Supporting Information for

Understanding Water Level Changes in the Great Lakes by an ICA-Based Merging of Multi-Mission Altimetry Measurements

Wei Chen [1], C. K. Shum [2], Ehsan Forootan [3], Wei Feng [4,*], Min Zhong [4], Yuanyuan Jia [2], Wenhao Li [5], Junyi Guo [2], Changqing Wang [6], Quanguo Li [1] and Lei Liang [7]

[1] College of Resource Environment and Tourism, Hubei University of Arts and Science, Xiangyang 441053, China
[2] Division of Geodetic Science, School of Earth Sciences, Ohio State University, Columbus, OH 43210, USA
[3] Department of Planning, Aalborg University, 9220 Aalborg, Denmark
[4] School of Geospatial Engineering and Science, Sun Yat-Sen University, Zhuhai 519082, China
[5] School of Geomatics Science and Technology, Nanjing Tech University, Nanjing 21180, China
[6] State Key Laboratory of Geodesy and Earth's Dynamics, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China
[7] School of Geographic Information and Tourism, Chuzhou University, Chuzhou 239099, China
* Correspondence: fengwei@mail.sysu.edu.cn

The supporting information includes: 1) the description of CICA in Text S1; 2) tips for the reconstructed process in Text S2; 3) the cumulative percentage of the total variance by EOF, CEOF, and CICA shown in figure S1; 4) spatial patterns and temporal components decomposed by EOF, CEOF, and CICA are shown in figure S2, S3, and S4, respectively; 5) relationships between the iteration numbers, the correlation coefficients, and RMSE values shown in figure S5; 6) the cumulative percentage of the total variance by MSSA shown figure S6; 7) the modes kept by MSSA shown in figure S7.

Text S1 The description of CICA.

In Appendix A, we presented the necessary formulation for the CICA. To make clearer about CICA, the detailed description is provided here to better understand and implement CICA. Similar to how CEOF extends EOF, CICA is developed to deal with non-stationarity of geophysical time series from ICA technique (Forootan et al., 2018). For CICA, it has been done by generating a new dataset that contains the observed time series in its real part. The out-of-phase patterns of these time series are estimated by applying a Hilbert transformation (Horel, 1984) and are considered to be the imaginary part of the new dataset. The derived complex dataset is used in the ICA procedure to extract the dominant independent space and time amplitudes and associated phase propagations (Forootan et al., 2012). Here, the tensorial approach of joint approximate diagonalization of eigenmatrices (Cardoso and Souloumiac 1993, 1996) is applied in the CICA to estimate the independent modes. Different from the standard EOF or CEOF, these three properties: non-Gaussianity, non-whiteness, and no circularity are all considered into CICA to approximate statistical independence. An advantage of CICA over the already existing EOF and CEOF techniques is that it incorporates higher-order statistical information, which likely reduces clustering of different physical modes within single extracted mathematical' modes (Forootan et al., 2018). In the section 3.2 of this manuscript, the first 12 modes decomposed by CICA can be reconstructed better than less modes from EOF/CEOF techniques, which denotes that more inherent information about lake level are included in the non-stationary and independent components from CICA.

Text S2 Some notes on the reconstruction.

In this reconstructed process, the first several modes are reserved when their cumulative percentage of the total variance is up to 90%. In general, 90% is the commonly used criterion in the reconstructed techniques based on EOF, ICA, and their extended type (Smith et al., 1996; Alvera-Azcárate et al., 2005; Erhardt et al., 2011). In figure S1, the first 5, 9, and 12 modes are kept to reconstruct the lake surface height in Lake Erie by EOF, CEOF, and CICA, respectively. Accordingly, their spatial patterns and temporal components from EOF, CEOF, and CICA are also shown in figure S2, figure S3, and figure S4, respectively. For the modes decomposed by EOF and CEOF, the 1st PC (principal components) probably presents lake surface changes when compared with the results of water level stations in Lake Erie. For CICA's modes, the 8th CIPC (complexed independent components) looks like the trend although contaminated with variation; the 9th CIPC shows low-quality results from 1981 to 1992 which probably presents the lake level change by altimetry; the 6th CIPC possibly shows that there is a multidecadal period which needs to be verified further. The other components decomposed by EOF, CEOF, and CICA need to study further to distinguish the true signal. In summary, more components from CICA present more inherent information about lake surface height which probably lay a good foundation for the later reconstruction.

We can obtain the 1st reconstructed result by the first several modes kept above. In order to improve the reconstructed results, the iterative process should be further implemented. The detailed procedure is presented in section 2.6. Here, the number of iterations is defined by the maximum correlation coefficients and minimum RMSE values which can be calculated by the reconstructed result and water level stations' data (see figure S5).

Reference

1. Forootan, E.; Kusche, J.; Talpe, M.; Shum, C.K.; Schmidt, M. Developing a complex independent component analysis (CICA) technique to extract non-stationary patterns from geophysical time series. *Surveys in Geophysics* **2018**, 39, 435-465.
2. Horel J D. Complex principal component analysis: Theory and examples. *Journal of Applied Meteorology and Climatology*, **1984**, 23(12): 1660-1673.
3. Forootan E, Kusche J. Separation of global time-variable gravity signals into maximally independent components. *Journal of Geodesy*, **2012**, 86(7): 477-497.
4. Cardoso J F, Souloumiac A. Blind beamforming for non-Gaussian signals. *Proceedings of the IEEE*, **1993**, 140(6): 362-370.
5. Cardoso J F, Souloumiac A. Jacobi angles for simultaneous diagonalization. *SIAM journal on matrix analysis and applications*, **1996**, 17(1): 161-164.
6. Smith T M, Reynolds R W, Livezey R E, et al. Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *Journal of Climate*, **1996**, 9(6): 1403-1420.
7. Alvera-Azcárate A, Barth A, Rixen M, et al. Reconstruction of incomplete oceanographic data sets using empirical orthogonal functions: application to the Adriatic Sea surface temperature. *Ocean Modelling*, **2005**, 9(4): 325-346.
8. Erhardt E B, Rachakonda S, Bedrick E J, et al. Comparison of multi‐subject ICA methods for analysis of fMRI data. *Human brain mapping*, **2011**, 32(12): 2075-2095.

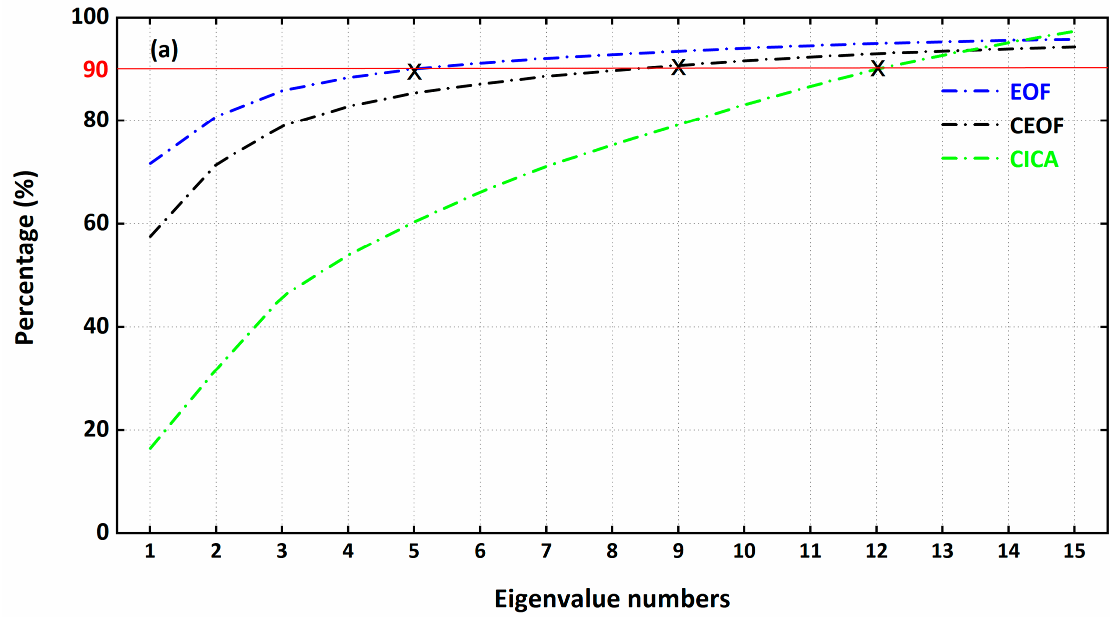

Figure S1. The cumulative percentage of total variance accounted by the first several modes, is presented above. Up to 90% of total variance accumulated from the first 5, 9, and 12 modes which are decomposed by EOF, CEOF, and CICA respectively, can usually represent the original signal of lake surface height in Lake Erie. According to the selected modes, the spatial patterns and temporal components from EOF, CEOF, and CICA are shown in **Figure S2, Figure S3, and Figure S4**, respectively.

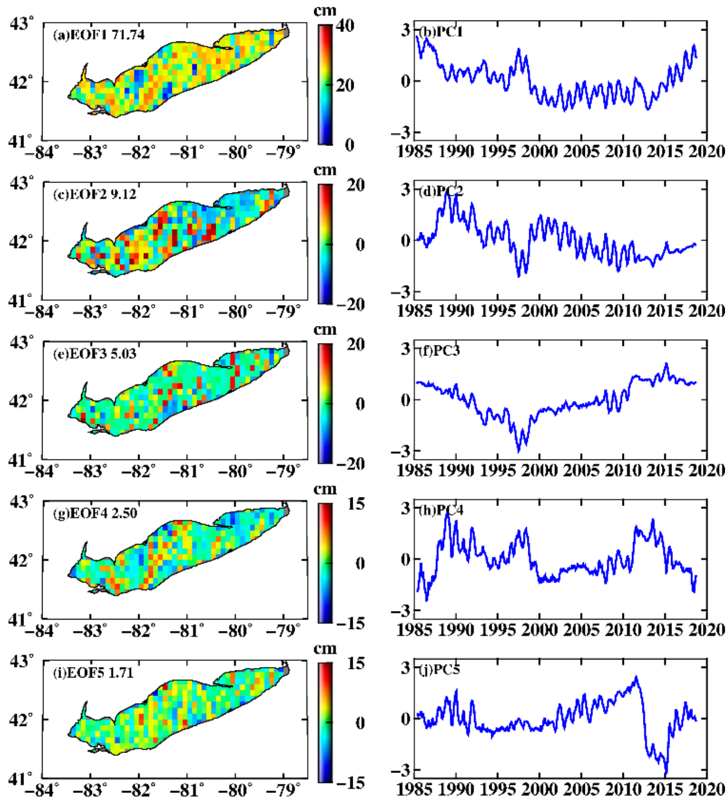

Figure S2. Spatial patterns (EOF) and principal temporal components (PC) are from EOF modes 1 to 5, respectively.

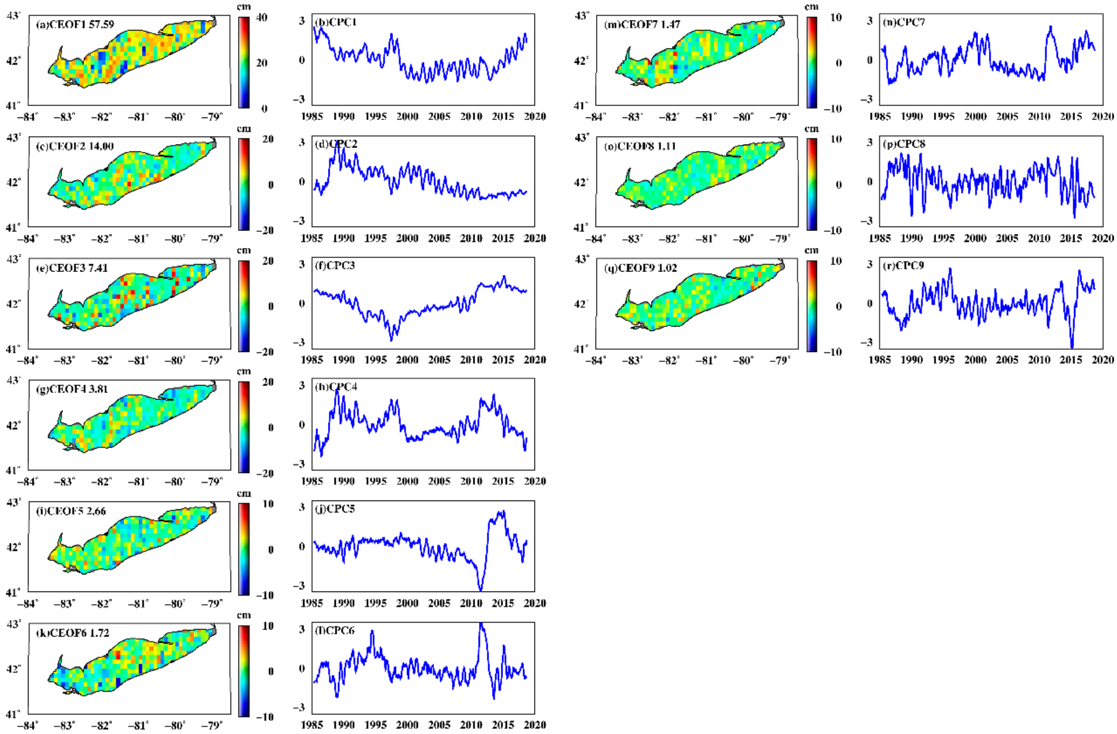

Figure S3. Spatial patterns (CEOF) and principal temporal components (CPC) are from CEOF modes 1 to 9, respectively.

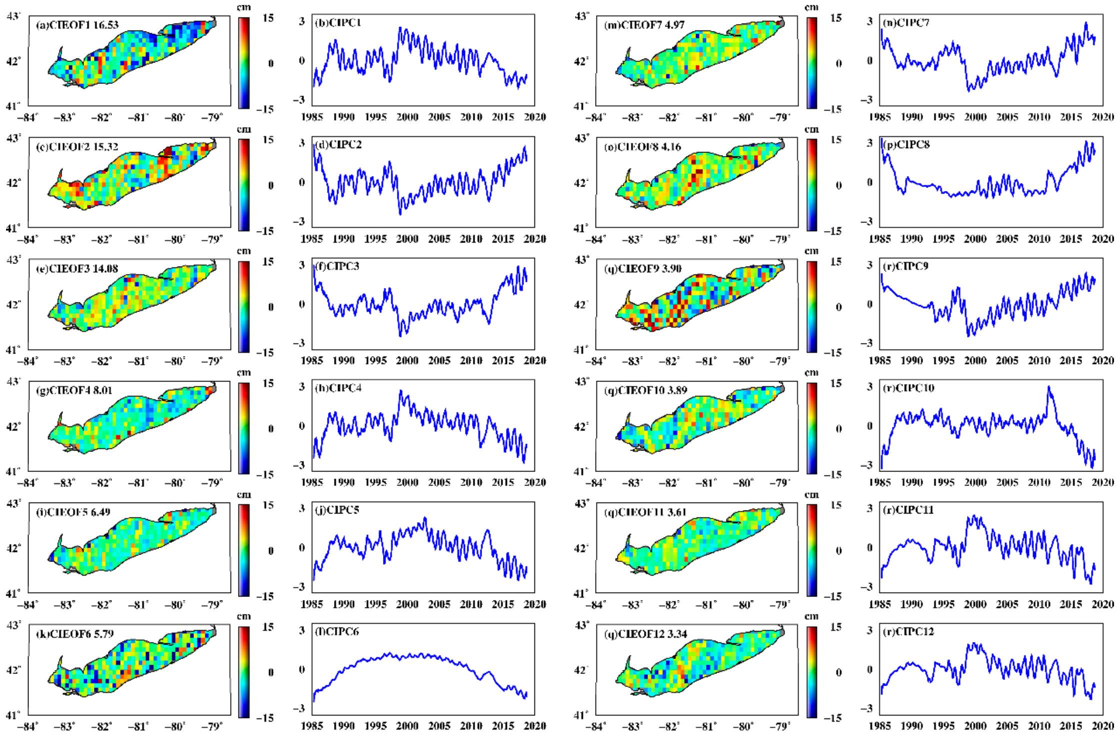

Figure S4. Spatial patterns (CIEOF) and principal temporal components (CIPC) are from CICA modes 1 to 12, respectively.

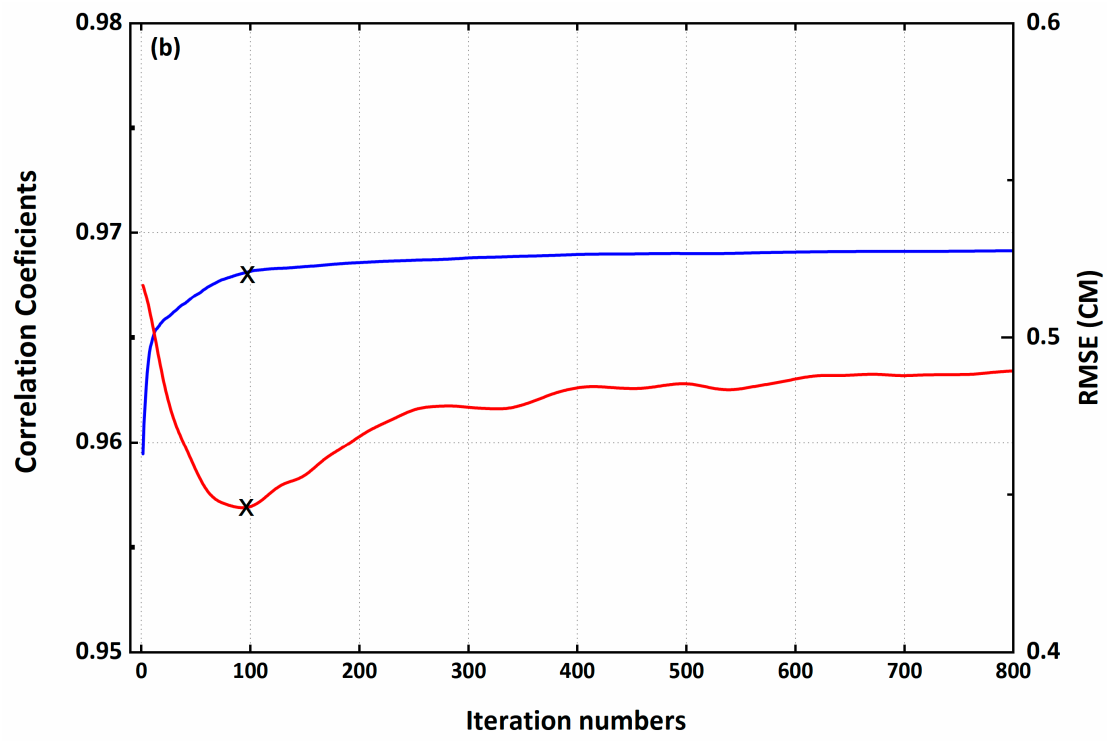

Figure S5. The iteration is introduced to decompose and reconstruct lake surface height by CICA. Here, iteration numbers can be defined by the maximum correlation coefficients and minimum RMSE values.

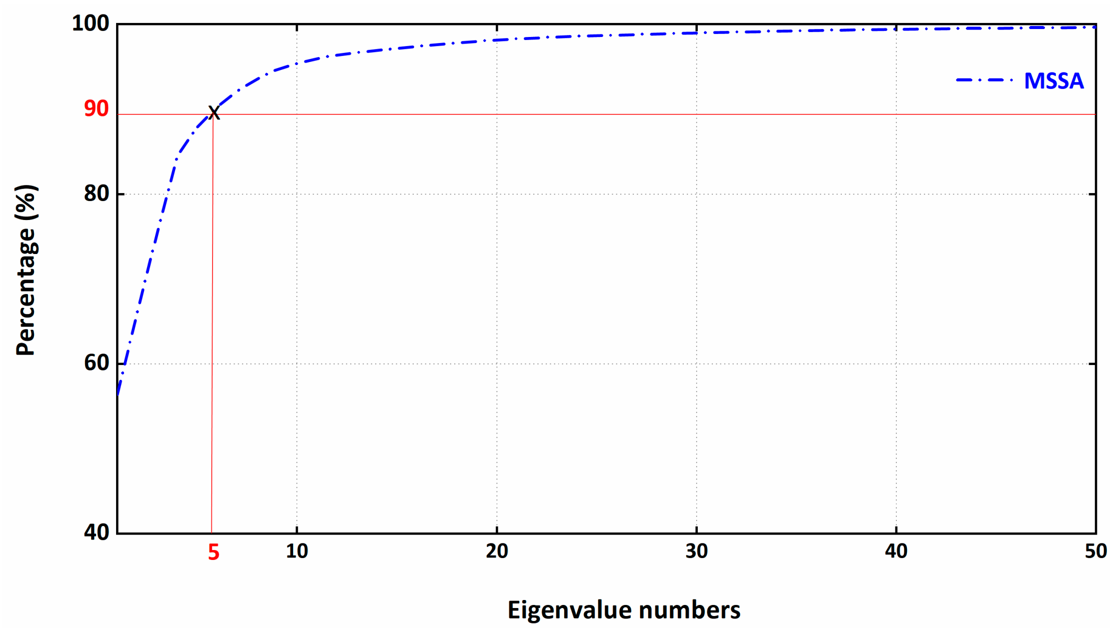

Figure S6. The cumulative percentage of total variance in MSSA, the first 5 components occupy > 90% total variance.

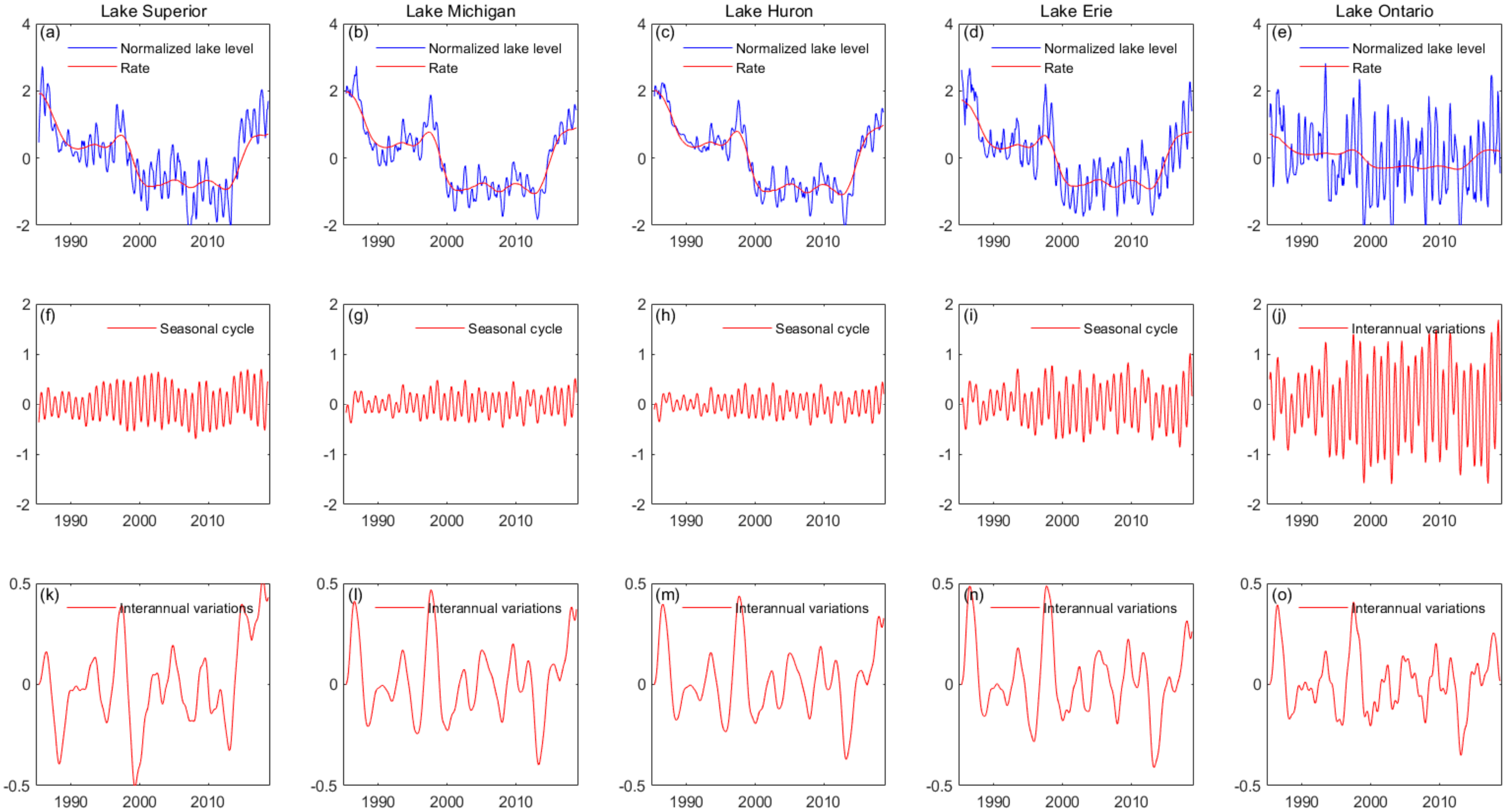

Figure S7. The rates (the first component in MSSA), seasonal cycle (including the third, fourth, fifth, and sixth components in MSSA), and interannual variations (the second component in MSSA) can be obtained by MSSA over the Great Lakes.