Tropospheric moisture in the Southwest Pacific as revealed by homogenized radiosonde data: Climatology and decadal trend

Ning Li[a], Gilles Bellon[a] and Glenn McGregor[b]

[a] *The University of Auckland*

[b] *Durham University*

ABSTRACT: Radiosonde data are valuable to the research community because they provide observations for multiple decades and can be used to validate model output. However radiosonde data often suffer from quality issues, which has undermined their credibility. Therefore, corrections for biases and changepoints are needed to remedy the situation. Homogenization of monthly radiosonde specific humidity (q) from the 1970s to the present has been performed on selected stations over the Southwest Pacific (SWP) at three pressure levels (i.e., 850, 700 and 500 hPa). A three-step procedure involving a) adjustments for two sampling biases, b) detecting secular changepoints (i.e., discontinuities) using both statistical techniques and metadata validation, and c) an innovative break size estimate approach, has been implemented to achieve this aim. In the last step, a discontinuity-free pseudo-q is constructed from saturated specific humidity q_s which itself is derived from an already homogenized temperature (T) time series. This pseudo-q serves as a reference that not only distinguishes artificial from natural changepoints but also helps estimate the magnitudes of the discontinuity.

On the decadal time scale, the adjusted q (q_{adj}) exhibits spatially more consistent moistening ~~in the lower atmosphere~~at the 850 hPa level over most of the region and a contrast at 500 hPa between the moistening in the tropics and a drying in the subtropical South Pacific. Mean regional trend estimates are ~ 1.8% (850 hPa), -0.2% (700 hPa) and 1.3% (500 hPa) per

Formatted: Font: Not Bold, Highlight

Formatted: Font: Not Bold, Highlight

Formatted: Font: Not Bold, Highlight

Formatted: Font: Not Bold, Highlight

Formatted: Highlight

Formatted: Font: Not Bold, Highlight

decade. A climatological comparison ~~to~~ with the three latest reanalysis products – CFSR, ERA-Interim and MERRA, suggests the reanalyses have significant negative biases over Southeast Asia (SEA) at all three levels. Over Australia the biases are negative at 850 hPa while positive at 500 hPa. The reanalysis products tend to be~~are~~ more similar amongst themselves in the estimates of q, ~~-~~as compared to the ~~than the raw~~ radiosonde measurements of q and q_{adj}. The homogenized radiosonde q, when assimilated into reanalysis, is likely to lead to a more realistic model of the hydrological cycle.

Formatted: Highlight

KEYWORDS: homogenization; radiosonde; specific humidity; changepoint detection; trends; Southwest Pacific

1. Introduction

Atmospheric moisture constitutes only a tiny fraction (~0.001%) of the total global water, but this says nothing about its paramount role in regulating the Earth's climate via modulation of energy and heat budgets (Trenberth and Stepaniak, 2003). Annually, some 40,000 km^3 of water is transported from ocean to land in the form of water vapour, providing water resources for a range of human activities. Atmospheric moisture is a major indicator of the status of the climate system as global warming has the potential to increase atmospheric humidity, because of the temperature satuaration vapour pressure ~~relationship~~ association as expressed by the Clausius-Clapeyron relationship ~~('CC') has the potential to increase atmospheric humidity~~. Furthermore, water vapour is a potent Greenhouse Gas ~~(GHG)~~ with the largest positive feedback on temperature (IPCC, 2007)~~; the radiative effect of water vapour is~~, with an effect comparable to the ~~radiative~~ initial forcing ~~of carbon dioxide~~ ~~increases~~ (Trenberth *et al.*, 2007). From a weather forecast point of view, it has been shown that assimilating observed humidity into numerical models helps improve forecast skills for wind and temperature (Andersson *et al.*, 2007).

Despite the critical role that atmospheric moisture plays in the climate system, at a range of time and space scales, interest in assessing water vapour concentrations has only been renewed in the last decade or so, far behind the efforts devoted to~~invested in~~ quantifying temperature changes (e.g., Free *et al.*, 2004; Haimberger, 2007; McCarthy *et al.*, 2008; Menne and Williams, 2009; Sherwood *et al.*, 2008), in part because well-calibrated observations are ~~scarce~~rare (Trenberth *et al.*, 2005). This, along with data quality and scarcity issues ~~in~~ related to precipitation, evaporation and wind fields has led to considerable uncertainties ~~imbalance~~ in the observed atmospheric moisture budget (Lenters *et al.*, 2000; Stickler and Bronnimann, 2011; Yeh and Famiglietti, 2008).

Climate variability over the SWP is primarily modulated by the monsoon system, modes of large-scale circulation and sea surface temperature (SST). In particular, the El Nino-Southern Oscillation exerts the greatest influence on global and especially low latitude interannual climate variability (Cai, van Rensch et al. 2010, Jiang, Griffiths et al. 2013); the Pacific Decadal Oscillation, a long-lived ENSO-like oscillation in the Pacific, has been linked to shifts in the climate regime in many parts of the world (Salinger and Mullan 1999, Mantua and Hare 2002); the monsoon system in concert with the Indian Ocean Dipole affects climate over southeast Asia and parts of Australia (Ummenhofer, Sen Gupta et al. 2011); the Southern Annular Mode (SAM) affects climate of the Southern Hemisphere extratropics (Rao, Do Carmo et al. 2003, Ummenhofer, Sen Gupta et al. 2009); and SST modulates all these circulation features via its influence on important hydro-climate fields such as surface wind, specific humidity and sea level pressure (Mullan 1998). Numerous studies have hinted at possible trends in circulation modes and SST. For example, Fogt, Perlwitz et al. (2009) and Marshall (2003) found a positive trend in the SAM for DJF and MAM. Given the significance of the circulation features in explaining climate variability, the reported trends will have far-reaching consequences on the enviroment and human society.

Amongst the various records of atmospheric moisture, observations from radiosondes represent the only historical data that span more than 100 years. These are considered the ~~long term~~long-term baseline against which other sources of data should be compared with. In addition, radiosonde data are essential inputs into climate reanalyses. Multidecadal climate variations in reanalyses can only be adequately resolved by assimilating such data (Karl *et al.*, 1995). Notwithstanding their potential for shedding light on atmospheric moisture processes, radiosonde humidity observations, traditionally used for operational weather forecasts ~~provide~~ present a number of challenges for developing climatologies of atmospheric moisture. This is because radiosonde records suffer from many discontinuities owing to often undocumented changes in instrument, sampling and observation practices (Gaffen *et al.*, 1991; Ross and Elliott, 1996; Wade, 1994). Although a complete listing of the causes of inhomogeneities is infeasible, some of the common ones include: (i) sensor wetting or icing, (ii) solar radiation heating in daytime measurements, (iii) slow response time under cold conditions, and (iv) non-zero ground check value to show the scope of the problem (Miloshevich *et al.*, 2009; Smit *et al.*, 2013). It appears that such issues have hindered the use of radiosonde humidity data in climate research compared to the seemingly more reliable and readily available data sources.

Studies that have adjusted radiosonde discontinuities tend to report a spatially more consistent pattern of upward trends in tropospheric moisture content. McCarthy *et al.* (2009) examined the homogeneity of monthly temperature (T), specific humidity (*q*) and relative humidity (RH) in the Northern Hemisphere. They used a multi-step procedure, including assessment of the sensitivity of humidity replacement values and instrument-specific adjustments. Durre *et al.* (2009) analyzed trends in homogenized surface-to-500-hPa precipitable water (PW) at approximately 300 stations for the period 1973-2006. They found that the Northern Hemisphere as a whole experienced tropospheric moistening at about 0.45 mm per decade with statistical significance. The rate of moistening reached 1.94 mm per decade in the western tropical Pacific. Dai *et al.* (2011) were the first to perform homogenization on daily radiosonde humidity observations at the global scale. Using the archived humidity variable dewpoint depression (DPD), they quantile-matched histograms of segments of inhomogeneous data to that of the latest segment measured by modern hygrometers, assuming DPD distributions are comparable throughout the record length~~s of the record~~. The adjusted DPD implies small changes in RH during 1973-2008.

With regard to the homogenization procedure, the studies from McCarthy *et al.* (2009) and Durre *et al.* (2009) made use of reference series that are constructed from selected neighbouring

Formatted: Highlight

stations, while the study ~~from~~ of Dai *et al.* (2011) relies purely upon statistical techniques such as a variant of the Kolmogorov-Smironov test for changes in distributions and the penalized maximal *F* test for mean shifts in the occurrence frequency for discretized DPD. Regionally, Agustí-Panareda *et al.* (2009) showed how radiosonde humidity bias correction in the West African region can improve short-term ~~NWP~~ numerical weather prediction~~forecast~~. Zhao *et al.* (2012) assessed long-term trends in *q*, PW and RH over China, all of which were derived from homogenized radiosonde T (Haimberger *et al.*, 2008) and DPD (Dai *et al.*, 2011) data. On average, the effects of the various correction procedures act to eliminate wet biases common in early records, leading to stronger upwards trends in specific humidty (*cf.* Agustí-Panareda *et al.* (2009) where the three investigated sonde types have dry biases). In contrast, RH trends remain close to zero, largely owing to the strong coupling in the variations between temperature and moisture content.

Despite having a much shorter time span, satellite observations provide critical information over global oceans that are complementary to the usually land-based radiosonde stations. Positive trends in atmospheric moisture shown in the short SSM/I record over oceans have been attributed to ~~be of~~ anthropogenic origins (Santer *et al.*, 2007). Using the 10-year long merged GOME and SCIAMACHY dataset, Mieruch *et al.* (2008) found that trends are regionally specific, but the global average is positive. Independent satellite observations for the ice-free ocean regions yield total column water vapour trend estimates that are in good agreement with each other (Mieruch *et al.*, 2014). Noteworthy though for satellite observations are the issues of questionable data quality under cloudy conditions and crude vertical resolution.

With regard to reanalysis products, the issue of assimilating multiple data sources which introduce numerious discontinuities in the final outputs has been treated with insufficient effort (Uppala *et al.*, 2005). When this is coupled with uncertainties in model physics, significant identifiable problems in reanalyses emerge (e.g., Bengtsson *et al.*, 2004). Having analyzed the homogeneity of 2m air temperature of the Twentieth Century Reanalysis, Ferguson and Villarini (2014) found that globally only ~ 40% of the grid points are free from non-climate shifts. Over the 30-60° S latitudinal band, the figure lowers to an astonishing ~ 5%. In a study that analyzes the contribution of urbanization to local warming, Wang, Yan et al. (2013) found that reanalysis products cannot fully reproduce multidecadal variability in the temperature time series present in station data. Trenberth *et al.* (2005), Chen *et al.* (2008) and Carvalho and Jones (2013) all report large discrepancies in tropospheric humidity trends between their choices of reanalysis products, recommending great care be taken when using reanalyses for variability and trend studies.

In light of the shortcomings in satellite observations and reanalyses, and the success of recent radiosonde data quality control initiatives (Durre *et al.*, 2006), it seems that further improvement in the quality of radiosonde humidity observations is crucial. In this paper, we present a homogenized monthly radiosonde based *q* dataset for the Southw~~-W~~est Pacific (SWP) region and use this to develop a climatology of *q* and assess trends in this atmospheric moisture variable. We also make comparisons of climatology and trends with three reanalysis products. No attempt is made to offer an explanation for any observed trends in atmospheric moisture as we consider this beyond the scope of this paper. The paper complements those with an exclusive Northern Hemisphere focus (Durre *et al.*, 2009; McCarthy *et al.*, 2009) and can be compared with the results from Dai *et al.* (2011) where a very different homogenization philosophy was adopted ~~(This this was not carried out because the data have not been made publicly available)~~. Section 2 introduces the major input datasets. Details of the homogenization procedure are given in section 3. Results are presented in section 4, followed by a discussion in section 5. ~~and c~~Conclusions are drawn in section ~~5~~6.

2. Data

2.1. Quality controlled radiosonde archive

Sub-daily radiosonde data ~~over the period from the 1970s to present~~ containing T and DPD for the surface, ~~and~~ mandatory as well as significant pressure levels from the Integrated Global Radiosonde Archive (IGRA; Durre *et al.*, 2006) over the period 1970s-the present are the core inputs to this study. The IGRA archive provides quality assured data that have gone through a) fundamental sanity checks to ensure observations are plausible and do not contain level duplication, b) checks on surface elevation, c) internal consistency checks, d) checks for the repetition of values in time and in the vertical, e) climatological checks, f) additional checks on temperature and g) checks for data completeness.

Also provided are station metadata which document secular changes (~~Sources~~ sources of metadata are originally from Gaffen (1993), among others). Albeit not complete, this feature of the archive enables better separation of non-climatic from climatic changepoints. Despite the rigorous checks applied, some errors remain in IGRA. These are quite noticeable in the daily time series of log-*q* such as that shown in Figure 1.

Sonde data measured at the mandatory levels for the lower to middle troposphere were considered initially, but due to a lack of data, the 1000 and 925 hPa levels were discarded. Above the 500 hPa level, the atmosphere is virtually devoid of water. Therefore only the 850, 700 and 500 hPa levels were retained for homogenization. The raw hourly data were firstly converted from UTC to local time, and then averaged into daily daytime (between 0900 and 2000) and nighttime (between 2100 and 0800) means. Secondly, standard equations were used to compute *q* from pressure, T and DPD. Thirdly, Tukey's biweight robust mean estimator, which is resistant to the influence of outliers and is more efficient in estimating sampling variability (Lanzante, 1996), was used to compute the monthly means of T and q given that there are at least 14 days of observations in each variable within a month.

A total of 42 stations that are located within the 95°° E-225°° E and 15°° N-70°° S region were selected (Figure 2). Among them, 33 stations have a minimum record of 25 years and no more than 15% missing observations (*Class A* in Figure 2) while the other 9, having on average 27% missing observations failed to meet these criteria but are still included (*Class B* in Figure 2). We stress that this decision is necessary, because the very limited radiosonde data in the SWP deems any available data valuable ~~failed to meet these criteria but are still included for a more even geographical coverage~~. Eighteen stations have both daytime and nighttime launches, 15/9 have only daytime/nighttime launches.

Formatted: Font: Not Bold

Formatted: Font: Not Bold

Formatted: Font: Not Bold

Formatted: Font: Not Bold

2.2. Homogenized radiosonde temperature data

The homogenization procedure relies on the use of homogeneous station q_s and hence T time series (see section 3.3). We found that t~~T~~he Radiosonde Observation Correction Using Reanalyses (RAOBCORE) is among the most suitable dataset for this purpose ~~developed and frequently updated by Haimberger and his team~~ (Haimberger, 2007; Haimberger *et al.*, 2012) ~~is deemed suitable for this purpose~~. ~~Not only is it spatiotemporally more homogeneous, but more importantly, it is based on the same source data – IGRA. Version 1.5.1 of the RAOBCORE data for the 42 stations is used here.~~ Over the years RAOBCORE has been through a series of algorithm and methodological improvements, and intercomparison with other sources of data has confirmed its general quality (Haimberger, Tavolato et al. 2008). Despite known unresolved discontinuities in the final products, such as unrealistic upper tropospheric T trends over the former Soviet Union and a few remote stations (Haimberger, Tavolato et al. 2012), most likely in the earlier part of the record, we still have faith in RAOBCORE because a) low-level tropospheric

T data (the focus of the study) are not particularly inhomogeneous compared with the scarcer and more instrument-sensitive upper tropospheric T data, and 2) the uncorrected T inhomogeneities should have a random structure which is unlikely to lead to systematic problems in pseudo-*q* and thus the broader homogenization procedure. The only caveat is that trend estimates for q_{adj} are likely to be conservative, meaning that the downward tende~~a~~ncies sometimes present in the radiosonde data are not completely adjusted, leaving what would otherwise be more positve trends flatter.

RAOBCORE data consist of 16 pressure levels for 00UTC and 12UTC. They are available as gridded files with 10°×5° resolution for the period 1958-2011, or as station files with time span depending on the length of the individual station data but generally covering the early 1970s to the present. RAOBCORE uses the IGRA radiosonde data as inputs, which is the same as the current study. ~~Not only is it spatiotemporally more homogeneous, but more importantly, it is based on the same source data – IGRA.~~ Version 1.5.1 of the RAOBCORE data for the 42 stations is used here.

Formatted: Font: Italic

2.3. Reanalysis products

Humidity fields from the three third-generation global reanalysis products have been analyzed in terms of differences in climatology and trends with respect to q_{adj}. ERA-Interim ('ERAI' hereafter), developed by the European Centre for Medium-Range Weather Forecasts (ECMWF), features a number of improvements compared to its predecessor ERA-40. These include a 4-dimensional variational analysis, finer spatial resolution of approximately 80 km on 60 vertical levels from the surface up to 0.1 hPa, and a better representation of low frequency variability (Dee *et al.*, 2011). The NCEP climate forecast system reanalysis (CFSR) was produced to replace the previous National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalyses. Its key strengths include the very fine horizontal resolution of ~ 38 km and vertical resolution of 64 levels, advanced assimilation schemes and assimilating satellite radiances instead of retrievals (Saha *et al.*, 2010). Modern-era Retrospective Analysis for Research and Applications (MERRA) from NASA's Global Modeling and Assimilation Office is on the 0.5° × 0.667° grid with the highest vertical resolution of 72 levels. It was undertaken with a special focus on the representation of the hydrological cycle. Due to the use of the GEOGS-5 data assimillation system that implements Incremental Analysis Updates that gradually adjusts the model state toward the observed state, unrealistic spin down in the hydrological cycle is reduced. So far it is the only reanalysis product with closed energy and water budgets (Rienecker *et al.*, 2011). Monthly means of q of ERAI, CFSR and MERRA were downloaded for the period 1979-present.

3. Homogenization procedure

A generic homogenization procedure usually involves four steps (WCDMP, 2003). Firstly screening is carried out to eliminate physically impossible values. This is followed by the construction of reference series from neighbouring stations (step 2), and identification of discontinuities in the difference or ratio series between the candidate and the reference using mainly statistical techniques (step 3). The rationale for a reference series is that it helps to distinguish natural fluctuations from spurious signals (Gaffen *et al.*, 2000). Ideally, reference series should experience similar climatic conditions to the candidate time series under investigation ~~(i.e., the candidate series)~~ but include no artificial biases (WCDMP, 2003). Where station metadata are available, they are often incorporated in the changepoint identification because they provide crucial information on the dates of secular changes and hence prevent the climate record from contamination of non-climatic signals. The procedure ends with adjustments of segment means to that of the latest homogeneous segment (step 4).

The homogenization procedure used herein is analogous to the generic homogenization procedure but only differs in minor details. ~~Generally a homogenization procedure involves four steps (Figure 3). Firstly screening is carried out to eliminate physically impossile values. Secondly, changepoint identification using statistical techniques is performed, sometimes accompanied by metadata analysis that can provide crucial information on the dates of secular changes and hence prevent the climate record from contamination by non-climatic signals. Step three is the construction of reference series from neighbouring stations. Ideally, reference series should experience similar climatic conditions to the time series under investigation (i.e., the candidate series) but include no artificial biases (WCDMP, 2003). The difference or ratio between the candidate and the reference series is then analyzed for discontinuities. This step is essential as omitting it prevents distinguishing natural fluctuations from spurious signals (Gaffen *et al.*, 2000). The procedure ends with adjustments of segment means to that of the latest homogeneous segment.~~

Since the IGRA provides pre-processed data, the first step is skipped, being instead replaced by corrections for two well known sampling issues in radiosonde humidity records (see section 3.1). Furthermore, ~~as the third step,~~ a new method of creating a reference time series that makes use of available homogenized radiosonde T data is introduced, for there is a sparse radiosonde network in the SWP ~~due to the absence of nearby stations~~ and possibly identical timing of changepoints within countries (see section 3.3). It is worth mentioning that it is not our intention to make instrument-specific corrections which are likely to be time, level and geographyically dependent (see Ciesielski *et al.* (2009), for example). Rather we make recourse to the performance of the changepoint detection method to fix these. The specifics of the current procedure have been refined multiple times so as to make it objective, automatic and applicable to incoming inhomogeneous radiosonde humidity data in the future (provided a pseudo-*q* reference series can be created).

3.1. Corrections for sampling biases

There are two biases – under representative sampling under dry or cold conditions – that have plagued radiosonde humidity records (McCarthy *et al.*, 2009). Between 1973 and 1993, many U.S. stations and some stations in the study region adopted the practice of recording RH less than 20% as 19%, or DPD as 30° C (Ross and Elliott, 1996). The implication of having no dry observations results in an apparent drying trend up until the practice was terminated. An example for the 700 hPa level nighttime observations at Brisbane is shown in Figure ~~4~~3. The abnormally high DPD=30° C occurrence during 1983-1998 indicates the prevalence of the practice in the region. Figure ~~5~~ 4 maps the spatial extent of this practice. All stations in Australia, a few island stations in the Pacific and one station in the Antarctic adopted the practice. It is unclear whether the three stations located at Tahiti, Rapa (French Polynesia) and Noumea (New Caledonia) adopted this ~~approach~~practice.

To account for the ~~this~~ missing dry observation sampling issue~~bias~~, ~~for each pre-1995 DPD=30° C occurrence,~~ we replaced each pre-1995 DPD=30° C occurrence~~it~~ with the median of a pool of 200 randomly sampled post-1995 DPDs that are greater than 30° C, allowing conservative variations in the replaced DPDs to be generated. However, extreme dry events cannot be recovered. In comparison with the before correction distribution, a shift of the centre toward lower values and a higher density of the very low *q* is observed, conforming to the expected distribution of a more homogeneous record.

Low humidity events are considered unreliable under cold conditions and hence are preferentially rejected in standard quality control procedures (McCarthy *et al.*, 2009). For

Formatted: Font: Not Bold

Formatted: Font: Not Bold

Formatted: Font: Not Bold

Formatted: Font: Not Bold

example, early hygristors ~~using~~ that use goldbeater's skin and films of lithium chloride, and carbon hygristors, perform poorly at tempertures below -20° C and ~~carbon hygristors below~~ -40° C, respectively (Smit *et al.*, 2013). Furthermore, DPD was reported as missing when temperatures were below -40° C (Ross and Elliott, 1996). Metadata show that a large number of the stations have adopted this practice (Figure ~~6~~5), although only the two Antarctic stations have recorded very low temperatures. Both issues have led to a warm sampling bias in the early part of the record.

To minimize the effects of missing cold observations, we used an approach similar to that in McCarthy *et al.* (2009), that is, firstly, to reject daily humidity reports when more than 5% of T<-40°° C occurs in a given month, and secondly, to regress the natural logarithm of monthly *q* against monthly T using the post-1995 data then apply the slope coefficients to estimate missing *q* for the pre-1995 period.

3.2. Changepoint detection

The Wild Binary Segmentation ('WBS') algorithm developed by Fryzlewicz (2014) was used for the search of multiple candidate changepoints. WBS is ~~an improvement over~~based on the standard Binary Segmentation (BS) method where changepoints are searched recursively on shorter and shorter and evenly split segments until a certain criterion is reached. Due to its low computation cost and conceptual simplicity, BS is arguably the most commonly used changepoint detection algorithm (Killick, Fearnhead et al. 2012). One advantage of WBS over BS is that a local rather than global cumulative-sum statistic ('CUSUM') is computed as a result of having random subsamples. In this way, the issue of the global CUSUM being unsuitable for certain multiple changepoints characteristics can be overcome. In addition, WBS is capable of finding

Formatted: Font: Not Bold

Formatted: Font: Not Bold

Formatted: Font: Not Bold

Formatted: Font: Not Bold

closely spaced changepoints. Readers may refer to Fryzlewicz (2014) for a thorough discussion on the computational, methodological and theoretical aspects of WBS. ~~Fryzlewicz (2014)~~

In short, the sequential steps in the search are listed below:

1) Remove seasonality in q;
2) Use the WBS function to obtain candidate changepoints;
3) Categorize the candidate changepoints as ‘documented artificial changepoints’ if they occur within a short timeframe of those documented in metadata;
4) As an interim measure, if all three information criteria (i.e., ‘strengthened Schwarz Information Criterion’, ‘Bayes Information Criterion’ and ‘modified Bayes Information Criterion’) implemented in the WBS ~~package~~ function agree on the changepoints, they are marked as ‘undocumented changepoints’ which can either be artificial or natural. For the sake of simplicity, the rest of the candidate changepoints are seen as false positives or secondary changepoints that entail minor level shifts;
5) Both the changepoints determined from steps 3 and 4 will form segments of series that are subject to level adjustments (see section 3.3).

Note that natural changepoints if any will be detected at this stage, but will not be adjusted as a result of the next step.

3.3. Creation of reference time series and break size estimates

Depending on the distributions of q, different linear regression models between q_s and properly transformed q over the post-2005 discontinuity-free segments are fitted. They are q~q_s for normally distributed q, $q^{1/2}$~q_s for right-skewed q, and q^2~q_s for left-skewed q. The two transformations have been shown to enhance the normality of data in practice. The coefficients from these regressions are then used to back construct so called 'pseudo-*q*' in time. The pseudo-*q* series serves as a reference.

The regressions are successful in most cases ~~cases as indicated by the strong and positive linearity of the relationships~~. The mean adjusted R^2 is 0.5, and the mean correlation coefficients for the positive and negative relationships are 0.73 and -0.63, respectively. One example of the diagnostics of the fit is given in ~~(~~Figure ~~7~~6~~)~~. Where no linear relationships can be established, level adjustments are carried out without using reference time series on the more certain documented artificial changepoints. We follow WCDMP's (2003) recommendation to adjust the data to match the conditions of its most recent homogeneous section. Specifically, this is done by calculating separate means for the difference (*q* minus pseudo-*q*) series for the segments defined by the changepoints. Subsequently, the obtained means are compared by calculating their difference and the obtained adjustment value is added to the inhomogeneous part. Noteworthy are the existence of inverse relationships seen mainly at the 700 and the 500 hPa levels for areas over the western tropical Pacific. Indeed, Ross *et al.* (2002) found that the observed positive/negative *q*~T correlations at the lower/middle troposphere over the Pacific region suggests displacement of convection to other parts of the tropics hence regionally enhanced subsidence. In the free troposphere, enhanced subsidence favours stronger adiabatic warming and drying, and subsequently higher temperature observations. The lack of convection on the other hand results in lower humidity observations.

4. Results

4.1. Homogenization

Results of changepoint detection for one station are illustrated in Figure ~~8~~7. At the 850 hPa level, three changepoints that have been documented in metadata are suggested by the WBS

Formatted: Default Paragraph Font, Font: (Default) +Body (Verdana), 12 pt, Character scale: 90%

Formatted: Default Paragraph Font, Font: (Default) +Body (Verdana), 12 pt, Character scale: 90%

Formatted: Default Paragraph Font, Font: (Default) +Body (Verdana), 12 pt, Character scale: 90%

algorithm. The nature of these change~~-~~points would be indeterminable without the metadata. At the other two pressure levels, the induced level shifts of the change~~-~~points are too small to be picked up by the algorithm and hence are left intact. Through visual inspection and the assistance of metadata analysis, ~~we feel~~it appears that the homogenization procedure is reasonably successful. On average, the 850, 700 and 500 hPa series have 4.9, 2.6 and 3.1 changepoints, respectively.

4.2. Climatology

The climatology of q_{adj} over the period 1970s-present has a number of features that conform to the general understanding of the spatial pattern of atmospheric moisture (Figure ~~9~~8). Firstly, higher moisture content ~~is higher~~ at low latitudes or altitudes confirms that temperature is the first-order control on the spatial configurations of moisture content, given that water is not limited~~, as a result of decreasing temperature poleward and upward~~. Secondly, it is moister over the maritime continent and the ocean than over the major land mass of Australia and the Antarctic. These two features reiterate the importance of energy and water availabilty as primary controls on atmospheric moisture content.

Climatological differences between the three reanalysis products and q_{adj} are calculated in order to evaluate the spatial and vertical discrepancies. Here a bilinear interpolation of reanalysis q to the station coordinates is implemented for fair comparison. Figure ~~10~~ 9 shows that reanalysis q at the 850 hPa and the 700 hPa levels have widespread dry biases especially over the equator. These biases exhibit a latitudinal gradient, approaching zero at the southernmost stations. At the 500 hPa level, the biases are positive over Australia and negative over SEA, indicating a possible overestimation of convective stability over Australia and an underestimate of moisture content over SEA. Having the largest negative climatological bias of 2.6g/kg (bottom left plot in Figure

9), MERRA is least ~~agreeable~~ similar to q_{adj} and slightly drier than CFSR and ERAI at the 850 hPa level. This is in contrast to the near-surface profiles of marine q in Kent *et al.* (2014) for ERAI and MERRA where they show global averages of q_2 and q_{10} ~~rose faster~~are higher in MERRA relative to ERAI over the period 1989-2007. If both statements are true, it would suggest the vertical structure depicted in MERRA or ERAI is problematic.

Figure 10~~1~~ shows seasonal cycles of q over the parts of SEA and Australia that are covered by a relatively dense radiosonde network and from the equivalent reanalysis grid cells for the period 1979-present. The seasonal cycles are remarkably alike, but the absolute bias is substantial. The most extreme bias is between q_{adj} and MERRA for SEA at the 850 hPa level reaching an average of 4 g/kg. Over SEA the bias appears to be relatively constant regardless of season. For Australia, the bias is noticeable at the 850 hPa level. At the 700 hPa level the bias becomes negligible during austral winter, and a subdued annual cycle of all reanalysis products is seen. At the 500 hPa level, the bias reverses in sign. ~~Also note that the biases are present even compared with the raw radiosonde q suggesting they are not due to the homogenization procedure.~~It is worth pointing out that even the homogenized q fails to narrow the gaps with the reanalysis q. Why the reanalysis products exhibit such large deviations from the observations in SEA deserves more attention.

~~Of interest is what accounts for the large differences seen for the two regions.~~ One possibility hinted at in a previous study is the change in observing system from time to time (Bosilovich *et al.*, 2011). Given this, the datasets for the two regions were split into three periods as follows: a) before the launching of SSM/I (prior to 1987), b) the period when SSM/I has been assimilated (1987-1997), and c) the period when AMSU-A has also been assimilated (1998-2010). As the differences remain almost invariant for the three periods, the observing systems do not seem to be the main cause of the differences (not shown).

4.3. Decadal trends

Linear trends over the period 1970s-present for the before and after correction series are displayed in Figures ~~12~~ 11 (daytime) and ~~13~~ 12 (nighttime). Overall the corrections result in statistically significant and more spatially consistent trends in an increase in atmospheric moisture for the 850 hPa level and reversal of trends from negative to neutral or positive for the 700 hPa and the 500 hPa levels. Stations that show ~~long term~~long-term drying seem to be located in the vicinity of the descending ~~limb~~ branch of the Hadley Cell. To what extent this relates to changes in large scale atmospheric circulation is worth exploring. Willett *et al.* (2010) study decadal changes in surface specific humidity. They believe that although there is evident drying over the arid areas in the extratropical Southern Hemisphere, it is not entirely attributable to temperature because the Clausius-Clapeyron~~CC~~ relationship is small and negative (but not significant) in this region. This conclusion implies that atmopsheric circulation which manifests itself as horizontal and vertical moisture transport may be a vital part in explaining the drying.

SEA appears to have experienced moistening at the 5% significance level at the majority of the stations throughout the lower to middle troposphere. Regarding trend estimates by level, Figure ~~14~~ 13 shows that the spread of trend increases from the lower to the upper levels. Plausible reasons for this are discussed in section 5.

Figure ~~15~~ 14 shows differences in trend estimates between the reanalysis products and radiosondes. Consensus amongst reanalyses is weak at the 850 hPa level. At a number of stations, they do not even agree in the sign of change. While differences in trend estimates are generally negative for ERAI and CFSR, they are positive for MERRA. In other words, in comparison with q_{adj}, ERAI and CFSR are characterized by either negative or weaker positive ~~long term~~long-term trends. In contrast, MERRA has steeper positive or less negative trends. From a regional

Formatted: Indent: First line: 0.5 cm perspective, the reanalysis mean trend differences compared to the baseline q_{adj} amount to no more than 0.1g/kg/10-yr, but the range of trend discrepancies can be as large as +/- 0.7g/kg/10-yr.

5. Discussion ~~and conclusions~~

Homogenization of observational data is a challenging task but is necessary because it is an integral element in ~~long term~~long-term climate monitoring, and it provides valuable validation for other sources of information such as reanalysis products. Many known deficiencies in reanalyses could have errors originating from poorly bias-adjusted observations, among other things. Therefore homogeneous time series of atmospheric moisture, in this case *q*, used as input to reanalysis ~~projects~~ products should better reconcile inter-~~narrow the gaps in inter-~~reanalysis inconsistencies~~y~~ and result in ~~better~~improved agreement with observations. To the best of our knowledge, homogenized radiosonde temperature records (but not humidity) have started to be assimilated in reanalyses (e.g., ERAI and MERRA; Andrae *et al.*, 2004; Dee *et al.*, 2011; Haimberger *et al.*, 2008). This is a welcom~~ing~~ed step, and ~~should pave the way~~we expect to see the same thing happening to radiosonde humidity observations ~~for assessment of atmospheric humidity fields in reanalyses~~ in the near future.~~because of the importance of this atmospheric physcial property for understanding climate system processes.~~

The motivation for conducting homogenization of radiosonde *q* observations over the SWP lies in the lack of studies in this region, and hence generally lack of data for constraining models. It is worth emphasizing that however prudent we may have been in ~~producing~~ adjusting the data set~~ on which this paper is based~~, some remaining inhomogeneities are ~~likely to remain~~inevitable because knowledge regarding observational errors in radiosonde-based humidity measurements is limited (e.g., other sampling biases might have been present but unknown to us at this time). It should also be acknowledged that errors are likely to be introduced as each step in the homogenization process makes assumptions about the configurations of the time series. For example, gradual drift in time series values cannot be easily detected, which introduces uncertainties in trend estimates. Indeed, none of the existing homogenization procedures that we are aware of deals with this problem. Moreover, although advanced statistical techniques could be applied to the homogenization problem their utility may be somewhat limited given an in-depth understanding of the characteristics of the series under investigation is required *a priori*. Notwithstanding these caveats, ~~after closely scrutinizing all the time series produced in this analysis~~the use of an objective changepoint detection technique, supplementary information to classify identifed changepoints, and reference series ~~we feel confident that~~has given us some confidence that the procedure has done a reasonable job in producing more reliable tropospheric q estimates~~the most outstanding changepoints have been identified by the WBS algorithm~~.

A unique feature of our analysis is the use of q_s which ~~has~~ has been derived from previously homogenized RAOBCORE T station data as predictors of q. Pseudo-q time series for individual stations and levels have been constructed and used as references ~~to~~ for the candidate q time series. In this way, we have been able to diagnose whether a particular candidate changepoint identified in the previous step is natural or artificial. Fitting a linear model between q (and its two transformations) and q_s ~~as shown in Figure 7~~ is in ~~most~~ many cases justified.

Formatted: Not Superscript/ Subscript

When non-linearity or no notable linear pattern is present, more conservative adjustments, that is, changepoints that have been documented in metadata, have been carried out. This decision prioritizes the preservation of climate shifts at the expense of leaving some changepoints in the time series uncorrected. In contrast, adjusting for all candidate changepoints would lead to the undesirable effects of filtering out natural variability along with non-climate signals. The overwhelming consequence of this is to remove trends that we want to detect.~~result in excessive smoothing of natural fluctuations, which should be done only if the reason behind the choice is~~

~~compelling.~~ Should there be inhomogeneities remaining in T or no associations between q and q_s, the accuracy of the labelling of the detected changepoints (natural or artificial) is compromised. As such, one should be cautious in interpreting values at individual stations.

The most intriguing result is perhaps the climatological biases of the reanalyses q. Not only are they present for different segments of the entire time period, but also systematically lower over the three pressure levels for the two regions considered herein ~~The fact that 1) they are present for different segments of the entire time period, 2) systematically lower over the three pressure levels for the two regions considered herein~~ (except for the 500 hPa level over ~~SEA~~Australia)~~;~~. In addition, the fact~~s~~ that~~3)~~ ~~there is a near complete~~the radiosonde q record over SEA contains little missing data (hence no interpolation), and that ~~4) the well-known difficulties of~~ modeling deep convection over the equatorial warm pool has been a well-known difficulty (Ricciardulli and Sardeshmukh, 2002) would imply issues in the reanalysis assimilation schemes such as eliminating or assigning small weighting to observations. Unfortunately, the ways in which radiosonde q was handled are not comprehensively documented in either CFSR or MERRA. ERAI excluded radiosonde q in extreme cold conditions and assimilated the rest without bias correction (Dee *et al.*, 2011). Therefore the speculation regarding the way in which observations have been handled in reanalyses remains to be verified by future studies.

The large humidity differences seen between the reanalyses and observations in the two regions have also been found in the near-surface marine context. A recent study by Kent *et al.* (2014) compared the quality of eight global marine surface q datasets including *in situ* observations, reanalyses and blended datasets. Although they show good agreement in the interannual variations and seasonal cycles, estimates of the mean values are substantially different, reaching a similar magnitude as in our study. Jin *et al.* (2015) show that near-surface q has the largest inter-dataset differences in regions with high absolute humidity. Results from these

Formatted: Font: Not Italic

Formatted: Font: Italic

Formatted: Font: Italic, Subscript

studies and those from the current study cast doubt on the legitimacy of reanalysis-based studies of atmospheric moisture ~~related analyses~~ over the equatorial western Pacifc.

Specific humidity at 500 hPa possesses a wide range of percentage trend estimates (Figure ~~14~~13). Given the strong coupling between temperature and humidity, it is logical to ask if the spatial spread of trend is also seen in temperature. However we cannot draw any conclusion from existing studies because virtually all of them focus on spatially-averaged temperature estimates. Another possible explanation is that episodic intrusions of dry air from the stratosphere may have impacted the humidity trend at 500 hPa. As those intrusions may preferentially affect some regions more than others, the large humidity trend differences in space may be physically grounded (G. Bodeker, personal communication, March 23, 2016). Finally the spatial changes in trends could be due to ~~This can have two interpretations, one is that the trends are less spatially uniform at the upper levels. More likely however, this is due to~~ the poorer data quality and larger data gaps at this level, rendering a much broader range of trend estimates. This explanation would imply that even though the between-dataset trend differences for the 500 level are small (Figure ~~15~~14), they are less reliable.

6. Conclusions

In summary, this study presents a new homogenized radiosonde q dataset for the SWP region where high-quality observations are lacking. We have implemented a procedure whch has a ~~been succesful~~reasonable performance ~~at~~ in identifying artificial changepoints and removing the corresponding shifts in moisture values. This has allowed some important characteristics of q to be recovered, providing much more credible long-term tropospheric humidity records. We believe that the homogenized radiosonde q, when assimilated into reanalysis, is likely to lead to a more realistic model of the hydrological cycle.

The main findings are: 1) the adjusted *q* observations have shown more consistent moistening trends ~~in the lower atmosphere~~at 850 hPa over most of the region and a contrast at 500 hPa between the moistening in the tropics and a drying in the subtropical Southern Pacific; 2) the mean regional trend estimates from the adjusted *q* are 1.8%, -0.2% and 1.3% per decade for the 850, 700 and 500 hPa levels, with the spread increasing from lower to upper levels, 3) compared with the adjusted radiosonde *q*, ERAI, CFSR and MERRA have negative biases over SEA at all three levels. Over Australia the biases are negative at 850 hPa while positive at 500 hPa with MERRA being least close to the adjusted radiosonde *q*. The magnitude of the bias is substantial over SEA for reasons unknown at present.

Future homogenization efforts should focus on understanding the characteristics of the inhomogeneous data records as inhomogeneities come in many forms and affect the statistical properties of the data in various ways. Mo~~ve~~reover, ~~wiser~~ good monitoring practices such as documenting operating procedures, instruments and other information pertinent to data interpretation should be encouraged so that climate scientists can capitalize on the rich information ~~to~~ and make sound conclusions ~~concerning~~ regarding climate variability and change ~~accurate~~.

Formatted: Indent: First line: 0.5 cm

Acknowledgements

We thank Professor Leonard Haimberger from University of Vienna for his sharing of data and kindness in providing information critical to the homogenization procedure.

Formatted: Indent: First line: 0.5 cm

References

Agustí-Panareda A, Vasiljevic D, Beljaars A, Bock O, Guichard F, Nuret M, Garcia Mendez A, Andersson E, Bechtold P, Fink A, Hersbach H, Lafore J, Ngamini J, Parker D, Redelsperger J, Tompkins A. 2009. Radiosonde humidity bias correction over the West African region for the special AMMA reanalysis at ECMWF. *Quarterly Journal of the Royal Meteorological Society* **135**(640): 595-617. DOI: 10.1002/qj.396.

Andersson E, Hólm E, Bauer P, Beljaars A, Kelly G, McNally A, Simmons A, Thépaut J, Tompkins A. 2007. Analysis and forecast impact of the main humidity observing systems. *Quarterly Journal of the Royal Meteorological Society* **133**(627): 1473-1485. DOI: 10.1002/qj.112.

Andrae U, Sokka N, Onogi K. 2004. The radiosonde temperature bias correction in ERA-40. *ERA-40 Project Report Series, ECMWF* **15**: 34.

Bengtsson L. 2004. Can climate trends be calculated from reanalysis data?. *Journal of Geophysical Research-Atmospheres* **109**(D11). DOI: 10.1029/2004jd004536.

Bosilovich M, Robertson F, Chen J. 2011. Global energy and water budgets in MERRA. *Journal of Climate* **24**(22): 5721-5739. DOI: 10.1175/2011jcli4175.1.

Cai WJ, van Rensch P, Cowan T, Sullivan A. 2010. Asymmetry in ENSO teleconnection with regional rainfall, its multidecadal variability, and impact. *Journal of Climate* **23**(18): 4944-4955.

Carvalho L, Jones C. 2013. CMIP5 simulations of low-level tropospheric temperature and moisture over the tropical Americas. *Journal of Climate* **26**(17): 6257-6286. DOI: 10.1175/jcli-d-12-00532.1.

Chen J, Del Genio A, Carlson B, Bosilovich M. 2008. The spatiotemporal structure of twentieth-century climate variations in observations and reanalyses. Part I: Long-term trend. *Journal of Climate* **21**(11): 2611-2633. DOI: 10.1175/2007jcli2011.1.

Ciesielski P, Johnson R, Wang J. 2009. Correction of humidity biases in Vaisala RS80-H sondes during NAME. *Journal of Atmospheric and Oceanic Technology* **26**(9): 1763-1780. DOI: 10.1175/2009jtecha1222.1.

Dai A, Wang J, Thorne P, Parker D, Haimberger L, Wang X. 2011. A new approach to homogenize daily radiosonde humidity data. *Journal of Climate* **24**(4): 965-991. DOI: 10.1175/2010jcli3816.1.

Dee D, Uppala S, Simmons A, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda M, Balsamo G, Bauer P, Bechtold P, Beljaars A, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer A, Haimberger L, Healy S, Hersbach H, Hólm E, Isaksen L, Kållberg P, Köhler M, Matricardi M, McNally A, Monge-Sanz B, Morcrette J, Park B, Peubey C, de Rosnay P, Tavolato C, Thépaut J, Vitart F. 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society* **137**(656): 553-597. DOI: 10.1002/qj.828.

Durre I, Vose R, Wuertz D. 2006. Overview of the Integrated Global Radiosonde Archive. *Journal of Climate* **19**(1): 53-68. DOI: 10.1175/jcli3594.1.

Durre I, Williams C, Yin X, Vose R. 2009. Radiosonde-based trends in precipitable water over the Northern Hemisphere: An update. *Journal of Geophysical Research-Atmospheres* **114**(D5). DOI: 10.1029/2008jd010989.

Ferguson C, Villarini G. 2014. An evaluation of the statistical homogeneity of the twentieth century reanalysis. *Climate Dynamics* **42**(11-12): 2841-2866. DOI: 10.1007/s00382-013-1996-1.

Fogt RL, Perlwitz J, Monaghan AJ, Bromwich DH, Jones JM, Marshall GJ. 2009. Historical SAM variability. Part II: Twentieth-century variability and trends from reconstructions, observations, and the IPCC AR4 models. *Journal of Climate* **22**(20): 5346-5365.

Free M, Angell J, Durre I, Lanzante J, Peterson T, Seidel D. 2004. Using first differences to reduce inhomogeneity in radiosonde temperature datasets. *Journal of Climate* **17**(21): 4171-4179. DOI: 10.1175/jcli3198.1.

Fryzlewicz, P. 2014. Wild binary segmentation for multiple change-point detection. *The Annals of Statistics* **42**(6): 2243-2281. DOI: 10.1214/14-Aos1245.

Gaffen D. 1993. Historical changes in radiosonde instruments and practices. IOM-50, WMO/TD-541. Available at https://googledrive.com/host/0BwdvoC9AeWjUazhkNTdXRXUzOEU/wmo-td_541_en.pdf.

Gaffen D, Barnett T, Elliott W. 1991. Space and time scales of global tropospheric moisture. *Journal of Climate* **4**(10): 989-1008. DOI: 10.1175/1520-0442(1991)004<0989:satsog>2.0.co;2.

Gaffen D, Sargent M, Habermann R, Lanzante J. 2000. Sensitivity of tropospheric and stratospheric temperature trends to radiosonde data quality. *Journal of Climate* **13**(10): 1776-1796. DOI: 10.1175/1520-0442(2000)013<1776:sotast>2.0.co;2.

Haimberger L. 2007. Homogenization of radiosonde temperature time series using innovation statistics. *Journal of Climate* **20**(7): 1377-1403. DOI: 10.1175/jcli4050.1.

Haimberger L, Tavolato C, Sperka S. 2008. Toward elimination of the warm bias in historic radiosonde temperature records — Some new results from a comprehensive intercomparison of upper-air data. *Journal of Climate* **21**(18): 4587-4606. DOI: 10.1175/2008jcli1929.1.

Haimberger L, Tavolato C, Sperka S. 2012. Homogenization of the global radiosonde temperature dataset through combined comparison with reanalysis background series and neighboring stations. *Journal of Climate* **25**(23): 8108-8131. DOI: 10.1175/jcli-d-11-00668.1.

IPCC. 2007. *Climate Change 2007: Synthesis Report.* Available at https://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr.pdf.

Jiang HB, Griffiths G, Lorrey A. 2013. Influence of large-scale climate modes on daily synoptic weather types over New Zealand. *International Journal of Climatology* **33**(2): 499-519.

Jin X, Yu L, Jackson D, Wick G. 2015. An improved near-surface specific humidity and air temperature climatology for the SSM/I satellite period. *Journal of Atmospheric and Oceanic Technology* **32**(3): 412-433. DOI: 10.1175/jtech-d-14-00080.1.

Karl T, Derr V, Easterling D, Folland C, Hofmann D, Levitus S, Nicholls N, Parker D, Withee G. 1995. Critical issues for long-term climate monitoring. *Climatic Change* **31**: 185-221.

Kent E, Berry D, Prytherch J, Roberts J. 2014. A comparison of global marine surface-specific humidity datasets from in situ observations and atmospheric reanalysis. *International Journal of Climatology* **34**(2): 355-376. DOI: 10.1002/joc.3691.

Killick R, Fearnhead P, Eckley IA. 2012. Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association* **107**(500): 1590-1598.

Lanzante J. 1996. Resistant, robust and non-parametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde station data. *International Journal of Climatology* **16**(11): 1197-1226. DOI: 10.1002/(sici)1097-0088(199611)16:11<1197::aid-joc89>3.0.co;2-l.

Lenters J, Coe M, Foley J. 2000. Surface water balance of the continental United States, 1963–1995: Regional evaluation of a terrestrial biosphere model and the NCEP/NCAR reanalysis. *Journal of Geophysical Research-Atmospheres* **105**(D17): 22393. DOI: 10.1029/2000jd900277.

Mantua NJ, Hare SR. 2002. The Pacific Decadal Oscillation. *Journal of Oceanography* **58**(1): 35-44.

Marshall GJ. 2003. Trends in the southern annular mode from observations and reanalyses. *Journal of Climate* **16**(24): 4134-4143.

McCarthy M, Thorne P, Titchner H. 2009. An analysis of tropospheric humidity trends from radiosondes. *Journal of Climate* **22**(22): 5820-5838. DOI: 10.1175/2009jcli2879.1.

McCarthy M, Titchner H, Thorne P, Tett S, Haimberger L, Parker D. 2008. Assessing bias and uncertainty in the HadAT-adjusted radiosonde climate record. *Journal of Climate* **21**(4): 817-832. DOI: 10.1175/2007jcli1733.1.

Menne M, Williams C. 2009. Homogenization of temperature series via pairwise comparisons. *Journal of Climate* **22**(7): 1700-1717. DOI: 10.1175/2008jcli2263.1.

Mieruch S, Noël S, Bovensmann H, Burrows JP. 2008. Analysis of global water vapour trends from satellite measurements in the visible spectral range. *Atmospheric Chemistry and Physics* **8**: 491-504. DOI: 10.5194/acp-8-491-2008.

Mieruch S, Schröder M, Noël S, Schulz J. 2014. Comparison of decadal global water vapor changes derived from independent satellite time series. *Journal of Geophysical Research-Atmospheres* **119**(22): 12489-12499. DOI: 10.1002/2014jd021588.

Miloshevich L, Vömel H, Whiteman D, Leblanc T. 2009. Accuracy assessment and correction of Vaisala RS92 radiosonde water vapor measurements. *Journal of Geophysical Research-Atmospheres* **114**(D11). DOI: 10.1029/2008jd011565.

Mullan AB. 1998. Southern hemisphere sea-surface temperatures and their contemporary and lag association with New Zealand temperature and precipitation. *International Journal of Climatology* **18**(8): 817-840.

Rao VB, Do Carmo AMC, Franchito SH. 2003. Interannual variations of storm tracks in the Southern Hemisphere and their connections with the Antarctic oscillation. *International Journal of Climatology* **23**(12): 1537-1545.

Ricciardulli L, Sardeshmukh P. 2002. Local time- and space scales of organized tropical deep convection. *Journal of Climate* **15**(19): 2775-2790. DOI: 10.1175/1520-0442(2002)015<2775:ltasso>2.0.co;2.

Rienecker M, Suarez M, Gelaro R, Todling R, Bacmeister J, Liu E, Bosilovich M, Schubert S, Takacs L, Kim G, Bloom S, Chen J, Collins D, Conaty A, da Silva A, Gu W, Joiner J, Koster R, Lucchesi R, Molod A, Owens T, Pawson S, Pegion P, Redder C, Reichle R, Robertson F, Ruddick A, Sienkiewicz M, Woollen J. 2011. MERRA: NASA's Modern-Era Retrospective Analysis for research and applications. *Journal of Climate* **24**(14): 3624-3648. DOI: 10.1175/jcli-d-11-00015.1.

Ross R, Elliott W. 1996. Tropospheric water vapor climatology and trends over North America: 1973–93. *Journal of Climate* **9**(12): 3561-3574. DOI: 10.1175/1520-0442(1996)009<3561:twvcat>2.0.co;2.

Ross R, Elliott W, Seidel D, Participating AMIP-II Modeling Groups. 2002. Lower-tropospheric humidity–temperature relationships in radiosonde observations and atmospheric General Circulation Models. *Journal of Hydrometeorology* **3**(1): 26-38. DOI: 10.1175/1525-7541(2002)003<0026:lthtri>2.0.co;2.

Saha S, Moorthi S, Pan H, Wu X, Wang J, Nadiga S, Tripp P, Kistler R, Woollen J, Behringer D, Liu H, Stokes D, Grumbine R, Gayno G, Wang J, Hou Y, Chuang H, Juang H, Sela J, Iredell M, Treadon R, Kleist D, Van Delst P, Keyser D, Derber J, Ek M, Meng J, Wei H, Yang R, Lord S, Van Den Dool H, Kumar A, Wang W, Long C, Chelliah M, Xue Y, Huang B, Schemm J, Ebisuzaki W, Lin R, Xie P, Chen M, Zhou S, Higgins W, Zou C, Liu Q, Chen Y, Han Y, Cucurull L, Reynolds R, Rutledge G, Goldberg M. 2010. The NCEP Climate Forecast System Reanalysis. *Bulletin of the American Meteorological Society* **91**(8): 1015-1057. DOI: 10.1175/2010bams3001.1.

Salinger MJ, Mullan AB. 1999. New Zealand climate: Temperature and precipitation variations and their links with atmospheric circulation 1930-1994. *International Journal of Climatology* **19**(10): 1049-1071.

Santer B, Mears C, Wentz F, Taylor K, Gleckler P, Wigley T, Barnett T, Boyle J, Bruggemann W, Gillett N, Klein S, Meehl G, Nozawa T, Pierce D, Stott P, Washington W, Wehner M. 2007. Identification of human-induced changes in atmospheric moisture content. *Proceedings of the National Academy of Sciences* **104**(39): 15248-15253. DOI: 10.1073/pnas.0702872104.

Sherwood S, Meyer C, Allen R, Titchner H. 2008. Robust tropospheric warming revealed by iteratively homogenized radiosonde data. *Journal of Climate* **21**(20): 5336-5352. DOI: 10.1175/2008jcli2320.1.

Smit H, Kivi R, Vömel H, Paukkunen A. 2013. Thin film capacitive sensors, in *Monitoring Atmospheric Water Vapour: Ground-based Remote Sensing and in-situ Methods.* Kämpfer N (ed). Springer: New York, 11-38.

Stickler A, Brönnimann S. 2011. Significant bias of the NCEP/NCAR and twentieth-century reanalyses relative to pilot balloon observations over the West African monsoon region (1940-1957). *Quarterly Journal of the Royal Meteorological Society* **137**(659): 1400-1416. DOI: 10.1002/qj.854.

Trenberth K, Fasullo J, Smith L. 2005. Trends and variability in column-integrated atmospheric water vapor. *Climate Dynamics* **24**(7-8): 741-758. DOI: 10.1007/s00382-005-0017-4.

Trenberth K, Smith L, Qian T, Dai A, Fasullo J. 2007. Estimates of the global water budget and its annual cycle using observational and model data. *Journal of Hydrometeorology* **8**(4): 758-769. DOI: 10.1175/jhm600.1.

Trenberth K, Stepaniak D. 2003. Covariability of components of poleward atmospheric energy transports on seasonal and interannual timescales. *Journal of Climate* **16**(22): 3691-3705. DOI: 10.1175/1520-0442(2003)016<3691:cocopa>2.0.co;2.

Ummenhofer CC, Sen Gupta A, Briggs PR, England MH, McIntosh PC, Meyers GA, Pook MJ, Raupach MR, Risbey JS. 2011. Indian and Pacific Ocean influences on Southeast Australian drought and soil moisture. *Journal of Climate* **24**(5): 1313-1336.

Ummenhofer CC, Sen Gupta A, England MH. 2009. Causes of late twentieth-century trends in New Zealand precipitation. *Journal of Climate* **22**(1): 3-19.

Uppala S, Kållberg P, Simmons A, Andrae U, Bechtold V, Fiorino M, Gibson J, Haseler J, Hernandez A, Kelly G, Li X, Onogi K, Saarinen S, Sokka N, Allan R, Andersson E, Arpe K, Balmaseda M, Beljaars A, Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins B, Isaksen L, Janssen P, Jenne R, Mcnally A, Mahfouf J, Morcrette J, Rayner N, Saunders R, Simon P, Sterl A, Trenberth K, Untch A, Vasiljevic D, Viterbo P, Woollen J. 2005. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society* **131**(612): 2961-3012. DOI: 10.1256/qj.04.176.

Wade C. 1994. An evaluation of problems affecting the measurement of low relative humidity on the United States radiosonde. *Journal of Atmospheric and Oceanic Technology* **11**(3): 687-700. DOI: 10.1175/1520-0426(1994)011<0687:aeopat>2.0.co;2.

Wang J, Yan ZW, Jones PD, Xia JJ. 2013. On “observation minus reanalysis” method: A view from multidecadal variability. *Journal of Geophysical Research-Atmospheres* **118**(14): 7450-7458.

WCDMP. 2003. Guidelines on climate metadata and homogenization. WCDMP-53, WMO/TD-1186. Llansó P (ed). Available at http://www.wmo.int/pages/prog/wcp/wcdmp/wcdmp_series/documents/WCDMP-53.pdf.

Willett K, Jones P, Thorne P, Gillett N. 2010. A comparison of large scale changes in surface humidity over land in observations and CMIP3 general circulation models. *Environmental Research Letters* **5**(2). DOI: 10.1088/1748-9326/5/2/025210.

Yeh P, Famiglietti J. 2008. Regional terrestrial water storage change and evapotranspiration from terrestrial and atmospheric water balance computations. *Journal of Geophysical Research-Atmospheres* **113**(D9). DOI: 10.1029/2007jd009045.

Zhao T, Dai A, Wang J. 2012. Trends in tropospheric humidity from 1970 to 2008 over China from a homogenized radiosonde dataset. *Journal of Climate* **25**(13): 4549-4567. DOI: 10.1175/jcli-d-11-00557.1.

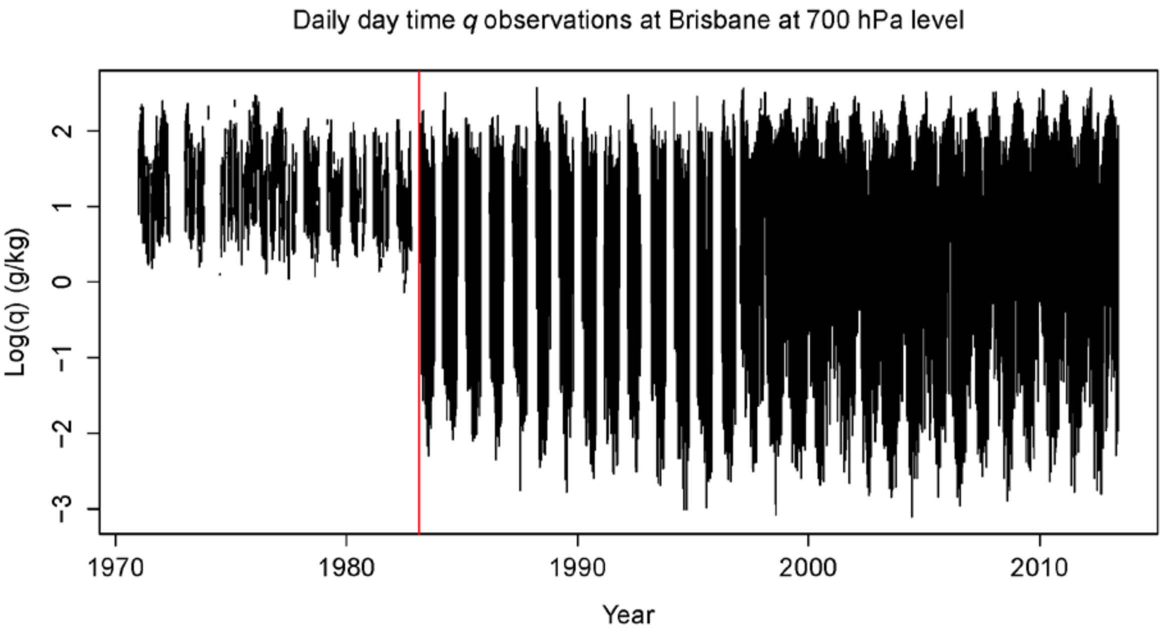

Figure 1 Time series of log-transformed daytime specific humidity at Brisbane at the 700 hPa level. Date of instrumental change as documented in the metadata is marked by the solid red line.

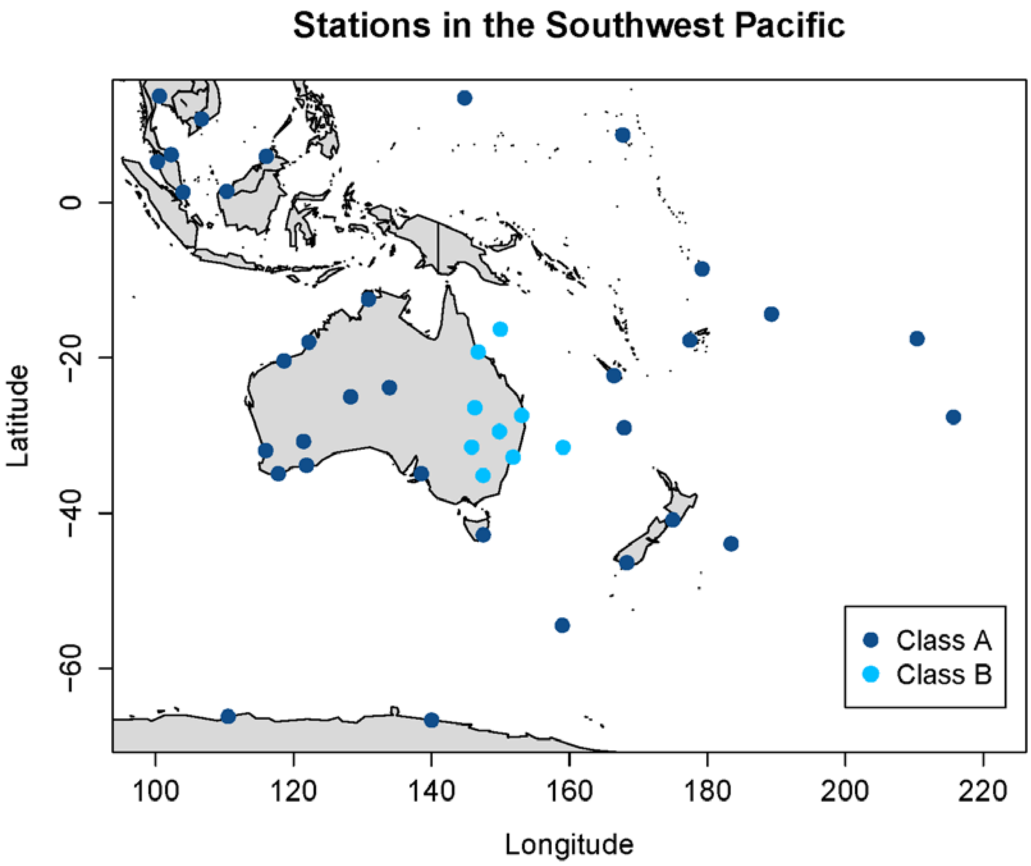

Figure 2 Stations in the Southwest Pacific. Stations that have more complete data are plotted as dark blue circles as opposed to those with light blue circles which have large gaps and/or shorter records.

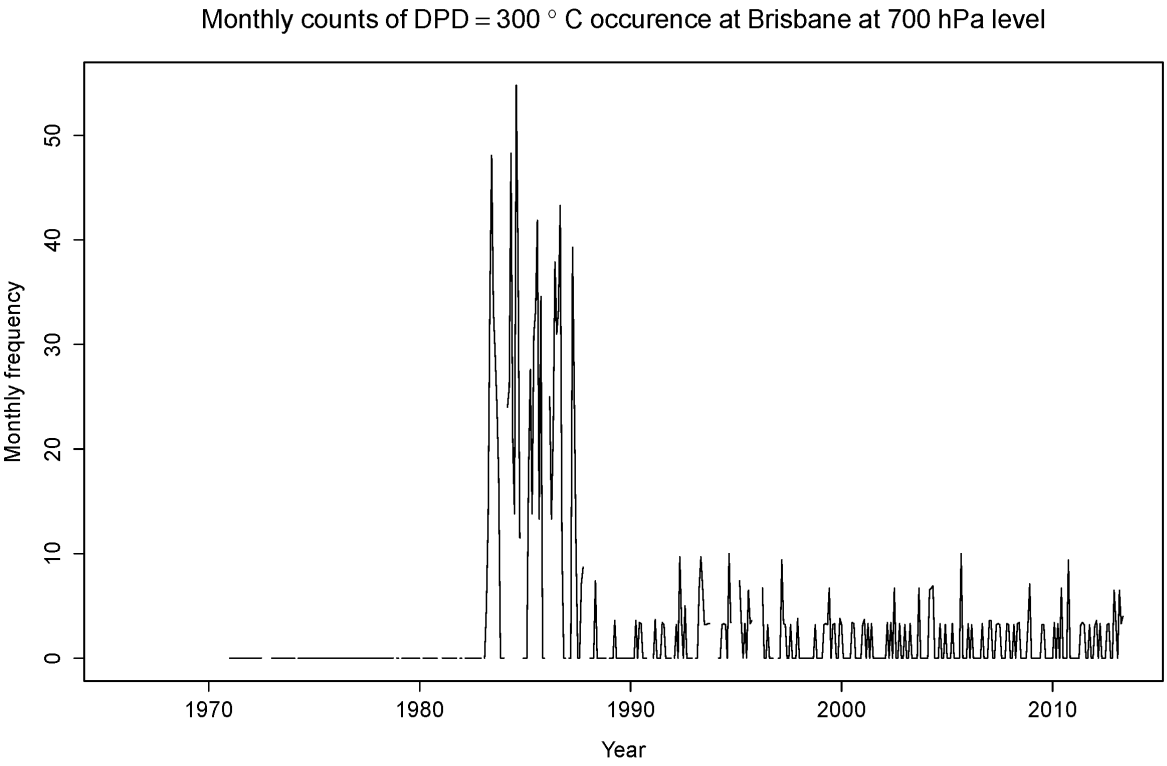

Figure 3 Monthly frequency of the occurrence of DPD=30° C at Brisbane at the 700 hPa level.

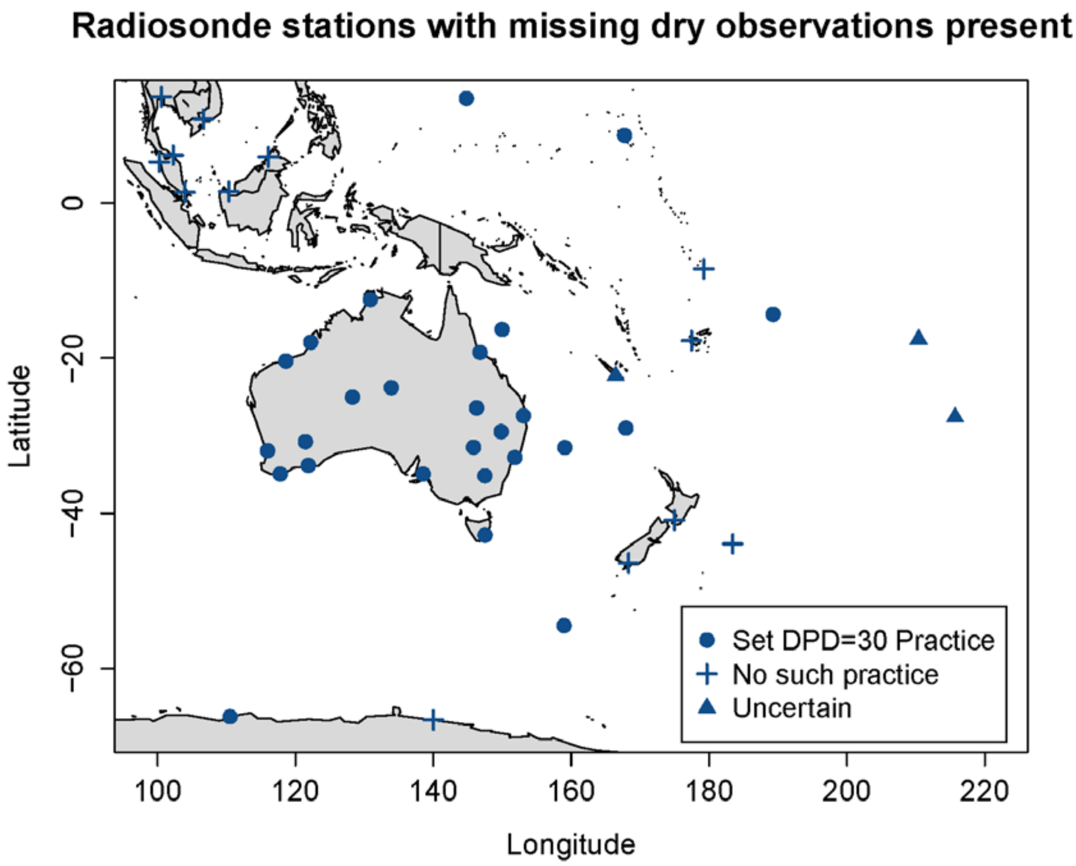

Figure 4 Map showing stations that had the practice (filled circles), did not have the practice (cross) of recording high DPD (or low RH) as 30° C. There are three stations (filled triangles) whereby the adoption of the practice cannot be determined under currently available information.

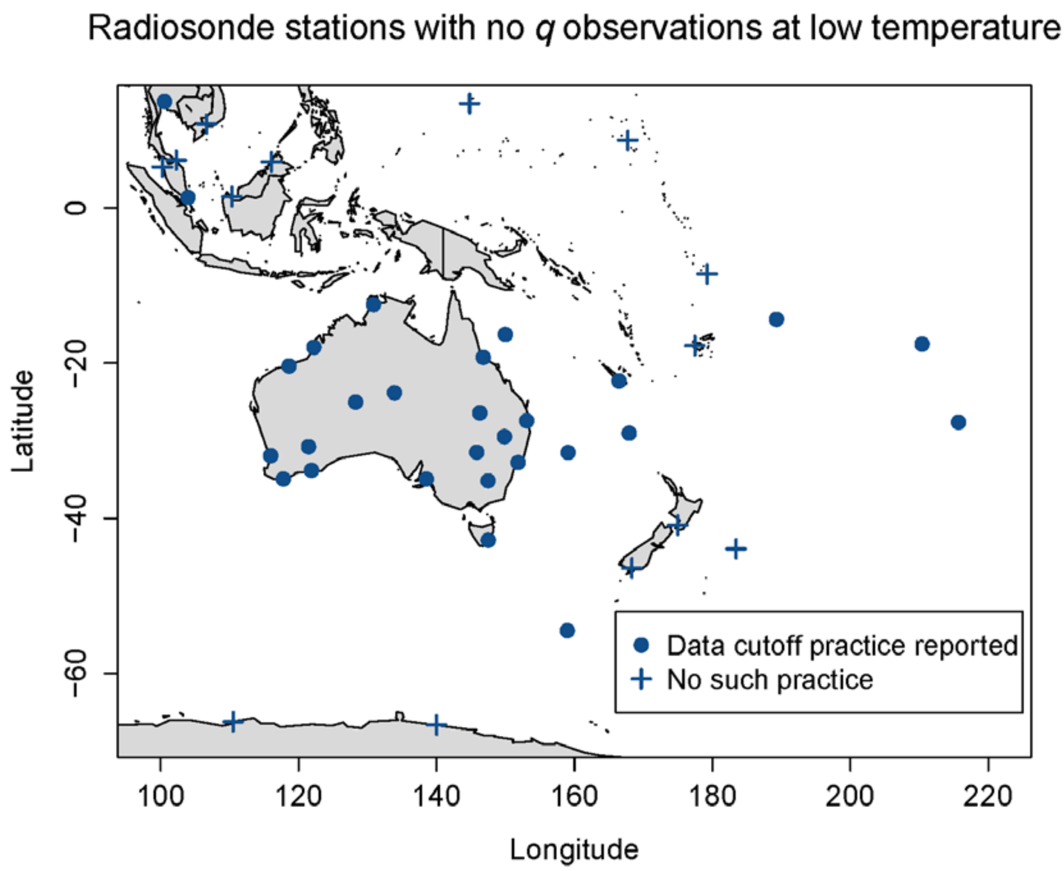

Figure 5 Map showing stations that adopted (filled circle) and did not adopt (cross) the data cutoff practice as documented in station metadata.

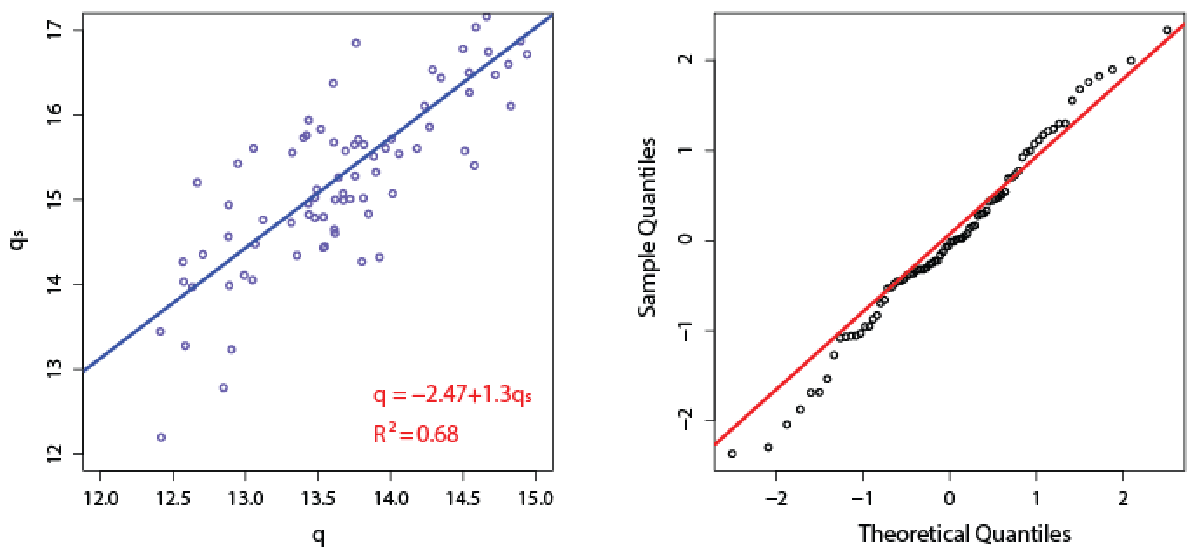

Figure 6 Scatterplot showing the relationship between q_s and q (left), and normal Q-Q plot of theoretical versus sample quantiles (right) for station 91765 at the 850 hPa level (lat: 14.33° S; lon: 170.72° W).

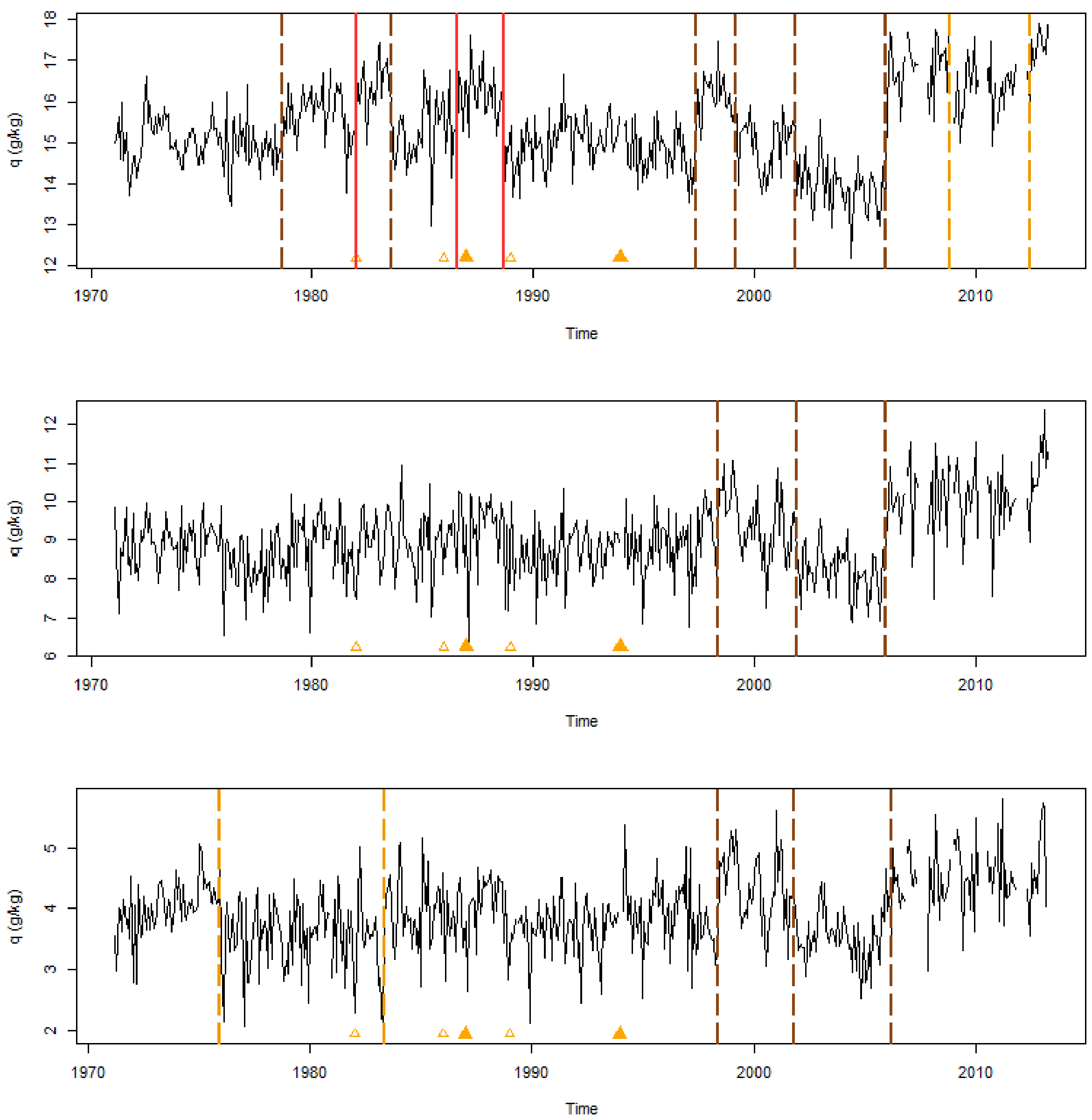

Figure 7 Daytime monthly q at three levels (top: 850 hPa; middle: 700 hPa; bottom: 500 hPa) and positions of changepoints for station Butterworth (lat: 5.47° N; lon: 100.38° E). Filled and hollow triangles denote dates of documented changepoints with high and low certainty, respectively. Solid red lines indicate that the candidate changepoints match those documented by metadata. Brown dashed lines indicate the candidate changepoints are agreed by all three information criteria, and the golden dashed lines indicate agreement by two information criteria only.

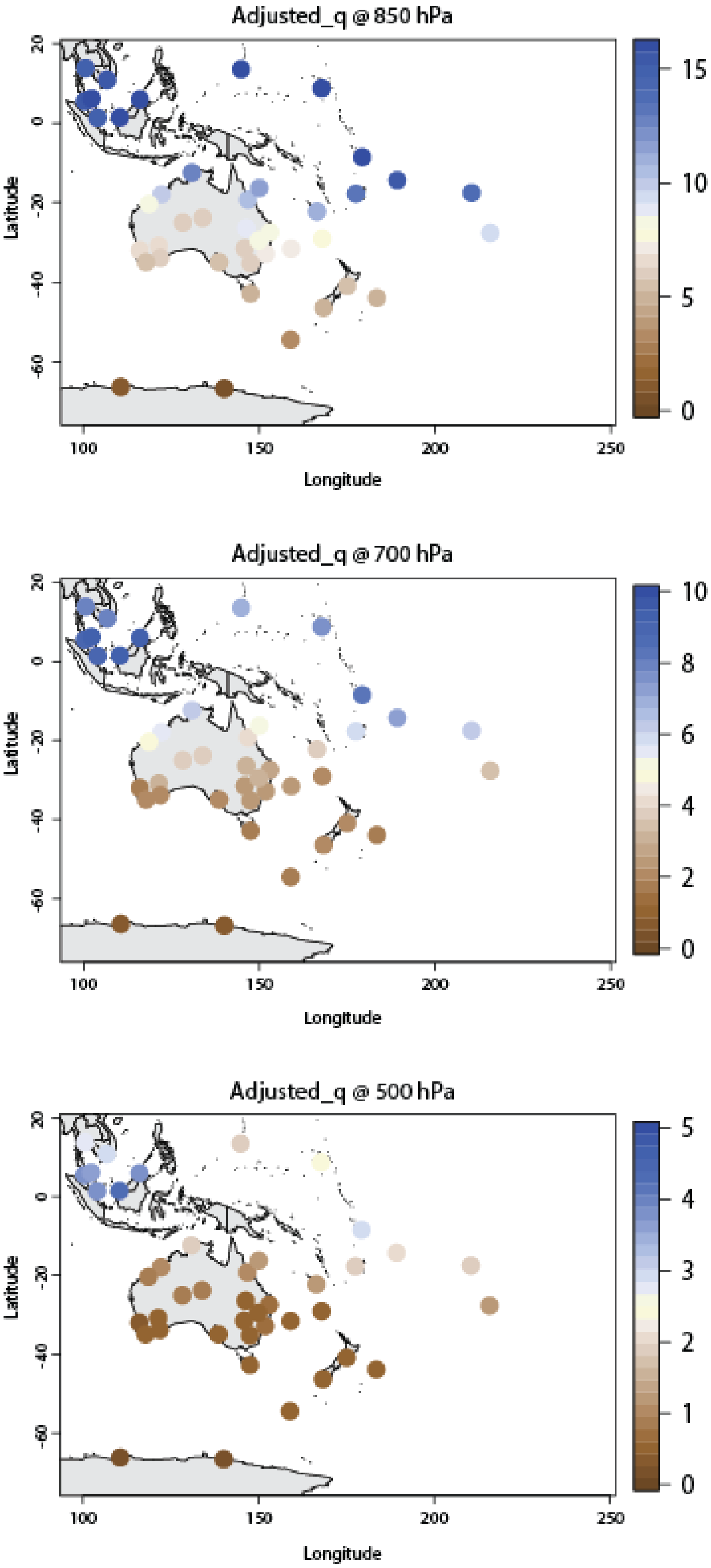

Figure 8 Climatology of q_{adj} (g/kg) from radiosondes at three pressure levels.

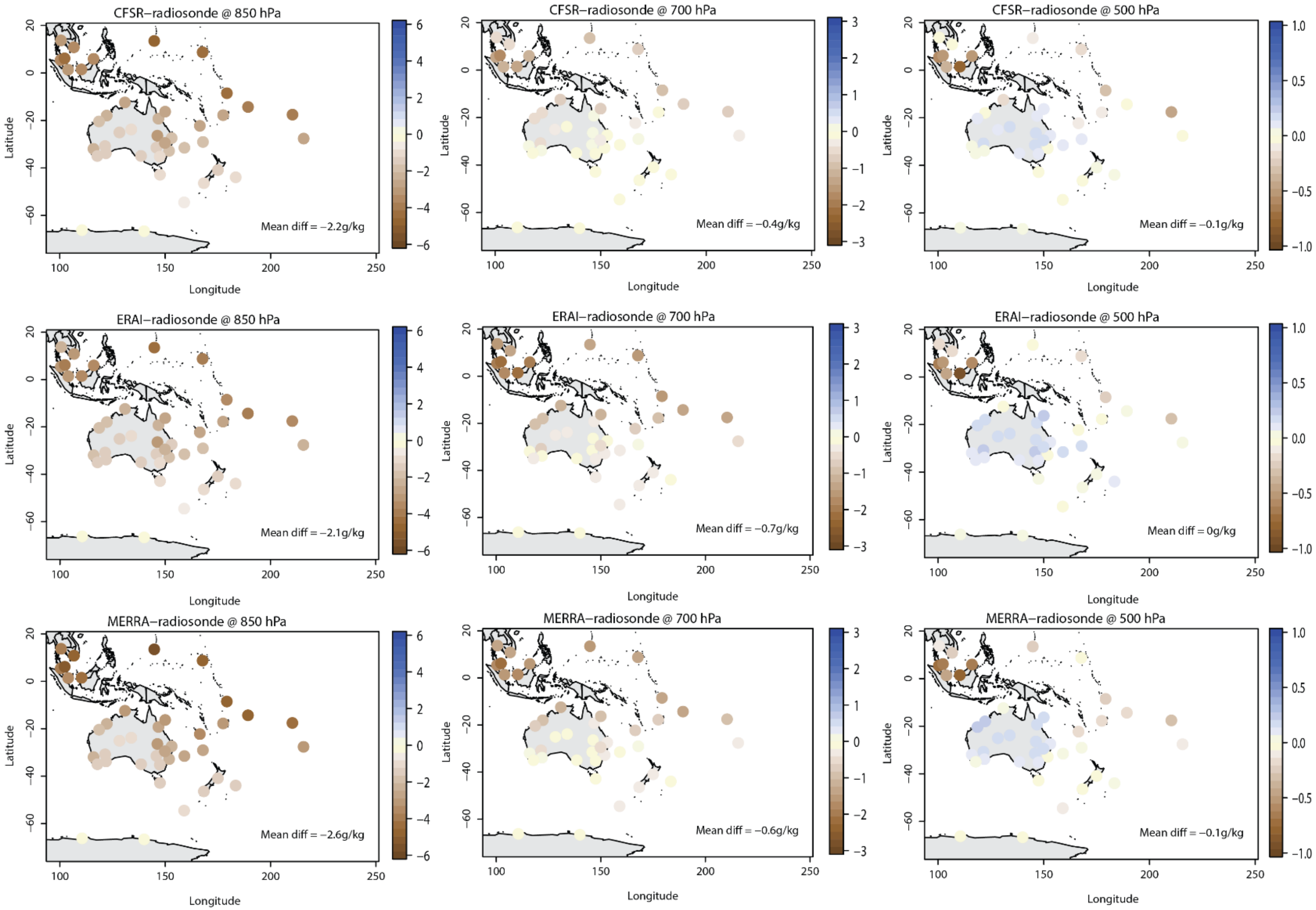

Figure 9 Mean climatological differences between CFSR (top), ERAI (middle), MERRA (bottom) and q_{adj}.

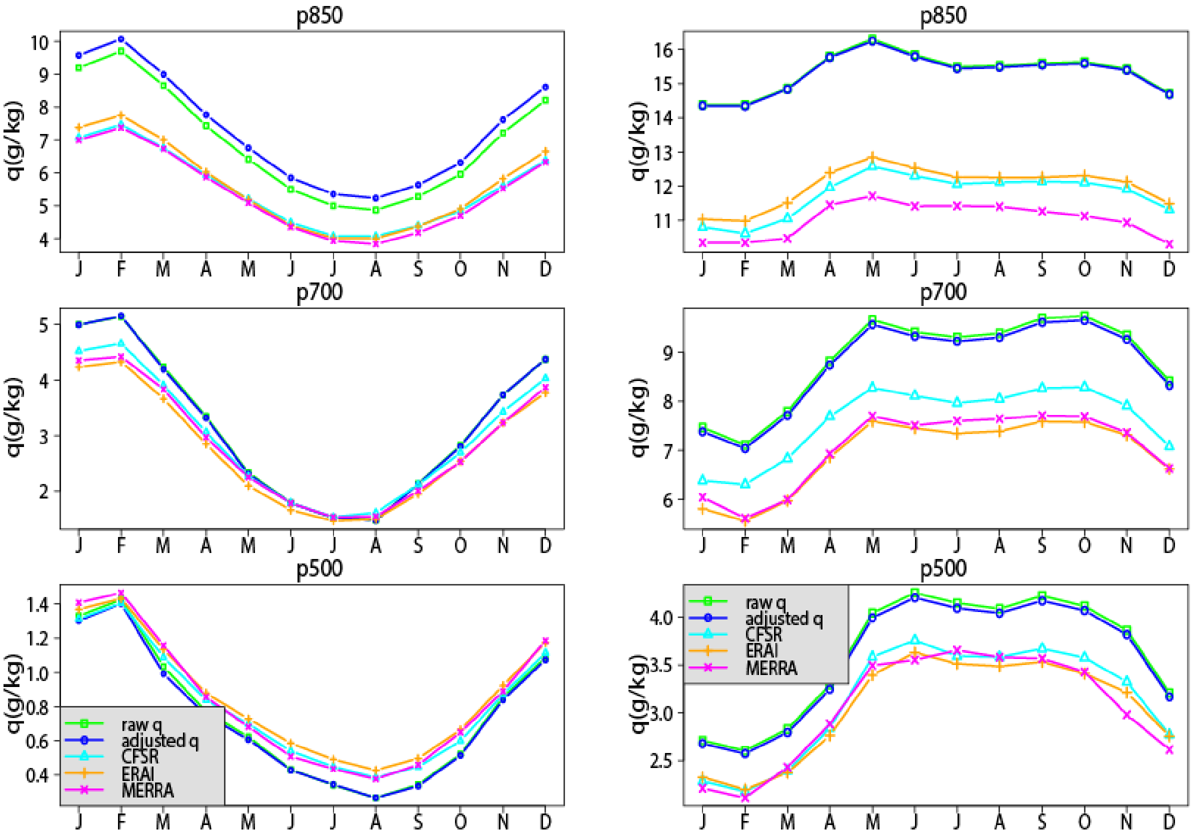

Figure 10 Seasonal cycle of *q* over Australia (left) and Southeast Asia (right).

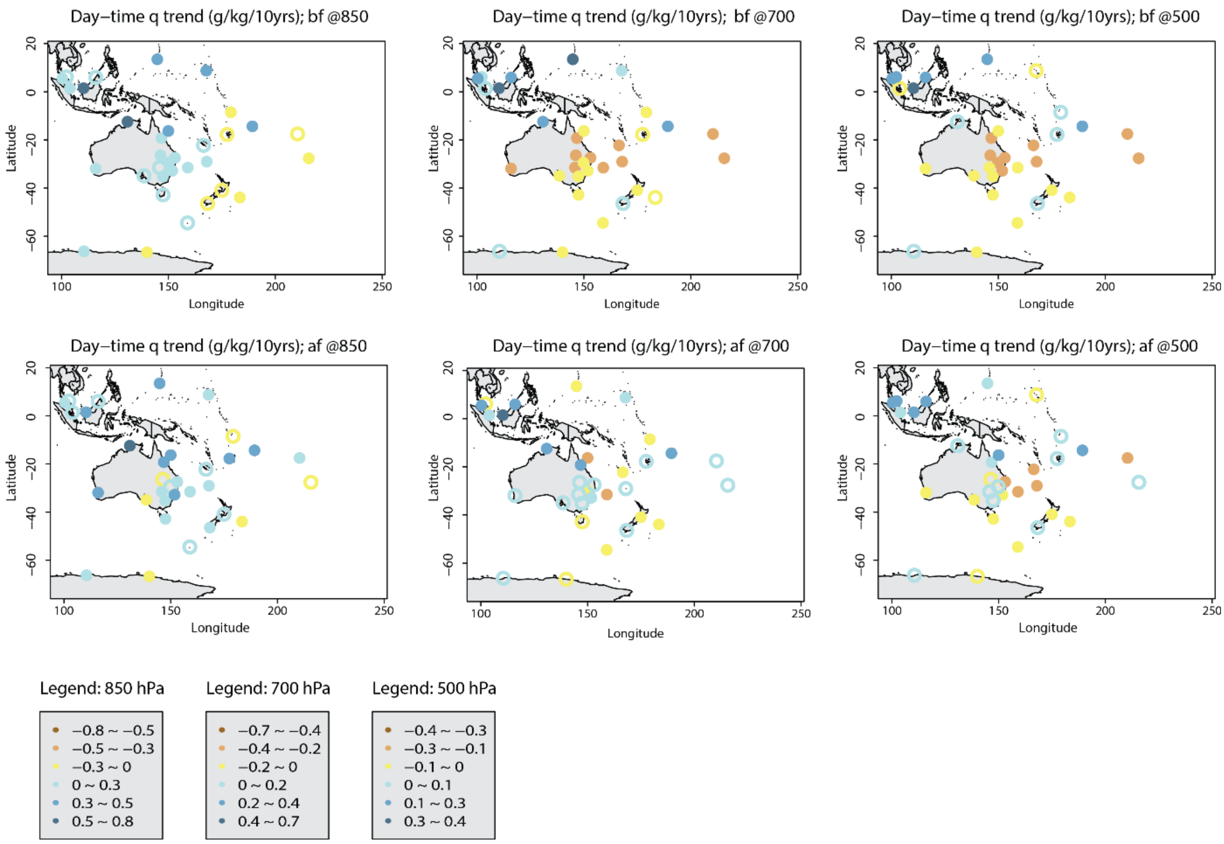

Figure 11 Long-term linear trends in monthly daytime *q* (g/kg/10-yr) for the before (top panel) and after (bottom panel) adjustment data. Filled circles indicate that the trends are statistically significant at the 5% level. Open circles indicate the trends are not significant.

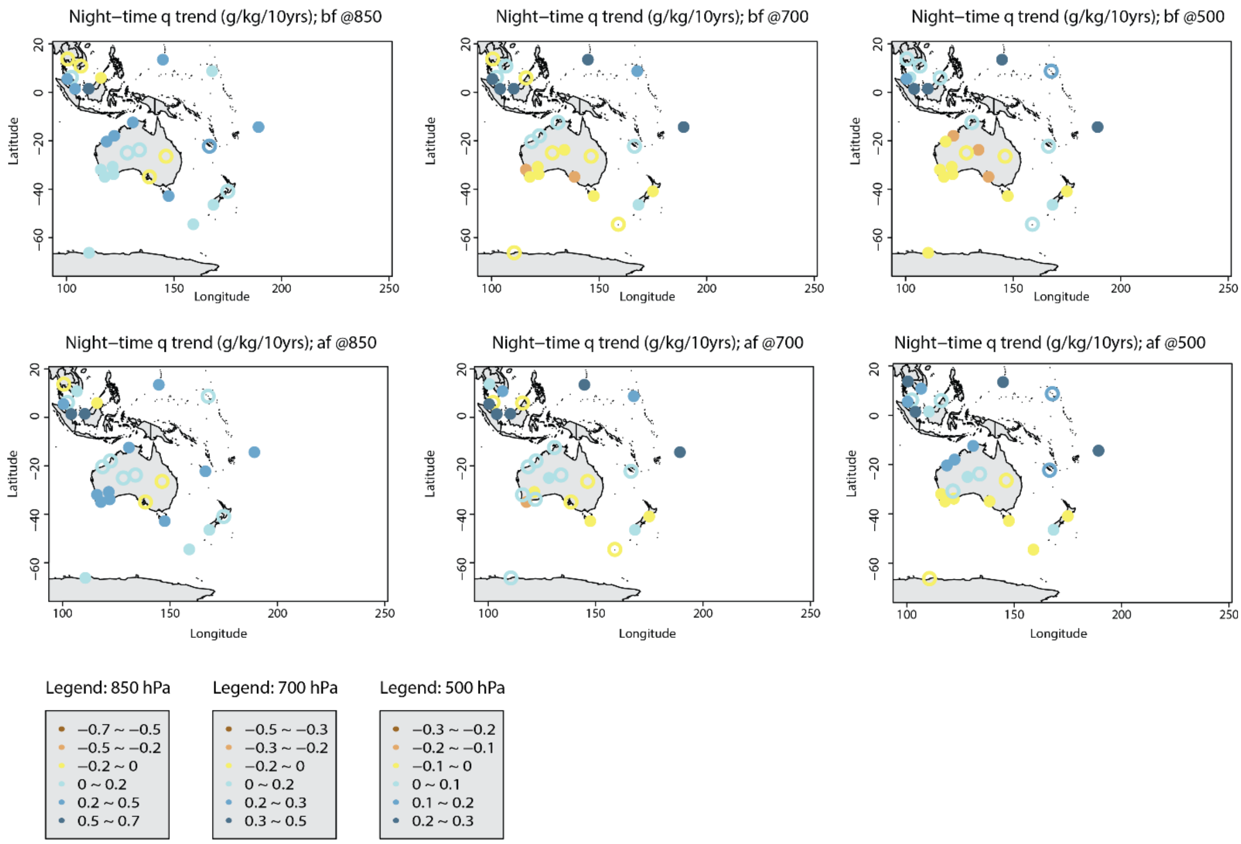

Figure 12 Same as Figure 11 except for nighttime data.

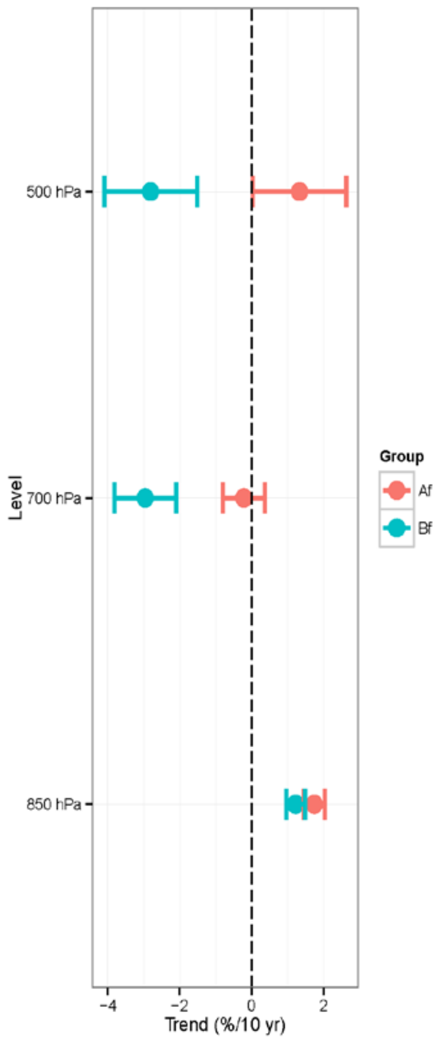

Figure 13 Trend estimates for the radiosonde data in percentage change per decade. The bars indicate the means (dots in the centre) and the standard errors (ends of the bars) of the trend estimate. Intervals coloured in blue and red are for the before and after correction estimates, respectively.

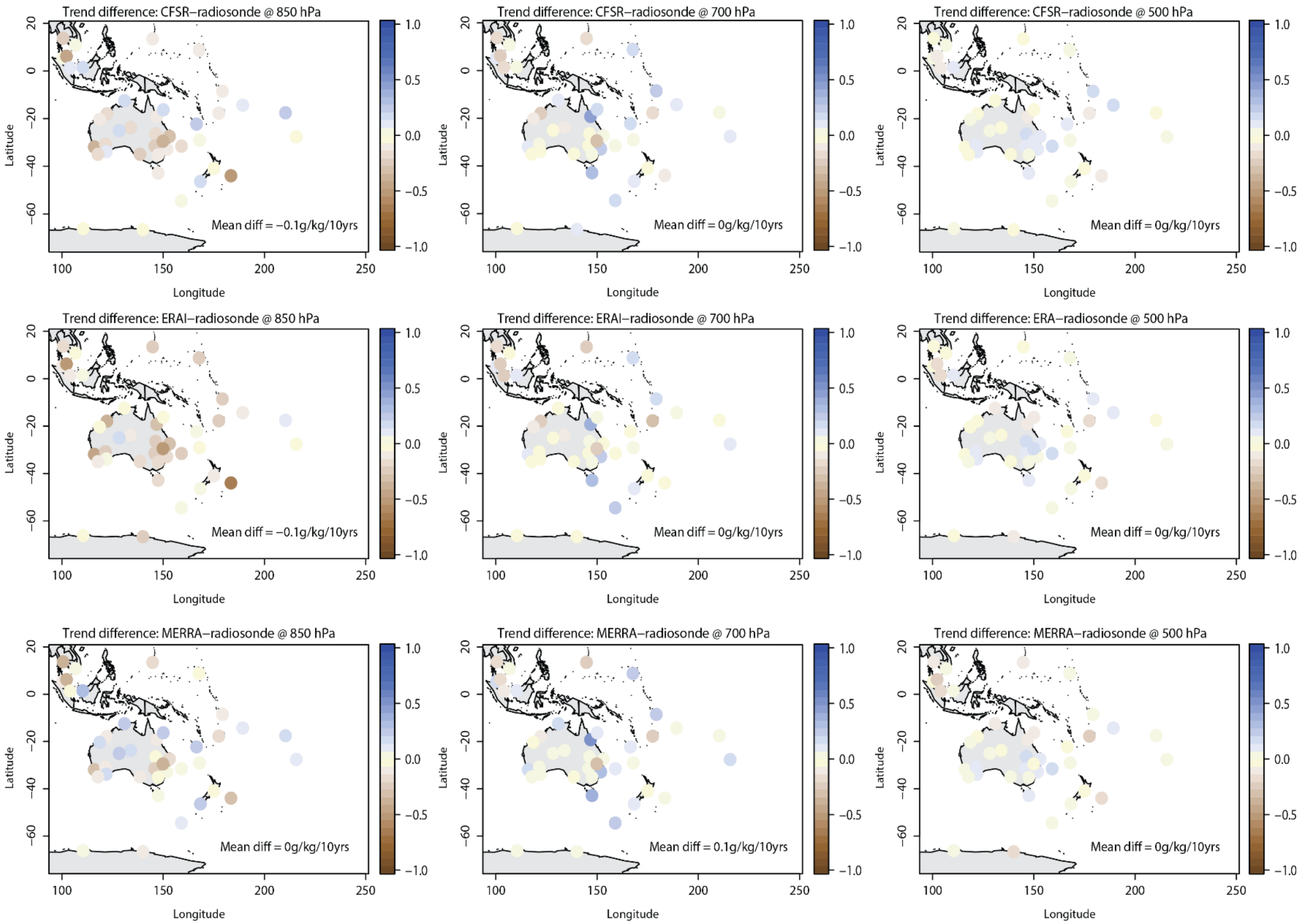

Figure 14 Differences in decadal trend estimates between CFSR (top), ERAI (middle), MERRA (bottom) and q_{adj}.