Environmental Research Letters

LETTER • **OPEN ACCESS**

Assimilating multi-satellite snow data in ungauged Eurasia improves the simulation accuracy of Asian monsoon seasonal anomalies

To cite this article: Peirong Lin *et al* 2020 *Environ. Res. Lett.* **15** 064033

View the article online for updates and enhancements.

This content was downloaded from IP address 144.173.6.94 on 16/07/2020 at 09:16

Environmental Research Letters

LETTER

Assimilating multi-satellite snow data in ungauged Eurasia improves the simulation accuracy of Asian monsoon seasonal anomalies

Peirong Lin[1,6], **Zong-Liang Yang**[1,2,8], **Jiangfeng Wei**[3], **Robert E Dickinson**[1,7], **Yongfei Zhang**[4] **and Long Zhao**[5]

[1] Jackson School of Geosciences, The University of Texas at Austin, Austin, TX, United States of America
[2] Key Laboratory of Regional Climate–Environment Research for Temperate East Asia, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, People's Republic of China
[3] Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters/Key Laboratory of Meteorological Disaster, Ministry of Education/International Joint Research Laboratory on Climate and Environment Change, Nanjing University of Information Science and Technology, Nanjing, People's Republic of China
[4] Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, United States of America
[5] Chongqing Engineering Research Center for Remote Sensing Big Data Application, School of Geographical Sciences, Southwest University, Chongqing, People's Republic of China

E-mail: liang@jsg.utexas.edu

Keywords: Asian monsoon, dynamical seasonal forecast, multi-satellite snow data assimilation, GRACE, MODIS

Supplementary material for this article is available online

OPEN ACCESS

RECEIVED
2 October 2019

REVISED
24 February 2020

ACCEPTED FOR PUBLICATION
18 March 2020

PUBLISHED
5 June 2020

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Abstract

Properly initializing land snow conditions with multi-satellite data assimilation (DA) may help tackle the long-standing challenge of Asian monsoon seasonal forecasts. However, to what extent can snow DA help resolve the problem remains largely unexplored. Here we establish, for the first time, that improved springtime snow initializations assimilating the Moderate Spectral Imaging Satellite (MODIS) snow cover fraction and the Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage data can improve the simulation accuracy of Asian monsoon seasonal anomalies. Focusing on the western Tibetan Plateau (TP) and mid- to high-latitude Eurasia (EA), two regions where multi-satellite snow DA is critical, we found that DA influences the monsoon circulation at different months depending on the regional snow–atmosphere coupling strengths. For the pre-monsoon season, accurate initialization of the TP snow is key, and assimilating MODIS data slightly outperforms jointly assimilating MODIS and GRACE data. For the peak-monsoon season, accurate initialization of the EA snow is more important due to its long memory, and assimilating GRACE data brings the most pronounced gains. Among all the Asian monsoon subregions, the most robust improvement is seen over central north India, a likely result of the region's strong sensitivity to thermal forcing. While this study highlights complementary snow observations as promising new sources of the monsoon predictability, it also clarifies complexities in translating DA to useful monsoon forecast skill, which may help bridge the gap between land DA and dynamical climate forecasting studies.

1. Introduction

The Asian monsoon affects more than 60% of the world's population with profound socio-economic implications [1], yet its skillful seasonal forecast, with a lead time of one to several months, remains very challenging [2–4]. Current state-of-the-art dynamical climate forecasts put emphasis on the memory of tropical oceans and sea surface temperatures (SSTs) data assimilation (DA) for the monsoon predictability [5–9]. However, properly intializing land conditions to fully utilize the weeks' to months' of land memory [10] for the monsoon forecast is largely unexplored. Despite increased satellites monitoring of global land snow, soil moisture, and vegetation conditions [11–13], land is still the least-constrained

[6] Current address: Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ, United States of America
[7] Current address: University of California, Los Angeles, CA, United States of America
[8] Author to whom any correspondence should be addressed.

© 2020 The Author(s). Published by IOP Publishing Ltd

earth system component by observations in dynamical climate models [14, 15]. This has been limiting our understanding of the intrinsic limit of seasonal monsoon predictability [10], as well as the additional skill potentially attainable from land DA [16, 17].

The link between Eurasian snow and the Asian monsoon has long been recognized [18]. Positive (negative) snow anomalies tend to be followed by negative (positive) monsoon anomalies [19], highlighting the need to properly initialize Eurasian snow for skillful monsoon forecasts. However, previous snow DA studies have mainly focused on assimilating satellite data to improve point- to basin-scale snow estimates for hydrological predictions [20–22], while global multi-satellite snow DA tailored to addressing dynamical climate forecasting problems have been relatively rare [23, 24]. Station-based snow depth and satellite snow cover have been assimilated in some climate reanalyses [13, 25], but they cannot be separated from these reanalyses to establish their contribution to climate forecasts. The lack of data has resulted in a lack of research efforts tailored to understanding monsoon forecast skill from the snow DA perspectives. Recently, Zhang *et al* [23] and Zhang and Yang [26] (hereafter Z1416) assimilated the Moderate Resolution Imaging Spectroradiometer (MODIS) snow cover fraction (SCF) and the disaggregated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage (TWS) data into a state-of-the-art land surface model (LSM). The resulting two global snow DA products potentially contain improved snow cover information from optical observations and improved snow mass information from gravity measurements, providing valuable data sources to tackle the unresolved challenge of dynamical monsoon forecast. Lin *et al* [16] used these products to initialize the land components in a suite of ensemble-based climate modeling experiments, and found the temperature prediction can be improved for the Tibetan Plateau (TP) and the high-latitude Eurasia (EA) regions. However, to the best of our knowledge, no existing studies have utilized similar satellite-constrained snow data for monsoon forecast studies.

This paper pioneers the use of multi-satellite snow DA data for seasonal forecast of precipitation in the Asian monsoon region. Our analyses put emphasis on snow DA over the TP and the EA, two regions featuring sparse rain gauges and pronounced snow DA updates, to demonstrate the mechanistic pathway for snow DA to influence the monsoon forecast. We also scrutinize the impacts of assimilating different satellite data, to understand how to maximize skill improvement from different sorts of snow-related satellite observations that may be assimilated in the future [24, 27]. Additionally, as the methodology development and refinement of global snow DA studies is an area of active research [22–24], we also offer a new perspective to assess the accuracy of DA products by independently analyzing the precipitation forecast skill, whereas conventional DA assessments commonly use highly uncertain snow observational data as references and thus have been proven problematical [24, 26]. Our analyses focus on springtime snow initialization (March 1st), because the role of snow in seasonal climate prediction is much less understood during the cold-to-warm season transition [16, 28]. This transitional time coincides with the so-called 'spring barrier' [29] when SSTs are limited in their predictability, thereby rendering it critical to seek for new sources of predictability. Although there is no guarantee that new sources of predictability exist, it is worth examining the issue since improved prediction would be quite important for people who live in the monsoon region.

This study shows, for the first time, that the simulation accuracy of Asian monsoon seasonal anomalies can be improved by assimilating complementary multi-satellite data on snow cover fraction and snow mass. The results suggest that multi-satellite snow DA may be an underutilized source of predictability, which can offer detectable improvements to the Asian monsoon seasonal forecast. Moreover, satellite-dependent, region-dependent, lead time-dependent, and the DA uncertainty-related skill improvement are also carefully examined through our experimental designs. This study may help bridge the gap between land DA and seasonal climate prediction studies, with potentially wide socio-economic implications.

2. Methods and materials

2.1. Models and experimental design

We design five suites of ensemble-based climate modeling experiments (table 1) to assess the impact of snow DA on monsoon forecast. OL, MOD, and GRAMOD stand for the 'open-loop' experiment, the experiment that has assimilated MODIS data, and the experiment that has assimilated both MODIS and GRACE data, respectively. The National Center for Atmospheric Research (NCAR) Community Earth System Model (CESM v1.2.1) is used to perform the coupled land–atmosphere 'hindcast' simulations. Each suite of experiments is initialized on 1 March of 2003–2009 (the overlapping 7 years for MODIS and GRACE) due to data availability of the snow DA products. The model is integrated for six months, i.e. two seasons' lead time. SSTs and sea ice are prescribed using the Hadley Centre monthly data [30]. The coupled atmosphere-land model uses the Community Atmosphere Model (CAM5) and the Community Land Model (CLM4), respectively, as CLM4 is the host land model for Z1416, and CAM5 is the most updated atmospheric model in CESM v1.2.1. The model resolution is $0.9^\circ \times 1.25^\circ$. Z1416 produced 40-member ensemble land model outputs, and the ensemble average is computed as the land

Table 1. Experimental design for the ensemble-based coupled land–atmosphere experiments. The model is initialized on 1 March of 2003 to 2009, and is integrated for six months through the end of August. SSTs and sea ice are prescribed using the Hadley Centre monthly data. Time-shifted ERA-Interim data are used to create eight-member ensemble for the atmospheric initialization. See more details in Methods and similar configurations in Lin *et al* [16]. Details of the DA implementations can be found in Z1416 [23, 26].

	Land initialization
Role of multi-satellite DA:	
OL (Exp1)	'Open-loop' experiment: land initialization without assimilating satellite data
MOD (Exp2)	"MODIS" DA experiment: land initialization with global assimilation of the MODIS SCF data [23]
GRAMOD (Exp3)	"GRAMOD and MODIS" DA experiment: land initialization with joint assimilation of the GRACE TWS and MODIS SCF data [26]
Role of regional DA (figure S2 (stacks.iop.org/ERL/15/064033/mmedia)):	
TP_only (Exp4)	Land initialization with GRAMOD applied to the TP only (figure 1); Other regions' land initialization is the same as OL
EA_only (Exp5)	Land initialization with GRAMOD applied to the EA only (figure 1); Other regions' land initialization is the same as OL

initial condition. Z1416 used the *Ensemble Adjustment Filter (EAKF)* for DA; the prior distribution for the model states (e.g. snow) is estimated using 40-member ensemble forcing to feed the model [23]. More DA details can be found in Z1416 [23, 26]. We caution the readers that these experiments are different from actual dynamical seasonal forecasts that initialize all components for fully coupled runs; instead, the experiments are designed to provide an isolated investigation of the land impacts only. Koster *et al* [31] designed similar modeling experiments to assess land impacts, and our experiments are different from theirs in that the SSTs are also prescribed with reanalysis data while Koster *et al* [31] used climatology SSTs.

To consider atmospheric chaos in the coupled forecast, the atmospheric component is initialized with time-shifted ERA-Interim data at eight time points (0000 UTC within ±4 d of 1 March), where the data are pre-processed to CAM5-consistent grids. This ensemble generation technique is commonly used in operational seasonal climate forecasting [32], producing eight members for ensemble-based experiments. The same land initialization is consistently used across the eight members. More details of the model configurations can be found in Lin *et al* [16].

2.2. Regional DA experiments

In addition to the first three experiments to examine the benefits of assimilating different satellite observations (MODIS and GRACE), we also design two other suites of experiments to understand the relative importance of regional snow DA (Exp 4–5, table 1). To apply DA updates to a specific region, the columns, plant functional types (PFTs), and land units in the DA land restart files are first modified where variables outside the predefined geographic bounding box are replaced with those from the OL restart file. To ensure the replaced variables are physically consistent with the OL snow conditions without imposing model imbalance errors, we also replace snow-related variables for water balance calculation, and soil/vegetation/temperature variables for energy balance calculation with OL. The geographic boxes for regional DA are defined in figure 1; the initial snow depth difference between regional DA and OL is shown in figure S2 to illustrate the effectiveness of regional DA. In total, 280 six-month model runs (i.e. 5 suites × 8-member ensembles × 7 years). The model time step is daily, and the total simulation data size is ~15 TB.

2.3. Statistical analyses and reference datasets

Due to the limited available snow DA products (7 years from 2003 to 2009) and possible noise from the chaotic atmospheric background, we jointly use the r^2 and *RMSE* metrics, bootstrapping statistical tests, and five reference precipitation datasets to ensure the robustness of the detected signal. Equations to calculate the skill metrics are shown below for one reference precipitation; the same equations are used for the other four additional reference datasets.

For each season (i.e. three month) into the forecast period, r_i^2 for each ensemble member is first computed at each grid cell using 21 data samples (3 months × 7 years) obtained from modeled and observed precipitation, denoted as $\boldsymbol{P}_m$ and $\boldsymbol{P}_o$, respectively (equation (1); $\delta\boldsymbol{P}_m$ and $\delta\boldsymbol{P}_o$ denote the standard deviation). r^2 is calculated by averaging r_i^2 for the 8 ensemble members (equation (2); n refers to the ensemble size), and the skill difference between DA and OL is computed using equation (3). Negative correlation is set to zero before the square operation in equation (1) to reduce sampling noise [33], although we found the influence of this procedure is minimal to our results.

$$r_i^2 = \left(\frac{cov(\boldsymbol{P}_m, \boldsymbol{P}_o)}{\delta\boldsymbol{P}_m \cdot \delta\boldsymbol{P}_o} \right)^2 \quad (1)$$

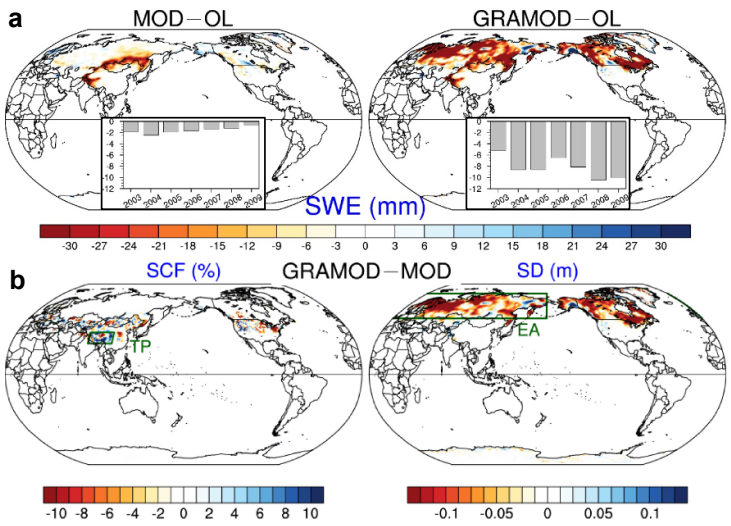

Figure 1. DA-introduced changes in snow initializations. (a) initial snow water equivalent (SWE) difference between MOD and OL (left), GRAMOD and OL (right); (b) initial snow cover fraction (SCF) and snow depth (SD) difference between GRAMOD and MOD (i.e. changes induced by assimilating GRACE data). The calculation is performed using 7 years averages on 1 March 2003 to 2009; bar charts show the year-to-year variation in the global area-weighted averages.

$$r^2 = \frac{1}{n}\left(\sum_{i=1}^{i=n} r_i^2\right) \tag{2}$$

$$r^2_{diff} = r^2_{DA} - r^2_{OL} \tag{3}$$

To test the null hypothesis that r^2_{diff} is not significantly different from zero, at each grid cell, a 1000-member bootstrapping is performed to randomly resample the 21 values in the $\boldsymbol{P}_m$ vector in equation (1). This leads to 1000 $r^2_{diff,shuffle}$ values with equations (2)–(3). The null hypothesis is rejected if the percentage of $r^2_{diff,shuffle}$ (α) is exceeded by r^2_{diff} with a two-tailed distribution. Since we used 21 data samples (3 months $\times$ 7 years), the skill in this analysis incorporates both *interannual* and *month-to-month* variability; greater r^2 suggests better explained total variance. Thus, the interpretation for the skill should differ from those using anomaly correlation skill, which only deals with *interannual* variance (see discussions in **Results** and **Text S3**).

RMSE is additionally computed with equation (4) (N = 21 for sample size; n = 8 for ensemble size). Compared to r^2, *RMSE* is more obscured with model and observational biases [31]. However, DelSole and Shukla [34] showed that model biases and skills are interrelated, which necessitates reduced *RMSE* as another target. Therefore, *RMSE* is jointly used in this study. To reduce the noise existing in the raw *RMSE* field, five reference datasets are used to compute *RMSE* with equations (4)–(5), and only the consensus percentage RMSE (*PRMSE*) is presented in figure S5—we define a consensus only when the signs of $PRMSE_{diff}$ are agreed among $\geq$4 references. Once a consensus is reached, $PRMSE_{diff}$ is averaged among the agreed observations as the final results. Regions with annual climatological precipitation $\leq$200 mm are masked (figure S5).

$$RMSE_s = \frac{1}{n}\left(\sum_{i=1}^{i=n}\left[\sqrt{\frac{\sum_{j=1}^{j=N}\left(\boldsymbol{P}^j_m - P^j_o\right)^2}{N}}\right]\right) \tag{4}$$

$$PRMSE_{diff} = (RMSE_{DA} - RMSE_{OL})/RMSE_{OL} \tag{5}$$

The five reference monthly precipitation datasets include data from the Global Precipitation Climatology Project (GPCP), the Global Precipitation Climatology Center (GPCC), the Tropical Rainfall Measuring Mission (TRMM), the University of Delaware gridded precipitation (UDEL), and the National Oceanic and Atmospheric Administration (NOAA) Precipitation Reconstruction over Land (PRECL). All datasets are interpolated to the model resolution before assessment.

3. Results

3.1. Multi-satellite snow DA is critical over ungauged Eurasia: the TP and the EA

We first delineate regions with prominent DA updates by showing the snow water equivalent (SWE) updates at the forecast starting date (figure 1, calculated as DA minus OL). For most terrestrial land surfaces, DA reduces the OL-simulated SWE (red colors, figure 1(a)), and only confined areas of the Rocky Mountains and the Alaska show increased SWE (blue colors). This DA behavior is in line with alleviating the positive snow bias at the high-latitude [35] and

the cold bias problem over the TP [36] as previously reported. Global area-weighted averages (bar plots, figure 1(a)) suggest DA is more impactful in certain years than others, such as 2004–2005 and 2008–2009.

It is noteworthy that assimilating the two satellite observations leads to different spatial patterns—GRAMOD brings prominent snow depth (SD) updates to the high latitudes (figure 1(b), right) while MOD does not. This is because GRACE uniquely measures the Earth's gravity change, and so it contains snow mass information that is absent in optical sensor observations, e.g. the MODIS snow cover fraction (SCF), especially when snow cover reaches ~100%. Over the TP, however, MOD more effectively reduces SCF than GRAMOD (figure 1(b), blue), which suggests inconsistent MODIS and GRACE DA updates as previously reported [16]. We note figure 1 and figure S1 only show the DA and OL difference, but it does not aim to address whether the snow estimates by DA are necessarily closer to the truth or not, or which satellite DA is better. This is partly to avoid repeating previous assessments that have reported MODIS DA issues over boreal forests [23] and GRACE DA issues over certain basins where TWS disaggregation is complex [26]. Moreover, due to the unknown 'truth', conventional DA assessments have been ambiguous when different observational references were used [24, 26]. Therefore, we will independently assess the DA uncertainty and effectiveness through the coupled modeling results of this study in Section 3.4.

Globally, the most prominent DA updates occur in the western TP, snow-covered and snow-free transitional areas (50°N–55°N), and the higher latitudes (figure 1). These geographic regions overlap with regions where models have high uncertainty in simulating snow, due to their relatively high seasonal snow variations and sparse rain gauge observations to serve as accurate meteorological inputs. DA mathematically weighs the relative uncertainty between model simulations and satellite observations [23], and hence greater DA updates are introduced there. The next sections focus on the TP and the EA (green boxes in figure 1) for further analyses due to their relatively well-known mechanistic link with the Asian monsoon [37]. Moreover, the TP and the EA, without constraints from satellite DA, were previously shown to have limited land-derived climate predictability [31]. As DA provides critical constraints to the land snow states there, we hypothesize that additional gains may be expected to complement the best-attainable skill reported in previous studies.

3.2. Contrasting thermodynamic effects introduced by snow DA in Tibet and EA

To facilitate physical understanding, DA-introduced SWE change is viewed as external forcing brought into the system, which persists into the forecast months. DA reduces snow in TP and EA (negative SWE_{diff}, figure 2(a)), leading to more absorbed shortwave radiation (reduced albedo) and less latent cooling (reduced warm-season snow melt). Hence, the net thermal effect of DA over these regions is equivalent to a positive sensible heating that warms up the local atmosphere (figure 2(b)). Interestingly, this sensible heating starts to warm up the atmosphere in different months over the two regions. During March and April, the dominant DA forcing is seen over the TP (figure 2(b), blue lines in MOD and GRAMOD), whereas in May and June, the dominant DA forcing is shifted to the EA (figure 2(b), red lines, only prominent in GRAMOD). This timing difference is explained by the regional snow-atmosphere coupling differences [16, 38]—the TP receives strong incident solar radiation to immediately translate changes in snow to atmospheric responses. By contrast, strong snow-atmosphere coupling in the EA only becomes established when solar radiation increases in warmer months to drive snow melts (although this EA forcing is only prominent and in GRAMOD still summer; see figure S3 for the forced air temperature responses to July). Evidently, although the springtime snow DA is applied globally, due to the snow-atmosphere coupling that differs from region to region, the period over which DA has a significant thermal effect to the atmosphere also differ, so as to the assimilation of optical versus gravity satellite data.

We next analyze the monsoon circulation changes because the thermal forcing by snow DA is expected to influence the land–sea thermal contrast (LSTC), which is the fundamental driver of the monsoon circulation [41] (additional assessments on the monsoon onset timing are provided in table S1 and text S1). To simplify the characterization of the highly complex Asian monsoon dynamics, two widely used monsoon circulation indices are computed (more in Text S2). One is the WY index [39] (defined as the zonal wind shear between U850 and U200) and the other is the monsoon Hadley circulation index, MH index [40] (defined as the meridional wind shear between V850 and V200). Among the three experiments, GRAMOD generally shows the least biased circulation indices (figures 2(c)–(d)) with the least uncertainty in May and June (figures 2(c)–(d)); this overlaps with the time when the EA forcing dominates (figure 2(b)) and the circulation shifts direction for monsoon onset. Compared to GRAMOD, the contribution of MOD is more consistently seen in the meridional (MH index) than in the zonal direction (WY index). This may be due to that the MOD signal is geographically confined to the TP, different from the widespread, zonally aligned EA forcing. Yet due to the high elevation of TP, the additional 4.8 W m^{-2} equivalent sensible heating (5%–6% of the total TP springtime sensible heating) can still exert strong controls on the monsoon meridional circulation [1, 42], although this signal quickly decreases after May (figure 2(b)). In June and July,

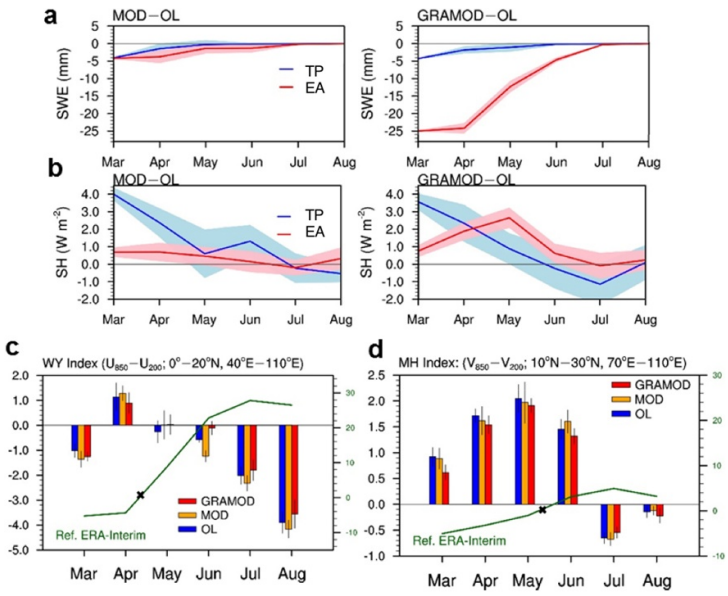

Figure 2. Persistence of DA-introduced thermal forcing and influence on the Asian monsoon circulation indices. (a) and (b) show persistence of the SWE changes and the sensible heating (SH) changes, respectively, for the TP (blue) and the EA (red). (c) and (d) show model-predicted monsoon circulation indices biased against the ERA-Interim reanalysis (zero: no bias). Crosses mark the time when the circulation shifts direction for monsoon onset in ERA-Interim (zero on the right *y*-axis). Vertical lines and shading denote one standard-deviation uncertainty. The WY index [39] and the MH index [40] respectively summarize the Asian monsoon broad-scale circulation in zonal and meridional directions.

GRAMOD becomes more effective in reducing the monsoon circulation bias than May and June, which may be ascribed to the latent heat release aloft the Indian subcontinent (figure S4). This condensation-related mid-troposphere warming was known to play important roles in monsoon strengthening in a positive feedback [43], which potentially explains the pronounced impact of GRAMOD on the monsoon circulation.

3.3. Robust skill improvement for monsoon forecast for central north India

We then quantify the monsoon forecast skill obtained from snow DA. To ensure the results robustness, we jointly use r^2 and *root-mean-square-error (RMSE)* metrics, five precipitation reference datasets, and bootstrapping techniques (**Methods**) for the analysis. Since the results are in general consensus in terms of spatial pattern (see figure S5), we only present r^2 with one reference data for simplicity in figure 3. Across four forecast seasons, significantly improved forecasts (*dotted*, $p < 0.05$) can be seen over the Southeast Asia (Myanmar and Thailand), southern China (near the Mei-Yu front), and the Indian subcontinent. The onset of the Asian monsoon typically occurs over the Indochina peninsula in late April, over southern China in mid-May, and over the Indian summer monsoon (ISM) in May to June [44]; the improvement for each of these regions covers a time span after the regional monsoon onset. The only exception is for the JJA ISM forecast in MOD, which is a satellite-dependent behavior that will be discussed in Section 3.4.

Among all the Asian monsoon sub-regions, central north India (CNI) features the most robust improvement independent of the metrics and precipitation references used (green box in figure S5). CNI is the region that receives the most rainfall in the ISM [46]. Hence, an improved area-averaged precipitation forecast by ~8% (figure S6(a)) could have profound agricultural and socio-economic implications. CNI is part of the South Asian Summer Monsoon, which was shown to exhibit the strongest sensitivity to LSTC changes among all global monsoons [45]. Therefore, it may be subject to less complex interplays among boundary forcings [43, 47] compared to the East Asian Summer Monsoon, potentially explaining its most robust improvements. For CNI, DA can reduce the *RMSE* by up to 30% and increase r^2 by up to 15% (Taylor diagram in figure S6), however, it is important to note that the increased variance consists of both *interannual* and *month-to-month variability* components (**Methods**). Further decomposition (figure S7) shows that DA increases the explained *interannual* variance only for the short-lead but not for the long-lead forecasts, while bias is consistently reduced at all lead times (figure S6). Although *interannual* variance is arguably the most critical target for skill improvement and our limited data samples prevent a clear separation of this, DelSole and Shukla [34] showed that model bias is intrinsically related to its forecast skill. To this end, our results are still promising for the CNI forecast as a pilot demonstration with snow DA, but future studies should use more DA samples to improve this skill quantification (see text S3 for a discussion).

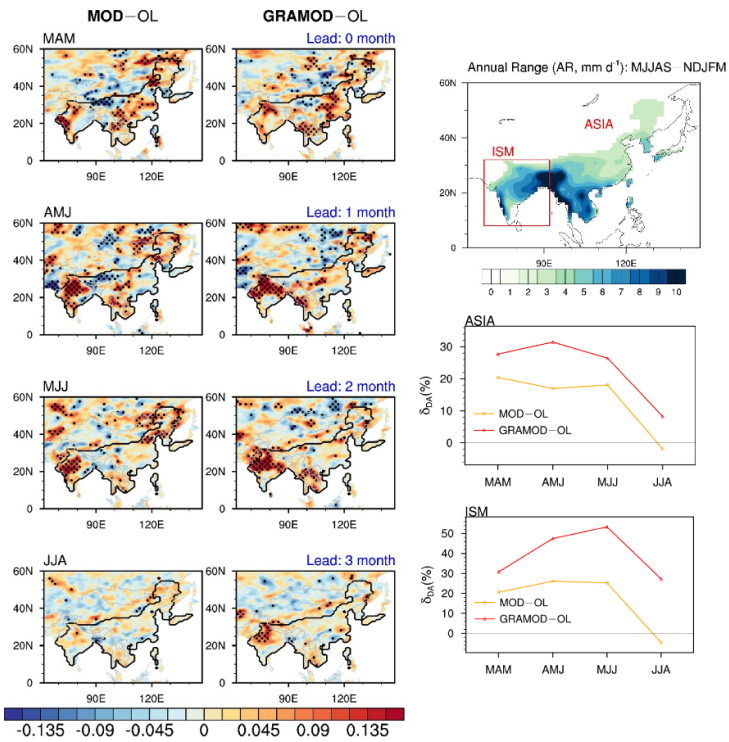

Figure 3. Precipitation forecast skill derived from snow DA (r^2_{DA} minus r^2_{OL}). (a) shows the skill from MOD (left) and GRAMOD (right) over the Asian monsoon region, in which the monsoon boundary (ASIA, black line) is outlined as the absolute precipitation annual range (MJJAS–NDJFM) exceeding 2.5 mm d^{-1}, following Kitoh *et al* [45]. (b) shows δDA for the Asian monsoon (ASIA) and the Indian Summer Monsoon (ISM); δDA is calculated as the percentage area with significant improvement minus that with significant degradation (dots: 95% confidence level with bootstrap, see Methods). The reference precipitation for figures 3–5 is the Global Precipitation Climatology Project (GPCP) product; results by using five reference precipitation datasets are shown in figure S5.

Some locations with degraded skill are also noted, such as the edge of the Asian monsoon boundaries (blue colors, figure 3(a)). This may be partly due to the known issues with Z1416 [23, 26], and partly be related to the possible atmospheric/sampling noise. Since a perfect improvement is by no means expected, we define δDA as percentage area with significant improvement minus that with degradation [28], to more objectively summarize the DA contribution. A positive (negative) δDA suggests DA improves (degrades) a forecast, while zero suggests no net contribution from DA. For the Asian monsoon with an area extent of 11.29×10^6 km^2 (figure 3(a)), δDA is +20% to +30%, with slightly higher values in GRAMOD (figure 3(b)). For the ISM (areal extent: 4.3×10^6 km^2), the strongest signal (δDA exceeds +50% in GRAMOD) is seen during the monsoon transitional period, overlapping with when the EA forcing signal dominates (figure 2). δDA is generally small for non-monsoonal areas; highlights for some other global monsoon regions are summarized in figure S8.

3.4. Relative importance of satellite data (MODIS and GRACE) in the ISM forecast

To understand the DA uncertainty and its impact on monsoon forecast, here we focus on the precipitation skill over the Indian subcontinent, where the most robust signal is located. In addition to MOD and GRAMOD (figure 4(a)), we isolate the exclusive contribution of assimilating GRACE data in figure 4(b), which shows GRACE DA brings significant improvement to the two- to three-month lead forecast (MJJ and JJA). However, GRACE DA barely offers improvement, or even slightly offsets the contribution from MODIS, at short-lead time in spring (blue in figure 4(b), MAM and AMJ). Since the springtime improvement is mostly related to the forcing over the TP (figure 2(b)), the slightly degraded skill indicates that GRAMOD may have provided inferior snow estimates over the TP, compared to MOD. In fact, when conducting multi-satellite snow DA, Zhang and Yang [26] applied no rules of preference if MODIS and GRACE data show opposite signs in updating snow (see figure 1(b) for the inconsistency problem over the TP). However, due to the unknowability of 'truth' [26], it was inconclusive as to which DA method produced better snow estimates. Here, by examining the precipitation forecast, and similarly the temperature forecast by Lin *et al* [16], we are able to infer that over the TP, only assimilating MODIS data produced better snow estimates than jointly assimilating GRACE and MODIS data. We further hypothesize that the inferiority of the latter approach may be related to the complex GRACE TWS composition over the TP, including glacier melt, lake

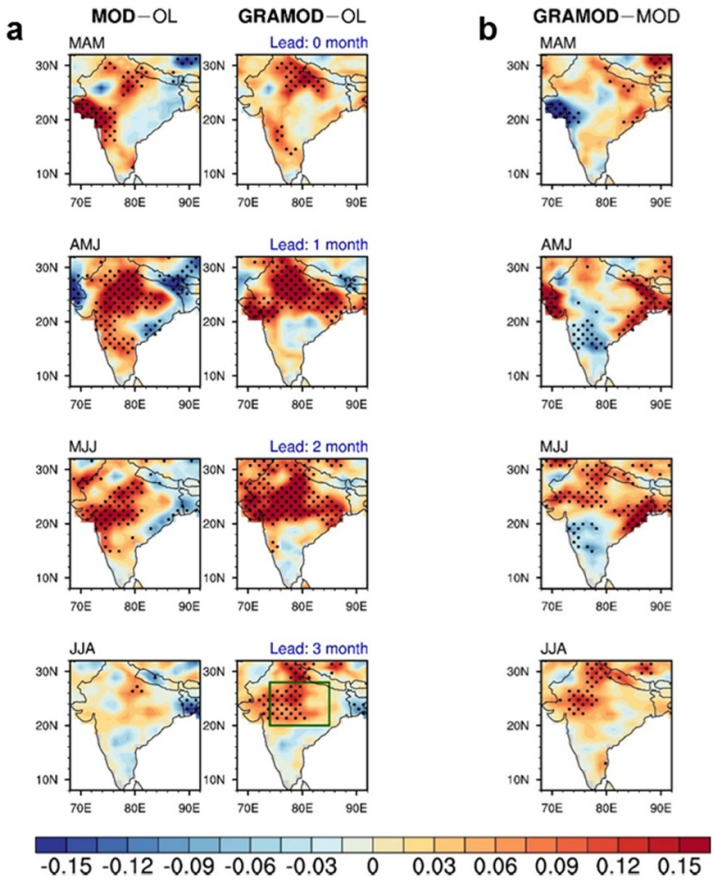

Figure 4. Precipitation forecast skill derived from snow DA over the Indian subcontinent (r^2_{DA} minus r^2_{OL}). (a) shows the skill from MOD (left) and GRAMOD (right), respectively. (b) shows the exclusive contribution from assimilating GRACE data.

changes, and crustal movement [48–50], while a simplified TWS disaggregation only considering snow, soil moisture, and groundwater variations was adopted by Z1416. Recent studies suggested that additional considerations for surface water variations in TWS disaggregation can yield better results for GRACE DA in some river basins [51, 52]. Our results suggest that similar region-specific modifications, or rules to reject unreliable observations [27], may be needed to improve the GRACE DA method over the TP, in order to achieve more skill for the pre-monsoon forecast.

Similarly, we can infer which DA method produced better snow estimates over the EA by focusing on the long-lead forecast in MJJ and JJA. As GRAMOD significantly outperforms MOD, the critical role of assimilating GRACE data in producing better snow estimates over the EA, and in improving the peak-monsoon ISM forecast is highlighted. This contribution of GRACE DA to the ISM forecast is outstanding, suggesting an overlooked source of the monsoon predictability, upon which better disaggregation of the GRACE TWS signal [51, 52] should further improve the forecast. Differing from conventional DA assessments, we use independent precipitation skill analysis to infer which satellite data/methods are more effective. This should have implications in guiding the multi-satellite DA methodology refinements in the absence of 'truth' snow conditions in the future.

3.5. Relative importance of regional snow DA in the ISM forecast

We also further assess the relative importance of regional snow DA in the ISM forecast by conducting two more modeling experiments (see **Methods** & figure S2 for model configuration details). These regional DA experiments (figure 5) show that the TP snow DA is more important for the CNI rainfall forecast in spring (pre-monsoon season), while the EA snow DA is more important for the long-lead ISM forecast (the peak monsoon season). This is consistent with figures 2(b) and 4, which is ascribed to the different regional snow–atmosphere coupling strengths, as well as the intrinsic memory difference of the TP and the EA snow. The elevated TP is a region of strong sensible heating [1, 42] receiving the primary research focus for the Asian monsoon [37, 53, 54]. Our results suggest that due to the relatively shorter memory of the TP snow (figure 2(b): forced atmospheric response disappears by May; see similar discussions on the TP snow by Orsolini *et al* [55]), the contribution of snow DA over the TP may be limited to pre-monsoon season. In comparison, properly initializing the EA snow conditions, especially with the

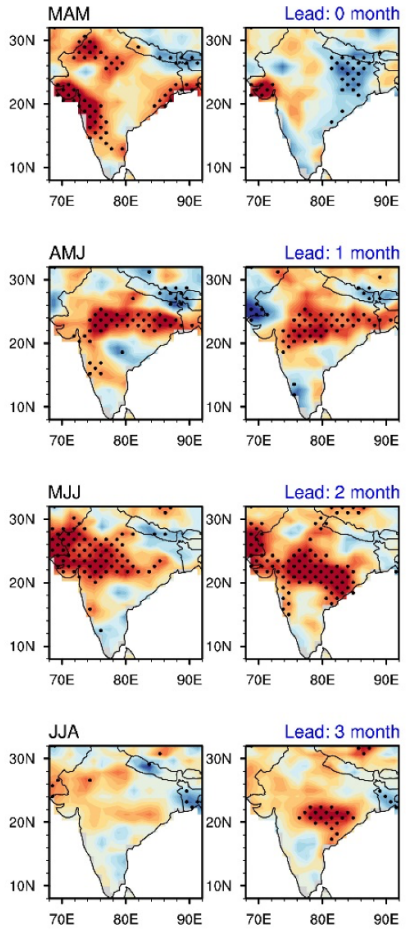

Figure 5. Precipitation forecast skill derived from regional DA (r^2_{DA} minus r^2_{OL}). Results from the TP only and the EA only experiments are shown on the left and right, respectively.

snow mass estimates from the GRACE DA, can offer the most gains to the peak-monsoon ISM forecasts. Despite the relatively simplified GRACE TWS disaggregation method used [26], assimilating GRACE data brings clear skill improvements for the long-lead ISM forecast. Therefore, our results here may provide a lower limit for the best attainable monsoon forecast skill, while future studies may improve the EA snow estimates through using improved GRACE DA methodology [51, 52] and/or incorporating other sensor observations for snow DA [24, 27].

Interestingly, beyond the Asian monsoon region, we also find some significant improvements in the North African Monsoon (NAF) region (see global monsoon delineations in figure S8), for which δDA can reach +35% in MJJ and JJA. Some recent studies reported the TP surface heating can influence its upstream NAF climate [56], while some viewed the Afro-Asian monsoon as a planetary-scale system that can respond to the same boundary forcings [57]. These suggestions may help explain the detectable improvements in the NAF forecast. However, given our snow DA was applied to a large Eurasian landmass beyond the elevated TP (figure 1(b)), future idealized modeling experiments tailored to the TP and the NAF are still warranted to better understand this complex interplay.

4. Conclusion and discussion

Overall, this study offers promising evidence as to the contribution of snow DA to the simulation accuracy of Asian monsoon seasonal anomalies, highlighting an underutilized source of the monsoon predictability. More importantly, we clarify the complexities associated with translating snow DA to useful monsoon forecast skill, including but not limited to the regional snow–atmosphere coupling, the utility/limitation of different satellite observations, and the DA uncertainty, which may be helpful to bridge the gap between the land DA and the dynamical seasonal forecasting studies. Despite such complexities, the snow DA contribution is still very clear, and thus the results presented may be arguably viewed as a lower limit for the attainable skill, upon which future studies could improve.

One immediate future step would be to increase the number of data samples (i.e. using longer time period's snow DA products to serve as the land initializations), to better quantify snow DA's contribution

to the interannual variability of the monsoon precipitation. To the best of our knowledge, our study is the pilot one and very few studies have been conducted along this line. To this end, future work is needed to better address this question which may have great implications for the huge population in the monsoon region.

Acknowledgments

This study is funded by the National Key Research and Development Program of China (2018YFA0606004) and the National Natural Science Foundation of China grant 91737307. PL was partially supported by the University Graduate Continuing Fellowship offered by the University of Texas at Austin. The authors appreciate helpful discussions with Drs Guoxiong Wu, Yimin Liu, Zhongfeng Xu, Qinjian Jin, Hui Zheng, and Randal Koster on the earlier versions of this manuscript. Particular thanks go to Dr Randal Koster whose earlier studies have motivated this work and who offered beneficial comments to us.

Author contribution:

PL, Z-L Y, and JW conceived this study; PL conducted the work with advice from Z-L Y JW, RED; YZ and LZ conducted the snow data assimilation runs and prepared the snow DA restart files; All authors contributed to the writing and interpretation of this paper. The authors declare no conflict of interests.

Data and code availability:

The model simulations were performed using the Texas Advanced Computing Center (TACC) supercomputing resources. The data that support the findings of this study (~15 TB) are available from the corresponding author upon reasonable request (liang@jsg.utexas.edu). The scripts to set up model simulations are also available upon request.

ORCID iD

Peirong Lin https://orcid.org/0000-0002-7275-7470

References

[1] Wu G *et al* 2012 Thermal controls on the asian summer monsoon *Sci. Rep.* **2** 404
[2] Webster P J *et al* 1998 Monsoons: processes, predictability, and the prospects for prediction *J. Geophys. Res. Oceans* **103** 14451–510
[3] Doblas-Reyes F J, García-Serrano J, Lienert F, Biescas A P and Rodrigues L R L 2013 Seasonal climate predictability and forecasting: status and prospects *Wiley Interdiscip. Rev. Clim. Change* **4** 245–68
[4] Wang B *et al* 2009 Advance and prospectus of seasonal prediction: assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980–2004) *Clim. Dyn.* **33** 93–117
[5] Goddard L *et al* 2001 Current approaches to seasonal to interannual climate predictions *Int. J. Climatol.* **21** 1111–52
[6] Palmer T N and Anderson D L T 1994 The prospects for seasonal forecasting—a review paper *Q. J. R. Meteorol. Soc.* **120** 755–93
[7] Wang B 2005 Fundamental challenge in simulation and prediction of summer monsoon rainfall *Geophys. Res. Lett.* **32**
[8] Chen D, Zebiak S E, Busalacchi A J and Cane M A 1995 An improved procedure for EI nino forecasting: implications for predictability *Science* **269** 1699–702
[9] Juricke S, MacLeod D, Weisheimer A, Zanna L and Palmer T N 2018 Seasonal to annual ocean forecasting skill and the role of model and observational uncertainty *Q. J. R. Meteorol. Soc.* **144** 1947–64
[10] Koster R D *et al* 2004 Regions of strong coupling between soil moisture and precipitation *Science* **305** 1138–40
[11] Koster R D, Mahanama S P P, Livneh B, Lettenmaier D P and Reichle R H 2010 Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow *Nat. Geosci.* **3** 613–6
[12] Mahanama S, Livneh B, Koster R, Lettenmaier D and Reichle R 2011 Soil moisture, snow, and seasonal streamflow forecasts in the United States *J. Hydrometeorol.* **13** 189–203
[13] Rodell M *et al* 2004 The global land data assimilation system *Bull. Am. Meteorol. Soc.* **85** 381–94
[14] Balmaseda M and Anderson D 2009 Impact of initialization strategies and observations on seasonal forecast skill *Geophys. Res. Lett.* **36** L01701
[15] Lahoz W A and Schneider P 2014 Data assimilation: making sense of earth observation *Front. Environ. Sci.* **2**
[16] Lin P, Wei J, Yang Z-L, Zhang Y and Zhang K 2016 Snow data assimilation-constrained land initialization improves seasonal temperature prediction *Geophys. Res. Lett.* **43** 11, 423–11, 432
[17] Santanello J A, Kumar S V, Peters-Lidard C D and Lawston P M 2015 Impact of soil moisture assimilation on land surface model spinup and coupled Land–atmosphere prediction *J. Hydrometeorol.* **17** 517–40
[18] Blanford H F 1884 On the connexion of the himalaya snowfall with dry winds and seasons of drought in India *Proc. R. Soc. London Ser. I* **37** 3–22
[19] Senan R *et al* 2016 Impact of springtime himalayan–tibetan plateau snowpack on the onset of the Indian summer monsoon in coupled seasonal forecasts *Clim. Dyn.* **47** 2709–25
[20] Slater A G, Martyn and Clark P 2005 Snow data assimilation via an Ensemble Kalman Filter *J. Hydrometeor.* **7** 478–493
[21] Su H, Yang Z-L, Dickinson R E, Wilson C R and Niu G-Y 2010 Multisensor snow data assimilation at the continental scale: the value of gravity recovery and climate experiment terrestrial water storage information *J. Geophys. Res. Atmos.* **115**
[22] Sun C, Walker J P and Houser P R 2004 A methodology for snow data assimilation in a land surface model *J. Geophys. Res. Atmos.* **109** D08108
[23] Zhang Y-F *et al* 2014 Assimilation of MODIS snow cover through the data assimilation research testbed and the community land model version 4 *J. Geophys. Res. Atmos.* **119** 7091–103
[24] Zhao L and Yang Z-L 2018 Multi-sensor land data assimilation_ Toward a robust global soil moisture and snow estimation *Remote Sens. Environ.* **216** 13–27
[25] Dee D P *et al* 2011 The ERA-Interim reanalysis: configuration and performance of the data assimilation system *Q. J. R. Meteorol. Soc.* **137** 553–97
[26] Zhang Y-F and Yang Z-L 2016 Estimating uncertainties in the newly developed multi-source land snow data assimilation system *J. Geophys. Res. Atmos.* **121** 8254–68
[27] Kwon Y *et al* 2016 Estimating snow water storage in North America using CLM4, DART, and snow radiance data assimilation *J. Hydrometeorol.* **17** 2853–74

[28] Thomas J A, Berg A A and Merryfield W J 2016 Influence of snow and soil moisture initialization on sub-seasonal predictability and forecast skill in boreal spring *Clim. Dyn.* **47** 49–65
[29] Webster P J and Hoyos C D 2010 Beyond the spring barrier? *Nat. Geosci.* **3** 152–3
[30] Hurrell J W, Hack J J, Shea D, Caron J M and Rosinski J 2008 A new sea surface temperature and sea ice boundary dataset for the community atmosphere model *J. Clim.* **21** 5145–53
[31] Koster R D *et al* 2011 The second phase of the global land–atmosphere coupling experiment: soil moisture contributions to subseasonal forecast skill *J. Hydrometeorol.* **12** 805–22
[32] Saha S *et al* 2013 The NCEP climate forecast system version 2 *J. Clim.* **27** 2185–208
[33] Koster R D *et al* 2010 Contribution of land surface initialization to subseasonal forecast skill: first results from a multi-model experiment *Geophys. Res. Lett.* **37**
[34] DelSole T and Shukla J 2010 Model fidelity versus skill in seasonal forecasting *J. Clim.* **23** 4794–806
[35] Toure A M *et al* 2015 Evaluation of the snow simulations from the community land model, version 4 (CLM4) *J. Hydrometeorol.* **17** 153–70
[36] Su F, Duan X, Chen D, Hao Z and Cuo L 2012 Evaluation of the global climate models in the CMIP5 over the tibetan plateau *J. Clim.* **26** 3187–208
[37] Turner A G and Slingo J M 2011 Using idealized snow forcing to test teleconnections with the Indian summer monsoon in the hadley centre GCM *Clim. Dyn.* **36** 1717–35
[38] Xu L and Dirmeyer P 2011 Snow-atmosphere coupling strength in a global atmospheric model *Geophys. Res. Lett.* **38**
[39] Webster P J and Yang S 1992 Monsoon and enso: selectively interactive systems *Q. J. R. Meteorol. Soc.* **118** 877–926
[40] Goswami B N, Krishnamurthy V and Annmalai H 1999 A broad-scale circulation index for the interannual variability of the Indian summer monsoon *Q. J. R. Meteorol. Soc.* **125** 611–33
[41] Chou C 2003 Land–sea heating contrast in an idealized Asian summer monsoon *Clim. Dyn.* **21** 11–25
[42] Wu G *et al* 2007 The influence of mechanical and thermal forcing by the tibetan plateau on asian climate *J. Hydrometeorol.* **8** 770–89
[43] Dai A *et al* 2013 The relative roles of upper and lower tropospheric thermal contrasts and tropical influences in driving Asian summer monsoons *J. Geophys. Res. Atmos.* **118** 7024–45
[44] Ding Y and Chan J C L 2005 The East Asian summer monsoon: an overview *Meteorol. Atmos. Phys.* **89** 117–42
[45] Kitoh A *et al* 2013 Monsoons in a changing world: A regional perspective in a global context *J. Geophys. Res. Atmos.* **118** 3053–65
[46] Jin Q and Wang C 2017 A revival of Indian summer monsoon rainfall since 2002 *Nat. Clim. Change* **7** 587–94
[47] Zhou T and Zou L 2010 Understanding the predictability of East Asian summer monsoon from the reproduction of land–sea thermal contrast change in AMIP-type simulation *J. Clim.* **23** 6009–26
[48] Zhang G, Yao T, Xie H, Kang S and Lei Y 2013 Increased mass over the tibetan plateau: from lakes or glaciers? *Geophys. Res. Lett.* **40** 2125–30
[49] Song C, Ke L, Huang B and Richards K S 2015 Can mountain glacier melting explains the GRACE-observed mass loss in the southeast tibetan plateau: from a climate perspective? *Glob. Planet. Change* **124** 1–9
[50] Shin Y H, Xu H, Braitenberg C, Fang J and Wang Y 2007 Moho undulations beneath Tibet from GRACE-integrated gravity data *Geophys. J. Int.* **170** 971–85
[51] Eicker A, Schumacher M, Kusche J, Döll P and Schmied H M 2014 Calibration/data assimilation approach for integrating GRACE data into the waterGAP global hydrology model (WGHM) using an ensemble kalman filter: first results *Surv. Geophys.* **35** 1285–309
[52] Kumar S V *et al* 2016 Assimilation of gridded GRACE terrestrial water storage estimates in the north american land data assimilation system *J. Hydrometeorol.* **17** 1951–72
[53] Li W *et al* 2018 Influence of tibetan plateau snow cover on east Asian atmospheric circulation at medium-range time scales *Nat. Commun.* **9** 4243
[54] Zhang Y, Zou T and Xue Y 2019 An arctic-tibetan connection on subseasonal to seasonal time scale *Geophys. Res. Lett.* **46** 2790–9
[55] Orsolini Y *et al* 2019 Evaluation of snow depth and snow-cover over the tibetan plateau in global reanalyses using in-situ and satellite remote sensing observations *Cryosphere* 1–28
[56] Lu M *et al* 2018 Possible effect of the tibetan plateau on the "upstream" climate over west Asia, North Africa, South Europe and the North Atlantic *Clim. Dyn.* **51** 1485–98
[57] Li Y, Ding Y and Li W 2017 Interdecadal variability of the Afro-Asian summer monsoon system *Adv. Atmos. Sci.* **34** 833–46