



Benchmarking global hydrological and land surface models against GRACE in a medium-sized tropical basin

Silvana Bolaños Chavarría^{1,2}, Micha Werner³, Juan Fernando Salazar¹, and Teresita Betancur Vargas¹

¹Grupo de Ingeniería y Gestión Ambiental (GIGA), Escuela Ambiental, Facultad de Ingeniería, Universidad de Antioquia, Medellín, Colombia

²Grupo de Investigación en Ingeniería Sostenible (GIS), Facultad de Ingeniería, Politécnico Colombiano Jaime Isaza Cadavid, Medellín, Colombia

³Water Resources and Ecosystems Department, IHE Delft Institute for Water Education, Delft, the Netherlands

Correspondence: Silvana Bolaños Chavarría (sbolanos@elpoli.edu.co)

Received: 15 June 2021 – Discussion started: 30 July 2021

Revised: 17 June 2022 – Accepted: 13 July 2022 – Published: 24 August 2022

Abstract. The increasing reliance on global models to address climate and human stresses on hydrology and water resources underlines the necessity for assessing the reliability of these models. In river basins where availability of gauging information from terrestrial networks is poor, models are increasingly proving to be a powerful tool to support hydrological studies and water resources assessments (WRA). However, the lack of in situ data hampers rigorous performance assessment, particularly in tropical basins where discordance between global models is considerable. Remotely sensed data of the terrestrial water storage obtained from the Gravity Recovery and Climate Experiment (GRACE) satellite mission can provide independent data against which the performance of such global models can be evaluated. However, how well GRACE data represents the dynamics of terrestrial water storage depends on basin scale and hydrological characteristics. Here we assess the reliability of six global hydrological models (GHMs) and four land surface models (LSMs) available at two resolutions. We compare the dynamics of modelled Total Water Storage (TWS) with TWS derived from GRACE data over the Magdalena–Cauca basin in Colombia. This medium-sized tropical basin has a well-developed gauging network when compared to other basins at similar latitudes, providing unique opportunity to contrast modelled TWS and GRACE data across a range of scales. We benchmark monthly TWS changes from each model against GRACE data for 2002–2014, evaluating monthly variability, seasonality, and long-term variability trends. The TWS changes are evaluated at basin level, as well as for selected

sub-basins with decreasing basin size. We find that the models poorly represent TWS for the monthly time series, but they improve in representing seasonality and long-term variability trends. The high-resolution GHM World-Wide Resources Assessment (W3RA) model forced by the Multi-Source Weighted Ensemble Precipitation (MSWEP) is most consistent in providing the best performance at almost all basin scales, with higher-resolution models generally outperforming lower-resolution counterparts. This is, however, not the case for all models. Results highlight the importance of basin scale in the representation of TWS by the models, as with decreasing basin area, we note a commensurate decrease in the model performance. A marked reduction in performance is found for basins smaller than 60 000 km². Although uncertainties in the GRACE measurement increase for smaller catchments, the models are clearly challenged in representing the complex hydrological processes of this tropical basin, as well as human influences. We conclude that GRACE provides a valuable dataset to benchmark global simulations of TWS change, in particular for those models with explicit representation of the internal dynamics of hydrological stocks, offering useful information for continued model improvement in the representation of the hydrological dynamics in tropical basins.