

Radiosonde comparison of ERA5 and ERA-Interim reanalysis datasets over tropical oceans

By MERI VIRMAN *, MARJA BISTER, JOUNI RÄISÄNEN, VICTORIA A. SINCLAIR, and HEIKKI JÄRVINEN, *Institute for Atmospheric and Earth System Research (INAR), University of Helsinki, Helsinki, Finland*

(Manuscript Received 10 September 2020; in final form 6 May 2021)

ABSTRACT

After the release of the ERA-Interim reanalysis, many changes have been made to the Integrated Forecasting System model and data-assimilation system, resulting in an improved reanalysis, ERA5. One of the changes in the model allows the model version in ERA5 to represent the moisture sensitivity of deep convection more realistically than the model version in ERA-Interim. A previous modeling study showed that this change alone improved the representation of the tropical atmosphere, e.g. tropical variability and precipitation distribution. Here we compare the vertical structures of average temperature and moisture over tropical oceans in ERA5, ERA-Interim and radiosonde observations to see whether ERA5 is also closer to observations for those regions and variables. Our results reveal that at many levels, temperature and relative humidity in ERA5 and ERA-Interim differ from observations, however ERA-Interim is generally closer to observations than ERA5 in the low-to-midtroposphere. Most notably, in many stations, ERA5 is on average colder than observations at ~550–800 hPa. Large vertical gradients occur in the profile of the mean temperature difference between ERA5 and observations at ~700–900 hPa, but are absent in ERA-Interim. Relative humidity differences are not as robust as temperature differences, however in many stations ERA5 is on average moister than observations at ~650–800 hPa while ERA-Interim is closer to observations there. Below the ~950 hPa-level ERA5 and ERA-Interim are generally colder and moister than observations. Our results indicate that ERA5 deviates more than ERA-Interim from tropical radiosonde observations in the low-to-midtroposphere. It seems plausible that this deviation is, at least partly, due to the newer formulation of organised deep entrainment in ERA5 and the associated mechanism for the moisture sensitivity. However, more extensive model evaluation is needed to understand the reasons for the differences between the reanalyses and radiosonde observations.

Keywords: radiosonde observations, ERA5, ERA-Interim, temperature, moisture

1. Introduction

Due to their outstanding global coverage, climate reanalyses are used extensively to understand atmospheric processes, yet they also depend on the underlying numerical model and the choices made in the representations of subgrid-scale physical processes. Recently, a new climate reanalysis dataset, ERA5 (Hersbach et al. 2020), was released. ERA5 follows the ERA-Interim (ERA-Interim, Dee et al. 2011) reanalysis dataset generated with a previous version of the same state-of-the-art numerical model (see Sect. 2 for the model versions), the Integrated Forecast System (IFS), maintained by the European Centre for Medium-Range Weather Forecasts (ECMWF).

Several improvements have been made in the physical parameterisations of subgrid-scale processes and data-assimilation system in the IFS since the release of ERA-Interim (see Hersbach et al. 2020, and references therein). One major improvement is in the formulation of organised deep entrainment in the convection parameterisation scheme (Bechtold et al. 2008; Hersbach et al. 2020). Namely, an earlier moisture-convergence-based entrainment formulation was replaced with an environmental relative humidity (RH)-dependent entrainment. Hirons et al. (2013a,b) studied the effect of the new entrainment, by comparing IFS experiments that differed only in their formulation of the entrainment, and showed that it led to a more realistic moisture-precipitation relationship, precipitation distribution in the equatorial belt and tropical

*Corresponding author. e-mail: meri.virman@helsinki.fi

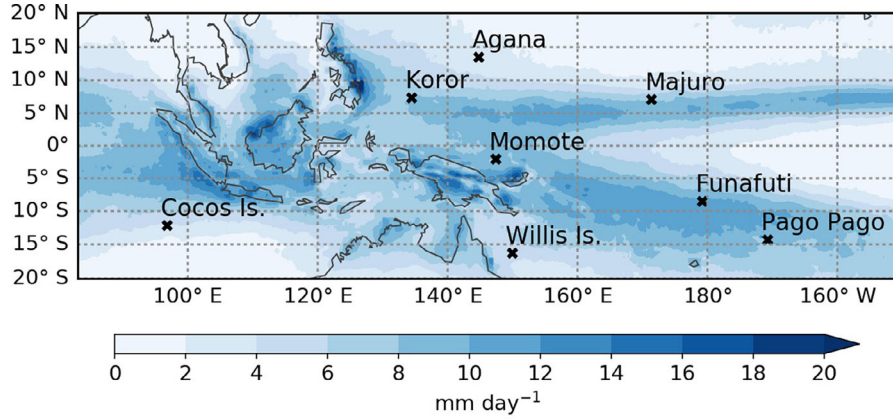


Fig. 1. Locations of the IGRA sounding stations. The coloured contours show the mean precipitation (in mm day^{-1}) during November-February 1998-2014, calculated using the Tropical Rainfall Measuring Mission (TRMM) 3B42 data.

variability. Because the ERA5 reanalysis was generated using a model version containing the new entrainment formulation, but ERAI was not, and because the new entrainment led to improvements particularly in the Tropics, it would seem likely that ERA5 also represents many properties of the tropical troposphere more realistically than ERAI. At least, it has been shown that the distribution of monthly-mean rainfall over most tropical oceans is better represented in ERA5 than in ERAI (Nogueira 2020).

In this study, we analyse whether the vertical profiles of average temperature and RH over tropical oceans are also in better agreement with observations in ERA5 than in ERAI. We conduct a comparison of the reanalyses and radiosonde observations towards this goal and speculate about the reasons of the differences but note that access to model versions would be needed to properly evaluate and understand the underlying reasons – such testing is not included here. The details of each dataset and analysis methods are described in Sect. 2 and the results are presented in Sect. 3. We discuss the results in Sect. 4. Finally, conclusions are provided in Sect. 5.

2. Materials and methods

In this study, we analyse reanalysis data and radiosonde observations from eight stations located over the western Pacific Ocean and eastern Indian Ocean (Fig. 1). We selected these stations as they have reasonably long records and are all located on small islands in tropical oceans. The radiosonde observations were downloaded from the Integrated Global Radiosonde Archive (IGRA) version 2 (Durre and Yin 2008) archive which reports quality controlled pressure, altitude, temperature, dew point depression and horizontal wind in thousands of sounding stations globally.

Atmospheric reanalyses combine observations and numerical models to produce a “best estimate” of the state of the atmosphere. The ERAI and ERA5 reanalyses were developed by ECMWF and generated using different versions of the IFS model, Cy31r2 and Cy41r2, that were used for operational forecasting in 2006-2007 and 2016, respectively. The main differences in ERA5 compared to ERAI are an increased horizontal, temporal and vertical resolution (0.25 degrees, 1 h and 137 levels in ERA5 and 0.75 degrees, 6 h and 60 levels in ERAI), newer data-assimilation system as well as several improvements in the physical parameterisation schemes (see e.g. Hersbach et al. 2020, for more details). Moreover, it should be noted that the IGRA radiosonde observations were assimilated into both ERA5 and ERAI.

We included only soundings that (1) reached the 500 hPa level, (2) had vertical resolution of at least 200 hPa and (3) included at least 5 observations. In one station, Funafuti, soundings with clearly erroneous dew point temperatures near the surface were excluded. These quality control criteria closely follow those used in Virman et al. (2018), which also analysed soundings from the same stations.

We selected only soundings made between 22-01 UTC (most often between 23-00 UTC). ERA5 and ERAI are valid at 00 UTC and we selected the grid points that were closest to the sounding stations. The time period was from November to February between 1998 and 2014 (i.e. 4 months from each year). We selected only the months from November to February to ensure that precipitating convection occurs frequently (Fig. 1). We selected the same sounding stations, times and months as in Virman et al. (2018) to facilitate possible comparisons. As there were time gaps in the sounding data and because all soundings did not meet the quality control criteria, the

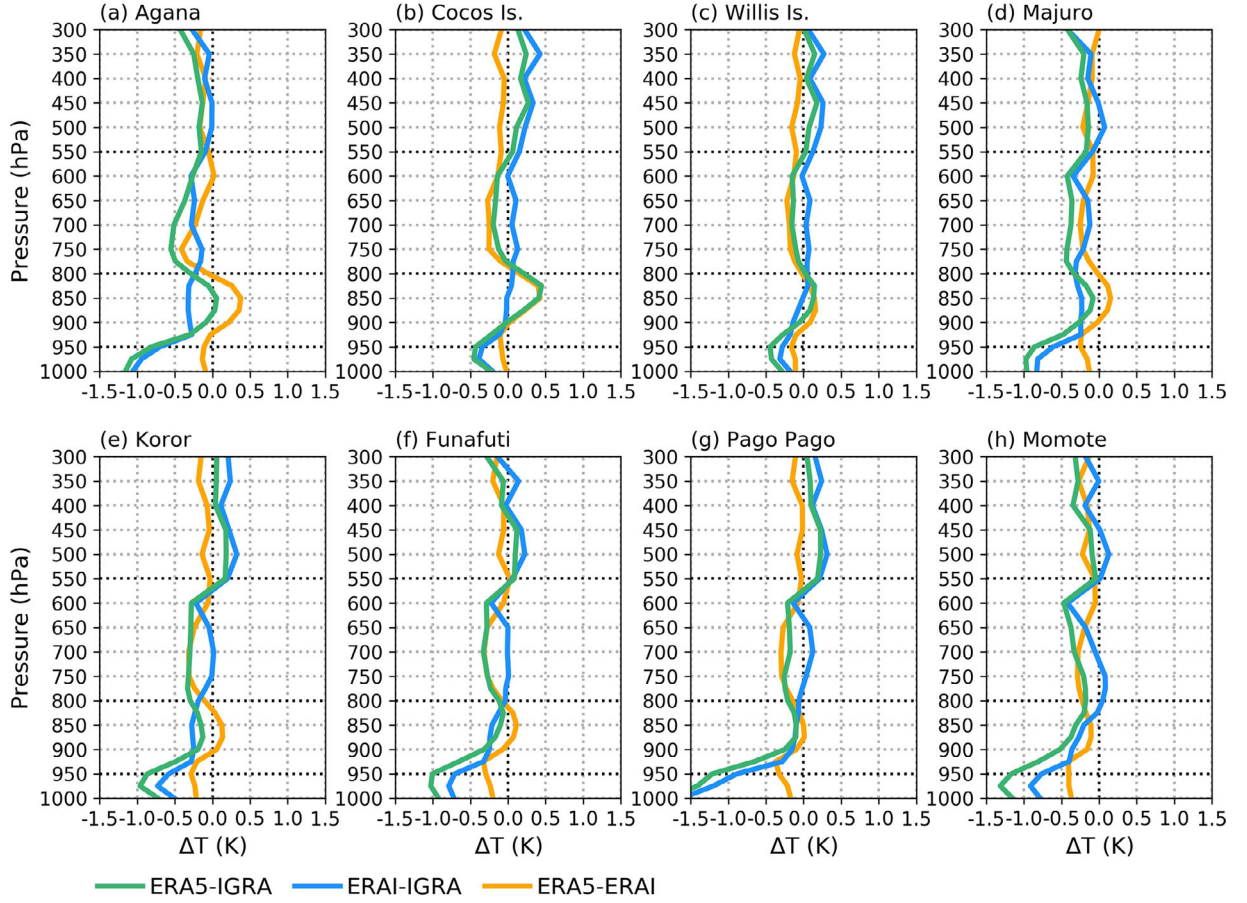


Fig. 2. Vertical profiles of the mean temperature difference (in units of Kelvin) between ERA5 and IGRA (green), ERAI and IGRA (blue) and ERA5 and ERAI (orange) in (a) Agana, (b) Cocos Island, (c) Willis Island, (d) Majuro, (e) Koror, (f) Funafuti, (g) Pago Pago and (h) Momote. The grid lines corresponding to some of the layers discussed in the text are highlighted in black. The stations are in order based on the climatological mean 500-700 hPa RH, from the driest (Agana) to the moistest (Momote) station, as in Virman et al. (2018).

total number of soundings in all stations was smaller than their theoretical maximum number. For consistency, we selected ERA5 and ERAI data only from days when a sounding exists. The total number of soundings and ERA5 and ERAI data was 1769 in Agana, 1716 in Cocos Island, 1458 in Willis Island, 1836 in Majuro, 1861 in Koror, 753 in Funafuti, 1838 in Pago Pago and 1066 in Momote. The pressure levels in IGRA varied between the soundings. Therefore, the soundings were interpolated to the same pressure levels as in ERA5 and ERAI, i.e. to every 50 hPa in the 300-700 hPa layer and to every 25 hPa in the 725-1000 hPa layer (the mean pressure difference between the interpolated value and the closest original level was ~ 5 -14 hPa, depending on the station). RH was calculated from the soundings, ERA5, and ERAI using equations from Emanuel (1994).

For each day, the temperature and RH difference between ERA5 and IGRA, ERAI and IGRA, and ERA5

and ERAI was calculated at every pressure level and station. Finally, the mean and standard deviation were calculated from the differences.

3. Results

We compare the vertical profiles of temperature and RH in IGRA to those in ERA5 and ERAI to see how close the reanalyses are to observations in months when all the stations experience at least some deep convection. The mean differences between the vertical temperature profiles of ERA5 and IGRA, ERAI and IGRA, as well as ERA5 and ERAI, are shown in Fig. 2 for each station.

At many levels and in most stations, ERA5 differs noticeably from IGRA (Fig. 2). Most notably, in the ~ 550 -800 hPa layer, ERA5 is on average ~ 0.2 -0.6 K colder than IGRA in all stations. In all stations except Momote, large vertical gradients at ~ 700 -900 hPa are

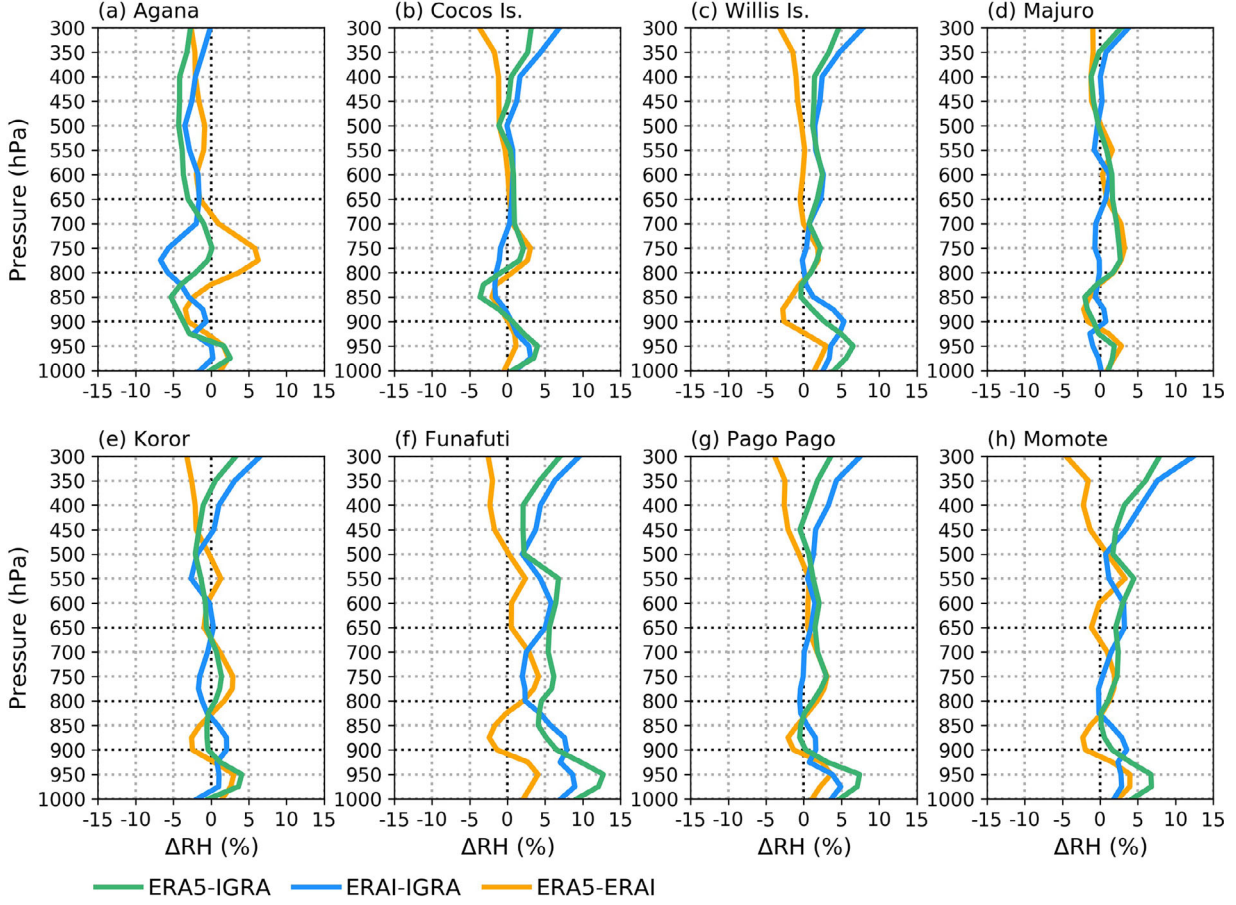


Fig. 3. As in Fig. 2, but for the mean relative humidity (RH) difference (in units of %). Please note that the grid lines differ from those in Fig. 2.

seen in the mean difference profile between ERA5 and IGRA, especially in Agana, Cocos, Willis and Majuro. Interestingly, the largest gradients are located over relatively dry regions of the tropical oceans whereas over relatively moist regions the gradients are smaller or absent (the stations in Figs. 2–4 are in order based on their climatological mean 500–700 hPa RH, from the driest, Agana, to the moistest station, Momote, see Virman et al. 2018 for more details). Moreover, below the ~ 950 hPa-level, ERA5 is ~ 0.2 – 1.5 K colder than IGRA but above the ~ 450 hPa-level, temperatures in ERA5 are relatively close to IGRA.

Figure 2 shows that at many levels below the ~ 500 hPa-level and in most stations, ERAI also differs from IGRA (Fig. 2), but generally not as much as ERA5. In most stations at ~ 550 – 800 hPa, the mean temperature difference between ERAI and IGRA is near-zero, whereas ERA5 is generally colder than IGRA in these stations (as discussed above). In three of the eight stations, ERAI is colder than IGRA in the ~ 550 – 800 hPa layer, however the mean temperature difference between

ERAI and IGRA in these stations is still generally smaller (~ 0 – 0.3 K) than between ERA5 and IGRA (~ 0.2 – 0.6 K). The large vertical gradients at ~ 700 – 900 hPa in the difference of ERA5 and IGRA are absent in the difference of ERAI and IGRA. Below the ~ 950 hPa level, ERAI, likewise to ERA5, is colder than IGRA. Above the ~ 450 hPa-level, the mean temperature difference between ERAI and IGRA is similar to that between ERA5 and IGRA, although temperatures in ERAI are slightly warmer. Moreover, in many stations below the ~ 500 hPa-level, the vertical structure of the mean temperature difference between ERA5 and ERAI (Fig. 2) resembles that between ERA5 and IGRA.

The mean RH differences (Fig. 3) between ERA5, ERAI and IGRA are not as robust as is the case for temperature differences (Fig. 2), but some systematic differences do occur. In all stations except Agana, ERA5 is generally moister than IGRA in the ~ 650 – 800 hPa layer, by ~ 2 – 6% . The mean RH difference between ERAI and IGRA at ~ 650 – 800 hPa is smaller than between ERA5 and IGRA in most stations (0 – 3%). In most stations, the

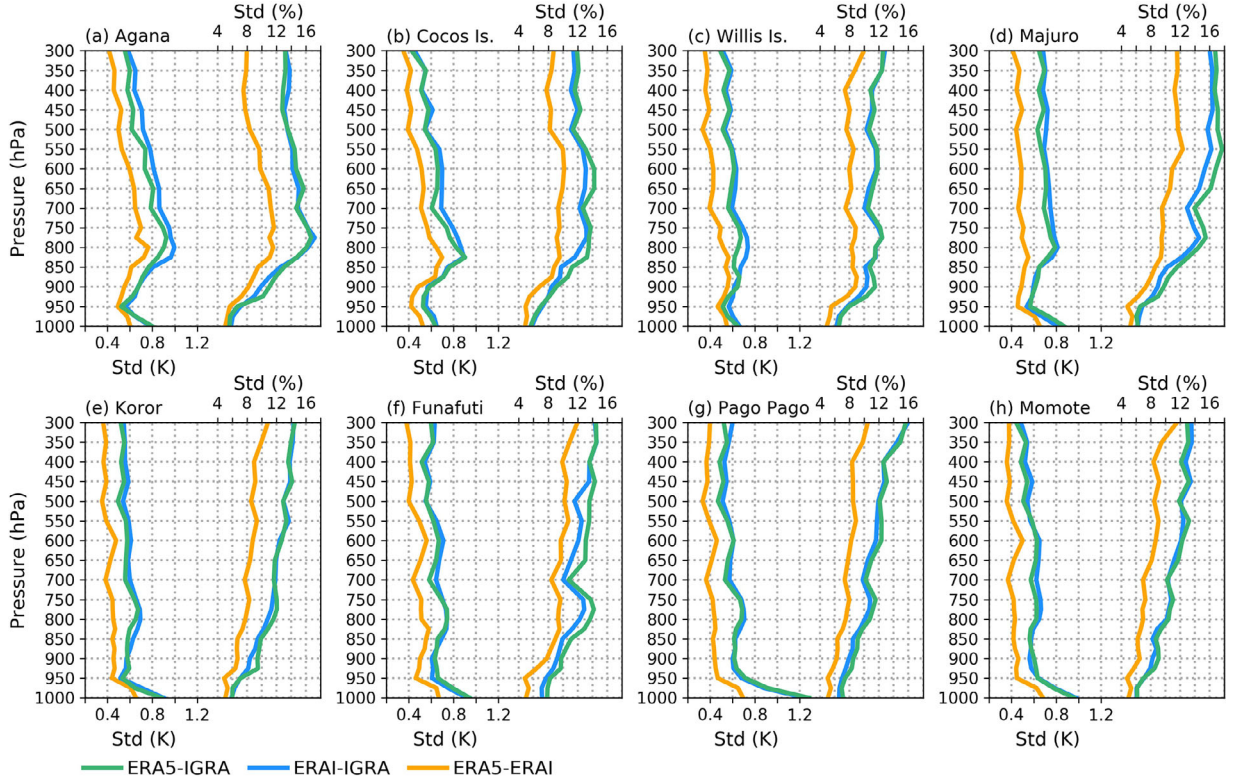


Fig. 4. Vertical profiles of standard deviation of the temperature (left-hand side curves) and RH (right-hand side curves) difference (in units of Kelvin and %) between ERA5 and IGRA (green), ERAI and IGRA (blue) and ERA5 and ERAI (orange) in (a) Agana, (b) Cocos Island, (c) Willis Island, (d) Majuro, (e) Koror, (f) Funafuti, (g) Pago Pago and (h) Momote. The standard deviation was calculated so that, first, the differences at each level were calculated for each day and, second, a standard deviation was calculated from the individual differences.

average RH in both ERA5 and ERAI is larger than in IGRA below the ~ 900 hPa-level, but the difference between ERA5 and IGRA is slightly larger than between ERAI and IGRA. Above the ~ 650 hPa-level, ERA5 and ERAI are moister than IGRA in some stations, however, in most stations RH in ERAI is even larger than that in ERA5. Moreover, below the ~ 500 hPa-level the vertical structure of the mean RH difference between ERA5 and ERAI somewhat resembles that between ERA5 and IGRA in Cocos, Majuro, Koror and Pago Pago.

To examine the variability of the temperature and RH differences, standard deviation of the differences between individual vertical profiles of temperature (left-hand side contours) and RH (right-hand side contours) was also investigated (Fig. 4). Figure 4 shows that large variability occurs in the temperature and RH difference in ERA5 minus IGRA, ERAI minus IGRA, as well as ERA5 minus ERAI. The standard deviations (Fig. 4) of temperature and RH differences are generally of similar magnitude to, or larger than, the mean differences (Figs. 2 and 3).

The standard deviations in Fig. 4 also show that there is somewhat less case-to-case variability in the ERA5 minus ERAI differences than in both ERAI minus IGRA and ERA5 minus IGRA. Furthermore, the vertical distribution of the standard deviations is generally different from the mean differences in Figs. 2-3, suggesting that the case-to-case variability is not dominated by the same dynamics that determine the mean differences in the temperature and RH profiles.

4. Discussion

Our results revealed that the vertical profiles of temperature and RH in ERA5 and ERAI generally differ from those in radiosonde observations (Figs. 2 and 3), however in the low-to-midtroposphere (at ~ 550 - 800 hPa for temperature and ~ 650 - 800 hPa for RH) ERAI was on average closer to IGRA than ERA5 was. ERA5 was up to 0.5 K colder than IGRA at ~ 550 - 800 hPa in most stations, whereas ERAI was closer to IGRA there. Moreover, large vertical gradients occurred in the mean

temperature difference between ERA5 and IGRA in the ~ 700 - 900 hPa layer, especially in relatively dry regions (see Sect. 3 and caption of Fig. 2 for more details), but were absent between ERAI and IGRA. Both ERA5 and ERAI were on average colder than IGRA near the surface below the ~ 950 hPa-level.

RH also differed between ERA5, ERAI and IGRA in some levels and stations, however the differences were not as robust as for temperature. Nevertheless, ERA5 was generally moister at ~ 650 - 800 hPa (roughly the same layer where it was too cold) than IGRA in most stations, whereas ERAI was closer to IGRA there. Below the ~ 900 hPa-level, RH in both ERA5 and ERAI was larger than in IGRA.

There are numerous differences in the IFS model versions used in ERA5 and ERAI, including the data assimilation system and use of remote sensing measurements, in the horizontal and vertical resolutions as well as in the parameterisations of subgrid-scale processes (Hersbach et al. 2020). Moreover, despite the quality control applied to IGRA, errors and biases may still be present in the sounding data. Therefore, we cannot point out any specific reason for the differences seen between ERA5, ERAI and IGRA in Figs. 2 and 3. However one can speculate based on existing literature and, below, we discuss some potential reasons for the differences.

The coarse horizontal resolution in ERA5 and ERAI likely partly explains why both are too cold and moist in the boundary layer below the ~ 950 hPa-level compared to IGRA. IGRA is likely warmer and drier in the boundary layer than reanalysis because the radiosonde is launched over land in a small island, whereas the coarser resolution numerical model used in the reanalysis may represent the corresponding grid point as more oceanic.

In most stations, a cold and moist anomaly occurs at ~ 600 - 800 hPa, and a warm and dry anomaly at ~ 800 - 900 hPa, in ERA5 compared to ERAI (Figs. 2 and 3). The differences between ERA5 and ERAI are qualitatively similar to the temperature and specific humidity changes caused by replacing a moisture-convergence-based organised deep entrainment formulation (also used in ERAI) with a RH-dependent entrainment (used in ERA5, see Sect. 1 for details). Namely, the newer entrainment formulation led to zonal-mean cooling and moistening at ~ 500 - 800 hPa and warming and mostly drying at ~ 800 - 900 hPa in the Tropics at 5-day forecast lead time (see Figs. 1c and 2c in Hirons et al. 2013b). It seems that the implementation of the RH-dependent entrainment formulation may, at least partly, contribute to the differences we see between ERA5 and ERAI.

Hirons et al. (2013b) suggested that the cooling and moistening seen in the low-to-midtroposphere resulted from decreased occurrence of deep convective clouds and

increased occurrence of cumulus congestus clouds, respectively. More specifically, Hirons et al. conclude that in dry conditions, the newer entrainment formulation led to dilution of convective updrafts at midlevels and, therefore, less deep convection and more congestus clouds. The congestus clouds detrained at midlevels, which led to moistening of air there, and thus made the troposphere more favourable for the formation of deep convection after some time had passed. The associated detrainment also causes cooling due to evaporation of cloud water.

Unlike Hirons et al. (2013b), we also compared the vertical structures of temperature and moisture to observed ones. This revealed that in the low-to-midtroposphere in many stations, first, ERAI is generally closer to observations than ERA5 is and, second, the vertical structure of the differences between ERA5 and IGRA qualitatively resemble the effects of the RH-dependent entrainment. Therefore, it can be speculated whether changes in the entrainment formulation may have led to somewhat unrealistic vertical structures of the temperature and RH climatologies in the tropical regions studied here. These facts also raise the question of whether congestus clouds might be too prominent in ERA5. Hirons et al. (2013b) also concluded that the modified precipitation-moisture relationship is "by no means perfect". For example, they showed that although the new entrainment formulation in the IFS model led to significant improvements in the precipitation distribution and tropical variability, i.e. the Madden-Julian Oscillation, it also led to too large precipitation rates compared to satellite-borne precipitation estimates, in humid conditions (Hirons et al. 2013b).

It therefore seems that the entrainment formulation used in ERA5 contributing to deep convection's sensitivity to humidity above the boundary layer may not be entirely correct in the model. It is possible that another mechanism, associated with evaporation of stratiform precipitation and subsidence warming below (see Figs. 5 and 9 in Virman et al. 2018, 2020, respectively), could at least partly explain this sensitivity. However, to understand the true reasons for the temperature and RH differences we see between ERA5, ERAI and IGRA, deeper knowledge of the effects of differences in the model versions and assimilation systems used in the reanalyses is needed. Changes in the parameterisation of other subgrid-scale processes, e.g. in the radiative, cloud and boundary layer parameterisation schemes, as well as changes in the data-assimilation system, also contribute to the differences in the vertical structures of temperature and RH we see. Lastly, it is worth noting that although the standard deviations of temperature and RH differences were large (Fig. 4), their vertical distributions were different from those of the mean differences (Figs. 2 and

3) suggesting that they are not dominated by the same dynamics.

5. Conclusions

In this study, the vertical profiles of temperature and relative humidity in ERA5 and ERAI reanalyses were compared to those in the IGRA sounding dataset in eight sounding stations over tropical oceans. The mean and standard deviations of the differences between these datasets were studied.

In most stations, a cold and moist anomaly occurs at ~600–800 hPa, and a warm and dry anomaly at ~800–900 hPa, in ERA5 compared to ERAI (Figs. 2 and 3). These differences between ERA5 and ERAI are qualitatively similar to the effect, seen in a modelling study by Hirons et al. (2013b), of changing the organised deep entrainment formulation from the one used in the model version of ERAI to the one used in the model version of ERA5 suggesting that the entrainment may be the main reason for these differences. The fact that ERAI is generally closer to observations than ERA5 is in the low-to midtroposphere (in the layer ~550–800 hPa for temperature and ~650–800 hPa for RH), suggests that the newer entrainment used in ERA5, although resulting in a better sensitivity of convection to moisture, may not be quite right. We suggest that another mechanism, in addition to entrainment, could at least partly cause the sensitivity of convection to moisture.

Acknowledgements

We would like to thank the anonymous reviewer, whose comments significantly improved the manuscript.

Disclosure statement

The authors declare that they have no conflict of interest.

Funding

We gratefully acknowledge the funding received for this research by the Academy of Finland (grant numbers 1333034, 307331, and 316704), the Doctoral Programme in Atmospheric Sciences in the University of Helsinki, and the Vilho, Yrjö and Kalle Väisälä Foundation.

Data availability statement

ERA5 is available in the Climate Data Store at <https://cds.climate.copernicus.eu>. ERA-Interim can be downloaded from the ECMWF web pages at <https://apps.ecmwf.int/datasets>. IGRA version 2 dataset is available at <https://www.ncdc.noaa.gov/data-access/weather-balloon/integrated-global-radiosonde-archive> and the TRMM 3B42 satellite precipitation data used in Fig. 1 at <https://gpm.nasa.gov/trmm>.

References

- Bechtold, P., Köhler, M., Jung, T., Doblas-Reyes, F., Leutbecher, M. and co-authors. 2008. Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Q. J. R. Meteorol. Soc.* **134**, 1337–1351. doi:10.1002/qj.289
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P. and co-authors. 2011. The ERA-Interim Reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**, 553–597. doi:10.1002/qj.828
- Durre, I. and Yin, X. 2008. Enhanced radiosonde data for studies of vertical structure. *Bull. Am. Meteor. Soc.* **89**, 1257–1262. doi:10.1175/2008BAMS2603.1
- Emanuel, K. A. 1994. *Atmospheric Convection*. Oxford University Press, New York.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A. and co-authors. 2020. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049. doi:10.1002/qj.3803
- Hirons, L. C., Inness, P., Vitart, F. and Bechtold, P. 2013a. Understanding advances in the simulation of intraseasonal variability in the ECMWF model. Part I: The representation of the MJO. *Q. J. R. Meteorol. Soc.* **139**, 1417–1426. doi:10.1002/qj.2060
- Hirons, L. C., Inness, P., Vitart, F. and Bechtold, P. 2013b. Understanding advances in the simulation of intraseasonal variability in the ECMWF model. Part II: The application of process-based diagnostics. *Q. J. R. Meteorol. Soc.* **139**, 1427–1444. doi:10.1002/qj.2059
- Nogueira, M. 2020. Inter-comparison of ERA-5, ERA-Interim and GPCP rainfall over the last 40 years: Process-based analysis of systematic and random differences. *J. Hydrol.* **583**, 124632. doi:10.1016/j.jhydrol.2020.124632
- Virman, M., Bister, M., Sinclair, V. A., Järvinen, H. and Räisänen, J. 2018. A new mechanism for the dependence of tropical convection on free-tropospheric humidity. *Geophys. Res. Lett.* **45**, 2516–2523. doi:10.1002/2018GL077032
- Virman, M., Bister, M., Sinclair, V. A., Räisänen, J. and Järvinen, H. 2020. Vertical temperature structure associated with evaporation of stratiform precipitation in idealized WRF simulations. *J. Atmos. Sci.* **77**, 1851–1864. doi:10.1175/JAS-D-19-0111.1