Comparison of SPEEDY calculations to observational data

Knop, Irmgard; Schubert, Natalie; Wolter, Ole; Aydın, Onurhan. 2021. Johann Wolfgang Goethe Universität, Frankfurt am Main.

Abstract: The SPEEDY model has been run over the time span of 10 years several times, including or excluding land-surface, sea-surface, and sea-ice temperature coupling. The model's results for the precipitation in the equatorial region are validated against observational data. Furthermore, the differences in the standard deviation between the model and the observations and the correlation between both sources are determined. Overall, the couplings only have a slight impact on the resulting precipitation data, however none of the runs are able to come very close to the observational data.

1.0 Introduction and setup

In this report, calculated values of a global climate model (GCM) are compared against observational data. The model used for these calculations is the SPEEDY (Simplified Parameterizations, privitivE-Equation DYnamics) model, which has been developed at the International Centre for Theoretical Physics (ICTP) located in the Castello di Miramare, Italy (**?**). The model has been set up to run over the ten year time span from January 1979 to December 1988.
A difficult task is modeling the precipitation in the equator region (25°S to 25°N). Running the SPEEDY model while having different coupling flags (land-surface temperature (ICLAND), sea-surface temperature (ICSEA), sea-ice temperature (ICICE)) turned on or off may significantly alter the results. To assess which coupling flags deliver results closest to the observed precipitation, the model is run a total of five times. Once all coupling parameters are off, three times where only one parameter is on, and once where all parameters are off. In this case, monthly averaged NCEP Reanalysis data of the precipitation for the ten year span was selected (**?**)
To achieve the optimal comparability, both the model results and the observation data set are spatially adapted to the same grid, in this case a 360 x 180 (longitude x latitude) grid.

2.0 Results (ISSTAN=0)

In this section, the data and some statistic measures of both the model and the observations are plotted in several ways seeking to validate the model performance against the observations. The respective ways are further described in the following subsections.
In this report, every season is to be understood as the season on the northern hemisphere (e. g. summer = winter in southern hemisphere).

2.1 Sea-surface temperature anomaly

SPEEDY can either be run using solely climatological sea-surface temperatures (SST) (ISSTAN=0) or including observed sea-surface temperature anomalies (ISSTAN=1). In order to determine the effect of ISSTAN, the model is run twice, once where ISSTAN=0 and in the other case for ISSTAN=1. In both runs, the three coupling flags are ICLAND=1 (land-surface temperature coupling), ICSEA=2 (full uncorrected SST from coupled ocean model), and ICICE=1 (ice-model coupling). In both cases, the resulting precipitation data sets are identical to one another, which is visible when comparing the resulting mean precipitation in the equator region over the ten year period, as illustrated in figure 1. Additionally, the difference of the two data sets is 0 for every grid point over the whole time period. Based on this fact, all further runs have ISSTAN set to 0, as this parameter doesn't seem to have an influence on the precipitation.

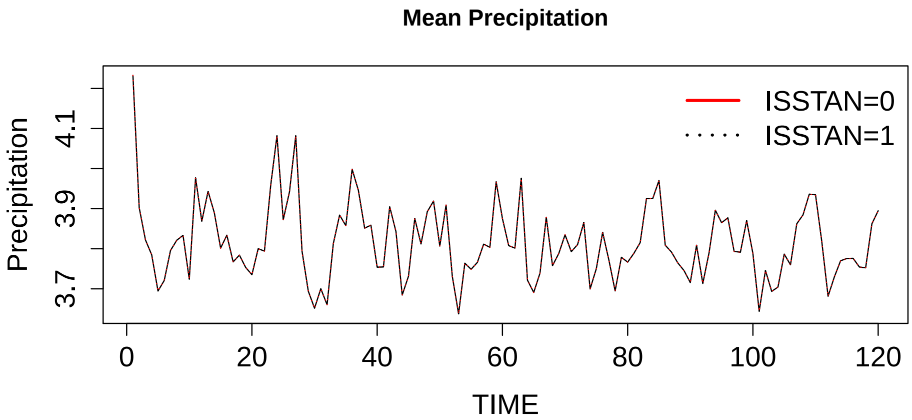

Figure 1: Comparison of the mean precipitation over a ten year time span (1979-1988) in the equator region from two SPEEDY runs, where ISSTAN=0 or 1. In both runs ISLAND=1, ICSEA=2, ICICE=1.

From now on, the SPEEDY runs will be S110-114, where

$$\begin{aligned}
\text{S110} &\,\widehat{=}\, \text{ICLAND} = 0\ ,\ \text{ICSEA} = 0\ ,\ \text{ICICE} = 0 \\
\text{S111} &\,\widehat{=}\, \text{ICLAND} = 1\ ,\ \text{ICSEA} = 0\ ,\ \text{ICICE} = 0 \\
\text{S112} &\,\widehat{=}\, \text{ICLAND} = 0\ ,\ \text{ICSEA} = 2\ ,\ \text{ICICE} = 0 \\
\text{S113} &\,\widehat{=}\, \text{ICLAND} = 0\ ,\ \text{ICSEA} = 0\ ,\ \text{ICICE} = 1 \\
\text{S114} &\,\widehat{=}\, \text{ICLAND} = 1\ ,\ \text{ICSEA} = 2\ ,\ \text{ICICE} = 1
\end{aligned}$$

If a coupling parameter is equal to 0, then the run will not include this coupling. If it is 1 or 2, then the run will include the coupling.

2.2 Comparison of the spatial mean

In this part, the coupling flags of SPEEDY and observational data will be analyzed statistically. The lines of figure 2 show the precipitation for the observation data set (black) and different coupling flags of the SPEEDY model (in order red, green, cyan, orchid, golden) in mm/day plotted against the time in months. The precipitation regime for the specific location (25°N - 25°S) is approximately between 3.1-4.3 mm per day. Since the observational location is the equatorial area, the seasonal effects are smaller. The difference of the results from SPEEDY runs using the various coupling flag combinations is between values of 4.3 and 3.6. Generally, it can be said that there is no huge fluctuation among the various SPEEDY runs. Although observational data has more outliers, unfortunately none of the configurations of the SPEEDY is able to efficiently forecast them. In addition to this, all the flags of the SPEEDY have inconsistent predictions in some specific time intervals. At the end of the plot, even though all SPEEDY run results are closer to the observation data, SPEEDY could not predict decreasing amount in precipitation as coherent as the line trending to increase.

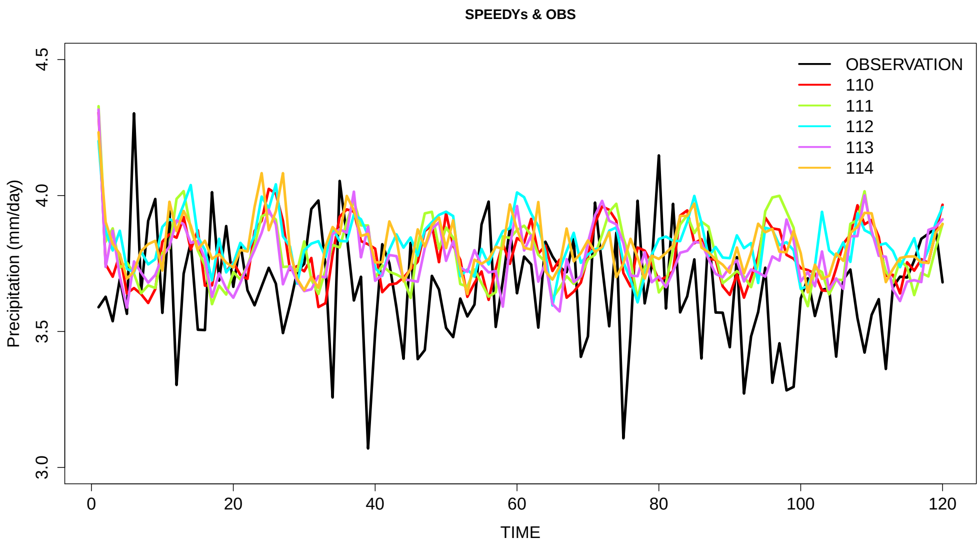

Figure 2: Comparison of the couples (in order red, green, cyan, orchid, golden) of the SPEEDY to the observations (black) for precipitation.

The box plots in figure 3 display some statistical measures for the coupling flag results and the observational data. The box size represents the average precipitation range, in which the respective data lies in, the line in bold print stands for the mean precipitation value over the whole time and the dashed upward lines represent the stochastic error. In the boxplots, the equatorial spatially and time averaged data of all sources is visualized. Red boxes represent the SPEEDYs and black one is observation.

The average stochastic global mean precipitation amount for the years 1979 to 1988 is about 3.8 mm/day for all SPEEDY runs and 3.7 mm/day for the observations. As already observed in figure 2 and now again visualized with the boxplots, the SPEEDY model is not able to predict large precipitation fluctuations. In terms of data size, IQR (Interquartile Range) which is the one of the best indicators of sample sizes is bigger in S111 and which means data range is much more than other flags. The mean value is smaller than the median value, so its average precipitation amount is near to minimum values. S112 has the lowest IQR range so values are closer to each other. Which means, mean and median is nearly same and its standard deviation (error) and variance would be smaller in the annual cycle of the precipitation amount. When considering all SPEEDY runs, S112 is the closest to the observational data set.

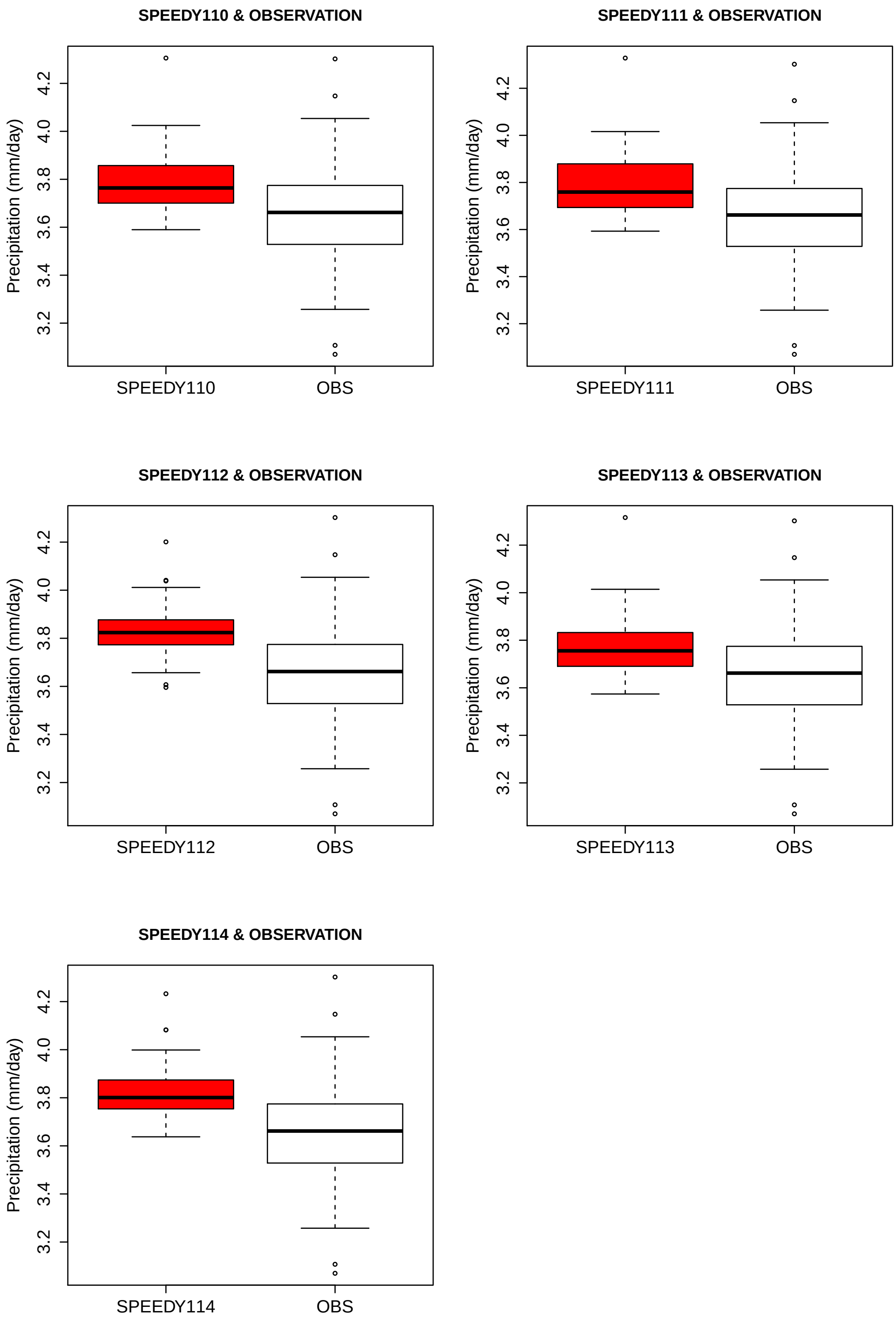

Figure 3: Comparison of the coupling flags of the SPEEDY (red) to the observations (black) for precipitation.

2.3 Tropic precipitation and standard deviation

In this subsection, the results from the five different SPEEDY model runs and the observational data are averaged over the full 10 year span from 1979 to 1988, and the data and the standard deviation of both sources and the respective differences are plotted spatially as a contour plot, i. e. the x-axis represents the longitude and the y-axis represents the latitude in each figure, while the colors represent different temperatures or differences between the model results and the observations. In addition, the outline of the land mass is shown to assign the data to identifiable locations. Due to the amount of plots the figures are split into three groups.

The first group of plots visualizes the average daily precipitation in the investigated time frame in 4a) and b). In both cases, most of the precipitation falls west of Central America, in the Indian ocean and north of Indonesia with over 12 mm/day. There are no big qualitative differences between both plots. Most notably, enabling the land surface temperature coupling (S111) slightly increases the region of heavy precipitation northeast of Indonesia.
It seems plausible, that since the differences between the predicted precipitations are marginal, so should be the respective distinctions in the differences between the model output and the observations (second row), in the standard deviations (third row) and in the differences between the model output and the observations regarding the standard deviation (fourth row). Both model runs seem to slightly overshoot on the precipitation, especially in the regions where it predicted the highest amount, as can be verified in the figures 4c) and 4d). For most of the investigated surface, the difference lies somewhere between 0 and 4 mm/day. The standard deviation for both model runs visualized in 4e) and 4f) is highest in the regions with the biggest precipitation, which roughly are also the places, where the standard deviation is bigger than in the observations, i. e. the difference in standard deviation displayed in the plots 4g) and 4h) is positive.

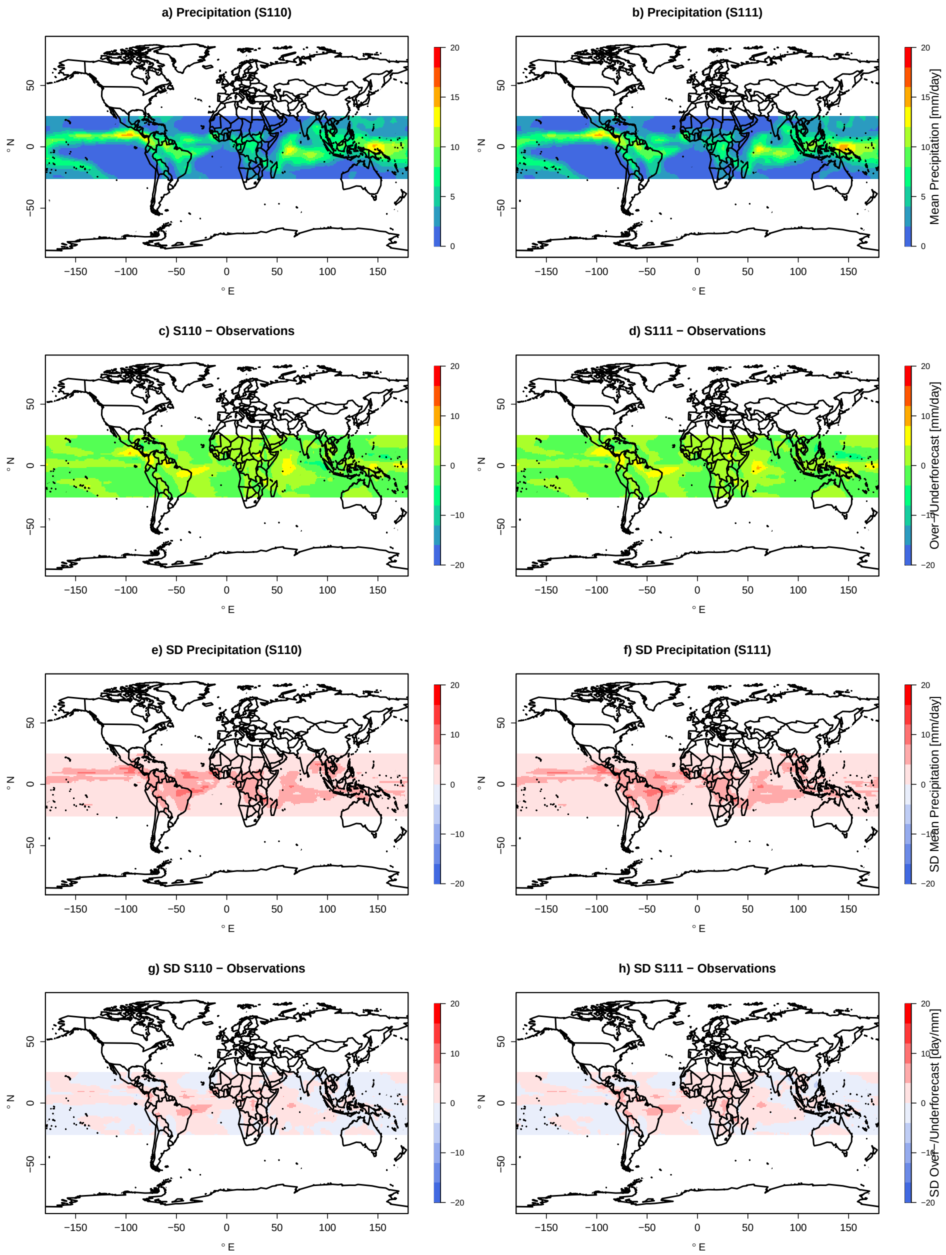

Figure 4: Precipitation data and standard deviation for SPEEDY data as contour plots. The latitude is plotted against the longitude. First line: Precipitation data; a) S110, b) S111. Second Line: c) Difference between S110 and observations, d) Difference between S111 and observations. Third line: standard deviation of precipitation; e) S110, e) S111. Fourth line: g) Difference in standard deviation between S110 and observations, h) Difference in standard deviation between S111 and observations.

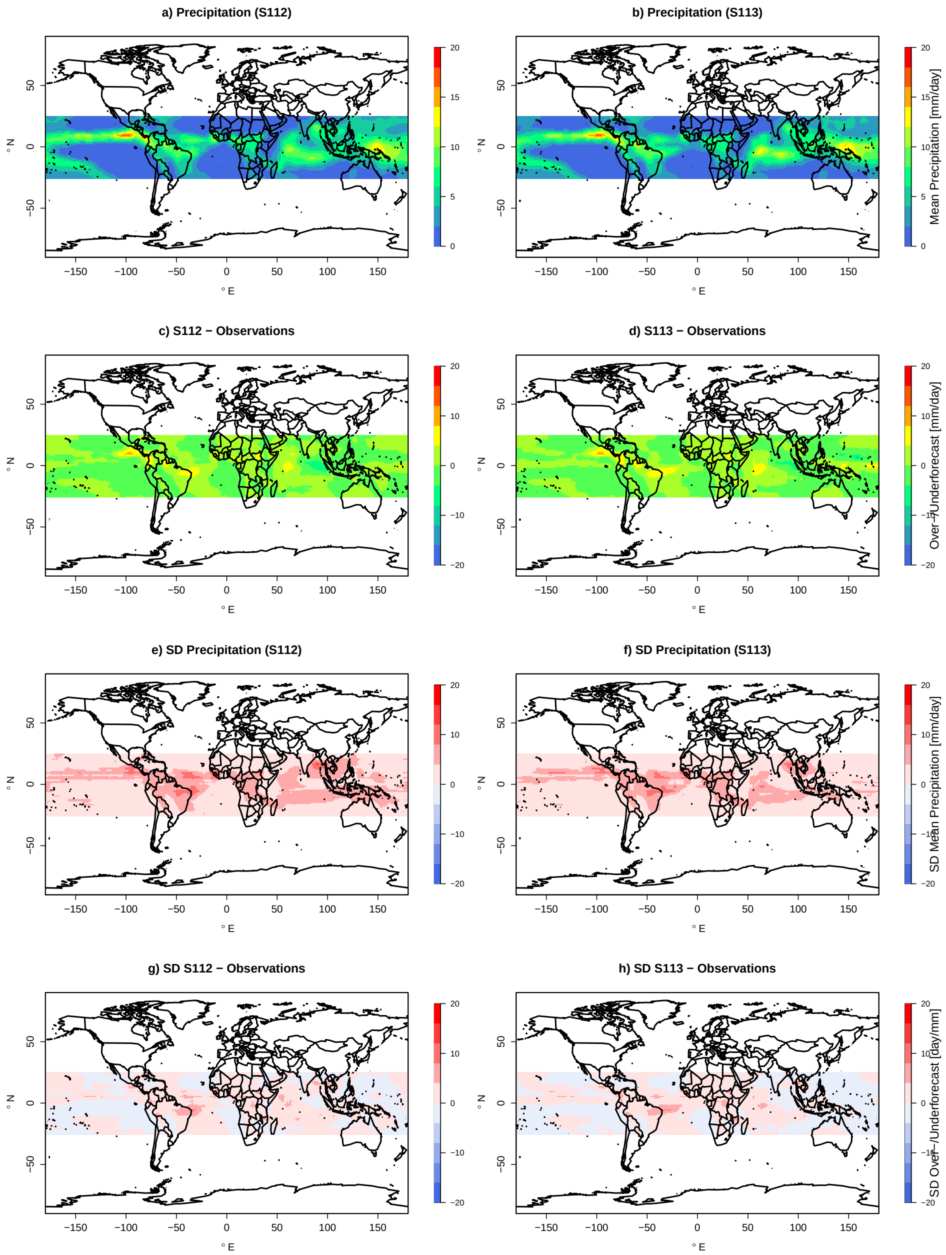

Figure 5: Precipitation data and standard deviation for SPEEDY data as contour plots. The latitude is plotted against the longitude. First line: Precipitation data; a) S112, b) S113. Second Line: c) Difference between S112 and observations, d) Difference between S113 and observations. Third line: standard deviation of precipitation; e) S112, e) S113. Fourth line: g) Difference in standard deviation between S112 and observations, h) Difference in standard deviation between S113 and observations.

In the second group, the data from the model runs S112 and S113 are illustrated. Again, the projected mean daily precipitation is shown in the plots 5a) and 5b). The precipitation northeast of Indonesia is lower for S112 and S113 than for S110 and S111, while the opposite tendency can be observed for the area west of Central America. Yet, in general, there seem to be no major discrepancies between the figures 5a) and b) and the corresponding plots in figure 4.
This justifies the assumption, that one will end up with similar results regarding the other six plots. In fact, neither for the case of the discrepancy between the model and the observation data set (figures 5c) and d)), nor in the standard deviation of the different runs S112 and S113 figures (5e) and f)), nor regarding the difference between the standard deviation of the model output and the standard deviation in the observation data. Again, the precipitation has been overforecasted for the runs S112 and S113 in the region of especially high precipitation, while it is lightly underestimated in many other places. The standard deviation for both model runs displayed in 5e) and 5f) is higher in the regions with the biggest precipitation, but again mostly smaller than 4 mm/day. This again coincides well with the places, where the standard deviation is bigger than in the observations, i. e. the difference in standard deviation visualized in the plots 5g) and 5h) is positive. The general deviation of the standard deviation is once again mostly smaller than 4 mm/day towards the model or the Observations.

Finally, the data of the final run S114 and the observational data set are displayed. The plots 6a), c), d) and e) show the mean precipitation of the model output, the difference between the model results and the observations for the data and the standard deviation of the data and the standard deviation of the S114 data, respectively. Once again, the differences to the previous model runs are rather small, i. e. the regions of high and low precipitation, the spatial distribution of the discrepancies in precipitation and the related standard deviation between the model output and the observation as well as the standard deviation of the computed data are almost identical.
The figure 6b) shows the averaged observations for the years 1979 to 1988. The main difference to all the model runs lies in the location of the regions of higher precipitation in the Indian Ocean and the western Pacific, which are observed more to the east compared to how they were projected. Also the precipitation is observed to be slightly lower than forecasted for Brazil and the region west of Central America. The higher values of the observation's standard deviation are accordingly shifted, as can be seen in plot 6f). In general, the standard deviation spreads less away from 0 than the computed data.

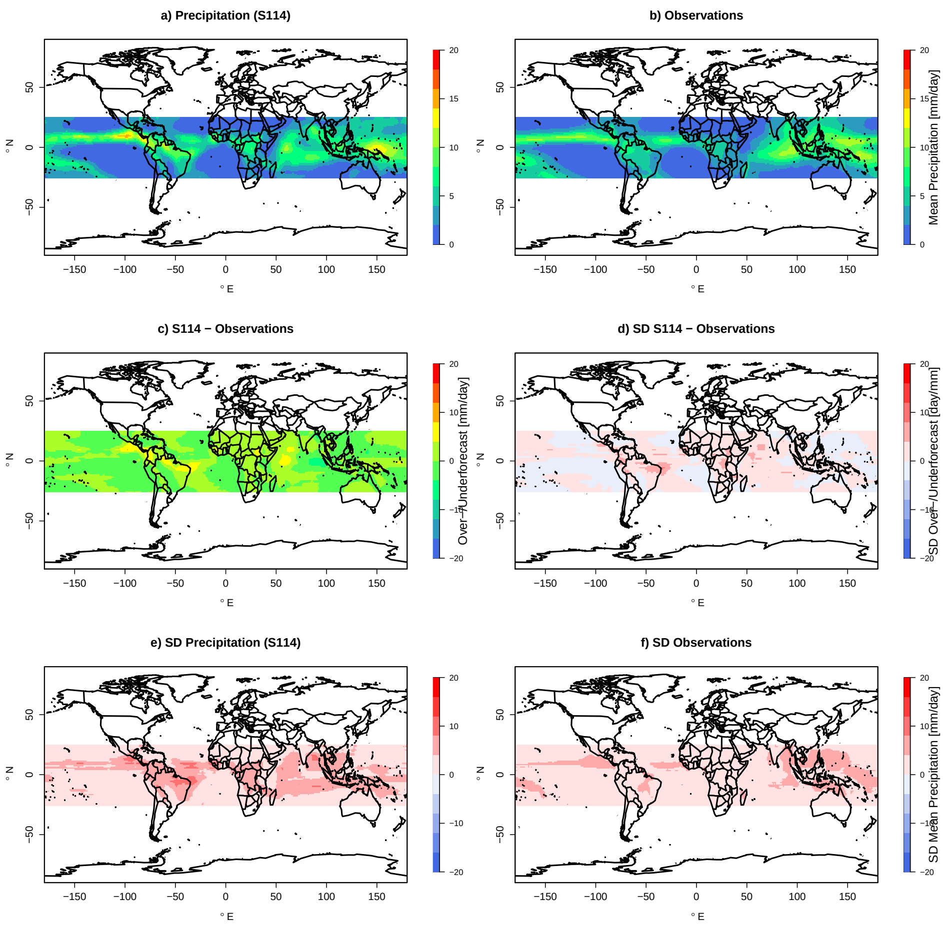

Figure 6: Precipitation data and standard deviation for SPEEDY data as contour plots. The latitude is plotted against the longitude. First line: Precipitation data; a) S114, b) OBSERVATIONS. Second Line: c) Difference between S114 and observations, d) Difference in standard deviation between S114 and observations. Third line: standard deviation of precipitation; e) S114, f) Observations.

2.4 Correlation

The correlation indicates if there might be a positive (correlation coefficient = 1) or negative (correlation coefficient = -1) linear dependence between two variables, or in this case between two data sets. If the correlation coefficient equals 0, there is definitely no linear dependence between two variables. In that case, either there is some other type of relation between the variables, or there is no dependence at all. Alas, the correlation cannot provide clarification in that matter.
Lastly, the correlation between the output data from SPEEDY and the observations is plotted in figure 7.

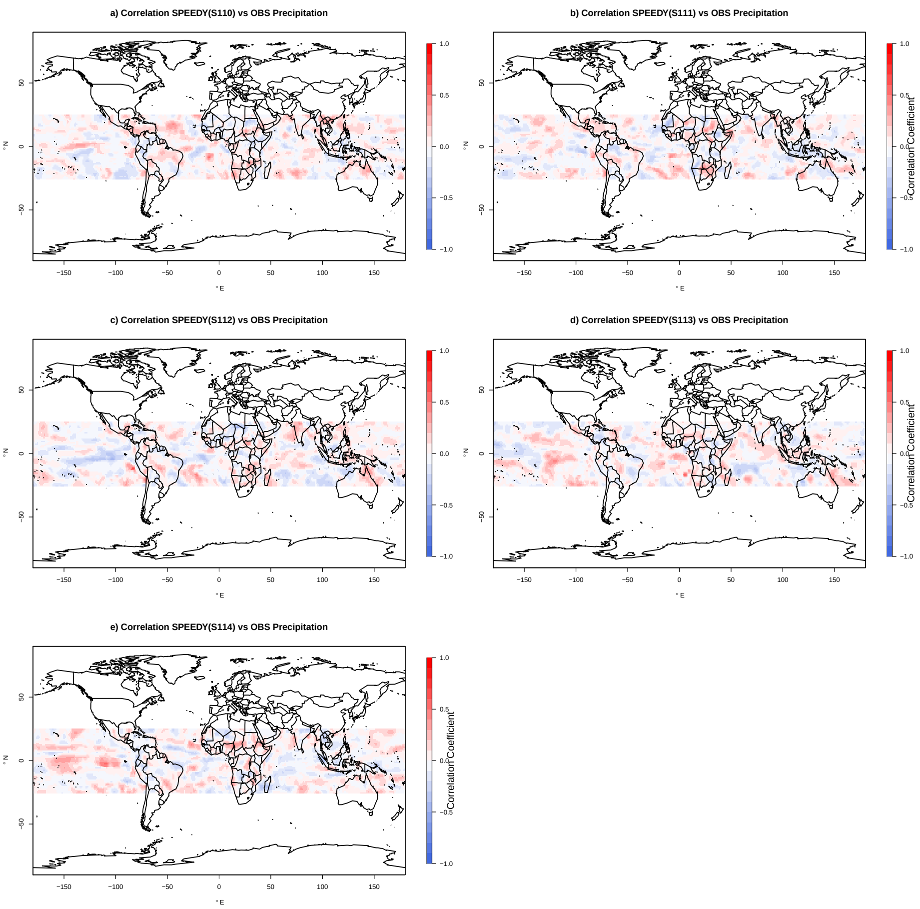

Figure 7: Correlation of the measured and computed Precipitation data sets visualized as a contour plot. The latitude is plotted against the longitude. Red stands for a possible positive linear dependence and blue represents a possible negative linear dependence. Light areas mean no linear correlation. First line: Correlation of precipitation; a) S110 vs. Observations, b) S111 vs. Observations, c) S112 vs Observations. Second Line: d) S113 vs. observations, e) S114 vs. Observations.

Figure 7a, 7b, 7c, 7d and 7e show the correlation between the different model runs (S110, S111, S112, S113 and S114) and the observational data. The plots look relatively similar as they show no big differences. The data lies between about -0.5 and +0.5 for the correlation coefficient. Figure 7e shows the most red areas, especially over the ocean compared to the other plots. The plots indicate a small positive or negative correlation between the SPEEDY data and the observational data. There are also a lot of small areas that show no correlation at all.
In General, there is no indication for a dependence between the data sets.

3.0 Results (ISSTAN = 1)

The parameter ISSTAN is only active if ICSEA = 0 or 1. Therefore, all runs where ICSEA = 0 are repeated, this time with ISSTAN = 1 (active). As in the previous section, the results will be compared to the observational data. These SPEEDY runs are S120, S121, and S123, where

$$\begin{aligned}
\text{S120} &\mathrel{\widehat{=}} \text{ICLAND} = 0\ ,\ \text{ICSEA} = 0\ ,\ \text{ICICE} = 0\\
\text{S111} &\mathrel{\widehat{=}} \text{ICLAND} = 1\ ,\ \text{ICSEA} = 0\ ,\ \text{ICICE} = 0\\
\text{S113} &\mathrel{\widehat{=}} \text{ICLAND} = 0\ ,\ \text{ICSEA} = 0\ ,\ \text{ICICE} = 1\ .
\end{aligned}$$

3.1 Comparison of the spatial mean

Again, we compare the spatial mean precipitation of the SPEEDY runs with the observation data, using line (figure 8) and box plots (figure 9).

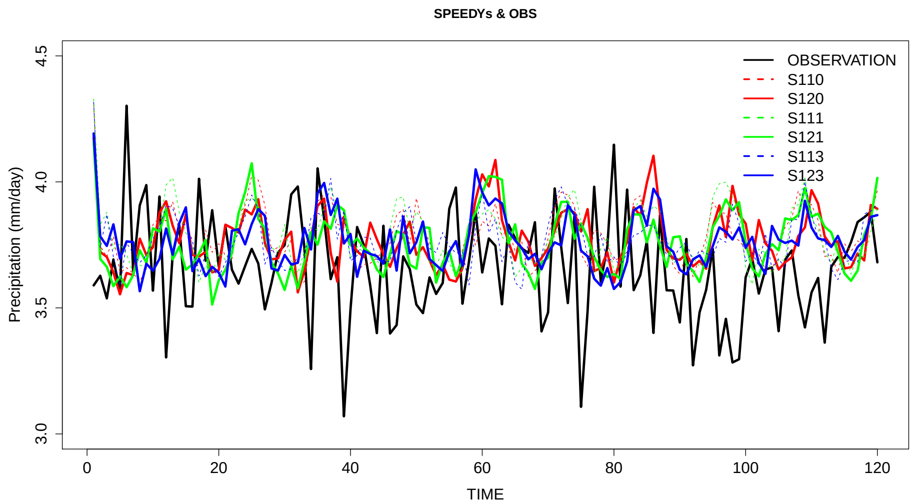

Figure 8: Comparision of the couples (in order dash red, red, dash green, green, dash blue, blue) of the SPEEDY to the observations (black) for precipitation.

The lines of figure 8 show the precipitation for the observational data set (black) and different coupling flags of the SPEEDY model (in order dashed red, red, dashed green, green, dashed blue, blue) in mm/day plotted against the time in months.
The precipitation regime for the specific location (25°N - 25°S) is approximately between 3.1 - 4.4 mm per day. Since the observational location is the equatorial area, seasonal effects is smaller. The intervals of the coupling flags are also between values of 4.4 and 3.6. Generally, it can be said that there is no huge fluctuation among the flags of the SPEEDY. Although observational data has more outliers, unfortunately none of the model runs is able to efficiently forecast them. In addition to this, all the flags of SPEEDY have inconsistent predictions in some specific time intervals. When different couplings are considered, as the relationship between the numerical model and observational data, noticeable differences can hardly be seen. In short, it can be said that both groups of SPEEDY have the same trend in the graphics.

The box plots in figure 9 display some statistical measures for the coupling flag results and the observational data. The box size represents the average precipitation range, which the respective data lies in, the line in bold print stands for the mean precipitation value over the whole time and the dashed upward lines represent the stochastic error. In the boxplots the equatorial spatially and time averaged data of all sources is visualized. Red boxes represent SPEEDY's results while the observations are displayed in black.

The average stochastic global mean precipitation amount for the years 1979 to 1988 is about 3.8 mm/day according to the coupling flags of the SPEEDY and 3.7 mm/day according to the observations. As in the previous box plots in section 2, none of the flags predict the outliers below the lower fence. S123 has outliers above the upper border. In terms of data size, IQR (Interquartile Range) which is the one of the best indicators of sample sizes is bigger in S111 and S121, which means the data range is much larger than for the other SPEEDY runs. The mean value is smaller than the median value so its average precipitation amount is closer to the minimal values. SPEEDY runs S113 and S123 have the lowest IQR so values are very close to each other. This means, mean and median are nearly identical and its standard deviation (error) and variance would be smaller in the annual cycle of the precipitation amount. When all the flags are considered, although the discrepancy between SPEEDY runs S111 and S121 are higher than for the other outputs, it is the closest to the observational data. In short, the SPEEDY outputs are very similar to each other.

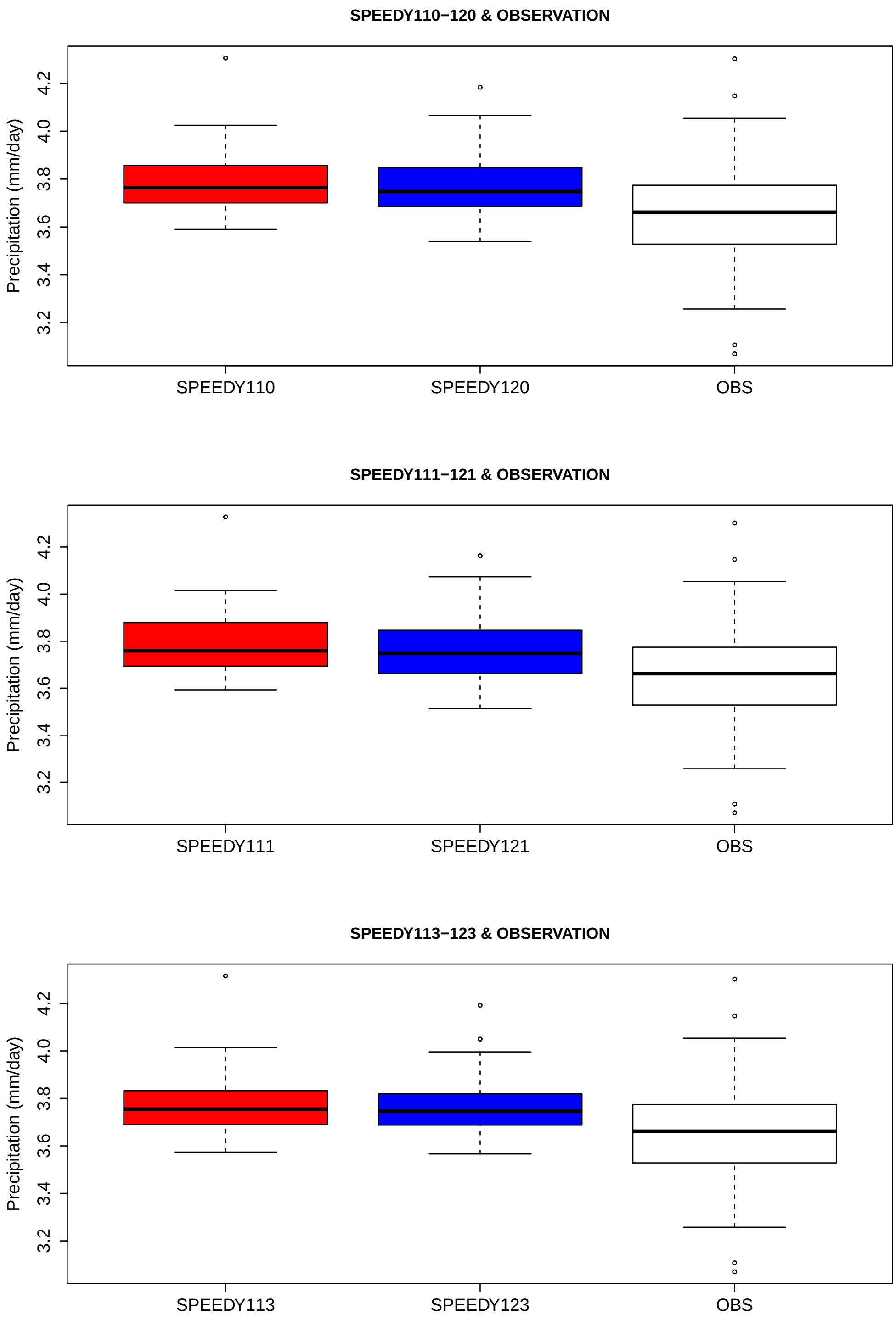

Figure 9: Comparision of the coupling flags of the SPEEDY (red and blue) to the observation (black) for precipitation.

3.2 Comparison of the temporal mean and standard deviation

In this subsection, the temporal mean and standard deviation of the SPEEDY results is compared to the observational data. The figures 10 to 13 illustrate the mean of the SPEEDY results and the observational data, the difference between the mean and the observational data, the standard deviation of all data sets, and the difference between the standard deviation of the SPEEDY results and the observational data, respectively.

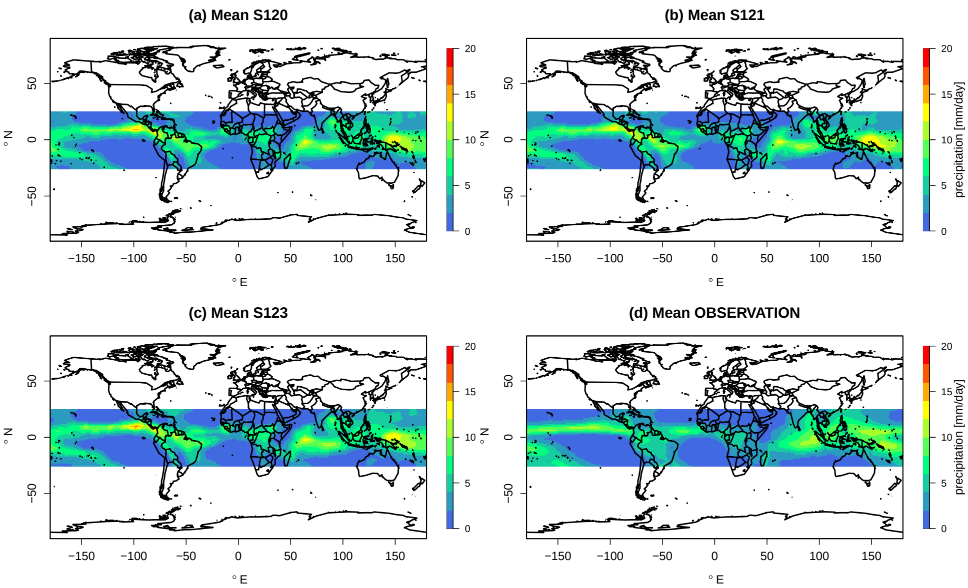

Figure 10: Temporal mean of the precipitation in the equatorial region for the SPEEDY runs (a) S120, (b) S121, (c) S123, and (d) the observational data.

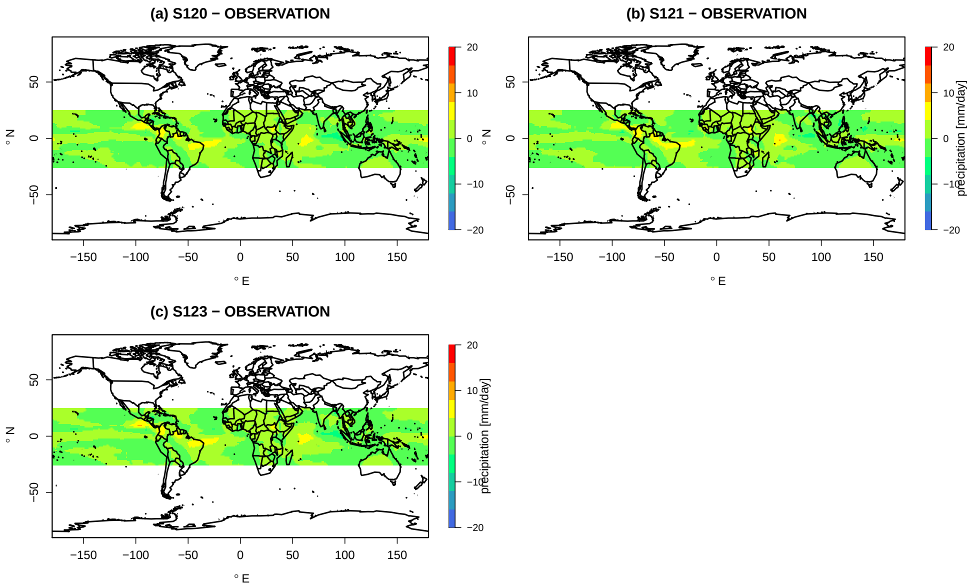

Figure 11: Difference between the precipitation mean modeled using SPEEDY and the observational data for the SPEEDY runs (a) S120, (b) S121, and (c) S123.

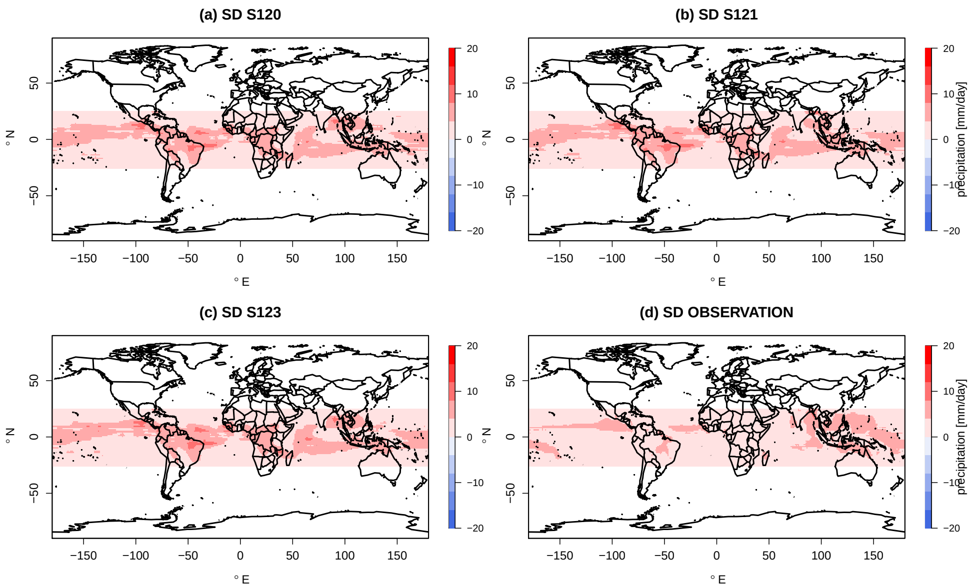

Figure 12: Standard deviation of the precipitation in the equatorial region of the SPEEDY runs (a) S120, (b) S121, (c) S123, and (d) the observational data.

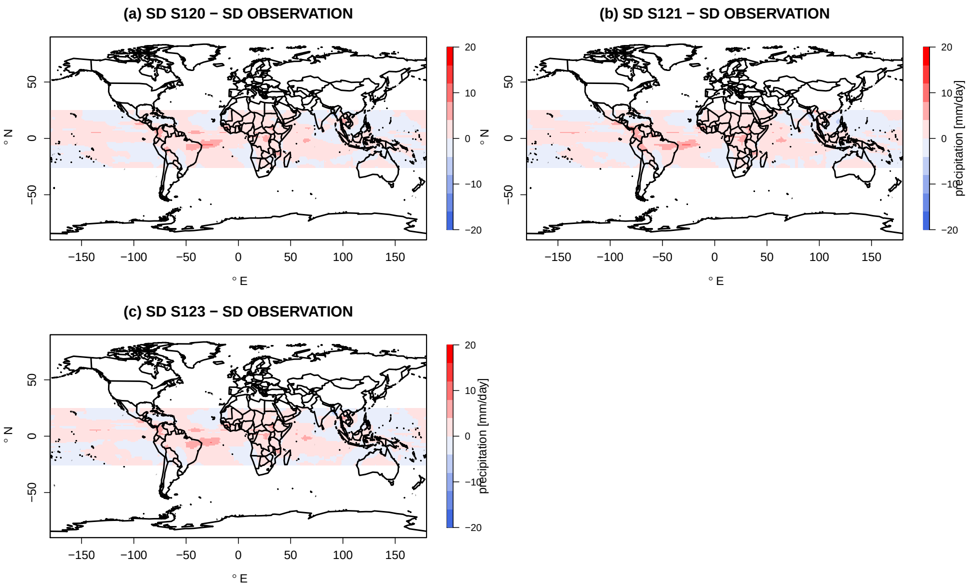

Figure 13: Difference between the standard deviation of the precipitation modeled using SPEEDY and the observational data for the SPEEDY runs (a) S120, (b) S121, and (c) S123.

As it was the case for the runs with ISSTAN deactivated, the three SPEEDY runs produce very similar results, which is visible not only when comparing the mean precipitation of the three runs (see figure 10) but also when comparing the standard deviation of the runs (see figure 12). In all three cases, SPEEDY often overpredicts precipitation, e.g. at approximately ±100 °E (see figure 11). SPEEDY also predicts precipitation with larger variance than in reality, e.g. at −50 °E (see figure 13). Overall, out of the three SPEEDY runs, there seems to be no runs with predicted precipitation significantly better than the other two runs.

3.3 Correlation between SPEEDY and observations

The correlation indicates if there might be a positive (correlation coefficient = 1) or negative (correlation coefficient = -1) linear dependence between two variables, or in this case between two data sets. If the correlation coefficient equals 0, there is definitely no linear dependence between two variables. In that case, either there is some other type of relation between the variables, or there is no dependence at all. Alas, the correlation cannot provide clarification in that matter.

Lastly, the correlation between the new output data from SPEEDY and the observations is plotted in figure 14.

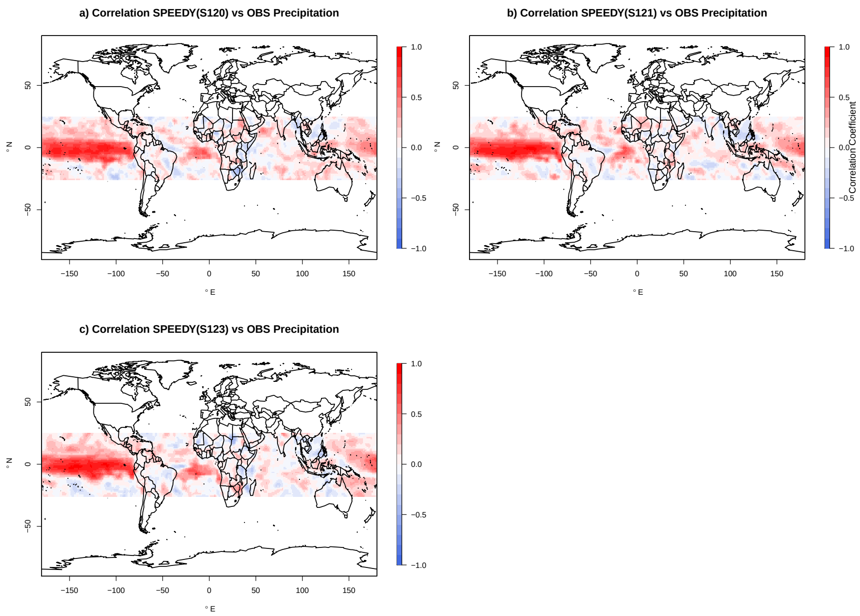

Figure 14: Correlation of the measured and computed precipitation data sets visualized as a contour plot. The latitude is plotted against the longitude. Red stands for a possible positive linear dependence and blue represents a possible negative linear dependence. Light areas mean no linear correlation. Correlation of precipitation; a) S120 vs. Observations, b) S121 vs. Observations, c) S123 vs Observations.

Figure 14a, 14b and 14c show the correlation between the different model runs (S120, S121 and S123) and the observational data. The plots look relatively similar as they show no big differences. The plots show a positive correlation coefficient especially over the pacific ocean. The coefficient of the other areas lies between about -0.5 and +0.5 but there are again also areas that show no correlation at all. As both data sets have been collected independently, they are not dependent on each other so there should not be an observable correlation. The positive correlation coefficient over the pacific ocean can be interpreted as a sign of good agreement between the data sets in this area.
Figure 15 shows the correlation between the old data sets (S110, S111 and S113) and the new data sets (S120, S121 and S123).

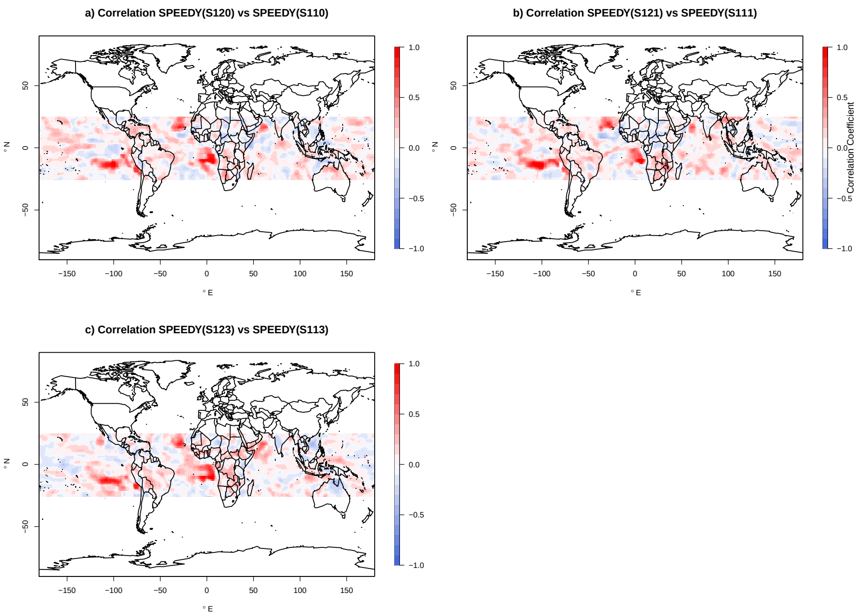

Figure 15: Correlation of the measured and computed precipitation data sets visualized as a contour plot. The latitude is plotted against the longitude. Red stands for a possible positive linear dependence and blue represents a possible negative linear dependence. Light areas mean no linear correlation. Correlation of precipitation; a) S120 vs. S110, b) S121 vs. S111, c) S123 vs S113.

The plots in figure 15 also show no big differences when comparing them. All three plots show a positive correlation in the Pacific Ocean near South America and west of Africa. The correlation coefficient in the other areas in figures 15a, 15b and 15c is again between -0.5 and +0.5 with small areas that show no correlation at all.
The biggest difference when comparing the S120, S121 and S123 with the old data sets and the observational data can be seen over the Pacific Ocean as the observational data show a positive correlation.

In General, aside from small areas over the Pacific Ocean, there is no indication for a dependence between the new data sets and the observational data as well as between the new and the old data sets.

4.0 Interpretation

Overall, having all coupling parameters turned on (ICLAND=1, ICSEA=2, ICICE=1) produces results, which are closest to the observational data. However, all SPEEDY runs produce very similar results when regarding the precipitation in the equatorial region. It seems as though the coupling parameters have no significant impact on the precipitation. Additionally, activating or deactivating ISSTAN has no significant impact on the quality of the SPEEDY results. To further understand the effect of these couplings, it would be interesting to analyze other SPEEDY climate variables.

References

[1] Qiu, Y., W. Cai, X. Guo, B. Ng, 2014: The asymmetric influence of the positive and negative IOD events on China's rainfall, Scientific Reports, 4, 4943, DOI: 10.1038/srep04943.

[2] Vinayachandran, P. N., P. A. Francis, S. A. Rao: Indian Ocean Dipole: Processes and impacts, Platinum Jubilee Special.

[3] Horii, T., I. Ueki, K. Ando, K. Mizuno, 2013: Eastern Indian Ocean warming associated with the negative Indian Ocean dipole: A case study of the 2010 event, JOURNAL OF GEOPHYSICAL RESEARCH: OCEANS, 118, 536–549, doi:10.1002/jgrc.20071.

[4] Ashok, K., Z. Guan, T. Yamagata, 2001: Impact of the Indian Ocean Dipole on the Relationship between the Indian Monsoon Rainfall and ENSO, GEOGRAPHICAL RESEARCH LETTERS, 28, 23, 4499-4502.

[5] Lu, B., Ren, H.-L. (2020). What caused the extreme Indian Ocean Dipole event in 2019? Geophysical Research Letters, 47, e2020GL087768, https://doi.org/10.1029/2020GL087768.

[6] Lu, B., and Coauthors, 2017: An extreme negative Indian Ocean Dipole event in 2016: dynamics and predictability, CrossMark, 51, 89-100, DOI 10.1007/s00382-017-3908-2.

[7] Yuan, C., T. Yamagata, 2015: Impacts of IOD, ENSO and ENSO Modoki on the Australian Winter Wheat Yields in Recent Decades, Scientific Reports, DOI: 10.1038/srep17252.

[8] Hidayat, R., K. Ando, Y. Masumoto, J. J. Luo, 2016: Interannual Variability of Rainfall over Indonesia: Impacts of ENSO and IOD and Their Predictability, IOP Conference Series: Earth and Environmental Science, 31, 012043, doi:10.1088/1755-1315/31/1/012043.