Geodesy and Cartography
ISSN 2029-6991 / eISSN 2029-7009
2022 Volume 48 Issue 4: 243–247
https://doi.org/10.3846/gac.2022.15171

UDC 528.8

PCA-SOM OF GRACE-FO TOTAL WATER STORAGE FOR GLOBAL CLIMATE DECISIONS

Omid MEMARIAN SORKHABI[1*], Iman KURDPOUR[2]

[1]Department of Geomatics Engineering, Faculty of Civil Engineering and Transportation, University of Isfahan, Isfahan, Iran
[2]School of Surveying and Geospatial Engineering, University of Tehran, Tehran, Iran

Received 21 June 2021; accepted 12 November 2022

Abstract. The gravity recovery and climate experiment (GRACE) and GRACE-Follow on (FO) data provide valuable information about dynamic total water storage (TWS). The complexity of the computational process and the influence of various parameters on TWS changes are complicated in their interpretation. Principal component analysis (PCA) has been used to identify key components to amplify signals and reduce noise in observations. For this purpose, in this research, the Self-organizing map algorithm (SOM) has been used to cluster TWS in 4 categories. The results show that the western regions of Greenland and part of Antarctica are in the critical cluster and have a TWS rate of about –0.2 m/year, which indicates the melting of ice in these regions. The advantage of PCA-SOM is the easy interpretation of TWS, which reduces the impact of seasonal parameters, observation noise and measurement error, and facilitates global policy decisions in the face of climate change.

Keywords: GRACE-FO, SOM, TWS, GRACE, climate.

Introduction

The GRACE mission was launched on March 17, 2002 (Liu et al., 2020). In more than 15 years, GRACE has provided groundbreaking observations of the world that have significantly contributed to our understanding of large-scale changes in polar ice, soil moisture, TWS and groundwater (Sasgen et al., 2020; Sorkhabi et al., 2022a). The GRACE-FO was launched on May 22, 2018, and its main mission goal is to continue tracking mass changes, especially water-related issues. One of the parameters obtained from GRACE observations is TWS that makes it possible to study water displacement in Earth's surface throughout the globe, which is of great importance in environmental and climate studies (Landerer et al., 2020).

Xu et al. (2019) are investigated spatio-temporal changes in China's groundwater storage from GRACE satellites and potential drivers. Six major TWS change regions have been identified, including negative trends in northwest China, northern China and the southeastern Tibetan plateau, positive trends in western China, northeastern China, and southern China. Jensen et al. (2020) are examined the assessment TWS at the 10-year coupled model intercomparison project phase 5 (CMIP5) by the global reconstruction of GRACE satellite data. Using GRACE reconstruction such as TWS, it has been shown that this type of data can be used to assess the proficiency of the decade-long prediction experiments available from various models of the Earth system. Khaki (2020) investigated the efficiency of GRACE TWS into hydrological models; the results showed that GRACE gaps could be filled by hydrological models. Sun et al. (2019) tested the combination of physical modeling and deep learning for a fusion of GRACE satellite data. The results show that the convolutional neural network models significantly improve GRACE TWS compliance, with an average country correlation coefficient of 0.94 and an efficient Nash-Sutcliff improvement of 14%.

TWS research is need of intelligent clustering to determine the water displacement behavior of global different regions and to conduct analyzes and studies based on it. Due to a large amount of data and the complexity of the results, it is necessary to use a tool such as PCA-SOM to simplify the output to make the correct interpretation. SOM is a kind of artificial neural network that has found many applications in recent years in the fields of engineering to medicine, biology and economics (Acevedo-Acosta et al., 2021; Sorkhabi et al., 2022b; Sorkhabi & Milani,

*Corresponding author. E-mail: *omidmemaryan@gmail.com*

2022). The SOM uses a competitive learning method for training and has been developed based on specific features of the human brain (Tan et al., 2019). The novelty of this research is the use of PCA-SOM to cluster areas with similar behavior. These Outputs help to analyze the overall water displacement behavior of the region with the same behaviors. In this study, GRACE-FO data were used to estimate TWS and then PCA-SOM was utilized for clustering.

1. Methods

1.1. Gravity recovery and climate experiment follow on (GRACE-FO)

The GRACE-FO project, a satellite mission designed to determine the Earth's gravitational field. The GRACE-FO mission was launched on May 22, 2018, with a SpaceX Falcon 9 rocket. GRACE-FO is a gravitational satellite and continuation of the GRACE mission and provides TWS information by measuring surface density changes that are mainly due to water displacement (Landerer et al., 2020). Principles of operation GRACE-FO satellite is a measure of the distance changes caused by the gravity of the front satellite with a laser. The TWS observed by GRACE-FO consists of the following formula.

$$\mathrm{TWS} = \mathrm{Pe} - \mathrm{Ev} - \mathrm{Sr}, \tag{1}$$

where Pe is perception, Ev is evaporation and Sr is surface runoff (Jiang et al., 2014). GRACE-FO observations are used in hydrology and water resources management.

1.2. Principal component analysis (PCA)

PCA is simply a way to extract important variables (in the form of components) from a large set of variables in data. PCA extracts a low-dimensional set of features from a high-dimensional set to help record more information with fewer variables (Godah, 2019). In this way, data visualization also becomes more meaningful. PCA is more useful when dealing with data with three or more dimensions. This method is always applied to the covariance or correlation matrix (Bryant et al., 2020); this means that the data must be numerical and standardized. The first function of this method is to determine the factors directly from the correlation matrix without estimating the commonalities.

In this research, the PCA method has been used for easier visual analysis and signal amplification and noise reduction. In this method, to explain the maximum amount of variance variables, their linear composition is estimated. Thus, the first component explains the greatest variance of the variables (Godah, 2019). The second component then explains the maximum amount of variance remaining in the variables after the first component and so on. Another function of PCA is to provide a set of measured variables that converts orthogonal linear combinations with maximum variance (Li et al., 2016).

1.3. Self organization map (SOM)

SOM is an efficient tool for data clustering and can turn nonlinear statistical relationships between input data into simple geometric relationships. The computations of this method are a non-parametric and non-variable regression process in which the specific set regression of model vectors into observable vector space. In an algorithmic form, clusters are organized in a competitive learning process relative to the input variables (Tan et al., 2019).

In this study, using GRACE-FO observations, TWS values were calculated and then clustered by SOM into 4 categories. SOM is an invariant regression relation that maps a set of m $m \in R^n$ vectors to the space of $x \in R^n$ vectors through steps. At each stage of the training, an x-sample vector is randomly selected from the input data set, and the distances between x and all prototype vectors are calculated. Based on minimizing the distance between one sample and other samples, the best matching unit is calculated by Equation 2 (Gholami et al., 2020).

$$\left\| x - m_b \right\| = \min_i \left\{ \left\| x - m_i \right\| \right\}. \tag{2}$$

Therefore, SOM is like a topology map that allows the display, interpretation and arrangement of clustering and can map the space degree of the input data to the two-dimensional network (Tan et al., 2019).

2. Results and discussion

In this study, to achieve TWS, the process of Liu et al. (2020) has been used. Figure 1a shows the TWS trend from 1 June 2018 to 31 May 2020. According to TWS results in western and southern Greenland, parts of Antarctica and Myanmar are trending around –0.2 m/year. Negative TWS trends in areas where there is ice, such as Greenland and Antarctica, indicate ice melting at rates of more than 0.20 m/year. TWS trends are around –0.04 m/year in the Caspian Sea, Eastern Europe, Alaska, South America, Argentina, Western Australia, Eastern China, Uzbekistan, Kyrgyzstan, northern South Africa, parts of Antarctica and northeastern Russia. TWS trends are around 0.14 in the Amazon and south-central Africa.

Figure 1b shows the PCA of TWS with 3 components from 1 June 2018 to 31 May 2020. PCA has been able to amplify the signal and mitigate noise. The results of PCA are similar to the results of the annual trend, except that in some areas, such as Antarctica, the signal is amplified and the annual rate is increased. PCA has also detected a declining TWS signal in Central Africa.

Figure 2a shows architecture of SOM. Inputs are 23 months PCA of TWS. The SOM output is selected as 4 clusters. Figure 2b shows the cluster neighbor distance; the cluster with the dark color showing the long distance and the cluster with light color showing the short distance. According to the results of clusters 3 and 4 are long distances, clusters 1 and 2 are close distances and the rest of the clusters are medium distances. Figure 2c shows the SOM sample hits, with cluster 4 having the highest

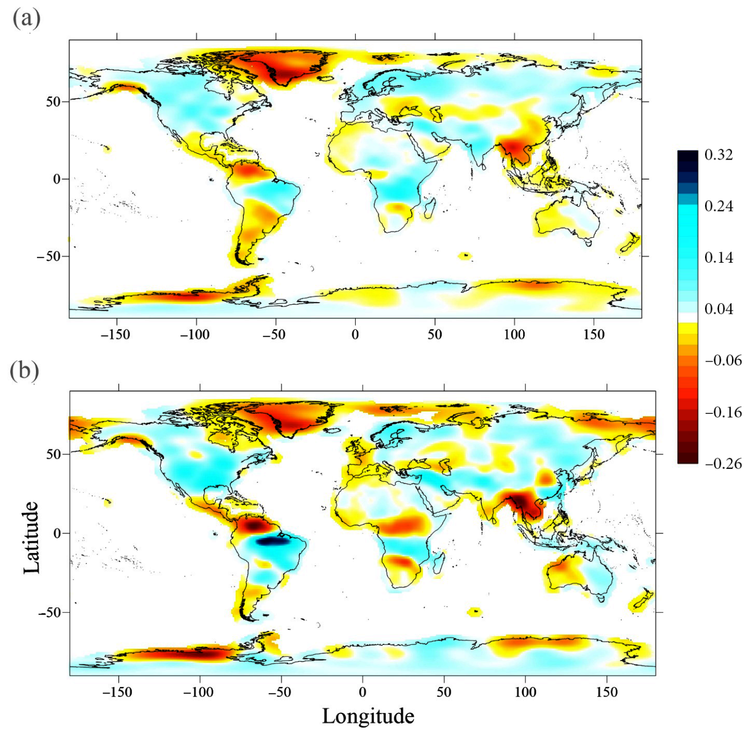

Figure 1. a – TWS annual trend (m/year) from 1 June 2018 to 31 May 2020; b – PCA of TWS (m) with 3 components

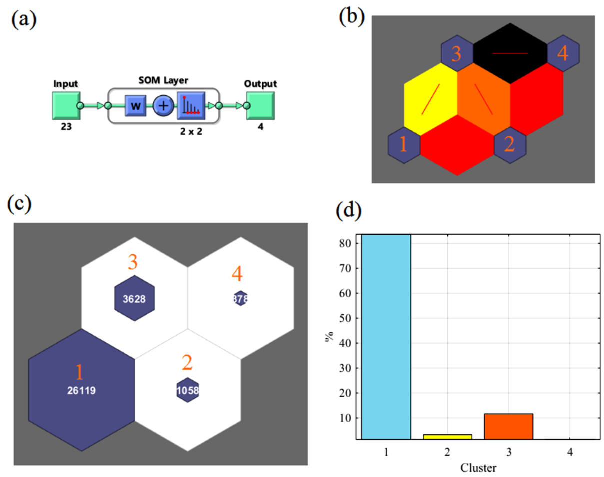

Figure 2. a – Architecture of SOM; b – SOM neighbor distance; c – SOM sample hits; d – Percentage of each cluster

number and cluster 3 having the lowest number. Figure 2d shows percentage of each cluster that the first cluster has the highest percentage. Table 1 shows the statistical characteristics of the results.

Table 1. Statistical characteristics of the results

Parameter	Trend (m/year)	PCA (m)
Minimum	–0.204	–0.263
Maximum	0.141	0.318
Mean	–0.001	–0.001
Standard deviation	0.036	0.051

Figure 3 shows the SOM cluster. According to the SOM clustering, the cluster 4 areas are marked in red, which are west of Greenland and part of Antarctica, which

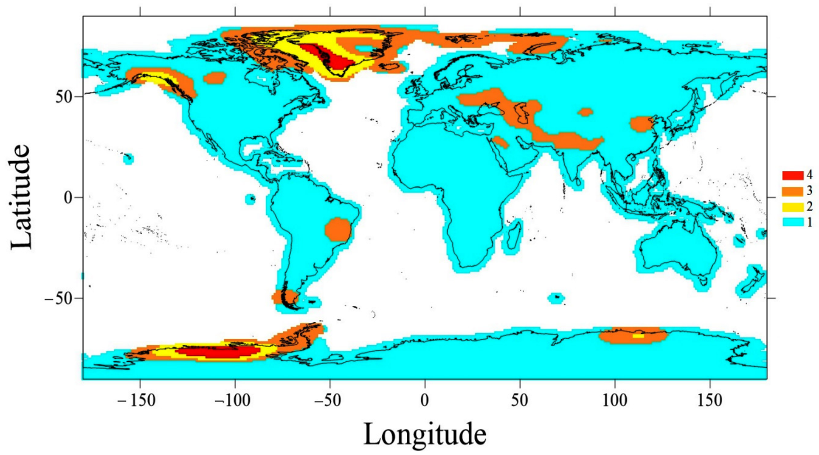

Figure 3. SOM cluster

indicates the loss of TWS or melting ice. Cluster 3 areas are marked in orange as part of Alaska, Greenland, Antarctica, North Pole, Caspian Sea, Turkey, Pakistan, North India, East China, East Africa and South Africa, and the TWS loss is secondary. Clusters 2 are marked in yellow as icy areas found in western Greenland, parts of Antarctica, and parts of Alaska, indicating TWS loss or melting ice. Cluster 2 is not much different from cluster 4, but it is less important. Cluster 1 is marked in blue, which represents more than 80% of all clusters, in which the TWS cluster has not changed much and is fixed. Since GRACE-FO data was used for two years, seasonal effects, noise, measurement error were reduced and the main and identical water displacement behavior of the regions was identified. PCA-SOM clustering provides an easy interpretation of regions that can identify regions with similar characteristics.

Figure 4 shows SOM Cluster for an important region. Western Greenland and parts of Antarctica have the most critical patterns. These areas are primarily sensitive to other regions due to ice melting in the GRACE-FO data. The pattern identified by SOM easily identifies insensitive areas and can inform macro decision makers about global conditions on a map. Melting ice is particularly important in western Greenland and parts of Antarctica due to rising sea levels in coastal areas. Other critical areas identified in Class 3 require further studies to determine the cause of the decline in TWS due to human or natural intervention.

One of the main applications of PCA is in dimensionality reduction operation. The PCA, as its name implies, can identify key components and help analyze a series of more valuable features rather than all the features. The PCA extracts those features that are more valuable. The basic premise in PCA is a linear relationship if nonlinear data is possible. The disadvantage of the SOM method is the weight of the neurons and the number of clusters inputs. The advantages of SOM-PCA are as follows:

- A powerful way to reduce the size of the data without losing a lot of information.
- It is a simple method and has a wide application in various fields of science.
- The principal components are always perpendicular to each other, so the problem of the variables correlation in this method does not matter.
- The PCA-SOM method can tolerate up to 25% of missing data.
- The PCA-SOM method provides a less dimensional view than larger dimensional data. It reveals the pattern of multidimensional data in two dimensions.

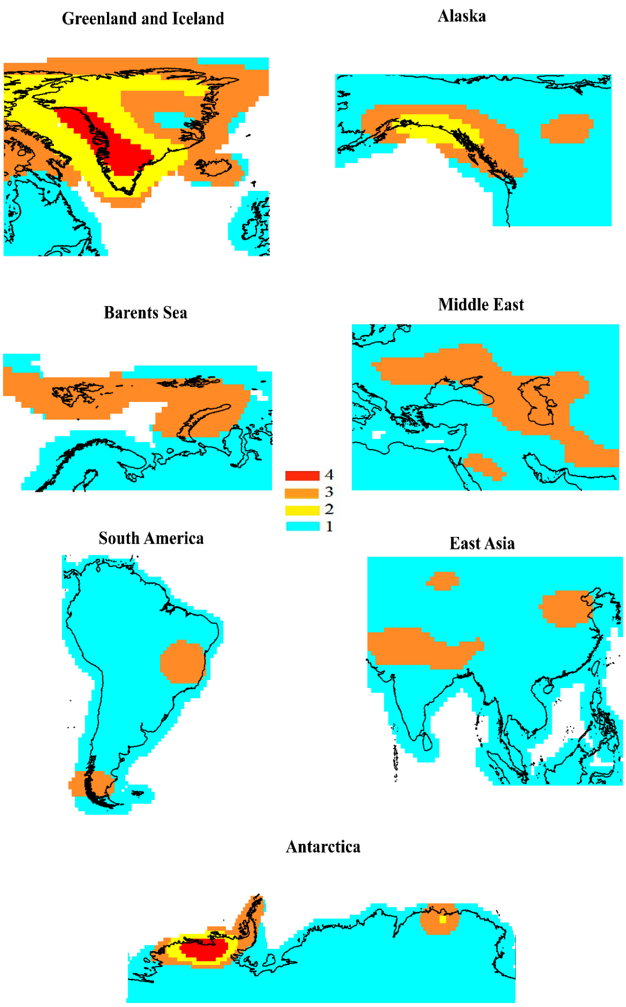

Figure 4. SOM cluster for important region

Conclusions

In this study, the biennial clustering of GRACE-FO TWS with PCA-SOM is investigated. West Greenland and Antarctica are in the TWS critical category at rates of about –0.2 cm/year. In terms of climate change, studying multi-year data can reveal the TWS behavior of areas like Greenland. The results show that the complexity of interpreting outputs is facilitated by SOM. One of the interesting results of TWS is the Caspian Sea, which is highly consistent with research based on sea-level measurements (Medvedev et al., 2019; Memarian Sorkhabi et al., 2021). PCA has been able to amplify the signal in some areas of Antarctica and also reveal the declining rate of TWS in Central Africa. The interpretation ease of complex observation outputs provides the tools to easily make global and regional decisions and planning. The PCA-SOM method has been able to identify areas where there have been few or constant changes to the TWS (which is more than 80%) reducing this concern. Critical areas are also easily detected with SOM. Further studies can be performed to predict changes in TWS with back propagations of artificial neural networks and deep learning.

References

Acevedo-Acosta, J. D., Martínez-López, A., Morales-Acoltzi, T., Albáñez-Lucero, M., & Verdugo-Díaz, G. (2021). Self-organization maps (SOM) in the definition of a "transfer function" for a diatoms-based climate proxy. *Climate Dynamics*, *56*, 423–437. https://doi.org/10.1007/s00382-020-05482-1

Bryant, R. J., Oxley, J., Young, G. J., Lane, J. A., Metcalfe, C., Davis, M., Turner, E. L., Martin, R. M., Goepel, J. R., Varma, M., Griffiths, D. F., Grigor, K., Mayer, N., Warren, A. Y., Bhattarai, S., Dormer, J., Mason, M., Staffurth, J., Walsh, E., Rosario, D. J., Catto, J. W. F., Neal, D. E., Donovan, J. L., Hamdy, F. C., & ProtecT Study Group. (2020). The ProtecT trial: Analysis of the patient cohort, baseline risk stratification and disease progression. *BJU International*, *125*(4), 506–514. https://doi.org/10.1111/bju.14987

Gholami, V., Khaleghi, M. R., & Salimi, E. T. (2020). Groundwater quality modeling using self-organizing map (SOM) and geographic information system (GIS) on the Caspian southern coasts. *Journal of Mountain Science*, *17*(7), 1724–1734. https://doi.org/10.1007/s11629-019-5483-y

Godah, W. (2019). IGiK–TVGMF: A MATLAB package for computing and analysing temporal variations of gravity/mass functionals from GRACE satellite based global geopotential models. *Computers* & *Geosciences*, *123*, 47–58. https://doi.org/10.1016/j.cageo.2018.11.008

Jensen, L., Eicker, A., Stacke, T., & Dobslaw, H. (2020). Predictive skill assessment for land water storage in CMIP5 decadal hindcasts by a global reconstruction of GRACE satellite data. *Journal of Climate*, *33*(21), 9497–9509. https://doi.org/10.1175/JCLI-D-20-0042.1

Jiang, D., Huang, S., & Han, D. (2014). Monitoring and modeling terrestrial ecosystems' response to climate change. *Advances in Meteorology*, *2014*, 429349. https://doi.org/10.1155/2014/429349

Khaki, M. (2020). Efficient assimilation of GRACE TWS into hydrological models. In *Satellite remote sensing in hydrological data assimilation* (pp. 51–74). Springer, Cham. https://doi.org/10.1007/978-3-030-37375-7_6

Landerer, F. W., Flechtner, F. M., Save, H., Webb, F. H., Bandikova, T., Bertiger, W. I., Bettadpur, S. V., Byun, S. H., Dahle, C., Dobslaw, H., Fahnestock, E., Harvey, N., Kang, Z., Kruizinga, G. L. H., Loomis, B. D., McCullough, C., Murböck, M., Nagel, P., Paik, M., Pie, N., Poole, S., Strekalov, D., Tamisiea, M. E., Wang, F., Watkins, M. M., Wen, H.-Y., Wiese, D. N., & Yuan, D.-N. (2020). Extending the global mass change data record: GRACE Follow-On instrument and science data performance. *Geophysical Research Letters*, *47*(12), e2020GL088306. https://doi.org/10.1029/2020GL088306

Li, J., Wang, S., & Zhou, F. (2016). Time series analysis of long-term terrestrial water storage over Canada from GRACE Satellites using principal component analysis. *Canadian Journal of Remote Sensing*, *42*(3), 161–170. https://doi.org/10.1080/07038992.2016.1166042

Liu, X., Feng, X., Ciais, P., Fu, B., Hu, B., & Sun, Z. (2020). GRACE satellite-based drought index indicating increased impact of drought over major basins in China during 2002–2017. *Agricultural and Forest Meteorology*, *291*, 108057. https://doi.org/10.1016/j.agrformet.2020.108057

Medvedev, I. P., Kulikov, E. A., Fine, I. V., & Kulikov, A. E. (2019). Numerical modeling of sea level oscillations in the Caspian Sea. *Russian Meteorology and Hydrology*, *44*(8), 529–539. https://doi.org/10.3103/S1068373919080041

Memarian Sorkhabi, O., Asgari, J., & Amiri Simkooei, A. (2021). Analysis of Greenland mass changes based on GRACE four-dimensional wavelet decomposition. *Remote Sensing Letters*, *12*(5), 499–509. https://doi.org/10.1080/2150704X.2021.1903608

Sasgen, I., Wouters, B., Gardner, A. S., King, M. D., Tedesco, M., Landerer, F. W., Dahle, C., Save, H., & Fettweis, X. (2020). Return to rapid ice loss in Greenland and record loss in 2019 detected by the GRACE-FO satellites. *Communications Earth* & *Environment*, *1*, 8. https://doi.org/10.1038/s43247-020-0010-1

Sorkhabi, O. M., & Milani, M. (2022). Deep Learning of Ionosphere Single-Layer Model and Tomography. *Geomagnetism and Aeronomy*, *62*(4), 474-481. https://doi.org/10.1134/S0016793222040120

Sorkhabi, O. M., Kurdpour, I., & Sarteshnizi, R. E. (2022a). Land subsidence and groundwater storage investigation with multi sensor and extended Kalman filter. *Groundwater for Sustainable Development*, *19*, 100859. https://doi.org/10.1016/j.gsd.2022.100859

Sorkhabi, O. M., Milani, M., & Seyed Alizadeh, S. M. (2022b). Investigating the efficiency of deep learning methods in estimating GPS geodetic velocity. *Earth and Space Science*, *9*(10), e2021EA002202. https://doi.org/10.1029/2021EA002202

Sun, A. Y., Scanlon, B. R., Zhang, Z., Walling, D., Bhanja, S. N., Mukherjee, A., & Zhong, Z. (2019). Combining physically based modeling and deep learning for fusing GRACE satellite data: Can we learn from mismatch? *Water Resources Research*, *55*(2), 1179–1195. https://doi.org/10.1029/2018WR023333

Tan, Y. C., Ke, K. Y., & Fang, H. T. (2019). Application of integrated SOM-and MOGA-SVM-based algorithms to forecast groundwater level in Choushui River Alluvial fan, Taiwan. In *AGUFM 2019*, H53Q-2052.

Xu, L., Chen, N., Zhang, X., & Chen, Z. (2019). Spatiotemporal changes in China's terrestrial water storage from GRACE satellites and its possible drivers. *Journal of Geophysical Research: Atmospheres*, *124*(22), 11976–11993. https://doi.org/10.1029/2019JD031147