




PySHbundle: A Python implementation of MATLAB codes SHbundle

Amin Shakya^{1*}, Vivek Yadav^{1*}, Tsungrojungla Walling², Maya Suryawanshi¹, and Bramha Dutt Vishwakarma^{1,3¶}

¹ Interdisciplinary Centre for Water Research, Indian Institute of Science, India ² Undergraduate Programme, Indian Institute of Science, India ³ Centre of Earth Science, Indian Institute of Science, India ¶ Corresponding author * These authors contributed equally.

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Open Journals](#) 

Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

GRACE or the Gravity Recovery and Climate Experiment¹, is a gravimetric satellite mission that can detect the Total Water Storage Anomaly (TWSA) in the earth system. The data from the satellite system has been used for various hydrological studies related to groundwater depletion, floods, droughts, etc. GRACE satellite products are typically released in different product levels. In this contribution, we have translated the existing matlab codes SHbundle into the python programming language. SHbundle is a Matlab code that converts GRACE level 2 (L2) Spherical Harmonics data products into Level 3 (L3) TWSA products. In addition, a GRACE data driven correction algorithm, firstly coded in Matlab, has also been translated into Python. With this contribution, we hope to enable further work on GRACE data analytics using the Python programming language. Further, we hope to develop synergies within the geodesy community using different programming languages to better collaborate with one-another through this common framework of SHbundle and PySHbundle packages in Matlab and Python programming languages respectively.

Introduction

GRACE stands for the Gravity Recovery and Climate Experiment, a joint satellite mission by NASA, the National Aeronautics and Space Administration and DLR, the German Aerospace Centre. Some details of the GRACE mission is provided in Table 1.

Table 1: Summary of GRACE satellite mission | Parameter | Details | | ———— | :————— |
| Start of Mission | 17 March 2002 | | End of Mission | 27 October 2017 | | Inclination | 89.0° |
| Period | 94.5 minutes |

GRACE consists of two identical satellites orbiting around the earth on the same orbital path. The basic principal of the GRACE satellite operation consists of the monitoring of the intersatellite distance between the twin satellites using microwave pulse measurements (Wahr et al., 1998). When the satellite system comes across a mass anomaly, each satellite accelerates or decelerates with a phase lag and the intersatellite distance changes. This change in intersatellite distance is later processed to obtain the magnitude of the mass anomaly. When it comes to the continental land surface, the hydrological processes consist of a major component of the mass anomaly over it. However various other signals such as oceanic and atmospheric variations, systemic correlated errors, etc. are also part of the obtained GRACE signals. These unwanted signals and errors necessitate application of various filtering and post-processing techniques. These post-processing steps however also introduce some errors as well as deteriorate the quality of the hydrological product (Humphrey et al., 2023). The hydrological signal estimated after the post-processing steps is the total water storage

anomaly (TWSA). TWSA is the sum of the total water components over a vertical extension of the grid area through the earth. Conventionally, it is represented in terms of the equivalent water height (m). GRACE has a successor, GRACE-FO, which was successfully launched on 22 May 2018.

Three different research centres provide GRACE data. These are the University of Texas Center for Space Research (CSR), Jet Propulsion Laboratory (JPL), and the German Research Center for Geosciences (GFZ). Further, GRACE data is available at different levels of processing. Level 1 data refers to the raw satellite data. These are further available as Level 1A and Level 1B, based on the level of processing done to the raw data. Level 2 are the spherical harmonic coefficients for the geospatial potential estimates. These may be accessed through the JPL's Physical Oceanography Distributed Active Archive Center (PO.DAAC)² or through the Information System and Data Center (ISDC)³. Level 3 consists of mass anomalies or other standardized products, such as the Monthly Ocean/Land Water Equivalent Thickness Surface-Mass Anomaly. Similarly, mass concentration blocks or mascons are also available. These directly provide the TWSA over gridded regions, and are available through the three GRACE data centers. More details on the mascon approaches for studying gravity fields and the approaches used by the different data centers for generating mascon products may be referred to in (Antoni, 2022). The mascon products from the various data centers have some differences, attributed to the difference in post-processing steps and corrections applied by the different data centers. An online tool exists developed by the Colorado Center for Astrodynamics Research⁴. This tool can be used for a quick visualization of the CSR and GSFC mascon products over all regions of the globe. While the mascon results make application of GRACE data easier to a wider audience, use of Level 2 data gives the user the freedom and the flexibility to choose their own post-processing algorithms. The choice of application of mascon data product or Level 2 data product may depend upon the purpose of the exercise and the expertise level of the user on the GRACE data post-processing. In this contribution, we enable the user to obtain the gridded TWSA data from Level 2 data.

Level 2 GRACE data products may be stored in various data formats. These include $|C\backslash S|$, $/S\backslash C$, clm , $k\backslash m$, vector, and Colombo format (Nico Sneeuw and Matthias Weigelt and Markus Antony and Matthias Roth and Balaji Devaraju, 2021). Our contribution in its current version can handle the $|C\backslash S|$, $/S\backslash C$, clm , and $k\backslash m$ data formats. A such, our contribution can be applied using L2 Spherical Harmonics data from any of the three research centers previously mentioned. In $|C\backslash S|$ format, the C_{lm} and S_{lm} coefficient are stored as lower triangle and upper triangle, respectively, in a matrix of dimension $(l + 1) \times (l + 1)$. In $/S\backslash C$ format, the coefficients are stored in a matrix of dimension $(l + 1) \times (2l + 1)$ with horizontally flipped triangular matrix of S_{lm} coefficients on the left half, triangular matrix of C_{lm} on the right half, and zeros on the rest of the matrix elements. In our contribution, conversion between the three data formats is made possible with the modules `cs2sc`, `sc2cs`, `clm2sc`, `clm2cs`, and `k\m2sc`.

Level 3 products are the catchment average hydrological estimates of TWSA. These are obtained through the further processing of Level 2 products. Level 3 products may further be processed to obtain catchment average timeseries data, labelled as Level 4 products. Various tools exist in the literature to process GRACE data and to analyze it. Some of these available in the Matlab programming language are: SHbundle (Sneew et al., 2021), GRACE Data Driven Correction (GDDC) (Vishwakarma et al., 2017), GRAMAT (Feng, 2019), SHADE (Piretzidis, D., & Sideris, M. G., 2018), GRACETOOLS (Darbeheshti et al., 2018), etc. Similarly, some GRACE data processing tools are also available based on the python programming language. These include gravity-toolkit (Sutterley, 2023), ggtools (Li, 2020) and GRACE-filter (Rietbroek, n.a.). General tools for spheric harmonic analysis are also available, such as SHTools (Wieczorek, M. A., & Meschede, M., 2018). SHBundle provides MATLAB-tools for spheric harmonic synthesis and spheric harmonic analysis. The earliest version of the code were developed in 1994 while the latest version with upgrades can be found dated 2018. GRAMAT provides a similar MATLAB-based tools for processing GRACE spherical harmonics data to obtain spatiotemporal global mass variations. The GRAMAT

toolbox includes Gaussian smoothening filter to remove North-South stripes, spherical harmonic analysis and synthesis routines, leakage effect reduction routines, harmonic analysis of times series over regions, and uncertainty analysis of GRACE estimates (Feng, 2019). SHADE provides a matlab-based toolbox for the empirical de-correlation of GRACE monthly spherical harmonics (Piretzidis, D., & Sideris, M. G., 2018). Gravity-toolkit is a python-based package meant to handle GRACE L2 data products. Its functionalities include visualization of GRACE and GRACE-FO L2 data products, and the estimation of GRACE and GRACE-FO L2 data product errors. Gg-tools too contain similar tools for signal correction and for conversion of GRACE L2 products to L3. GRACE-filter provides tool for filtering of GRACE L2 product using DDK filter based on Kusche et al. (2009).

Statement of need

A MATLAB code bundle already exists called SHbundle developed by Sneew et al. (2021) and distributed under the GNU license. The code bundle can be freely used and modified by anyone giving proper credit to the original developers. However, MATLAB being a proprietary software may have some limitations in terms of accessibility. Brief description of impact of SHBundle package here

On the other hand, a strong community of programmers also exists for Python, an open-source programming language. In this contribution, we have translated the MATLAB codes from the SHbundle into the Python programming language. In addition to the SHBundle codes, we have further translated the GRACE Data Driven Correction (GDCC) codes from Matlab to Python. GDCC allows the correction of filtered GRACE Level 2 products and restore the signal loss, independent of the catchment size (Vishwakarma et al., 2017). It is hoped the contribution will make GRACE L2 data processing more accessible to a wider audience of programmers. Our python package is titled PySHbundle and the working code can be accessed in GitHub : <https://github.com/mn5hk/pyshbundle>

Mathematic Background

GRACE works on the principal of gravimetric changes. Level 2 GRACE data consists of the spherical harmonic coefficients $C_{l,m}$ and $S_{l,m}$. Gravimetric potential function $V(r, \theta, \lambda)$ can be represented by the spherical harmonic coefficients in the frequency domain with the help of the following relation (Vishwakarma, 2017; Kaula, 1996; Chao & Gross, 1987; Wahr et al., 1998):

$$V(r, \theta, \lambda) = \frac{GM}{r} \sum_{l=0}^{\infty} \left(\frac{a}{r}\right)^l \sum_{m=0}^l \bar{P}_{l,m}(\cos \theta) [C_{l,m} \cos m\lambda + S_{l,m} \sin m\lambda], \quad (1)$$

where G is the Gravitational constant, M represents the total Earth mass, a is the average radius of the Earth, $\bar{P}_{l,m}$ represents the the fully normalized Legendre functions of the first kind, $C_{l,m}$ and $S_{l,m}$ represent the fully normalized spherical harmonic coefficients, and l and m represent the degree and order, respectively.

It should be noted that equation 1 does not deal with the variability of gravimetric potential function over time. However, a major application of the GRACE satellite system is to retrieve the time-variable gravity information. This is achieved by taking the variation of the spherical harmonic coefficients over time. To obtain this, a long-term mean of the monthly values of the spherical harmonic coefficients is removed from the monthly spherical harmonic coefficients obtained from GRACE L2 data. This can be denoted by $\Delta C_{l,m}$ and $\Delta S_{l,m}$. Thus, equation 1 can be modified to obtain the change in gravimetric potential function over time.

Since we are interested in the change in mass in our system, we need to obtain the change in density from the change in gravity potential function. It is further assumed that the

redistribution of the mass of the earth takes place within a thin layer close to the surface of the Earth. Furthermore, this mass redistribution takes place with a deformation. The mass deformation is accounted for by the load Love numbers k_l (Wahr et. al., 1998). As such, equation 1 further resolves to:

$$\Delta\sigma(\theta, \lambda) = \frac{a\rho_{avg}}{3} \sum_{l=0}^{\infty} \sum_{m=0}^l \bar{P}_{l,m}(\cos\theta) \frac{2l+1}{1+k_l} [\Delta C_{l,m} \cos m\lambda + \Delta S_{l,m} \sin m\lambda], \quad (2)$$

Here, $\Delta\sigma(\theta, \lambda)$ represents the change in surface density of the Earth, and ρ_{avg} represents the average density of the Earth (5517 kg / m³). As the mass redistribution on Earth over a monthly time scale is dominated by the hydrological processes, the density change $\Delta\sigma(\theta, \lambda)$ relates to the Equivalent Water Height (EWH) by: $\Delta\sigma(\theta, \lambda) = EWH(\theta, \lambda) \cdot \rho_{water}$. Thus, equation 2 can be rewritten in terms of EWH as:

$$EWH(\theta, \lambda) = \frac{a\rho_{avg}}{3\rho_{water}} \sum_{l,m} \bar{P}_{l,m}(\cos\theta) \frac{2l+1}{1+k_l} [\Delta C_{l,m} \cos m\lambda + \Delta S_{l,m} \sin m\lambda], \quad (3)$$

Thus, we can obtain the hydrological parameter EWH from GRACE Level 2 data using equation 3. The accuracy and precision of the EWH computed depends upon the accuracy and precision of the $\Delta C_{l,m}$ and $\Delta S_{l,m}$, obtained from GRACE. However, these GRACE products are both noisy and coarse in resolution (Wahr et. al., 1998). A tradeoff exists between the noise and resolution of the spherical harmonic products. To capture the spherical harmonic products at a higher spatial resolution, their values at higher degree and order needs to be used. However, noise increases with the increase in degree and order, making the computed EWH also noisy. Similarly, if the spherical harmonics are truncated at a lower degree and order, the noise in the computed EWH decreases, however, the spatial resolution of the obtained EWH also reduces.

To improve the signal-to-noise ratio of the obtained EWH, various filtering techniques have been used. An ideal filter retains all of the signal while filtering out all of the noise. A popular filter used for GRACE applications is the Gaussian filter. The weights, w , for the Gaussian spatial averaging is given by:

$$\omega(\psi) = \frac{\beta}{2\pi} \frac{\exp[-\beta(1 - \cos\psi)]}{1 - \exp^{-2\beta}}, \quad (4)$$

where, $\beta = \frac{\ln(2)}{(1 - \cos(\frac{r_{fil}}{a}))}$. Here, r_{fil} is the averaging radius of the filter. Thus, the Gaussian filter obtained in the spectral domain is written as (Wahr et. al., 1998):

$$\bar{\sigma}(\theta, \lambda) = \frac{2a\rho_{avg}\pi}{3} \sum_{l,m} W_l \bar{P}_{l,m}(\cos\theta) \frac{2l+1}{1+k_l} [\Delta C_{l,m} \cos m\lambda + \Delta S_{l,m} \sin m\lambda], \quad (5)$$

Equation 5 is similar to equation 3, but for an additional multiplication factor, W_l , defined as $W_l = \int_0^\pi w(\psi) P_l(\cos\psi) \sin\psi d\psi$ and $P_l = \frac{\bar{P}_l}{\sqrt{2l+1}}$. Equation 5 defines a Gaussian filter that decays with only degree. However, for our GRACE spherical harmonics, the decay occurs with the location as well as with the degrees and orders. Thus, equation 5 is further generalized as (Wahr et. al., 1998; Devaraju, 2015):

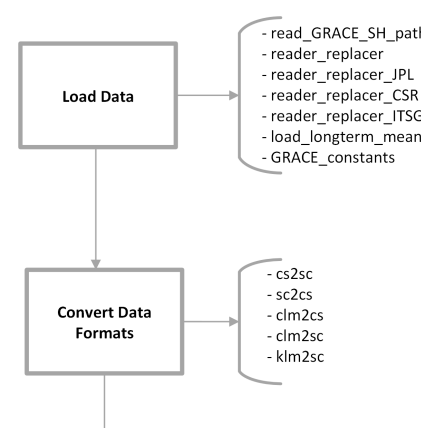
$$\bar{\sigma}(\theta, \lambda) = \frac{a\rho_{avg}}{12\pi} \sum_{l,m} \sum_{n,k} W_{lm}^{nk} \bar{P}_{l,m}(\cos\theta) \frac{2l+1}{1+k_l} [\Delta C_{l,m} \cos m\lambda + \Delta S_{l,m} \sin m\lambda], \quad (6)$$

where W_{lm}^{nk} represents the spectral weight in its general form. Equation 6 is the final result we obtain after spectral harmonic analysis and application of Gaussian filter. More details on the mathematical description presented in this section can be referred to in Vishwakarma (2017).

Methodology

In this contribution, tools to implement the spherical harmonic analysis and filtering application has been developed in the python programming language. In addition, complementary analytical tools such as spherical harmonic synthesis and GRACE data driven correction have also been included. To achieve this, we have implemented the preexisting matlab codes SHbundle into the python programming language. More details on the SHbundle package may be referred to at Sneew et al. (2021). In addition, GRACE Data Driven Corrections algorithm (Vishwakarma et al., 2017) has also been translated from matlab to python. The naming of the modules and the workflow between the modules has been preserved as much as possible in the PySHbundle python implementation. This is to ensure smooth communication between user communities of the two packages and/or the two different programming language communities. Further, our code has been tested using the SHbundle implementation results for validation.

Implementation



A schematic diagram of the code workflow is presented in the Fig 01.

Fig 01: Schematic Diagram of the Code Workflow

The module codes can be categorized into four categories: load data, convert data formats, core functionality and auxillary codes. The load data codes can read data from either of the JPL, ITSG or CSR GRACE data centers. The codes further performs the necessary preprocessing, including the conversion of data formats, replacement of some Legendre coefficients as well as removing the longterm mean. The convert data format codes can convert L2 GRACE Spherical Harmonics data from one format to another. These codes are invoked by the load data codes for data format conversion. Further, these codes can be independently invoked as well by the user for their needs.

The core functionalities of the module are the gsha, gshs, gddc, tws_cal, and basin_avg codes. GSHA module inputs the GRACE L2 spherical harmonic coefficients and performs the

GRACE Spherical Harmonics Analysis (GSHA) algorithm. The algorithm converts the input L2 spherical harmonic coefficients into gridded values at the user-desired grid resolution. An inverse module is also provided, called the gshs.py module. This module performs the GRACE Spherical Harmonics Synthesis (GSHS) algorithm. The algorithm converts the gridded TWSA values into the GRACE L2 spherical harmonics coefficients. In addition to the translation of the SHbundle matlab package, this contribution further includes the GRACE Data Driven Correction function, detailed and first coded in matlab in Vishwakarma et al. (2017). The implementation is done via the gddc module. More details on the gddc implementation can be referred to in the paper cited above. tws_cal computes the total water storage values at each grid from the GRACE L2 spherical harmonics by first applying the Gaussian filter on the L2 data products, and then calling the gshs module. basin_avg module computes the L4 GRACE basin average timeseries product, for any basin input by the user as a GIS shapefile.

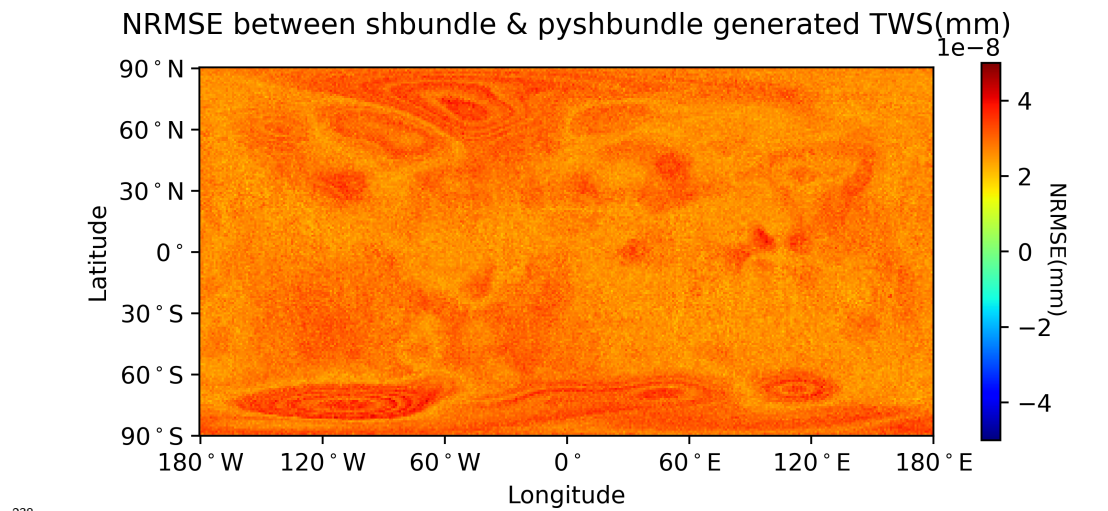
The rest of the codes are bundled together as auxiliary codes in Fig 01. An important part of the GSHS algorithm implementation is the implementation of the PLM algorithm. The PLM algorithm inputs degree, order and co-latitude values and computes the Legendre functions. The plm.py module can also provide the first and second derivatives of the Legendre functions. The implementation of the integrals of the Legendre functions is also done; this is available through the iplm.py module. IPLM inputs the degree, order and co-latitude, and returns the integrated Legendre functions.

Some important modules for the spherical harmonic synthesis step are normalklm, eigengrav, and ispec. normalklm.py returns the hydrostatic equilibrium ellipsoid for the earth surface based on Lambeck (1988) "Geophysical Geodesy", p.18. eigengrav.py provides the isotropic spectral transfer to obtain the equivalent water thickness (m). Lastly, ispec.py inputs the sine and cosine coefficients and returns the field function F.

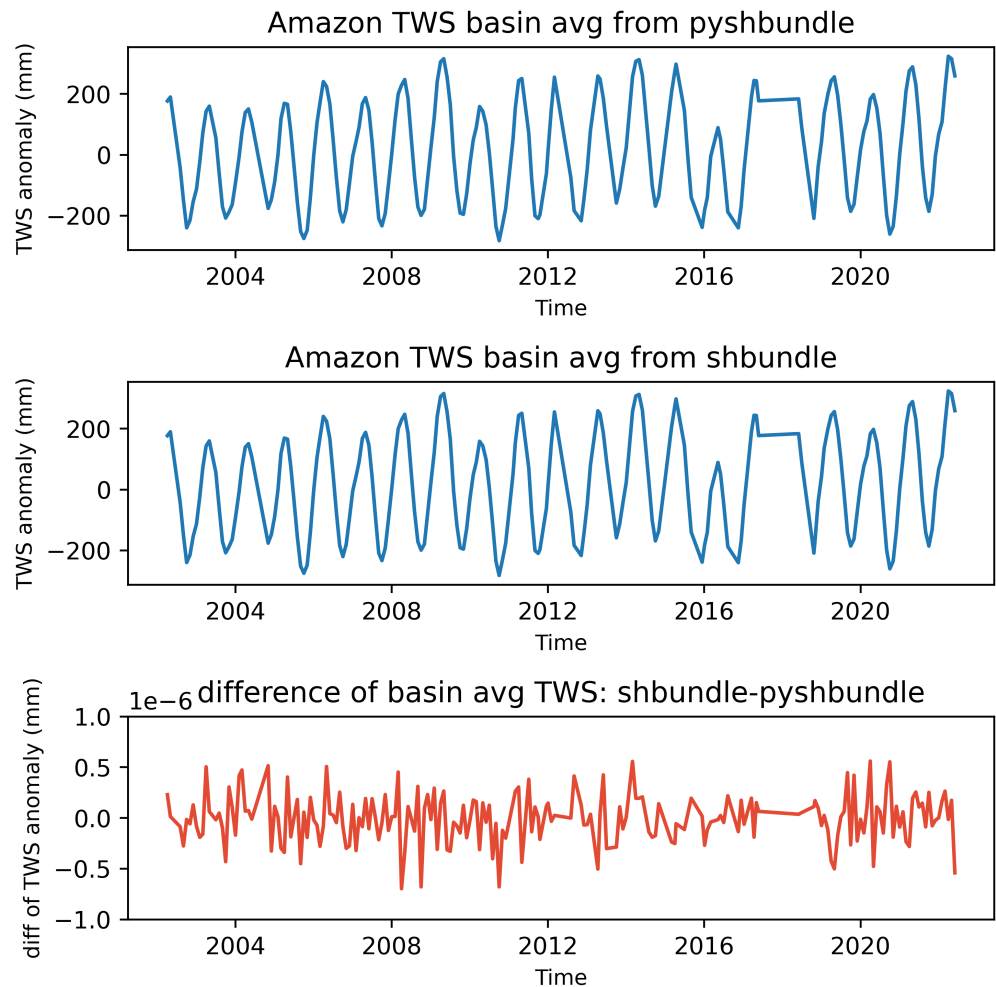
The Global Spherical Harmonic Analysis code depends upon neumann along with the IPLM and sc2cs. neumann.py returns the weights and nodes for Neumann's numerical integration scheme on the sphere. The gshs.py code provides options for spherical harmonic synthesis to compute the sine and cosine components of the Legendre function. These include least squares, weighted least squares, approximate quadrature, first neumann method, second neumann method and block mean values. The neumann.py code is required for the implementation of the first neumann method and second neumann method.

Validation

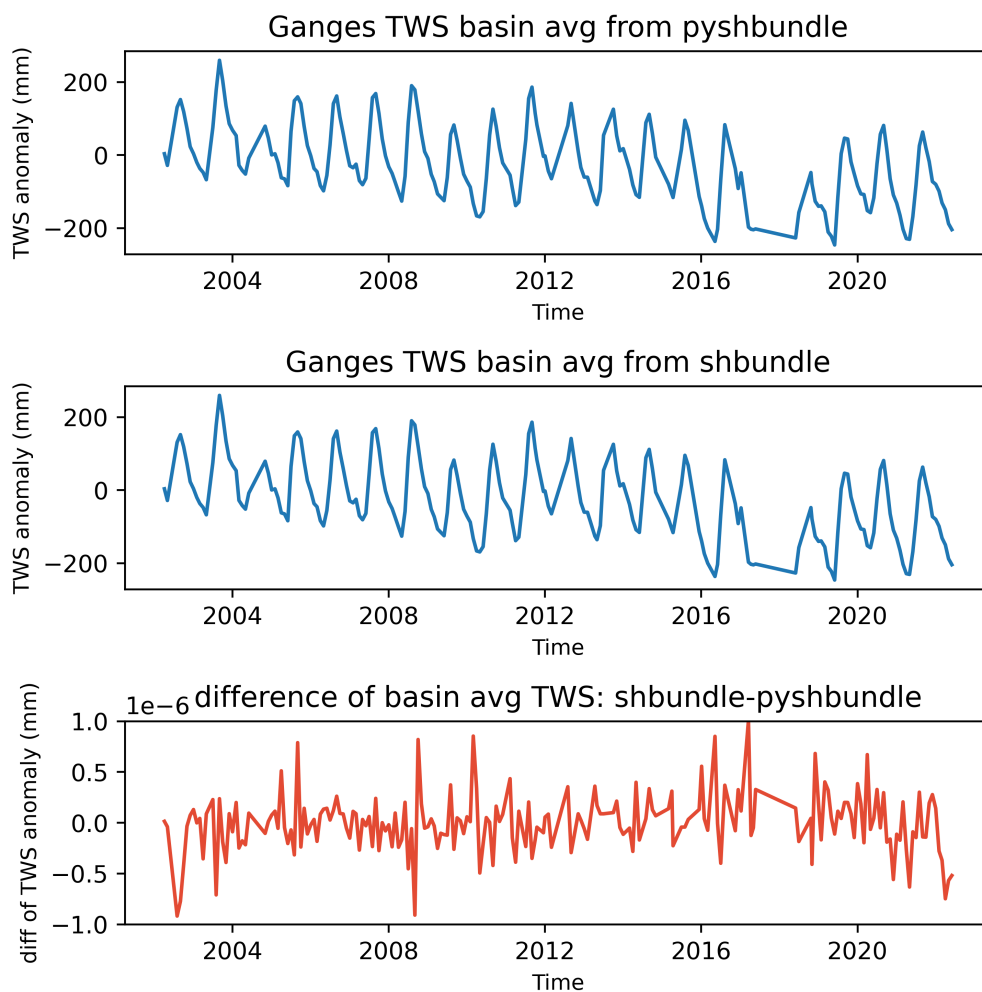
The results of the PySHbundle TWS computation has been validated with respect to TWS computation using SHbundle and presented in Fig 02. The NRMSE values are in the order of $e-8$. Timeseries plots for the Amazon and the Ganges basins have been plotted in Fig 03 and Fig 04, respectively. In both the cases, the order of magnitude of the signal is $e2$, while the error is in the order of $e-6$. Additionally, water budget closure timeseries for the world is provided in Fig 05. The magnitude of difference between the errors and the signal is of the order $e-4$. As such, the errors are likely computational artifacts; and of very small order that can be neglected. Thus, the python package PySHbundle is deemed to give the desired performance in the processing of GRACE L2 Spherical Harmonics to obtain L3 TWS anomalies over land grids.



238
239 Fig 02: RMSE and NRMSE of TWS computation for PySHbundle with respect to SHbundle
240 results.

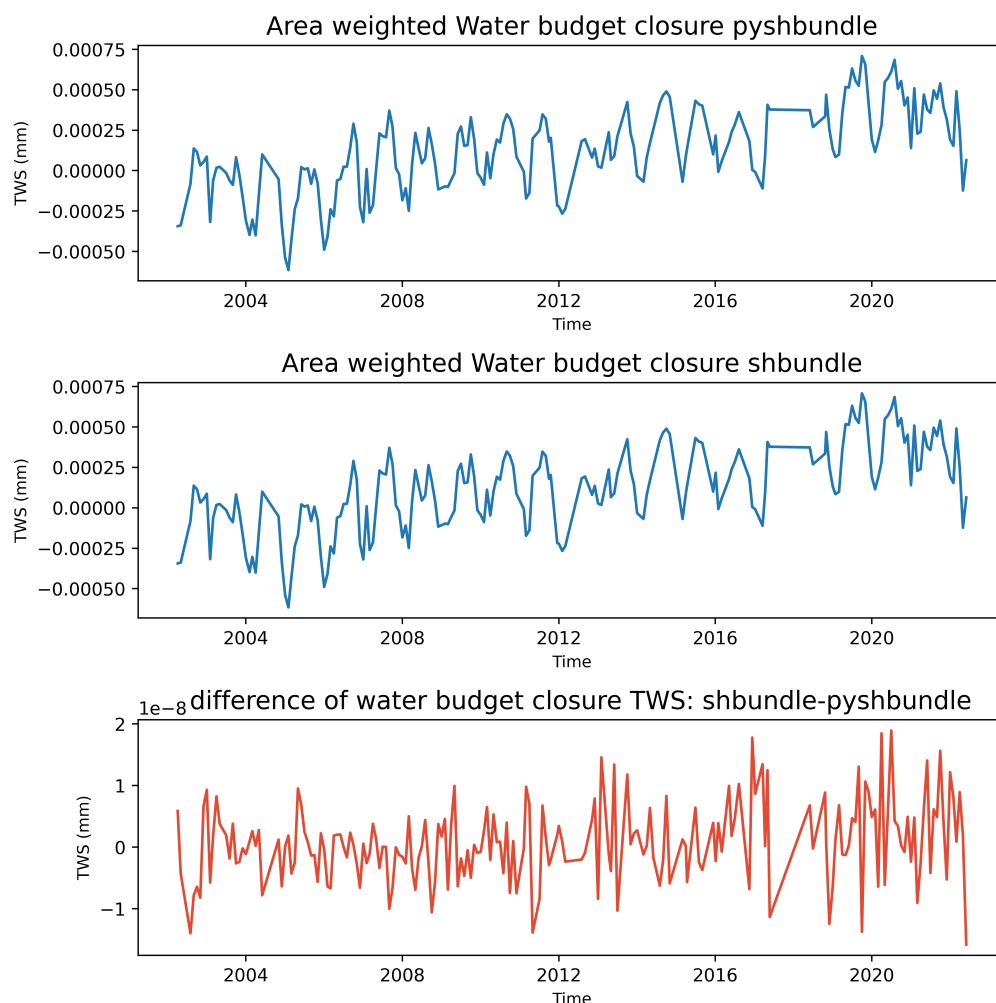


241
242 Fig 03: Timeseries plot of TWS signal from pyshbundle, shbundle and error signal for the
243 Amazon basin



244

245 Fig 04: Timeseries plot of TWS signal from pyshbundle, shbundle and error signal for the
246 Ganges basin



247
248 Fig 05: Water budget closure timeseries plot of TWS signal from pyshbundle, shbundle and
249 error signal

250 References

- 251 Antoni, M. (2022). A review of different mascon approaches for regional gravity field modelling
252 since 1968. *History of Geo-and Space Sciences*, 13(2), 205–217.
- 253 Humphrey, V., Rodell, M., & Eicker, A. (2023). Using satellite-based terrestrial water storage
254 data: A review. *Surveys in Geophysics*, 1–29.
- 255 Nico Sneeuw and Matthias Weigelt and Markus Antony and Matthias Roth and Balaji Devaraju.
256 (2021). *SHBUNDLE 2021*. <http://www.gis.uni-stuttgart.de/research/projects/Bundles/>.
- 257 Wahr, J., Molenaar, M., & Bryan, F. (1998). Time variability of the earth's gravity field:
258 Hydrological and oceanic effects and their possible detection using GRACE. *Journal of*
259 *Geophysical Research: Solid Earth*, 103(B12), 30205–30229.