www.reading.ac.uk/centaur

CentAUR
Central Archive at the University of Reading
Reading’s research outputs online
EARLY ONLINE RELEASE

This is a preliminary PDF of the author-produced manuscript that has been peer-reviewed and accepted for publication. Since it is being posted so soon after acceptance, it has not yet been copyedited, formatted, or processed by AMS Publications. This preliminary version of the manuscript may be downloaded, distributed, and cited, but please be aware that there will be visual differences and possibly some content differences between this version and the final published version.

The DOI for this manuscript is doi: 10.1175/BAMS-D-13-00212.1

The final published version of this manuscript will replace the preliminary version at the above DOI once it is available.

© 2014 American Meteorological Society
Challenges in quantifying changes in the global water cycle

Gabriele C. Hegerl\textsuperscript{1}, Emily Black\textsuperscript{2}, Richard P. Allan\textsuperscript{2}, William J. Ingram\textsuperscript{3,4}, Debbie Polson\textsuperscript{1}, Kevin E. Trenberth\textsuperscript{5}, Robin S. Chadwick\textsuperscript{3}, Phillip A. Arkin\textsuperscript{6}, Beena Balan Sarojini\textsuperscript{2}, Andreas Becker\textsuperscript{7}, Aiguo Dai\textsuperscript{8,5}, Paul J. Durack\textsuperscript{9,16}, David Easterling\textsuperscript{10}, Hayley J. Fowler\textsuperscript{11}, Elizabeth J. Kendon\textsuperscript{3}, George J. Huffman\textsuperscript{12}, Chunlei Liu\textsuperscript{2}, Robert Marsh\textsuperscript{13}, Mark New\textsuperscript{14}, Timothy J. Osborn\textsuperscript{15}, Nikolaos Skliris\textsuperscript{13}, Peter A. Stott\textsuperscript{3}, Pier-Luigi Vidale\textsuperscript{2}, Susan E. Wijffels\textsuperscript{16}, Laura J. Wilcox\textsuperscript{2}, Kate M. Willett\textsuperscript{3}, Xuebin Zhang\textsuperscript{17}

1. School of GeoSciences, University of Edinburgh, Grant Institute, Kings Buildings, UK EH9 3JW.
2. Department of Meteorology, University of Reading/NCAS Climate
3. Met Office Hadley Centre, Exeter, UK.
4. Department of Physics, Oxford University, Oxford, UK
5. NCAR, Boulder, CO, USA.
6. ESSIC, University of Maryland, College Park, College Park, MD
7. Deutscher Wetterdienst, Offenbach, Germany
8. Department of Atmospheric and Environmental Sciences, University at Albany, SUNY, Albany, NY, USA.
9. Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, California, USA.
10. NOAA/National Climatic Data Center, Asheville, NC.
11. School of Civil Engineering and Geosciences, Newcastle University, UK.
12. NASA Goddard Space Flight Center, Greenbelt, MD, USA.
13. Ocean and Earth Science, University of Southampton, UK
14. University of Cape Town, South Africa.
15. Climatic Research Unit, School of Environmental Sciences, University of East Anglia, UK.
17. Climate Research Division, Environment Canada, Canada.

Corresponding Author: Gabriele Hegerl, gabi.hegerl@ed.ac.uk, GeoSciences, Grant Institute, Kings Buildings, W Mains Rd, Edinburgh, EH9 3JW, UK
CAPSULE (35 words):

Human influences have likely already impacted the large-scale water cycle but natural variability and observational uncertainty are substantial. It is essential to maintain and improve observational capabilities to better characterize changes.

Abstract

Understanding observed changes to the global water cycle is key to predicting future climate changes and their impacts. While many datasets document crucial variables such as precipitation, ocean salinity, runoff, and humidity, most are uncertain for determining long-term changes. In situ networks provide long time-series over land but are sparse in many regions, particularly the tropics. Satellite and reanalysis datasets provide global coverage, but their long-term stability is lacking. However, comparisons of changes among related variables can give insights into the robustness of observed changes. For example, ocean salinity, interpreted with an understanding of ocean processes, can help cross-validate precipitation. Observational evidence for human influences on the water cycle is emerging, but uncertainties resulting from internal variability and observational errors are too large to determine whether the observed and simulated changes are consistent. Improvements to the in situ and satellite observing networks that monitor the changing water cycle are required, yet continued data coverage is threatened by funding reductions. Uncertainty both in the role of anthropogenic aerosols, and due to large climate variability presently limits confidence in attribution of observed changes.
1. Introduction

Climate change, alongside increased demand for water (World Water Development Report 2003; WHO/UNICEF 2011), is projected to increase water scarcity in many regions over the next few decades (e.g., Arnell et al. 2013; Kundzewicz et al. 2007). Extremes linked to the water cycle, such as droughts, heavy rainfall and floods, already cause substantial damage (e.g. Lazo et al. 2011; Peterson et al., 2012; 2013) and such events are expected to increase in severity and frequency (Dai 2011a, 2013a; IPCC 2012, Collins et al. 2013a).

Better management of water resources and adaptation to expected changes require reliable predictions of the water cycle. Such predictions must be grounded in the changes already observed. This requires quantification of long-term large-scale changes in key water cycle variables, and estimation of the contribution from natural climate variability and external forcings, including through studies that are referred to as detection and attribution (see Stott et al., 2010; Hegerl and Zwiers 2011). Successful examples of detection and attribution are reported in Bindoff et al. (2013).

We discuss how well the available observing capability can capture expected changes in the global water cycle, including the increasing water content of the atmosphere, strengthening of climatological precipitation minus evaporation (P-E) patterns, the pronounced spatial structure and sharp gradients in precipitation change, and increases of extreme precipitation. We also discuss the challenges inherent in combining an incomplete observational record with imperfect climate models, to detect anthropogenic changes in the water cycle.
Drawing on discussions from a workshop held at the University of Reading, U.K. in June 2012, we focus on long-term large-scale changes in a few key variables that are both potentially related to climate change, and essential for diagnosing changes in the global water cycle. These include humidity, precipitation, P-E, and salinity. We also give recommendations that will lead towards more robust predictions and identification of the human influence on recent observed changes. It is beyond the scope of this paper to provide a full review of water cycle changes, or to discuss regional changes (see Parker 2013; Collins et al. 2013b), changes in the biosphere and cryosphere, river discharge (see Dai et al. 2009), or drought (see Dai 2011a, 2011b, 2013; Trenberth et al. 2014).

We briefly describe the expected physical changes, before discussing the challenges of observing such changes with present observational capabilities, globally, as well as over ocean and land separately. We also discuss how physically consistent a picture these observations draw, and conclude with recommendations to ensure continued and improved ability to document the changing water cycle. The supplement provides more information on available observational data and quality control procedures.

2. Expected changes in the global water cycle

Changes in the hydrological cycle are an expected consequence of anthropogenic climate change. The Clausius-Clapeyron relationship suggests a strong quasi-exponential increase in water vapor concentrations with warming at about 6-7%/K near the surface. This is consistent with observations of change over the ocean (e.g., Trenberth et al. 2005; Dai 2006a; Chung et al., 2014) and land (Dai 2006b; Willett et al. 2010), and with simulations of future changes (e.g., Allen and Ingram 2002) and
assumes that on large scales the relative humidity changes little, as generally expected (see Sherwood et al. 2010; Allen and Ingram, 2002) and approximately seen in models (Richter and Xie 2008; Collins et al. 2013a). Locally, however, relative humidity changes may arise where large-scale circulation patterns alter, or when moisture sources are limited over land (e.g., Dai 2006; Vicente-Serrano et al. 2013).

Changes in global mean precipitation are limited by the energy budget, both through evaporation and the ability of the atmosphere to radiate away the latent heat released when precipitation forms (e.g., Trenberth 2011; O’Gorman et al. 2012). This largely explains why global mean precipitation increases by only 2-3% per K of warming in climate models (the ‘hydrological sensitivity’; see Figure 1). Broadly, the radiative effect of greenhouse gas forcing reduces the global precipitation increase driven by warming itself (e.g., Bony et al., 2013), while the direct radiative effect of aerosols that scatter rather than absorb sunlight does not influence the rate at which precipitation increases with warming. Figure 1 illustrates this for climate models run under the Coupled Model Intercomparison Project 5 (CMIP5) protocol (Taylor et al. 2012) for the 20th century, and for 4 standard scenarios for the 21st century. These range from RCP8.5, a high-emissions scenario, to RCP2.6, a low-emissions scenario (see Collins et al. 2013a). With stronger greenhouse gas forcing, global-mean temperature and precipitation both increase more, but the hydrological sensitivity becomes slightly smaller (see also Wu et al. 2010; Johns et al. 2011). Pendergrass and Hartmann (2014) show that the spread in CMIP5 model response of precipitation to increases in carbon dioxide is related to differences in atmospheric radiative cooling, which are in turn related to changes in temperature profiles and water vapor amounts. Forced changes in global-mean
precipitation are expected to be relatively small at present (Fig. 1b) and are therefore hard to distinguish from natural variability.

Spatial patterns are important both for identifying fingerprints of forced changes in precipitation and for impacts. Since global-mean evaporation and precipitation are expected to increase more slowly with temperature than implied by water vapor content, this implies slightly increased water vapor residence times and reduced atmospheric mass convergence (Vecchi et al. 2006; Held and Soden 2006). However, increasing water vapor more than offsets the weakened atmospheric wind convergence in the tropics (Vecchi et al. 2006; Held and Soden 2006; Allan 2012; Kitoh et al. 2013). Thus, where E exceeds P in the mean (such as over the sub-tropical oceans), it would do so even more, while areas where P exceeds E (such as the Intertropical Convergence Zone, ITCZ, and high latitudes) would receive yet more precipitation excess (Manabe and Wetherald 1980; Held and Soden 2006; Seager and Naik 2012; Bengtsson et al. 2011, Bintanja and Selton, 2014). Simulations of future climate changes broadly confirm this, particularly when zonally averaged (see Fig. 2, bottom panel) and show rainfall generally increasing at latitudes and seasons that currently have high rainfall and less in dry regions (Collins et al. 2013a). This ‘wet get wetter, dry get drier’ paradigm involves a range of atmospheric processes, including an increased vertical gradient of atmospheric water vapor, which leads to intensified convective events in the deep tropics (see Chou et al. 2009).

However, simple P-E enhancement does not necessarily apply to dry land, where moisture is limited (Greve et al. 2014). It also does not hold true at regional scales, where atmospheric circulation changes may displace the geographical positions of
"wet" and "dry" regions (Xie et al., 2010; Chadwick et al., 2013; Allan 2014). GCMs generally simulate an expansion of the Hadley Cells as the globe warms, with associated poleward migration of subtropical aridity and storm tracks, but the size varies, and there is limited agreement on the mechanisms (Yin 2005; Lu et al. 2007; Seidel et al. 2008; Scheff and Frierson 2012a, 2012b).

Anthropogenic aerosol effects counteract some of the anticipated greenhouse-gas driven warming, and hence the associated increase in precipitation (Liepert et al., 2004; Wu et al., 2013). Aerosols reduce the available energy for evaporation, and absorbing aerosols such as black carbon locally heat the atmosphere, effectively short-circuiting the hydrological cycle. Pendergrass and Hartmann (2012) show how black carbon forcing influences the inter-model spread in global-mean precipitation change in CMIP3 models. The aerosol indirect effect may account for almost all aerosol cooling in models (Zelinka et al. 2014), and so be key to the aerosol-driven decrease in precipitation (Liepert et al., 2004; Levy et al. 2013), although this is model-dependent (e.g., Shindell et al., 2012). The radiative effect of anthropogenic aerosols is also expected to affect the spatial pattern of precipitation and evaporation changes. As surface emissions of aerosol are spatially heterogeneous, and atmospheric residence times are relatively short, the direct radiative impact of aerosol is geographically variable, with the largest concentrations in the Northern Hemisphere (NH). The geographical heterogeneity of aerosol distribution is expected to affect the interhemispheric temperature gradient, and hence the atmospheric circulation – which should shift the ITCZ (e.g., Rotstayn et al. 2000; Ming and Ramaswamy 2011; Hwang et al. 2013) and change the width of the Hadley cell (Allen et al. 2012). Models' representation of aerosols, and their interactions with clouds in particular, affect their ability to reproduce trends in the
interhemispheric temperature gradient (e.g. Chang et al., 2011; Wilcox et al. 2013).

Modeling studies also suggest that aerosols may have contributed to the drying of the
Sahel from 1940 to 1980 (Rotstyn and Lohmann, 2002; Ackerley et al. 2011; Hwang et
al. 2013; Dong et al. 2014), and influence the East Asian monsoon (e.g. Lau et al. 2006;
Meehl et al. 2008; Bollasina et al. 2011; Guo et al. 2012), and mid-latitude precipitation
(Leibensperger et al. 2012; Rotstyn et al. 2012).

Stratospheric aerosols from explosive volcanic eruptions also influence the water cycle.
Sharp reductions in observed global-mean land precipitation and stream flow were
observed after the Mt Pinatubo eruption in 1991 (Trenberth and Dai 2007) and other
20th century eruptions (Gu et al. 2007). This effect is particularly evident in
climatologically wet regions, where the observed reduction in precipitation following
eruptions appears significantly larger than simulated (Iles et al. 2014). Volcanoes may
also contribute to regional drought by influencing the inter-hemispheric energy budget
(e.g., Haywood et al. 2013).

3. Observing and attributing changes in the global-scale water cycle

Increases in atmospheric moisture are a key fingerprint of climate change. Surface
specific humidity at global scales is reasonably well observed over land since 1973
(HadISDH; Willett et al., 2013), and over ocean since 1971 (NOVSv2.0; Berry and Kent
2009, 2011) using in situ data (for measurement techniques and more background as
well as dataset information, see supplement); and results are quite robust across
different data products (e.g., Dai 2006; Willett et al. 2007, 2013). Combined land and
ocean surface specific humidity over the 1973-1999 period shows widespread
increases. This change has been attributed mainly to human influence (Willett et al. 2007). As expected, globally, changes in relative humidity between 1973 and 1999 are small or negative (Hartmann et al., 2013). Since 2000, however, a decrease has been observed over land, likely related to the greater warming of land relative to the ocean (Joshi et al., 2008; Simmons et al., 2010; Willett et al., 2014).

In situ measurements of atmospheric humidity from radiosonde data provide time-series of Total Column Water Vapor (TCWV) from the 1950s. Increasing water vapor is apparent although spatial sampling is limited and temporal inhomogeneities are problematic (Dai et al. 2011; Zhao et al. 2012). Global-scale patterns of change became observable only when the satellite era began. Since the 1980s, near-global satellite-based estimates of TCWV over the ice-free oceans and of clear-sky upper tropospheric relative humidity have allowed variability in tropospheric water vapor to be explored (e.g., Trenberth et al. 2005; Chung et al. 2014). The satellite-based Special Sensor Microwave Imager (SSMI) TCWV data for 1988-2006 has enabled a robust anthropogenic fingerprint of increasing specific humidity to be detected over the oceans (Santer et al. 2007; 2009).

Satellite-based sensors, in combination with in situ data for best results, provide the only practical means for monitoring precipitation over land and ocean combined (e.g., Fig 1). Satellite precipitation passive retrievals are restricted to the thermal infrared (IR) and microwave (MW) spectral bands. IR-based estimates are available from geostationary satellites at high frequency, but have modest skill at instantaneous rainfall intensity (e.g., Kidd and Huffman, 2011). Passive MW data, available since mid-1987, have made precipitation retrievals more reliable, and are particularly successful
over oceans. Retrievals over land are more approximate, since coasts and complex
terrain increase uncertainty, and the accuracy of current algorithms deteriorates
polewards of 50°. The latter is because these algorithms are tuned to lower-latitude
conditions and because they cannot identify precipitation over snowy/icy surfaces.

Combined-satellite algorithms have been developed to merge individual estimates,
either as relatively coarse-resolution, long-period climate data records (the Global
Precipitation Climatology Project, GPCP, monthly dataset on a 2.5°x2.5°
latitude/longitude grid begins in 1979; Adler et al. 2003), or, alternatively, as high-
resolution precipitation products that start with the launch of the Tropical Rainfall
Measuring Mission (TRMM) in late 1997 and will be continued with the successful
launch of the Global Precipitation Mission (GPM) in early 2014. A recently released
high-resolution dataset covers a somewhat longer period (Funk et al, 2014). Some
products use rain-gauge data, where available, as input and to calibrate satellite-based
rainfall estimates (Huffman et al. 2007). Therefore, satellite-derived products are not all
independent of in situ data, and trends based on the satellite record may be affected by
inhomogeneities in both the satellite and the surface data used (Maidment et al, 2014).

The satellite record has been very useful for understanding precipitation changes. A
study sampling blended satellite observations of the wet and dry regimes as they shift
spatially from year to year indicates enhanced seasonality (Chou et al. 2013), while Liu
and Allan (2013) found tropical ocean precipitation increased by 1.7%/decade for the
wettest 30% of the tropics in GPCP data, with declines over the remaining, drier, regions
of -3.4%/decade for 1988-2008. Polson et al. (2013b) detected the fingerprint of a
strengthening contrast of wet and dry regions in the GPCP satellite record since 1988,
and attributed this change largely to greenhouse gas increases. Marvel and Bonfils (2013) arrive at a similar conclusion, explicitly accounting for circulation changes and using the full record. Some of the changes detected in observations were significantly larger than modelled, for example, in wet regions over ocean (Polson et al. 2013b; see also Chou et al. 2013; Liu and Allan 2013).

Atmospheric reanalyses provide a global 3-dimensional and multi-decadal representation of changes in atmospheric circulation, fluxes and water vapor by assimilating observations (satellite, in situ, radiosondes, etc) into numerical weather prediction models. Notably, global quasi-observed P-E estimates are available only from reanalyses. Reanalyses, however, are affected by biases in the models and by long-term inhomogeneity of the observations, particularly, changing input data streams (Trenberth et al. 2005, 2011; Dee et al. 2011; Allan et al. 2014). These factors lead to inconsistencies between reanalyses and substantial uncertainties in their long-term trends; uncertainties that can be explored by using water budget closure constraints (e.g., Trenberth and Fasullo 2013a, b). The issues of long-term homogeneity will be improved in future developments (e.g. ERA-CLIM, http://www.era-clim.eu).

In conclusion, the satellite record is essential for monitoring the changing water cycle on a near-global scale, while future climate quality reanalyses hold considerable promise. Uncertainty estimates on long-term trends are difficult to provide (see supplement) but would be very useful.

4. Interpreting changes over ocean
Changes in P-E and precipitation by climate models are particularly consistent over the oceans (Fig. 1b; Meehl et al. 2007; Bony et al. 2013). In terms of observations, in addition to the satellite record, limited in situ records are available, such as evaporation analyses (although fraught with discontinuities and global lack of closure) (Yu and Weller 2007; Yu et al. 2008) and precipitation from island stations and buoys (e.g., CRU, precipitation data as used in Josey and Marsh 2005). Overall, however, the in situ observations lack the spatial and temporal coverage needed to measure global changes (see Xie and Arkin 1998 for precipitation), and satellite and reanalysis data are consequently indispensable.

Both evaporation and precipitation affect local sea surface salinity. Thus, patterns and changes in the net freshwater flux, P-E, contribute to its temporal variations, and long-term changes to ocean salinity provide an important independent measurement from which the water cycle can be monitored. It should be noted, however, that in-situ ocean salinity is strongly influenced by changes to the ocean’s circulation (which is influenced by ocean warming and surface wind changes), and thus that care must be taken when using in-situ salinity to infer P-E (Durack and Wijffels 2010; Skliris et al. 2014).

Ocean salinity observations have been made since the late 19th century by research cruises. Historical observational coverage is, however, sparse in the early part of the record, with near-global coverage achieved only recently (Supplementary Fig. 1), largely due to the Argo network of 3600 free-drifting floats initiated in 1999 (Freeland et al. 2010). These floats measure the salinity and temperature of the upper 2000 m of the global ocean almost in real time. The Aquarius and Soil Moisture Ocean Salinity (SMOS)
satellite missions have provided global estimates of ocean surface salinity since late 2009 and June 2011 respectively.

The observed pattern of salinity change at high latitudes and in the subtropics is broadly consistent with the expected changes in P-E, although the observational uncertainty is also clear (Fig. 3). These observed changes, broadly speaking, reflect an amplification of the climatological pattern of salinity – with salty regions getting saltier, and fresh regions getting fresher (Durack et al. 2012; Skliris et al. 2014). Observed salinity changes in the Atlantic and Pacific Ocean since the mid-20th century have been found to be outside the range of internal climate variability in model simulations, and have been attributed to anthropogenic influences (e.g. Stott et al. 2008; Terray et al. 2012; Pierce et al. 2012). The attribution of salinity changes to anthropogenic factors was important evidence for the Intergovernmental Panel on Climate Change (IPCC)’s conclusion that there has been ‘likely’ a human contribution to the changing water cycle (see Bindoff et al., 2013). However, further work is required to better understand the effects of unforced variability on ocean salinity and their influence on the patterns of reported long-term changes.

It is essential that satellite-based, ship-based and Argo float measurements continue to monitor the ocean. Reliance on a single record type would hamper the identification of errors introduced by changes in coverage and measurement methods.

5. Interpreting changes over land

Over land, in situ data provide a long-term record of changing humidity and precipitation. However, the lack of reliable homogeneous terrestrial evapo-
transpiration data hampers studies of changes in the terrestrial water balance. Flux
towers provide direct measurements of water, energy and carbon fluxes at a few points,
but only for short periods (typically 5-15 years – e.g., Blyth et al. 2011). Pan evaporation
can easily be diagnosed from general circulation climate models (GCMs; as “potential
evaporation”) and effectively measures evaporative demand, which is very relevant to
some crops and natural ecosystems. Long time-series would therefore be valuable (e.g.
Greve et al. 2014), but measurements are sparse, and as it is not part of the actual
energy or moisture budget it cannot be deduced from other measurements. Pan
evaporation has decreased in many regions studied (related, at least partly, to wind
stilling; McVicar et al. 2012), in contrast to actual evapotranspiration measured at
Fluxnet sites, which increased until recently (Hartmann et al. 2013). Inferring
evaporation from the atmospheric moisture budget in reanalyses (Trenberth et al.
2011; Trenberth and Fasullo 2013b) is the most realistic option to analyse large-scale
changes in P-E over land. As was mentioned above, however, reanalyses are affected by
model error and their trends by changing data streams, and thus reanalysis evaporation
data should be treated with caution.

The most widely used record of the changing water cycle over land is from long-term
precipitation station data (e.g. Peterson and Vose 1997; Menne et al. 2012). Several
gridded products are available (see Supplementary Table 1; Harris et al. 2014; Becker et
al. 2013; Zhang et al. 2007), of which this paper shows three that have been processed
differently, some completely interpolating precipitation over land (GPCC, Becker et
al. 2013; CRU; Harris et al., 2014; with information on support available), or only providing
values where long-term stations are available (Zhang et al., 2007). An additional dataset
(VASCLIMO, Beck et al. 2005) uses a subset of GPCC stations that are considered long-
term and homogeneous. Figure 4 shows the density of the station network used in the CRU dataset, supplementary Fig. 2 for GPCC. Generally, data availability increased until 1990, but has dropped since, especially in the tropics. For the GPCC this dramatic drop occurs a decade later. Country-specific readiness to share data is the biggest constraint for data density in the most recent decade.

The gridded precipitation datasets available vary also in their methods of quality control and homogenization (see Supplementary Material). This diversity leads to substantial differences in trends and discrepancies between datasets, and contributes to our uncertainty in how drought has changed (Trenberth et al. 2014).

Figure 5 illustrates similarities and differences in precipitation change from these datasets for high latitudes, and Figure 2, upper panel, for zonal mean changes. The zonal mean increase in northern high latitudes shown by most datasets (with the exception of the GPCC Full Data V6 dataset, which was not constructed with long-term homogeneity as a priority) agrees with expectation (see Fig. 2, lower panel), and is supported by Arctic regional studies (Rawlins et al. 2010). Min et al. (2008) detected the response to anthropogenic forcing in the observed moistening of northern high latitudes, using the Zhang et al. (2007) dataset. Figure 5, however, suggests substantial observational uncertainty, which may be partly due to coverage and data processing, and may contain a small contribution by changing liquid-to-solid ratio of precipitation (see discussion in supplement).

A substantial fraction of the differences between zonal changes recorded in different datasets can be explained by differences in spatial coverage (Polson et al. 2013a). The IPCC 5th Assessment report concluded that there is ’medium’ confidence in precipitation
change averaged over land after 1951 (and lower confidence before 1951) due to data uncertainty (Hartmann et al. 2013). Simulated changes in land precipitation are also uncertain, as evident from Fig. 1 (right panel).

The incomplete spatial coverage of precipitation changes in observations tends to increase noise and hence delay detection of global and large-scale changes (e.g., for precipitation changes, Balan Sarojini et al. 2012; Trenberth et al. 2014; note that in detection and attribution, only regions covered by observed data are analysed in both models and observations). Since station-based records are point measurements and precipitation tends to be highly variable spatially (e.g., Osborn, 1997), many stations are required to correctly reflect large-scale precipitation trends (e.g., Wan et al. 2013). In general, the variability in grid cells based on few stations is higher than if a larger number of stations are used, and changes may be recorded incompletely (see Zhang et al., 2007).

Despite these difficulties, zonal-mean precipitation changes agree better with the expected response to forcing than expected by chance, and show detectable changes for boreal winter and spring data (Polson et al. 2013a), as well as for annual data (see Fig. 2; Zhang et al. 2007; Polson et al. 2013a) for most datasets. These findings contributed to the IPCC 5th assessment’s conclusion of ‘medium confidence’ that a human influence on global-scale land precipitation change is emerging (Bindoff et al. 2013). Wu et al. (2013) argue that the lack of an increase in Northern Hemispheric (NH) land precipitation over the last century is because aerosols induce a reduction in precipitation that counteracts the increase in precipitation expected from increases in greenhouse gases.
Due to data uncertainty, it is currently difficult to decide whether observed precipitation changes are larger than model simulated changes (Polson et al. 2013a). Averaging across mis-located precipitation features in models may reduce the magnitude of multi-model mean simulated precipitation change. This bias can be reduced by expressing changes relative to climatological precipitation (Noake et al., 2011; Liu and Allan, 2013; Polson et al. 2013b; Marvel and Bonfils, 2013), or by morphing model changes onto observed features (Levy et al. 2013a). However, in some cases, results still show observed changes that are large compared to model simulations (e.g., Polson et al. 2013a,b).

In summary, the record over land is extensive in time, but has serious limitations in spatial coverage and homogeneity. The drop in availability of recent in situ precipitation data (Fig. 4; supplementary Fig. 2) is of real concern. Data are particularly sparse in the tropics and subtropics, where substantial and spatially variable changes are expected. In addition to improving gauge density, more data-rescue funding and improved data-sharing practices and capabilities would help to address this problem.

6. Intensification of precipitation extremes

Since storms are fuelled by moisture convergence, storm-related extremes are expected to increase in a moister atmosphere (Emanuel 1999; Trenberth et al. 2003). It is less clear how large this increase will be, as limited moisture availability over land and possible stabilization of atmospheric temperature profiles tend to reduce the empirically derived response in precipitation extremes below the Clausius-Clapeyron-based increase in water vapor of 6-7%/K, while feedbacks of increased latent heat
release on storm intensity may amplify the response for sub-daily precipitation extremes (Lenderink and van Meijgaard 2008; Berg et al. 2013; Westra et al. 2014). Overall, under global warming, a substantial increase in the intensity of the stronger storms and precipitation events is expected. This increase is expected to be larger for more intense events (see Allen and Ingram 2002; Pall et al. 2011; Kharin et al. 2013; IPCC 2012), and is a robust fingerprint for the detection of climate change (Hegerl et al. 2004).

This larger increase in intense precipitation than annual total precipitation implies light or no rain must become more common, suggesting longer dry spells and increased risk of drought, exacerbated by increased potential evapotranspiration (Trenberth et al. 2003). How this intensification of extremes of the water cycle will be expressed is uncertain, as climate models still struggle to properly depict the diurnal cycle, frequency, intensity, and type of precipitation (see Flato et al. 2013), a problem which may be improved in part with the use of higher resolutions (e.g. Kendon et al. 2012; Strachan et al. 2013; Demory et al. 2014; Arakawa et al. 2011). Accurate representation of local storm dynamics may be an essential requirement for predicting changes to convective extremes (Kendon et al. 2014).

Worldwide in situ data for analysing changes in daily precipitation extremes have been collected by the CLIVAR Expert Team on Climate Change Detection and Indices (Donat et al. 2013). However, the record is far from complete in covering the global land masses, and is particularly sparse in key tropical regions. Increases in precipitation intensity have been identified in observations over many land regions (Fowler and Kilsby 2003; Groisman et al., 2005; Min et al. 2011; Zolina et al. 2010). Analysis of
observed annual maximum 1-day precipitation over land areas with sufficient data samples indicates an increase with global mean temperature of about 6-8%/K; Westra et al. 2013). Min et al. (2011) and Zhang et al. (2013) report detection of human influence on widespread intensification of extreme precipitation over NH land, although with substantial uncertainty in data and estimates of internal variability. Observed responses of daily precipitation extremes to interannual variability (e.g., Liu and Allan 2012) potentially offer a constraint on climate change projections for future changes in extremes (O’Gorman 2012).

Characterizing sub-daily precipitation variability is difficult on large scales, given the limitations of the satellite record (see above), and agreement is poorer on short timescales than for multi-day averages (Liu and Allan 2012). However, a number of regional studies show recent increasing sub-daily precipitation intensities in response to rising temperatures (e.g., Lenderink and van Meijgaard 2008; Utsumi et al. 2011; see Westra et al., 2014). In the future, radar data exchanged globally show promise, if remaining technical and administrative problems can be resolved (e.g., Winterrath et al. 2012a, 2012b; Michelson et al. 2013; Berg et al. 2013).

In short, it is essential to observe precipitation extremes to understand changing precipitation characteristics and quantify human-induced changes. However, uncertainties are substantial, and temporal and spatial scales reliably observable at present fall short of what is necessary for characterizing global changes.

7. The challenge of climate variability
Natural variability generated within the climate system can cause multi-decadal features in precipitation that are difficult to separate from the response to long-term forcing – especially in view of the relatively short observational record (e.g., Dai 2013). When determining if an observed change is significant relative to climate variability, a large sample of variability realizations from climate model simulations is generally used, since the observed record is short. However, discrepancies between simulated precipitation variability and that estimated from observations are substantial, particularly in the tropics (Zhang et al. 2007, see supplement) because of a combination of observational and model limitations. This introduces substantial uncertainty in detection and attribution results, even when model estimates of variance are doubled (as is often done; e.g., Zhang et al. 2007; Polson et al. 2013a). Long-term observed data obtained, for example, through data rescue are critical when evaluating simulations of multi-decadal variability (www.oldweather.org; www.met-acre.org, Allan et al. 2011).

Figure 6 illustrates how natural modes can induce apparent trends in precipitation over large regions (after Dai 2013). The Inter-decadal Pacific Oscillation index (IPO; closely related to the Pacific Decadal Oscillation, Liu 2012), for example, corresponds to an index of Southwest U.S. precipitation in observations and model experiments forced by sea surface temperatures (e.g. Schubert et al. 2009). This suggests that both an increase in Southwest U.S. precipitation from the late 1940s to early 1980s, and a subsequent decrease are largely caused by internal variability. El Niño and the IPO also influence precipitation patterns globally (Gu and Adler 2012; Dai 2013), which can influence trends over short periods such as those from satellites (Polson et al. 2013b; Liu and Allan 2013). This strong climate variability makes it difficult to detect the expected
long-term regional precipitation response to greenhouse gas forcing using historical
data (see also Deser et al. 2012).

For understanding and attributing changes in the water cycle it is therefore important
to account carefully for natural decadal climate variability, be it internally generated or
volcanically forced. This is particularly true when using short records. Because un-
forced internal variability is realization-dependent, discrepancies between model-based
and observed records of variability should be expected and need to be accounted for in
comparing models with observations for climatology, variability and trends.

8. Conclusions and Recommendations

There is strong evidence that changes are underway in aspects of the water cycle, which
are consistent with theoretical expectations of the hydrological response to increased
greenhouse gases and a warming planet. Many aspects of water cycle change, however,
remain uncertain owing to small expected signals relative to the noise of natural
variability, limitations of climate models, and short and inhomogeneous observational
datasets.

Uncertainty may be reduced by cross-validating changes between multiple datasets and
across variables, by putting these comparisons in the context of the theoretical
expectation of the response of the water cycle to global climate change, and by exploring
closure constraints. The observations, for example, suggest increases in high latitude
precipitation, global-scale atmospheric humidity, and precipitation extremes that are
consistent with expected changes. Furthermore, satellite data show signals of
precipitation increases over wet regions and decreases over dry regions, corroborated
by in situ data over land, and physically consistent with an amplification of salinity
patterns over the global ocean. The consistency in the evidence of changes of
precipitation over land and from changes in ocean salinity is reflected in the IPCC’s
conclusion that human activity has ‘likely’ influenced the global water cycle since 1960
(Bindoff et al. 2013), even though confidence in individual lines of evidence, such as
attribution of precipitation changes to causes, is lower.

Observational uncertainty and a low signal-to-noise ratio pose serious difficulties when
determining the magnitude of the human contribution to observed changes. Several
studies report observed changes that are significantly larger than those simulated by
climate models. However, these findings were generally not robust to data uncertainty.
The uncertainty arises because the satellite record is short compared to decadal climate
variability, and affected by calibration uncertainty; and because the available in situ
record has many gaps, particularly in the tropics and subtropics, and is sparse on sub-
daily timescales. Thus while observations can place constraints on future temperature
changes, this is not yet possible for future precipitation projections (see Collins et al.
2013 and Bindoff et al. 2013).

To improve the situation, we recommend:

1) The satellite record is vital, particularly to capture the strong changes over ocean
that are robustly predicted by models. Only the full constellation can capture the
intermittent nature of precipitation and capture extremes. The new GPM mission
has exciting prospects for better calibration of space-based observations. Improved
sampling by the constellation should enable the intermittency of precipitation to be
better handled. Planning for future missions, providing continuity and temporal
overlap of measurements is essential to be able to reliably determine long-term trends.

2) *In situ stations* are vital both for cross-validating and calibrating satellite datasets and for long-term monitoring. However, the drop in available in situ data in recent decades, as illustrated for precipitation (Fig. 4), is alarming and needs to be addressed. Many observations are not made available for analysis, while some remain in paper form only and are not catalogued. It is necessary to strengthen efforts to rescue, scan and digitize data. Also, impediments to data sharing need to be overcome, and data delivery needs to be more timely in order to monitor the changing water cycle in near-real time, as is done for temperature.

3) There is need for better global coverage and higher time resolution data to capture *changing precipitation extremes*. Hourly datasets are needed to track and identify changes in short-term extremes, which are another important fingerprint of anthropogenic changes, and critical for flood management.

4) *Gridded products* of in situ precipitation change show substantial differences (Figs. 2, 5), related to numbers of stations used, their homogeneity, manner of analysis, quality control procedures and treatment of changing data coverage over time. This uncertainty needs to be better characterized and best practices developed.

5) *Observations* in key regions are still sparse, particularly in the tropics, where the observing system is insufficient to record the anticipated changes in the water cycle. For the Asian monsoon, data sparsity is partly related to practical and administrative issues with data sharing. An improved international capacity to monitor all aspects of observed changes is important.
6) Ocean salinity observations provide an independent insight into the changing water cycle. Continued maintenance and improved coverage of the Argo Program, along with the development of satellite missions to follow Aquarius/SMOS for ocean salinity will strongly improve our understanding of global water cycle changes.

7) Key diagnostics, such as P-E, are not directly observable on large scales. Therefore, reanalysis data are vital, and their homogeneity in time and reliability for study of long-term changes need to be improved. Climate quality reanalysis will be very useful and are strongly encouraged. Closure of the water cycle using multiple variables provides a physical constraint that should be exploited to help quantify uncertainties.

8) Analyses of observed changes are more powerful if they make use of and diagnose physical mechanisms which are responsible for the atmospheric and oceanic change patterns. Studies need to investigate the robustness of results across data products, and evaluate the physical consistency of recorded changes across water cycle variables. Process studies may be able to constrain and better understand the fast circulation response to CO$_2$ forcing, which is a source of uncertainty.

9) Uncertainty in the role of aerosols on precipitation is central when quantifying the human contribution to observed changes. Aerosols vary enormously in space and time and in composition. Covariability with water vapor and clouds remain issues. Interactions between aerosol and cloud microphysics need to be better understood and represented in models, and the role of aerosol on precipitation changes needs to be better understood. This requires scientists from aerosol and water cycle communities to work together.
Variability generated within the climate system, particularly regionally on interannual to multidecadal timescales, has a large effect on water cycle variables and delays detection and emergence of changes. There is substantial uncertainty in present understanding about the magnitude and structure of variability in the water cycle which, if addressed, will improve the reliability of detection and attribution studies, and help societies in managing the impacts of decadal variability and change.

Acknowledgements:

We thank several anonymous reviewers for their helpful and perceptive suggestions. The workshop that formed the basis for this paper was funded by NERC (the UK Natural Environment Research Council (NERC grant NE/I006672/1). We thank Eleanor Blyth and David Parker for comments. Robin Chadwick, Kate Willett, William Ingram, Lizzie Kendon and Peter Stott were supported by the Joint DECC/Defra Met Office Hadley Centre Climate Programme (GA01101). Paul Durack is supported by the Lawrence Livermore National Laboratory which is funded by the U.S. Department of Energy under contract DE-AC52-07NA27344, and William Ingram is partly supported by NERC grant NE/I00680X/1. Gabi Hegerl and Richard Allan are partly supported by NCAS. Kevin Trenberth is supported by NASA grant NNX11AG69G. NCAR is sponsored by The National Science Foundation. P.L. Vidale, R.P. Allan, B. Balan Sarojini, C. Liu, E. Black, P. Stott, G. Hegerl, D. Polson, R. Marsh, N. Skliris and Laura Wilcox are supported by the PAGODA project of the Changing Water Cycle programme of the UK Natural
A. Dai acknowledges the support of NSF grant #AGS-1353740. H.J. Fowler is supported by the CONVEX project of the Changing Water Cycle Programme of NERC under grant NE/I006680/1) and European Research Council funded INTENSE (ERC-2013-CoG).


Hartmann, D., and coauthors, 2013: Observations: atmosphere and surface. Chapter 2, *Climate Change 2013; The physical science basis. Contribution of Working Group 1 to the Fifth Assessment report of the Intergovernmental Panel on Climate Change* [Stocker T. et al. (eds.)], Cambridge University Press, 159-254.


Figure Captions

**Figure 1 left panel:** Projected global-mean precipitation change (mm/day) against global-mean 2m air temperature change (K) from CMIP5 models, for four representative concentration pathways (RCP) scenarios. Values are means over successive decades between 2006 and 2095 and all ensemble members of each model. Anomalies are relative to mean values over 1986-2005 in the CMIP5 historical runs.

Right panel: Precipitation sensitivity for future (RCP scenarios) and past (Historical and Atmospheric Model Intercomparison Project, AMIP) change in precipitation amount [%] per degree global-mean warming. Trends are calculated from the linear least squares fit of annual global-mean precipitation change (%) against temperature (K) change relative to the period 1988-2005 (without decadal smoothing). Crosses indicate ensemble means for each CMIP5 model, circles indicate multi-model mean. Precipitation sensitivity is also shown for historical periods; comparing GCMs with GPCP, GPCC and CRU data (see text), using temperature changes from HadCRUT4 (Morice et al., 2012; note that land and ocean dP/dT values use global-mean temperature). Whiskers indicate 95% confidence intervals for observed linear trends (model trend confidence intervals are not shown, but are often large).

**Figure 2:** Observed and model simulated annual and zonal mean precipitation change (%/decade) for: top, observations where they exist over land; bottom, GCMs, all gridboxes. Top panel: Observed 1951-2005 changes (solid colored lines) from 4 datasets CRU TS3.0 updated, Harris et al. 2014; Zhang et al. 2007 updated; GPCC VasClimO, Beck et al. 2005; and GPCC Full data V6, Becker et al. 2013). Range of CMIP5 model simulations (grey shading, masked to cover land only) and multi-model ensemble
mean (black dashes, 'MM'). Blue shading shows latitudes where all observed datasets show positive trends and orange shading shows where all show negative trends. Interpolated data in the CRU dataset are masked out. Bottom panel: Trends based on global coverage from climate models from the Historical simulations (grey dashed lines are individual simulations, black dashed line multi-model mean; blue dashes multi-model mean from simulations forced by natural forcing only) compared to the 2006-2050 trend from the RCP4.5 multimodel simulations (green shading: 5-95% range, green dashes: multimodel mean). Blue (orange) shading indicates where more than two thirds of the historical simulations show positive (negative) trends.

**Figure 3:** Three observed estimates of long-term global and basin zonal-mean near-surface salinity changes, nominally for the 1950-2000 period. Positive values show increased salinities and negative values freshening. Changes are expressed on the Practical Salinity Scale (PSS-78) per 50-years. The data coverage, as used in Durack and Wijffels (2010), is shown in Supplementary Figure 1. Reproduced from Durack et al. (2013).

**Figure 4:** Number of in situ stations over time for the CRU TS 3.21 gridded precipitation dataset (updated from Harris et al., 2014). Evolution over decades of the latitudinal density of stations per zonal band for the Americas (orange), Europe/Africa (green) and Asia/Australasia (blue), stacked to indicate the zonal total. Incomplete data series are included as a fraction of available data. The black line indicates the number of stations per zonal band required to obtain an average zonal coverage of 1 station per (100km)$^2$ of land at that latitude. This figure shows the station numbers in absolute terms and in
relation to the latitudinally-varying land area. Other datasets have similar differences in coverage over time (see supplementary figure 2 for GPCC).

**Figure 5**: High latitude (55-90N) annual mean precipitation trends [mm/decade] from 1951-2005 for three observational datasets: Zhang et al. (2007; updated; 5x5 degree grid); GPCC Full data V6 (Becker et al., 2013), CRU TS3.0, updated (Harris et al., 2013; grid points with CRU station data available for >95% of the time are stippled) compared to the CMIP5 multimodel mean trend of Historical runs with all external forcings ('Multi-model Mean'). Note that both GPCC and CRU use spatial interpolation to varying extents, while Zhang et al., 2007 average a subset of stations only, considered to be homogeneous in the long-term within grid-boxes.

**Figure 6**: Top: The 2nd EOF of global sea surface temperature (3-yr running mean) data from 1920-2011 based on the HadISST data set. The red line is a smoothed index representing the inter-decadal Pacific Oscillation (IPO). The bottom panel shows smoothed precipitation anomalies averaged over the Southwest U.S. (black line) compared with the IPO index, scaled for comparison. (Reproduced from Dai 2013b).
Figure 1 left panel: Projected global-mean precipitation change (mm/day) against global-mean 2m air temperature change (K) from CMIP5 models, for four representative concentration pathways (RCP) scenarios. Values are means over successive decades between 2006 and 2095 and all ensemble members of each model. Anomalies are relative to mean values over 1986-2005 in the CMIP5 historical runs.

Right panel: Precipitation sensitivity for future (RCP scenarios) and past (Historical and Atmospheric Model Intercomparison, AMIP) change in precipitation amount [%] per degree global-mean warming. Trends are calculated from the linear least squares fit of annual global-mean precipitation change (%) against temperature (K) change relative to the period 1988-2005 (without decadal smoothing). Crosses indicate ensemble means for each CMIP5 model, circles indicate multi-model mean. Precipitation sensitivity is also shown for historical periods; comparing GCMs with GPCP, GPCC and CRU data (see text), using temperature changes from HadCRUT4 (Morice et al., 2012; note that land and ocean dP/dT values use global-mean temperature). Whiskers indicate 95% confidence intervals for observed linear trends (model trend confidence intervals are not shown, but are often large).
Figure 2: Observed and model simulated annual and zonal mean precipitation change (%/decade) for: top, observations where they exist over land; bottom, GCMs, all gridboxes. Top panel: Observed 1951-2005 changes (solid colored lines) from 4 datasets CRU TS3.0 updated, Harris et al. 2013; Zhang et al. 2007 updated; GPCC VasClimO, Beck et al. 2005; and GPCC Full data V6, Becker et al. 2013). Range of CMIP5 model simulations (grey shading, masked to cover land only) and multi-model ensemble mean (black dashes, 'MM'). Blue shading shows latitudes where all observed datasets show positive trends and orange shading shows where all show negative trends. Interpolated data in the CRU dataset are masked out. Bottom panel: Trends based on global coverage from climate models from the Historical simulations (grey dashed lines are individual simulations, black dashed line multi-model mean; blue dashes multi-model mean from simulations forced by natural forcing only) compared to the 2006-2050 trend from the RCP4.5 multimodel simulations (green shading: 5-95% range, green dashes: multimodel mean). Blue (orange) shading indicates where more than two thirds of the historical simulations show positive (negative) trends.
Figure 3: Three observed estimates of long-term global and basin zonal-mean near-surface salinity changes, nominally for the 1950-2000 period. Positive values show increased salinities and negative values freshening. Changes are expressed on the Practical Salinity Scale (PSS-78) per 50-years. The data coverage, as used in Durack and Wijffels (2010), is shown in Supplementary Figure 1. Reproduced from Durack et al. (2013).
Figure 4: Number of in situ stations over time for the CRU TS 3.21 gridded precipitation dataset (updated from Harris et al., 2013). Evolution over decades of the latitudinal density of stations per zonal band for the Americas (orange), Europe/Africa (green) and Asia/Australasia (blue), stacked to indicate the zonal total. Incomplete data series are included as a fraction of available data. The black line indicates the number of stations per zonal band required to obtain an average zonal coverage of 1 station per (100km)$^2$ of land at that latitude. This figure shows the station numbers in absolute terms and in relation to the latitudinally-varying land area. Other datasets have similar differences in coverage over time (see supplementary figure 2 for GPCC).
Figure 5: High latitude (55-90N) annual mean precipitation trends [mm/decade] from 1951-2005 for three observational datasets: Zhang et al. (2007; updated; 5x5 degree grid); GPCC Full data V6 (Becker et al., 2013), CRU TS3.0, updated (Harris et al., 2013; grid points with CRU station data available for >95% of the time are stippled) compared to the CMIP5 multimodel mean trend of Historical runs with all external forcings ('Multi-model Mean'). Note that both GPCC and CRU use spatial interpolation to varying extents, while Zhang et al., 2007 average a subset of stations only, considered to be homogeneous in the long-term within grid-boxes.
**Figure 6**: Top: The 2nd EOF of global sea surface temperature (3-yr running mean) data from 1920-2011 based on the HadISST data set. The red line is a smoothed index representing the inter-decadal Pacific Oscillation (IPO). The bottom panel shows smoothed precipitation anomalies averaged over the Southwest U.S. (black line) compared with the IPO index, scaled for comparison. (Reproduced from Dai 2013b).