	global exposition term	1.3e-7 (0)
expo, lat, lon	ti	5	5	spatially varying exposition term	8.6e-22 (0)
slope	s	15	estimated	global slope term	1.9e-2 (1)
slope, lat, lon	ti	5	5	spatially varying slope term	7.2e-34 (0)
TRI	s	10	estimated	global TRI term	8.6e-2 (2)
t2m	s	20	estimated	global 2 m temperature term	1.0e-71 (0)
uw_hd	s	10	estimated	global upwind height difference term	1.2e-41 (0)
dw_hd	s	10	estimated	global downwind height difference term	8.8e-76 (0)
uw_cd	s	10	estimated	global upwind coast distance term	4.2e-45 (0)
prox_LCi	s	10	estimated	global term for local effect of each LC class	1.6e-4 (0), 5.1e-5 (0), 1.45e-13 (0), 8.2e-3 (0), 1.5e-1 (3), 1.2e-14 (0), 1.4e-4 (0), 2.2e-2 (1), 2.1e-8 (0)
prox_LCi, lat, lon	ti	3	5	spatially varying local effect of each LC class	2.6e-2 (1), 2.3e-6 (0), 1.3e-2 (1), 1.7e-1 (1), 1.1e-2 (1), 5.4e-1 (3), 2.7e-2 (1), 2.1e-4 (0), 9.8e-1(4)
prox_LCi, t2m, alt	ti	5	5	term varying with mean climate for local effect of each LC class	7.3e-27 (0), 1.3e-20 (0), 2.1e-9 (0), 5.0e-6 (0), 1.9e-2 (2), 1.7e-15 (0), 5.3e-12 (0), 3.2e-2 (1), 5.4e-8 (0)
uw_pct_LCi	s	10	estimated	global upwind term for each LC class	1.9e-15 (0), 5.5e-18 (0), 4.4e-11 (0), 1.5e-8 (0), 1.1e-24 (0), 2.9e-58 (0), 5.4e-19 (0), 2.0e-21 (0), 2.0e-2 (1)
uw_pct_LCi, lat, lon	ti	3	5	spatially varying upwind term for each LC class	6.0e-1 (2), 3.2e-2 (3), 1.8e-5 (0), 1.5e-9 (0), 2.2e-2(1), 3.6e-5 (0), 0.34e-2 (1), 2.2e-2 (2), 2.4e-1 (2)
uw_pct_LCi, t2m, uw_ct	ti	5	5	upwind term varying with mean climate for each LC class	4.9e-12 (0), 6.7e-7 (0), 1.9e-44 (0), 4.5e-33 (0), 5.0e-3 (0), 2.0e-60 (0), 1.1e-3 (0), 5.3e-3 (0), 2.0e-11 (0)
lat, lon	ti	15	5	spatial term to capture large scale variation of precipitation	0 (0)



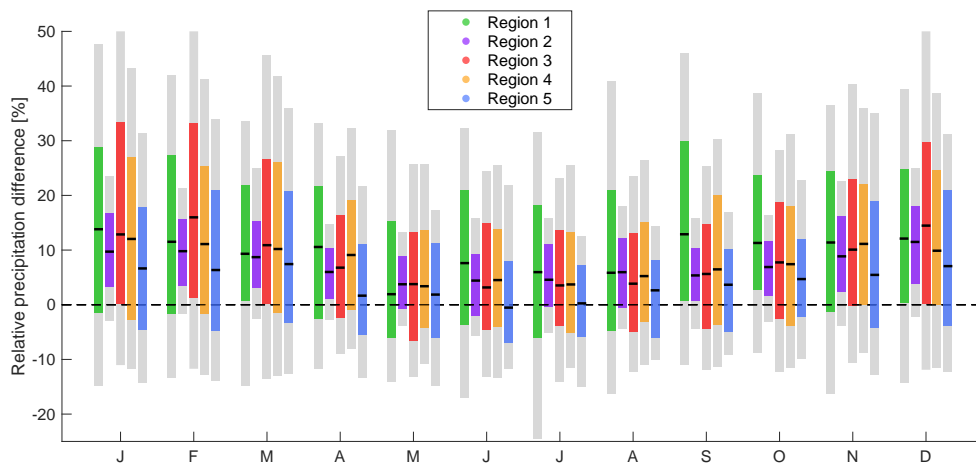
**Extended Data Figure 2. Examples of trajectory-based fields.** Upwind forest fraction in January (a) and August (b). Panel c shows the upwind fraction of  $AL_r$  in August. Panel d shows the upwind distance to coast, panel e the upwind height difference, and panel f the downwind height difference in January.



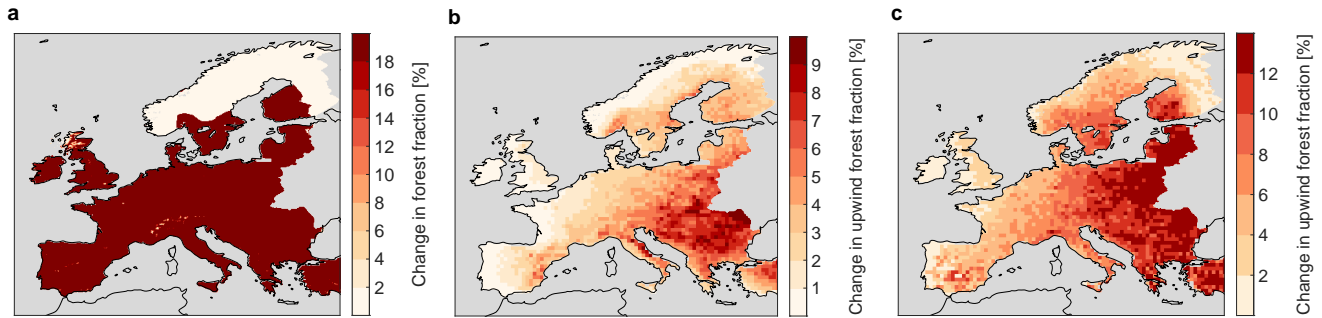
**Extended Data Figure 3.** Location of the GSDR (purple) and GHCN (orange) rain gauge stations included in the site pair analysis.



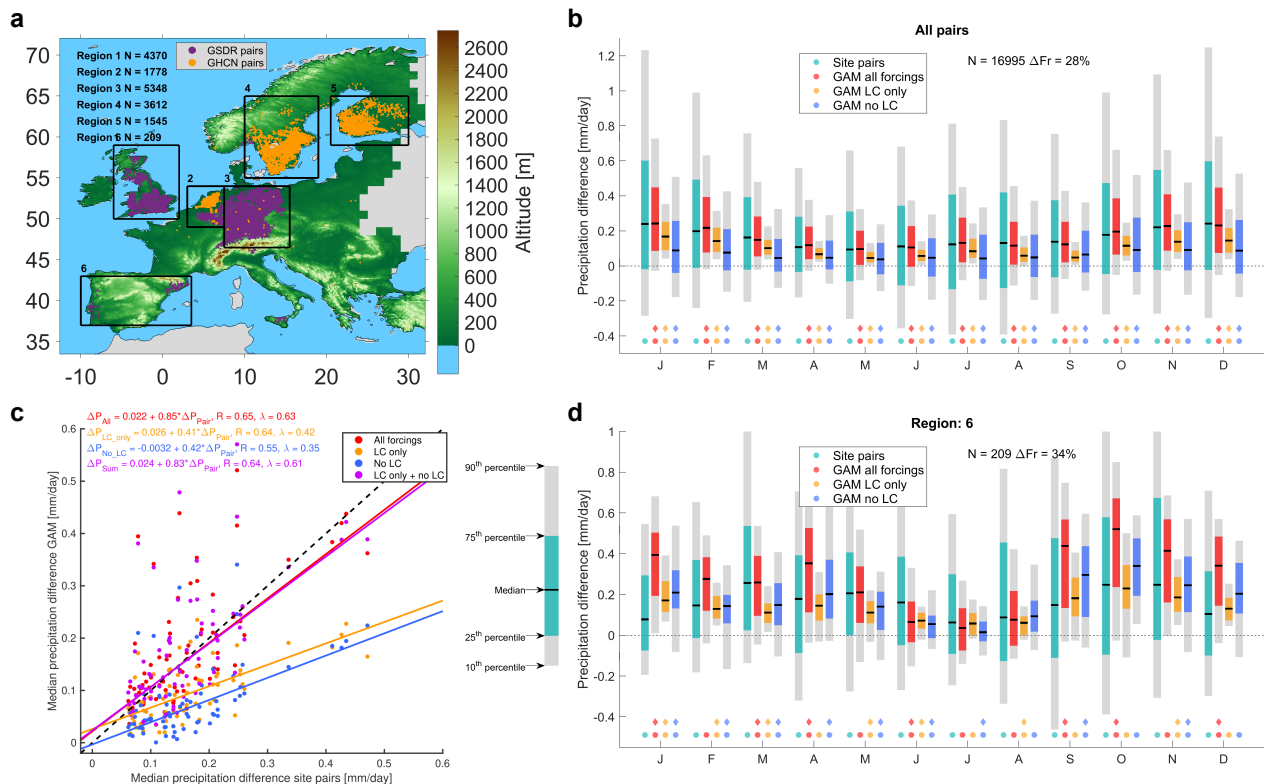
**Extended Data Figure 4. Local precipitation difference between forest and agricultural land at site pairs.** As Fig. 1 d but for Regions 1, 2, 4, and 5.



**Extended Data Figure 5. Relative local precipitation difference between forest and agricultural land at site pairs.** The median (black line), interquartile range (colored shading), and range between 10<sup>th</sup> and 90<sup>th</sup> percentile (grey shading) of  $\Delta P_{loc}$  in the rain gauge data sets as a fraction of the precipitation at the site with more AL<sub>r</sub> over the five regions in Fig. 1 a.



**Extended Data Figure 6. Changes in LC fractions for uniform forestation scenario.** Panel a, change in local  $AL_r$  and forest fractions used to predict the local effect of forestation with the GAM (calculated from Eq. 3). To the right, the change in the upwind forest fraction following 20% forestation in January (b) and July (c), retrieved from recalculating the upwind LC fractions in the trajectories with the altered LC maps.



**Extended Data Figure 7. Local precipitation difference between forest and agricultural land at site pairs with relaxed selection criteria.** As Fig. 1 but with relaxed selection criteria (ID 10 in Supplement D). Note that Panel d shows Region 6 instead of Region 3.