Detectability of groundwater storage change within the Great Lakes Water Basin using GRACE

J. Huang,[1] J. Halpenny,[1] W. van der Wal,[2] C. Klatt,[1] T. S. James,[3] and A. Rivera[4]

Received 20 September 2011; revised 30 May 2012; accepted 5 June 2012; published 1 August 2012.

[1] Groundwater is a primary hydrological reservoir of the Great Lakes Water Basin (GLB), which is an important region to both Canada and US in terms of culture, society and economy. Due to insufficient observations, there is a knowledge gap about groundwater storage variation and its interaction with the Great Lakes. The objective of this study is to examine the detectability of the groundwater storage change within the GLB using the monthly models from the Gravity Recovery And Climate Experiment (GRACE) satellite mission, auxiliary soil moisture, snow and lake (SMSL) data, and predictions from glacial isostatic adjustment (GIA) models. A two-step filtering method is developed to optimize the extraction of GRACE signal. A two dimensional basin window weight function is also introduced to reduce ringing artifacts caused by the band-limited GRACE models in estimating the water storage change within the GLB. The groundwater storage (GWS) as deviation from a reference mean storage is estimated for the period of 2002 to 2009. The average GWS of the GLB clearly show an annual cycle with an amplitude range from 27 to 91 mm in water thickness equivalent (WTE), and a phase range of about two months. The estimated phases of GWS variations have a half year shift with respect to the phase of SMSL water storage variations which show peaks in March and April. The least squares estimation gives a GWS loss trend of from 2.3 to 9.3 km^3/yr within the GLB for the period of study. This wide range of the GRACE GWS results is caused largely by the differences of soil moisture and snow storage from different land surface models (LSMs), and to a lesser extent by the GRACE commission and omission errors, and the GIA model error.

Citation: Huang, J., J. Halpenny, W. van der Wal, C. Klatt, T. S. James, and A. Rivera (2012), Detectability of groundwater storage change within the Great Lakes Water Basin using GRACE, *J. Geophys. Res.*, *117*, B08401, doi:10.1029/2011JB008876.

1. Introduction

[2] Groundwater storage is a primary reservoir of the hydrological cycle in the Great Lakes Water Basin (GLB). However, groundwater recharge and discharge patterns are not well understood compared with other primary reservoirs such as soil moisture, surface waters, and ice and snow, which are operationally gauged by in situ and remote sensors at various spatiotemporal resolutions for the development and management of water resources. *Grannemann et al.* [2000] suggest that more work needs to be done to define and quantify the interactions between regional groundwater flow and groundwater discharge into the Great Lakes to understand the GLB system, which supports a regional economy that is critical to both Canada and US [see http://www.ec.gc.ca/grandslacs-greatlakes].

[3] In 1995 groundwater supplied drinking water for about 21.8 million people on the southern side of the GLB. The total withdrawal of groundwater was 2.1 km^3/yr which included 1.3 km^3/yr for drinking water [*Solley et al.*, 1998]. The Great Lakes Commission reported the same level of groundwater withdrawal for 2004 in its annual water use report (see http://www.glc.org/wateruse/). On the Canadian side, all groundwater withdrawal for drinking purposes comes from aquifers located within the GLB in the province of Ontario. Here, there are 3.6 million people relying on groundwater with a total withdrawal of 0.2 km^3/yr.

[4] The total estimated groundwater storage in the GLB is more than 4,000 km^3 [*Grannemann et al.*, 2000] in contrast to the total lake water storage of about 23,000 km^3. With economic and societal development, the demand for water continues to increase in the region. Groundwater will remain a primary source to meet the demand, and should be managed in such a way as to sustain its supply. *Rivera* [2008] suggests the management of groundwater resources should be supported by strong science-based programs with scientific

[1]Geodetic Survey Division, Canada Centre for Remote Sensing, Natural Resources Canada, Ottawa, Ontario, Canada.
[2]Faculty of Aerospace Engineering, Delft University of Technology, Delft, Netherlands.
[3]Pacific Geoscience Centre, Geological Survey of Canada, Natural Resources Canada, Sidney, British Columbia, Canada.
[4]Geological Survey of Canada, Natural Resources Canada, Quebec City, Quebec, Canada.

Corresponding author: J. Huang, Geodetic Survey Division, Canada Centre for Remote Sensing, Natural Resources Canada, 615 Booth St., Ottawa, ON K1A 0E9, Canada. (jianhuan@nrcan.gc.ca)

Table 1. Gauge Stations Used for the Lake Water Storage Estimation

Station	Location
10220	Rossport, ON
10750	Michipicoten, ON
11375	Parry Sound, ON
11690	Tobermory, ON
11940	Point Edward, ON
12065	Kingsville, ON
12865	Port Colborne, ON
13320	Toronto, ON
13988	Kingston, ON
11070	Thessalon, ON
9087096	Port Inland, MI
9099044	Ontonagon, MI
9099090	Grand Marais, MN
9075059	Harrisville, MI
9075014	Harbor Beach, MI
9087023	Ludington, MI
9087057	Milwaukee, WI
9063020	Buffalo, NY
9063063	Cleveland, OH
9052030	Oswego, NY
9052058	Rochester, NY
9087044	Calumet Harbor, IL

knowledge of groundwater availability, vulnerability and sustainability and should be incorporated into water laws and regulations to sustain groundwater supplies in urban and rural areas. For the GLB, the knowledge of groundwater recharge and discharge, fluxes and interactions with surface water is essential to support the management of groundwater. However, obtaining this knowledge is extremely challenging because ground infrastructure and techniques are inadequate to monitor the large-scale dynamics of groundwater in the GLB.

[5] The joint US-German Gravity Recovery And Climate Experiment (GRACE) mission, launched on March 17, 2002, is a new type of remote sensor that can infer changes in terrestrial Total Water mass Storage (TWS) from its monthly gravity solutions [*Tapley et al.*, 2004]. GRACE cannot measure absolute TWS. GRACE TWS is given as deviation from an unknown mean TWS for a selected reference period. TWS changes are then inferred from GRACE TWS deviations (in short GRACE TWS). Estimates of the GRACE TWS have a spatial resolution of better than 500 km and an accuracy of about 2 cm in Water Thickness Equivalent (WTE) [see, e.g., *Chen et al.*, 2009a]. With auxiliary measurements of surface water, soil moisture, and snow and ice, it is possible to detect a groundwater change of 2–3 cm in WTE over a water basin greater than the GRACE spatial resolution of about 300 km. The GRACE solutions have been used to investigate the groundwater storage (GWS) change in many large water basins. Comparisons between the GRACE-inferred groundwater storage variations and the in-situ groundwater observations generally show good agreement in both phase and amplitude [see, e.g., *Yeh et al.*, 2006; *Rodell et al.*, 2007; *Swenson et al.*, 2008; *Strassberg et al.*, 2009; *Henry et al.*, 2011]. A few investigations have successfully applied the GRACE solutions in estimating the trend of groundwater depletion within large water basins in India and USA [see, e.g., *Rodell et al.*, 2009; *Tiwari et al.*, 2009; *Famiglietti et al.*, 2011].

[6] Compared with other major water basins on Earth, the Great Lakes Basin is more complex and challenging in terms of using GRACE data to infer groundwater changes. First, the Great Lakes are the largest system of fresh surface water on Earth, containing roughly 18% of the world's fresh surface water. It discharges about 240 km^3/yr into the St Lawrence river [see http://www.ec.gc.ca/stl/] while being replenished by approximately the same amount from all influxes which include precipitation, river, and groundwater to maintain the lake levels. Accurate estimation of the total lake water storage time series requires demanding and rigorous data compilation, analysis and volume corrections for seasonal temperature change. Secondly, the soil moisture over the land which is two-thirds of the GLB area is an equally dynamic component of surface water. The land surface models have demonstrated success and usefulness in providing soil moisture corrections to the GRACE TWS estimates to separate the groundwater storage signals [e.g., *Rodell et al.*, 2007; *Strassberg et al.*, 2009]. However, it is not trivial to quantify the uncertainty of the soil moisture correction due to lack of sufficient in-situ validation data. Thirdly, the dominant un-corrected geodynamic effect in GRACE data is glacial isostatic adjustment (GIA) in the Great Lakes region. There are a number of GIA models which can provide GIA corrections, but like land surface models, their uncertainty is usually not well-quantified. Since GIA directly affects the estimation of trend in water storage, any error in the GIA models translates into apparent groundwater change. These three factors pose major difficulty for a groundwater study using GRACE in the GLB.

[7] The goal of this study is to examine the GRACE detectability of groundwater storage change within the GLB with supplementary surface water storage (lake water, snow and soil moisture) information and existing GIA models. First, a two-step filtering method is developed to maximize the refinement of GRACE signals. Secondly, a two dimensional window weight function is suggested to improve the estimation of the total water storage within a basin. Thirdly, the GIA effect on the GRACE result is estimated and corrected. Its uncertainty is discussed through the comparison of four GIA models. Fourthly, the soil moisture, snow and lake water storage is estimated using lake level observations and output of land surface models. Finally, total and groundwater storage as deviations from their reference means along with uncertainties are estimated to understand their seasonal and trend patterns and interactions with lakes water storage.

2. Hydrological Data and Land Surface Models

2.1. Water Levels in the Great Lakes

[8] To estimate the groundwater storage, the surface water storage has to be subtracted from the GRACE TWS. The Great Lakes are major surface water reservoirs. Water levels in the Great Lakes are monitored by well-distributed gauge stations which are operated by Canada's Department of Fisheries and Oceans and the National Oceanic and Atmospheric Administration (NOAA) in the USA. The water level data for the period of 2002–2009 at 22 stations shown in Table 1 and Figure 1 have been acquired online from the two agencies to estimate the water storage in the lakes. The lake levels are observed with an accuracy of one centimeter at each tide gauge station. The water storage variation for each

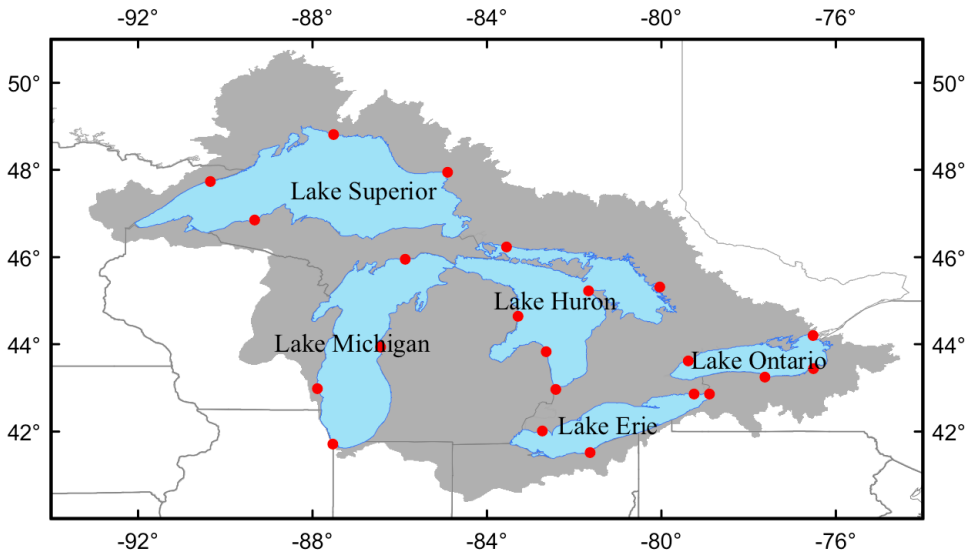

Figure 1. The Great Lakes Water Basin and gauge stations (dots).

lake is computed from a number of stations on both sides of the lakes so that the vertical GIA uplift effect on gauge stations is reduced. Gauge stations close to outlets are excluded to make the storage estimates more realistic.

[9] Lake water volume is affected by water temperature. Thermal expansion and contraction of water needs to be accounted for when the water level is used to estimate the lake water storage. *Meredith*'s [1975] method is used to estimate the thermal expansion and contraction with respect to the lake state at the temperature of 0°C. It provides dimensionless temperature profiles which can be scaled with the lake surface temperature to get temperatures at different depths. With this temperature and the volume for each depth, the expansion of the entire water column can be computed.

[10] The monthly lake surface temperature data are retrieved from NOAA CoastWatch Great Lakes Program (http://coastwatch.glerl.noaa.gov). *Schwab et al.* [1999] describe how these surface temperature data are produced. The mean difference between the buoy temperature and these satellite-derived temperature estimates is less than 0.5°C for all buoys. The root mean square differences range from 1.10 to 1.76°C.

[11] Figure 2a shows the temperature effect on the water levels for the Great Lakes. Lake Superior shows the largest effect which peaks at more than 100 mm while the other four lakes shows a similar magnitude of the effect that is less than 50 mm at peaks. Those peaks happen during summers due to the highest temperature. Evidently, the temperature effect on the water levels is significant enough to be considered.

[12] Figure 2b shows the temperature-corrected water levels variation with respect to the 5-year mean of 2005–2009 for the Great Lakes. Of them, Lake Ontario has the strongest annual variability with a peak-to-peak amplitude of about 1000 mm, while the other four lakes shows a smaller and similar annual variation with a peak-to-peak amplitude of about 500 mm. Lakes Huron and Michigan show the same phase of annual variation which differs from the other three lakes. The average water level variation of all five lakes (thick red line) clearly shows both annual and interannual changes. The latter can be seen by a declining trend of the average water level from 2004 to 2008 and a return from 2008.

2.2. Land Surface Models

[13] The second largest surface water reservoir is moisture contained within the topsoil layer above the groundwater table. The most useful method to determine the soil moisture over a large area is to apply a land surface model system which assimilates in situ and space-based observations into a sophisticated land surface state model [see, e.g., *Georgakakos and Baumer*, 1996; *Moran et al.*, 2004]. One of such systems is NASA's Land Data Assimilation System (LDAS) operated by Goddard Space Flight Center (GSFC) NASA. It has two implementations: North America LDAS (NLDAS) and Global LDAS (GLDAS). Soil moisture and snowfields from LDAS have been widely used in GRACE applications [see, e.g., *Rodell et al.*, 2007; *Chen et al.*, 2009a; *Tiwari et al.*, 2009]. The soil moisture fields from NLDAS have been compared with in-situ observations available in a few regions. For instance, *Strassberg et al.* [2009] suggest a correlation of 0.82 between NOAH model and in situ measurements in the High Plains aquifer. *Fan et al.* [2011] show a correlation of 0.81 between NOAH model and in situ measurements in Illinois.

[14] In this study, soil moisture content over the GLB is retrieved from Global LDAS. GLDAS produces optimal fields of land surface states and fluxes in near-real time by combining satellite- and ground-based observations into four land surface models: CLM, Mosaic (MOS), VIC and NOAH [*Rodell et al.*, 2004]. The monthly NOAH soil moisture states in a 0.25° by 0.25° grid (NOAH025) are used to estimate the water storage in the top 2 m soil layer along with the NOAH snowfield on the surface within the GLB. The maximum mean snow water equivalent is 22 mm in December over the GLB for the period of study. The snow water storage is small compared with the soil moisture storage variation that can reach more than 100 mm in WTE.

[15] To understand the uncertainty associated with the NOAH025 soil moisture and snowfield, five other models have been used for comparisons. Four of them are monthly 1°

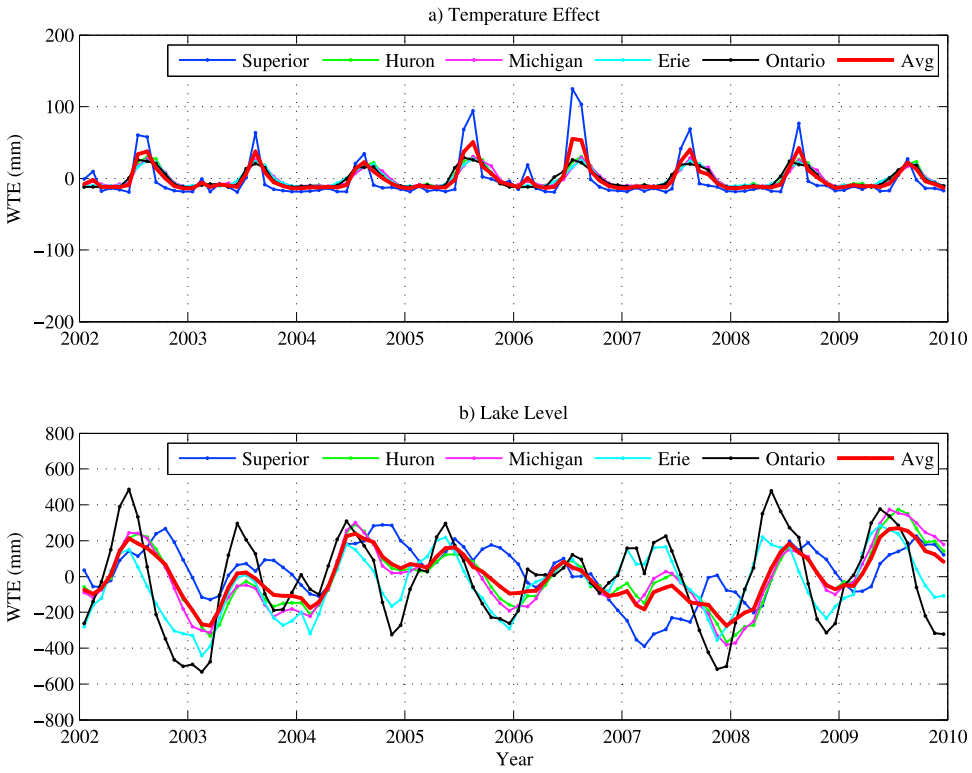

Figure 2. Water level variations in the Great Lakes. (a) The thermal expansion effect on lake water levels. (b) Lake level changes with respect to the mean for the period of 2005–2009.

GLDAS models: CLM10, MOS10, NOAH10 and VIC10. The fifth model is the WaterGAP Global Hydrology Model (WGHM) of version 2.1h in a 0.5° by 0.5° grid [*Döll et al.*, 2003, 2012]. The soil moisture and snow water storage variations averaged over the land part of GLB are shown in Figure 3. These variations are referred to their respective mean fields for the period of 2005–2009. All models show similar annual cycles but considerably different amplitudes. Among them, CLM10 shows the weakest annual amplitude while the MOS10 and WGHM show the strongest. NOAH10 model is close to VIC10 in amplitude with an RMS difference of 24 mm. A mean model averaged from the four 1° GLDAS models has been derived as AVG10. Its standard deviation time series ranges from 3 to 59 mm with a mean of 27 mm for the period of study. The RMS difference is 19 mm between NOAH025 and AVG10, and is 44 mm between NOAH025 and WGHM. If these values are averaged over the whole GLB, they need to be re-scaled by a factor of 0.68 because the land area accounts for about two-thirds of the GLB area. These statistical comparisons somewhat characterize the uncertainty of the present soil moisture models, which is comparable with the GRACE uncertainty.

[16] In addition, we have also included the Climate Prediction Center (CPC) global monthly soil moisture model [*Fan and van den Dool*, 2004]. *Fan et al.* [2011] show a correlation of 0.74 between this model and the in-situ data in Illinois. The CPC model shows a close phase to other models but a relatively weak amplitude. Due to the lack of a snowfield in this model, it will not be used to derive the groundwater storage change in this study.

[17] It is worth to point out that the GLDAS models adopted here have a maximum soil moisture depth of 2 to 3.5 m, while WGHM has mostly 1 m. This implies that the water storage at depths exceeding the soil moisture depth is interpreted as groundwater storage if no additional soil moisture information is available.

3. Refinement of the GRACE Signal

3.1. GRACE Models

[18] The GRACE satellite system, launched in March 2002, consists of two co-orbiting satellites which maintain a distance of about 220 km between each other along a nearly polar orbit at an altitude of about 450 km above the Earth's surface. It detects changes of 1 billionth of the Earth's gravity field with resolution in time and space that approximately corresponds to the monthly changes of 2.4 cm in water thickness equivalent at a spatial resolution of 300–400 km. The gravity changes with time are caused by mass movement within the whole Earth system including atmosphere, liquid and solid parts of the Earth. In the standard of GRACE data processing (http://podaac.jpl.nasa.gov/gravity/grace), gravity effects of secular variations of the low-degree harmonics, solid, ocean and polar tides, atmospheric and non-tidal oceanic variations are combined into the so-called background models. They are removed from GRACE observations before the GRACE gravity models are derived. Thus, the measured gravity changes from the GRACE models are dominantly caused by water mass changes on and below the Earth's surface as well as solid mass changes associated with non-tidal geodynamic processes within the Earth. Accordingly the errors in the background models and measurements are also propagated into the resulting GRACE models.

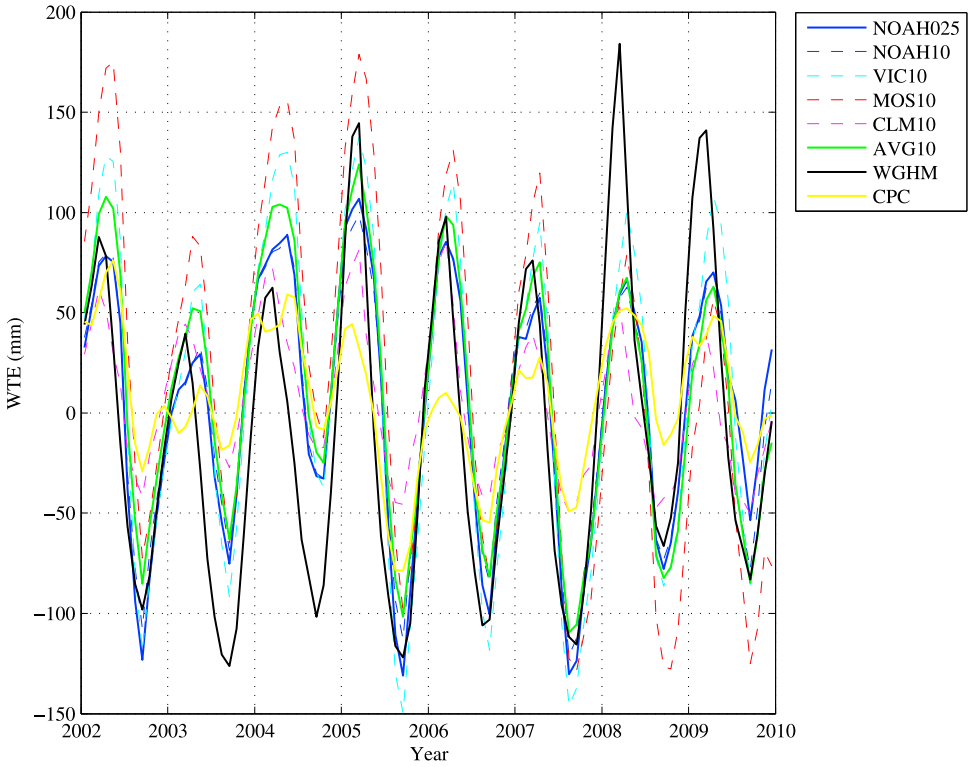

Figure 3. The total soil moisture and snow water storage variations in WTE over the land part of GLB. Note that CPC model does not include the snow contribution.

[19] In physical geodesy, the Earth's gravity field is conventionally expressed as the anomalous gravitational potential with respect to the normal potential defined by a geodetic reference ellipsoid. The temporal change of the Earth's anomalous potential with respect to the long-term mean potential field can be computed from a GRACE spherical harmonic (SH) model of epoch t_i as follows:

$$\Delta T(\phi, \lambda, t_i) = \frac{GM}{R} \sum_{n=0}^{N} \sum_{m=-n}^{n} \Delta \bar{C}_{nm}(t_i) \bar{Y}_{nm}(\phi, \lambda) \qquad (1)$$

where ϕ and λ are latitude and longitude of a computational point, t_i are the epochs of GRACE gravity models, R is the mean radius of the Earth, N is the maximum SH degree of the GRACE models, GM is the geocentric gravitational constant, $\Delta \bar{C}_{nm}$ are time-variable components of the fully normalized SH coefficients, and $\bar{Y}_{nm}(\phi, \lambda)$ are normalized surface SHs, defined as follows:

$$\bar{Y}_{nm}(\phi, \lambda) = \bar{P}_{n|m|}(\sin \phi) \begin{cases} \cos m\lambda & m \geq 0 \\ \sin |m|\lambda & m < 0 \end{cases} \qquad (2)$$

where $\bar{P}_{n|m|}$ are the fully normalized associated Legendre functions of the first kind, n and m are the degree and order of SHs, respectively.

[20] The potential change ΔT can be represented by the potential of a mass layer on the Earth's surface with the water thickness equivalent [*Wahr et al.*, 1998]

$$\Delta h(\phi, \lambda, t_i) = \frac{R\rho_a}{3\rho_w} \sum_{n=0}^{N} \frac{2n+1}{1+k_n} \sum_{m=-n}^{n} \Delta \bar{C}_{nm}(t_i) \bar{Y}_{nm}(\phi, \lambda) \qquad (3)$$

where ρ_w (= 1000 kg/m^3) and ρ_a (= 5517 kg/m^3) are the water and mean Earth densities, respectively, k_n is the load Love number of degree n. Due to the high altitude of the GRACE satellite, it does not have vertical resolution. Considering most water storage is within a range of a few hundred meters below the Earth surface, GRACE TWS can be accurately modeled by a surface mass layer.

[21] *Chao* [2005] proves that on a 2-D spherical shell the inverse solution of the surface density distribution is unique, and further states that this solution applies quite readily in the inversion of time-variable gravity signals observed by the GRACE space mission where the sources largely come from the Earth's surface over a wide range of timescales. Therefore given sufficiently dense data coverage, equation (3) can model the mass layer up to a spatial resolution of interest by choosing the maximum SH degree N high enough.

[22] Due to the limited coverage of the GRACE observations in both time and space, errors from its observation system, errors in the background models, data processing methods and the limited sensitivity of satellite gravimetry, the resulting mass estimates are subject to both commission and omission errors. Among all the commission errors, the stripe-like correlation error is dominant and stretches in the north-south direction globally. It is often many times as strong as the Earth's mass change signal, and is unevenly distributed across SH components of the models. The total commission error increases rapidly with increasing SH degrees and orders. The SH components beyond a certain degree are too noisy to recover any signal and therefore have to be truncated.

[23] This truncation of the GRACE models causes the omission error. The severity of the omission error depends

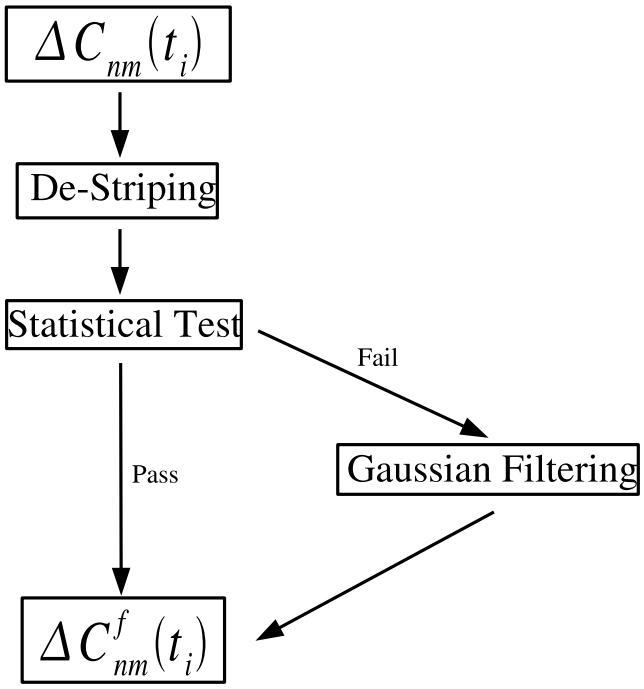

Figure 4. Method for filtering the time-variable GRACE coefficients.

on the maximum truncation degree N and the magnitude of local variability of gravity changes. Current GRACE models contain recoverable gravity change signals up to maximum SH degree 60 that approximately represents a spatial resolution of 350 km at low latitude. In general, the presence of strong local variability leads to a poor representation by the GRACE model.

[24] For this study, the monthly GRACE models of Release 4 by the Center for Space Research (CSR) for the period of 2002–2009 are used. There are a total of 90 monthly models for this period. The 84 unconstrained monthly models of Release 4 by the GeoForschungsZentrum (GFZ) Germany are used for comparisons. Note that GFZ released two sets of monthly models for the period of October 2008 to December 2009 complete to degree and order 120 and 60, respectively. The latter set is used for this study. The secular variations of low-degree harmonic coefficients are restored (http://podaac.jpl.nasa.gov/gravity/grace). The C_{20} values in the GRACE models are replaced by the estimated values from the Satellite Laser Ranging by *Cheng and Ries* [2007].

3.2. A Two-Step Filtering Method for GRACE Models

[25] A number of filtering methods have been suggested to smooth the stripe noise of the GRACE gravity field. *Werth et al.* [2009] provide a comprehensive review of commonly used methods. In particular, six frequently used filter methods are compared, and optimized parameter values are derived for major water basins worldwide. One of these basins is the St. Lawrence Basin (SLB) which includes the GLB as its western sub-basin. They suggest that no optimized parameter values can be achieved for the SLB.

[26] A two-step method is developed in this study. It combines and improves the de-striping method [*Swenson and Wahr*, 2006], the statistical filtering method [*Davis et al.*, 2008] and non-isotropic Gaussian filtering [*Wahr et al.*, 1998; *Han et al.*, 2005] as described in Figure 4.

[27] The time-variable SH coefficients $\Delta\bar{C}_{nm}$ in Figure 4 are the differences between the coefficients of monthly GRACE models $\overline{C}_{nm}$ and the mean coefficients $\bar{C}^{Mean}_{nm}$:

$$\Delta\bar{C}_{nm}(t_i) = \bar{C}_{nm}(t_i) - \bar{C}^{Mean}_{nm} \tag{4}$$

[28] The mean coefficients $\bar{C}^{Mean}_{nm}$ can come from a static mean GRACE gravity model, or the mean of monthly GRACE models weighted by coefficient standard deviations. The latter is used in this study. Like monthly coefficients, the mean coefficients are also contaminated by various errors, in particular, the stripe-like error so that the time-variable components inevitably contain the errors from the mean components when the mean model is subtracted from a monthly model.

[29] The first step is de-striping coefficients of the same order with even and odd degrees in each monthly model using *Swenson and Wahr*'s [2006] method. The de-striping process can cause the loss of signal. In order to minimize the signal loss, a criterion is introduced to decide if a set of even- or odd-degree coefficients of the same order is de-striped in terms of the ratio of the root-mean-squares (RMS) of the time-variable components $\Delta\bar{C}_{nm}$ before and after the de-striping, i.e.,

$$RT = \left(\sum_{n_{e/o}=p/q}^{N} (\Delta\bar{C}_{nm})^2 \Big/ \sum_{n_{e/o}=p/q}^{N} (\Delta\bar{C}^{\mathrm{DS}}_{nm})^2 \right)^{1/2} \tag{5}$$

where $n_{e/o}$ stands for even or odd degree, p/q is a chosen minimum SH even/odd degree for de-striping, $\Delta\bar{C}^{\mathrm{DS}}_{nm}$ are the de-striped coefficients. The minimum degree p/q is experimentally determined so that no significant striping errors are shown within the low degree band below it. *Chambers* [2006] keeps the lower 7×7 portion of the coefficients unchanged when filtering the CSR GRACE monthly models of Release 2, while we keep the lower 11×11 portion unchanged in this study for the Release 4 models. This difference of choices reflects the improvement of the Release 4 models over the Release 2. In addition, coefficients higher than order 50 are also kept unchanged because of too few even or odd degree coefficients available for polynomial fitting.

[30] A noisy and striped field as represented in the numerator of equation (5), has more power than a de-striped field as represented in the denominator of equation (5). The larger the value returned by RT, the more the de-striping. A small RT value means that the set of coefficients is less affected by the stripe-like error, and the de-striping can remove the signal under this situation. The criterion introduced for de-striping is whether RT is greater than a chosen critical value RT_C, i.e. when $RT > RT_C$, the set of de-striped coefficients will be adopted. In the case of $RT_C = 0$, each set of de-striped coefficients will be adopted regardless of its statistical significance. The greater the critical value RT_C,

the less the de-striping. The assumption behind this criterion is that the correlative part of the coefficients should be significantly greater than the signal to be identified as the stripe noise. Ideally, the critical value RT_C should be chosen so that the stripe noise is removed the most while the signal is removed the least.

[31] A critical value RT_C is determined by a realistic simulation. Twelve monthly simulation grids in WTE are created using the NOAH soil moisture and snowfield for year 2007 from GLDAS [*Rodell et al.*, 2004]. A mean grid over the year is removed to form twelve monthly time-variable grids. The nodes without data are approximated by predictions of the de-striped GRACE monthly models. A Gaussian filter with a radius of 450 km is used to create the WTE grids from the GRACE models. The twelve grids are expanded into twelve SH models complete to degree and order 720, then truncated to degree and order 60 for de-striping. The grids from these truncated SH models vary from 34 mm in June to 76 mm in September in the area-weighed RMS with a total RMS of 59 mm over the GLDAS covered global land area. When being restricted to the land part of GLB, they vary from 19 mm in June and 88 mm in September with a total RMS of 67 mm. The de-striped grids using $RT_C = 0$, 1, 2 have total RMSs of 56, 56 and 59 mm in WTE over the GLDAS covered global land area, respectively, in contrast to the true RMS of 59 mm. Over the land part of GLB, the de-striped grids have total RMSs of 57, 57 and 67 mm in WTE in contrast to the true RMS of 67 mm. As expected, a larger RT_C results in a smaller signal loss. With $RT_C = 2$, there is no significant loss of power in the de-striped grids. The total RMS difference between the simulation grids and the de-striped grids with $RT_C = 2$ is 2 mm in WTE over the land part of GLB.

[32] However the larger RT_C also tends to retain more stripe noise. The question is then: how much stripe noise is acceptable for further Gaussian filtering, or removable using the statistical filtering? The GRACE results for $RT_C = 2$ after Gaussian filtering with the radius of 450 km do not show significant stripe-like patterns, especially over the GLB region. It suggests that the choice of $RT_C = 2$ strikes a good balance between keeping the signal and removing the stripe noise.

[33] The second step involves a statistical test. A parameter model for a coefficient time series is postulated as

$$\begin{aligned}\Delta\bar{C}^{DS}_{nm}(t_i) = {} & \Delta\bar{C}^{DS}_{nm}(t_0) + v_{nm}(t_i - t_0) + 0.5a_{nm}(t_i - t_0)^2 \\ & + A_{nm}\cos(2\pi(t_i - t_0) + \phi_A) + B_{nm}\cos(4\pi(t_i - t_0) \\ & + \phi_B) + \delta_{nm}(t_i) \end{aligned} \quad (6)$$

where v_{nm} and a_{nm} stand for velocities and accelerations of the coefficient changes in time by applying scale factors of 1 and 0.5 to these two terms, respectively, A_{nm} and B_{nm} are amplitudes of respective annual and semi-annual variations while ϕ_A and ϕ_B are their phases, and δ_{nm} are un-modeled residuals. It is assumed that physical signals approximately have time behavior according to equation (6). The semi-annual term is used to correct for the inability of the sinusoidal annual term in representing a realistic time series.

[34] Unlike *Davis et al.*'s method [2008], this parameter model is used to test whether the signal in a coefficient time series is statistically significant at a selected confidence level. Different confidence levels have been tried and compared in terms of the strength of remaining stripe patterns in the synthesized gravity change maps. It is shown that the level of 99.9% can ensure the coefficients passing the test do not contain significant stripe errors. When the series passes the test, residuals of all coefficients in the series are then tested against 3 times the RMS of the residuals. The coefficients which pass both tests above are directly used to compute the mass storage estimates without Gaussian filtering. Otherwise they are filtered using the non-isotropic Gaussian filter [*Han et al.*, 2005]. Note that this non-isotropic filter alone can effectively reduce the stripe error when the filter parameters are properly chosen. The extra statistical step is to identify the signal-dominant coefficients and to apply the Gaussian filter only to the rest.

[35] The least squares fitting can be performed with different time windows. In order to determine the best time window, equation (6) was fitted to the monthly coefficients of SH degrees 2 to 60 from the Center for Space Research with time windows of 1 to 5 years for 2004–2008. The weighted cumulative RMS residuals are 1.93×10^{-10}, 2.39×10^{-10}, 2.47×10^{-10}, 2.48×10^{-10} and 2.54×10^{-10} for 1 to 5 years, respectively. They represent a range of 4.4–5.6 mm in WTE and 1.2–1.5 mm in geoid undulation. Even though the 1-year RMS is the smallest, its maximum degree of freedom is 5 which is not redundant enough to yield a reliable least squares solution. Thus a 3-year window is chosen to produce stable analysis results and partially account for the inter-annual change. The time window moves forward one year at a time, and monthly models at the end year are statistically tested at each move. The coefficients passing the tests contribute 75% of the total mass storage estimates on average.

[36] After the two-step filtering, the SH coefficients $\overline{C}_{nm}$ in equation (4) change to

$$\overline{C}^{\mathrm{f}}_{nm}(t_i) = \bar{C}^{Mean}_{nm} + \Delta\bar{C}^{\mathrm{f}}_{nm}(t_i) \quad (7)$$

[37] The equivalent spectral filtering weights for the two-step filter can be written as

$$f_{nm}(t_i) = \frac{\Delta\bar{C}^{\mathrm{f}}_{nm}(t_i)}{\Delta\bar{C}(t_i)} \quad (8)$$

[38] They are equivalent to a filter function that is anisotropic and two-point symmetric in space and non-stationary in time (see Appendix A).

3.3. Comparisons of Filtering Methods

[39] The two-step method as sketched in Figure 4 is numerically compared with the de-striping-Gaussian (isotropic and non-isotropic) filtering methods in Figure 5. The criterion $RT_C = 2$ is used for de-striping. The 300 km Gaussian filtering results in evident north-south-stretching patterns as shown in the upper-left map most likely due to significant residuals of the stripe-like error that requires further filtering. The 450 km Gaussian filtering removes more effectively the north-south stretching patterns as demonstrated in the upper-right map in contrast to the upper-left one, and also predictably leads to a smoother result in terms of magnitude and details. The non-isotropic Gaussian filtering performs significantly better than the Gaussian ones in

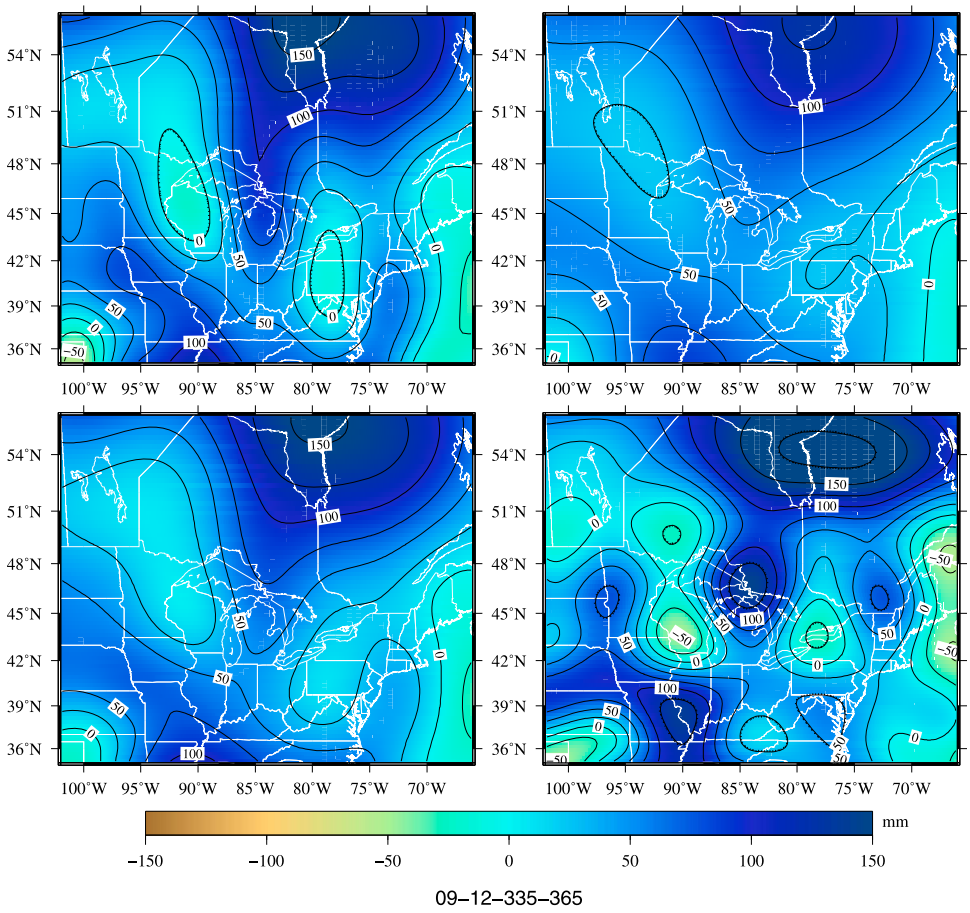

Figure 5. The total mass changes for December 2009 with respect to the mean model for the period of 2005–2009. (top left) De-striping and 300 km Gaussian filter. (top right) De-striping and 450 km Gaussian filter. (bottom left) De-striping and non-isotropic Gaussian filter with r0 = 300 km, r1 = 450 km and ml = 60 [*Han et al.*, 2005]. (bottom right) The two-step method with r0 = 300 km, r1 = 450 km and ml = 60.

removing the residual stripe-like error and retaining the magnitude and features simultaneously as shown in the lower-left map. The two-step method results in more features and stronger magnitudes as shown in the lower-right map than the non-isotropic one due to less smoothing.

[40] Naturally, the next question will be whether the two-step filtering results in an improvement in terms of the refinement of GRACE signals. In order to answer this question, a correlation analysis has been performed between the filtered GRACE results and the combined soil moisture, snow and lake (SMSL) water storage within the Great Lakes Basin. The hypothesis for this analysis is that if a filtering method works better, it should extract a GRACE total mass storage result that better correlates with the combined SMSL water storage. The groundwater storage over the land part of GLB is missing in these combined SMSL validation data. They affect the correlation analysis at a certain level.

[41] The combined SMSL water storage data are created from the NOAH025 model in GLDAS land cover areas from 60°S to 90°N and the lake water storage estimated from water level observations at 22 stations around the Great Lakes [see section 2.1]. The grids of soil moisture and polygon water level data are geographically merged to form monthly grids of 15 by 15 arc-minutes that match the NOAH025 resolution [see section 2.2]. Less than 1 percent of the grid nodes have no values for the region of study. They represents a few small lakes where water level data are not available. The two-step-filtered and GIA-corrected (see section 5) GRACE grid values are used to fill up the void nodes in the region of study along with oceans and southern pole region to achieve global coverage. Since the Great Lakes region is distant from oceans and the southern pole region where GRACE grid values have been used to fill up data gaps, the filling effect on the correlation analysis is considered negligible. These monthly global grids span from 2002 to 2009 covering 90 months giving the SMSL storage as deviations from a reference mean storage of 2005–2009. They are used to produce 90 monthly SH models truncated at SH degree 60. These truncated models are used to create the SMSL grids without further filtering for the correlation analysis.

[42] The correlation analysis is shown in Figure 6. The results from all the four methods display a mean correlation of greater than 0.56. Among them, the two-step filtering result demonstrates the best correlation, thus is inferred to be the best refinement of the GRACE signal. However, some detailed features shown in the lower-right map of Figure 5 may not be realistic considering the uncertainty of 2 cm

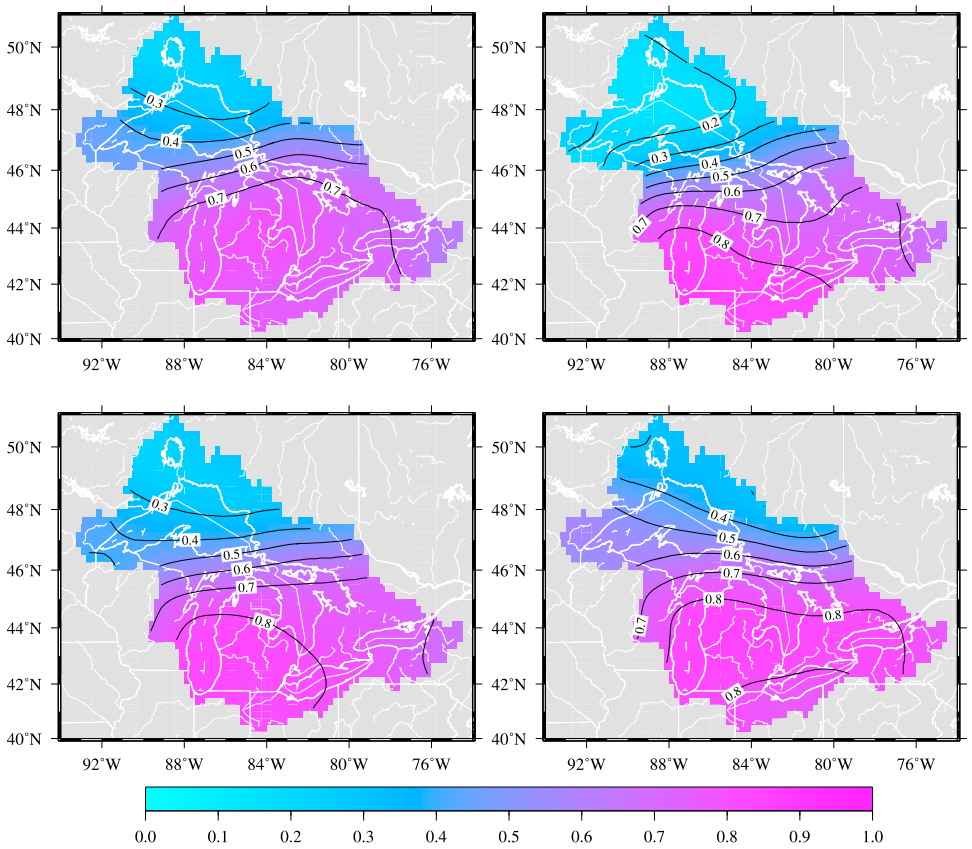

Figure 6. The correlation between the combined soil moisture, snow and lake water storage and the GRACE TWS for the period of 2002–2009 using four different filtering methods within the Great Lakes Basin. (top left) De-striping and 300 km Gaussian filter; the mean correlation coefficient is 0.60. (top right) De-striping and 450 km Gaussian filter; the mean correlation coefficient is 0.56. (bottom left) De-striping and non-isotropic Gaussian filter with r0 = 300 km, r1 = 450 km and ml = 60; the mean correlation coefficient is 0.62. (bottom right) The two-step method with r0 = 300 km, r1 = 450 km and ml = 60; the mean correlation coefficient is 0.67.

with the GRACE results. It should be noted that correlation is an indicator of phase and periodicity. One remarkable feature among the four maps is that the correlation shows a general decreasing trend toward North. It has the fingerprint of glacial isostatic adjustment, which will be discussed in section 5.

4. Methods for Estimation of the Total Water Storage Within a Water Basin

4.1. Integration of the Total Water Storage

[43] The average GRACE TWS in WTE as deviation from a reference mean storage within a water drainage basin can be estimated as

$$\Delta\bar{h}_{GRC}(t_i) = \frac{1}{A}\int\int_D w(\phi,\lambda)\Delta h(\phi,\lambda,t_i)\cos\phi d\lambda d\phi \qquad (9)$$

where

$$A = \int\int_D w(\phi,\lambda)\cos\phi d\lambda d\phi \qquad (10)$$

w is the exact basin averaging weight function which equals to 1 inside and 0 outside the water basin D. The weight function can be expanded into the SH form as follows

$$w(\phi,\lambda) = \frac{1}{4\pi}\sum_{n=0}^{\infty}\sum_{m=-n}^{n} w_{nm}\overline{Y}_{nm}(\phi,\lambda) \qquad (11)$$

[44] Substituting it into equation (9) and considering the two-step filtering, we have

$$\Delta\bar{h}_{\mathrm{GRC}}(t_i) = \frac{R\rho_a}{3A\rho_w}\sum_{n=0}^{\infty}\frac{2n+1}{1+k_n}\sum_{m=-n}^{n} f_{nm}\Delta\bar{C}_{nm}(t_i)w_{nm} \qquad (12)$$

[45] The monthly GRACE models are only complete to a limited SH degree and order N, therefore the estimate from equation (3) is subject to omission error. The missing components beyond the maximum degree N introduce the omission error into the TWS estimate $\Delta\bar{h}$ of equation (9), and can be written as

$$\Delta\bar{h}^{\mathrm{Omission}}_{GRC}(t_i) = \frac{1}{A}\int\int_D w(\phi,\lambda)\Delta h^{\mathrm{Omission}}(\phi,\lambda,t_i)\cos\phi d\lambda d\phi \qquad (13)$$

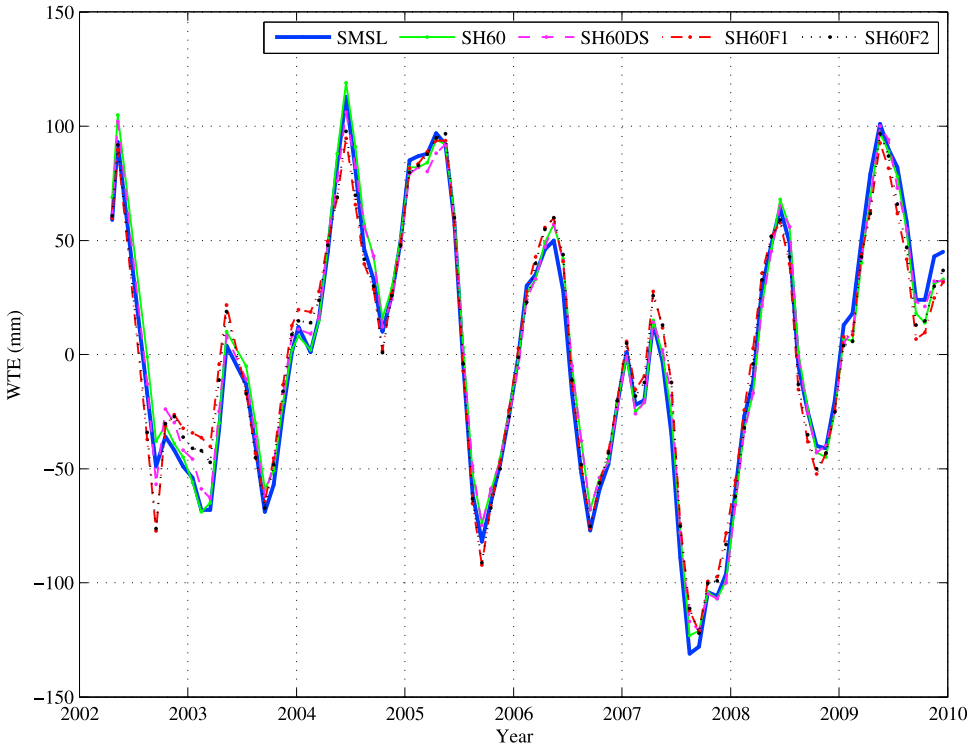

Figure 7. The combined soil moisture, snow and lake water storage variations, their estimates of the truncated SHs and GRACE equivalent estimates using two different basin averaging weight functions.

where

$$\Delta h^{\text{Omission}}(\phi,\lambda,t_i) = \frac{R\rho_a}{3\rho_w}\sum_{n=N+1}^{\infty}\frac{2n+1}{1+k_n}\sum_{m=-n}^{n}\Delta\bar{C}_{nm}(t_i)\overline{Y}_{nm}(\phi,\lambda) \tag{14}$$

[46] The omission of high-degree SH components mathematically causes a ringing artifact of Δh around its true value within the water basin, and therefore affects the average TWS estimate for the basin [see, e.g., *Swenson and Wahr*, 2002]. This resulting error is called the leakage error which equals to the sum of the leakage-in and -out errors for the basin. For fixed maximum degree N, the larger the basin area, the smaller the omission error in equation (13). One extreme case is that when the study area covers the entire Earth, the maximum required degree of the SH model is 0 to compute a true average value. The averaging integral of equation (13) equivalently acts as a low-pass filter [see, e.g., *Jekeli*, 1981; *Huang et al.*, 2008]. It tends to smooth out the ringing artifact within the basin. As a rule of thumb, if the dimension of a basin is larger than the GRACE resolutions in both the north-south and west-east directions, an approximate water storage can be estimated from the GRACE model.

[47] *Swenson and Wahr* [2002] propose an optimized averaging kernel to minimize the leakage error with a constraint on the value of satellite measurement error. *Chen et al.* [2009b] use forward modeling to account for the omission error.

[48] To estimate how the omission error affects the TWS estimate in equation (12), the SMSL water grids and their monthly SH expansions as described in section 3.3 are used to simulate the TWS grids. Since the SMSL storage constitutes a major part of the TWS in the GLB, the SMSL grids can be used to analyze the error of the TWS estimated by equation (12) due to the truncation at degree N and the two-step filtering. The results are shown in Figure 7. The thick blue solid line (SMSL) represents the result from the 15 by 15 arc-minute combined SMSL grids. It is used as the true TWS for this comparison. The green solid line (SH60) gives the TWS result for the truncated SMSL SH models complete to degree and order 60 without any filtering. The green line differs from the blue line by 7 mm in RMS suggesting the level of truncation error for the TWS estimation within the GLB. We have estimated the "leakage in" error by zeroing the SMSL values within the basin in the simulation. The resulting error ranges from −26 to 17 mm with a RMS value of 10 mm. We have also estimated the "leakage out" error by zeroing the SMSL values outside the basin. The resulting error ranges from −23 to 18 mm with a similar RMS value of 10 mm. The sum of leakage in and out errors is 6 mm in RMS which is consistent with the mean omission (truncation) errors of 7 mm above considering computational errors. The cancellation of the leakage in and out errors tends to reduce the total leakage error in the GLB.

[49] The filtering of monthly GRACE models also affects the TWS estimate in equation (12). The same SMSL monthly models truncated at degree/order 60 are used to estimate the signal loss due to the filtering by the two-step method. First, the signal loss due to the de-striping is estimated. The analysis in section 3.2 suggests that $RT_C = 2$ is a balanced choice between removing the stripe like error and keeping the signal. Using this criterion, the SMSL models are de-striped.

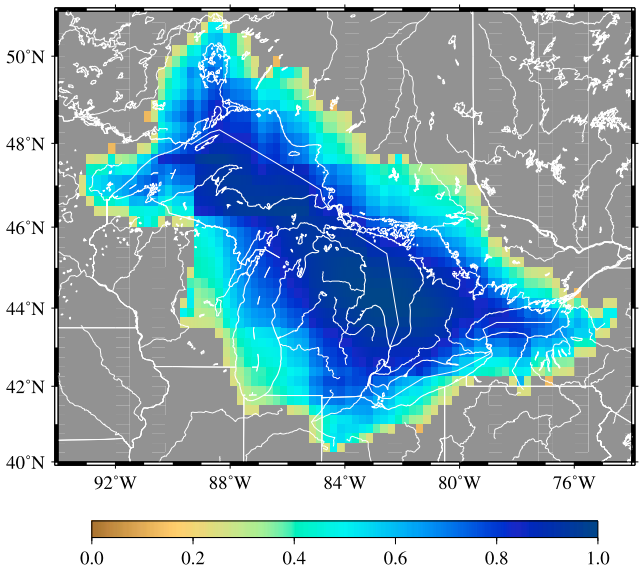

Figure 8. The 2D basin weight window over the Great Lakes Basin.

The TWS estimate for the de-striped models is shown as a magenta dash line (SH60DS). The signal change due to the de-striping ranges from −19 mm to 10 mm with a RMS value of 5 mm in WTE. This RMS change is larger than the RMS change using the soil moisture and snowfield for simulation in section 3.2, and is more representative for the GLB because of the inclusion of lakes in the SMSL grids.

[50] Second, the de-striped SMSL models are further filtered by the GRACE-equivalent filtering coefficients at step two of the two-step method

$$f'_{nm}(t_i) = \frac{\Delta\bar{C}^{\mathrm{f}}_{nm}(t_i)}{\Delta\bar{C}^{\mathrm{DS}}_{nm}(t_i)} \tag{15}$$

The resulting TWS estimate is shown as a red dot-dash line (SH60F1). The signal loss due to step two ranges from −22 mm to 25 mm with a RMS value of 11 mm in WTE. Comparing this TWS estimate (SH60F1) with the true one (SMSL), the total signal loss due to the omission and the two-step filtering ranges from −31 mm to 29 mm with an RMS value of 12 mm. It is below the GRACE noise level. Note that a rescaling is not applied here, when it is usually required for the Gaussian filtering method [see, e.g., *Velicogna and Wahr*, 2006; *Chen et al.*, 2007]. The actual omission effect may exceed this value because this result does not include the groundwater component.

4.2. A Two Dimensional Window Weight Function for Water Basin

[51] Applying a window function is an effective method of smoothing the ringing artifact in Δh due to the omission of high-degree SHs. *Harris* [1978] compares the most commonly used window functions for harmonic analysis. The Gaussian window performs reasonably well in terms of reducing side lobes. It is selected to form a discrete two dimensional (2D) window weight function to compute a weight for each grid node within the basin replacing the uniform weight 1 as follows:

$$w^{\mathrm{G}}(\phi_i, \lambda_j) = w_{\lambda_j}(\phi_i) w_{\phi_i}(\lambda_j) \tag{16}$$

where

$$w_{\lambda_j}(\phi_i) = \begin{cases} e^{-a_j\left[1-\cos\ (\phi_i - \phi_0^j)\right]} & (\phi_i, \lambda_j) \in D \\ 0 & \text{other} \end{cases} \tag{17}$$

$$w_{\phi_i}(\lambda_j) = \begin{cases} e^{-a_i\left[1-\cos\ (\lambda_j - \lambda_0^i)\right]} & (\phi_i, \lambda_j) \in D \\ 0 & \text{other} \end{cases} \tag{18}$$

ϕ_0^j is the central latitude of meridian grid line j within the basin to make the weights at all nodes along the line j symmetric with respect to ϕ_0^j. Similarly λ_0^i is the central longitude of parallel grid line i. In the case of co-existence of multiple disconnected sections separated by nodes outside the basin in one grid line, each section is treated independently, and associated with one distinct Gaussian weight function.

[52] The 2D window weight function in equation (16) is a mathematical extension of a one dimensional window function to the two dimensional averaging. It inherits the most important property of a window function: weight decreases toward the boundary of the basin to smooth the ringing effect. This property is also consistent with the definition of a water drainage basin where water always flows toward main water bodies in the interior of basin. In the case of the GLB, the main water bodies are the Great Lakes which contribute the largest water storage. Errors from reduced weight near the GLB edges are therefore insignificant. A suitable choice of parameter sets a_i and a_j exists to lead to a realistic estimate of the total water storage. It can be determined by

$$a = -\frac{\ln(w_c)}{1-\cos(hd)} \tag{19}$$

where w_c defines the weight at the first and last nodes of a (meridian or parallel) grid line. In this study, it is set as 0.5. The quantity hd is half the latitude/longitude difference between the first and last nodes. Accordingly the maximum weight is 1 when a node is located at a central location for both the meridian and parallel grid lines passing that node. Conversely the minimum weight is 0.125 when a node connects to the basin only by one node. In most cases, a node on the boundary connects to the basin by two nodes and is assigned a weight of 0.25.

[53] Figure 8 shows the derived 2D window weight map over the GLB. It is derived from a 15 by 15 arcmin basin mask map. A major feature is lower weights along the edge of basin and higher weights deep inside the basin. Figure 9 shows its truncated SH representation along with that of the exact weight function according to equation (11). They represent the actual weight being applied to the GRACE TWS result globally. As expected, the 2D function is more concentrated on the interior of the basin and tends to reduce the edge ringing effect better.

[54] The 2D function in equation (16) is used in equation (9) to compute the simulated TWS using the same filtered SMSL models in section 4.1. The resulting TWS is shown in Figure 7 as the SH60F2. The RMS difference between this estimate and the true one is 10 mm in contrast to 12 mm in

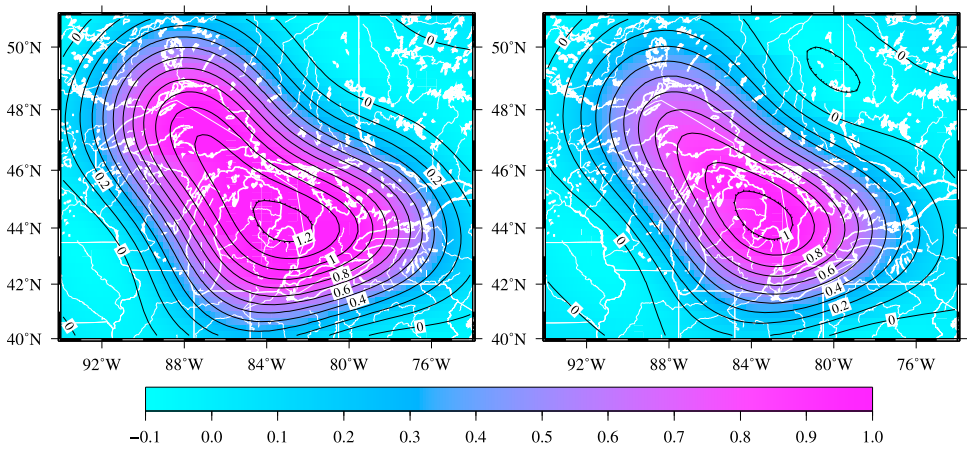

Figure 9. The truncated SH expansion (degrees 2 to 60) of the averaging weight map over the Great Lakes Basin for (left) the exact weight window, i.e., 1 inside the basin and 0 outside and (right) the 2D weight window.

the case of SH60F1. It suggests that the 2D averaging function performs slightly better than the exact averaging function. Another conclusion drawn from this simulation is that the GRACE detectability of the TWS over the GLB is not significantly affected by the GRACE omission error. In other words, the truncated and filtered GRACE monthly models can be used to detect the TWS change over the GLB with one-centimeter accuracy in WTE if their commission errors are negligible.

5. Glacial Isostatic Adjustment Effect

[55] The GLB is located at the margin of a region with a very significant glacial isostatic adjustment effect (GIA). The GIA effect on the GRACE observation is equivalent to the secular mass increase in the GLB region, and must be removed to estimate the TWS change from GRACE. Its modeling has been an active research area for many decades [*Peltier*, 2004]. Many different models have been developed, and they can be significantly different from each other.

[56] In this study, we use and compare four GIA models recently discussed in the literature.

[57] *GIA-ICE-3G*: A GIA model which uses the ICE-3G loading history [*Tushingham and Peltier*, 1991], and an Earth model with a uniform mantle viscosity of 1×10^{21} Pa s. This model agrees well with geodetic data over the Great Lakes [*Mainville and Craymer*, 2005]. Of a range of models investigated in *Braun et al.* [2008], this model provides one of the best fits to uplift rate data in the Great Lakes region.

[58] *GIA-ICE-4Ga*: A GIA model which uses the ICE-4G loading history [*Peltier*, 1994], and an Earth model which has upper and lower mantle viscosities of 0.4×10^{21} Pa s and 12.8×10^{21} Pa s, respectively. It fits best to GRACE data in North America out of a range of mantle viscosities and for ice models ICE-3G, ICE-4G and ICE-5Gv1.2 [*van der Wal*, 2009].

[59] *GIA-ICE-4Gb*: A GIA model which uses the ICE-4G loading history, and an Earth model which has upper and lower mantle viscosities of 0.8×10^{21} Pa s and 3.2×10^{21} Pa s, respectively. It best fits the GPS uplift rate of *Sella et al.* [2007] in North America out of a range of mantle viscosities for ICE-4G and ICE-5Gv1.2 [*van der Wal et al.*, 2009].

[60] *GIA-ICE-5G*: A GIA model which uses the ICE-5G loading history [*Peltier*, 2004], and an Earth model which has upper and lower mantle viscosities of 0.8×10^{21} Pa s and 1.6×10^{21} Pa s, respectively. It best fits the GPS uplift rate in North America out of a range of mantle viscosities for ICE-5Gv1.2 [*van der Wal et al.*, 2009].

[61] The GIA corrections for the four models above are shown in Figure 10 in mm(WTE)/yr. The correction for GIA-ICE-3G is similar to the one for GIA-ICE-4Ga, but they differ from the other two models considerably. To analyze their GIA prediction skills and find the best one, the similar correlation comparisons to the ones shown in Figure 6 are used to compare the four GIA models within the GLB region. The hypothesis of these comparisons is that the SMSL water storage is only correlated with the total water storage contained in the GRACE results, i.e., the better a GIA correction, the larger the correlation. The four GIA models above are used to correct for the GIA effect in the GRACE models before the two-step filtering described in Figure 4.

[62] The correlation results between the GRACE TWS grids and the SMSL grids are shown in Figure 11. On one hand, it is evident that the correlations utilizing GIA-ICE-3G, GIA-ICE-4Ga and GIA-ICE-4Gb have been significantly improved with respect to the best case without the GIA correction seen in Figure 6. The mean correlations for GIA-ICE-4Ga and GIA-ICE-4Gb are identical, and slightly better than that of GIA-ICE-3G. But it is not evident which one is the best. On the other hand, the spatial distributions of correlation for GIA-ICE-3G and GIA-ICE-4Ga are relatively more homogeneous than the one for GIA-ICE-4Gb which shows higher correlation over a zone across Lakes Michigan and Heron but a decreasing trend toward northwest of the GLB. Considering the differences of the ice models (ICE-3G versus ICE-4G) and the constraints (vertical uplift by GPS versus gravity change by GRACE), the consistency of correlations and corrections between GIA-ICE-3G and GIA-ICE-4Ga indicates that they are likely more realistic in the GLB region. On the contrary, the model GIA-ICE-5G worsens the correlation within the GLB suggesting a need for

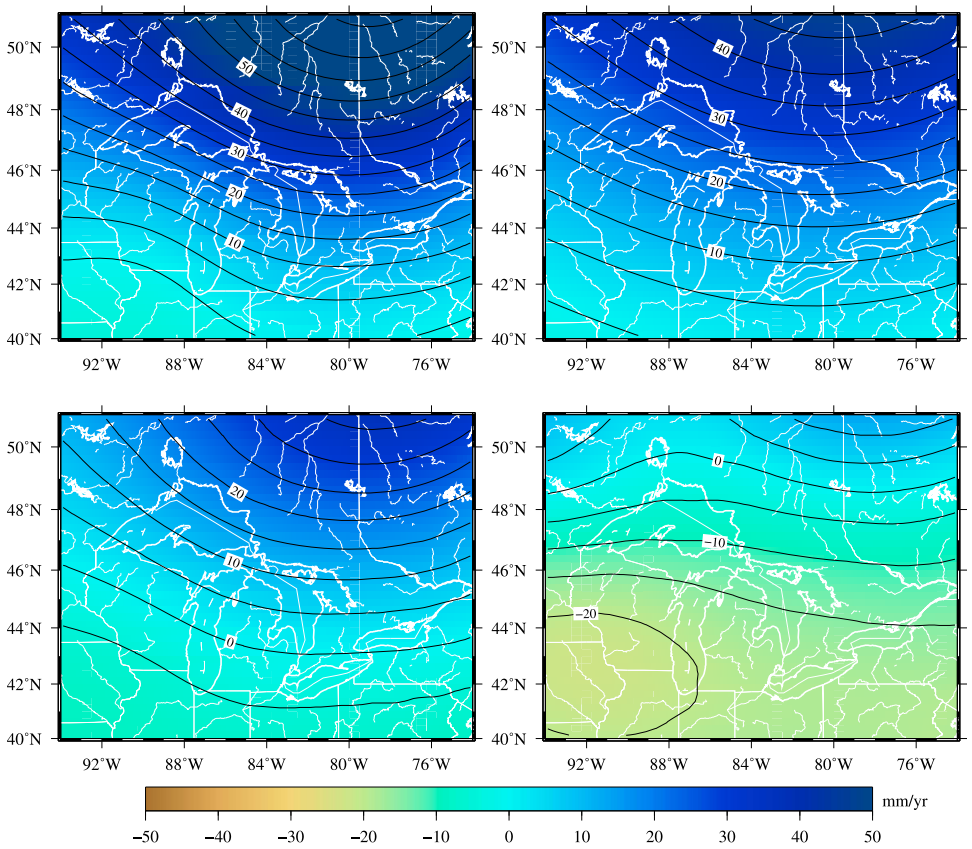

Figure 10. GIA effects in WTE for four different GIA models. (top left) GIA-ICE-3G. (top right) GIA-ICE-4Ga. (bottom left) GIA-ICE-4Gb. (bottom right) GIA-ICE-5G.

improvement in this region. It will not be further discussed in this study. It should be pointed out that *Paulson et al.* [2007] derived a GIA model using the ICE-5G loading history plus GRACE and relative sea level data around Hudson Bay. Their model is different from the GIA-ICE-5G used in this study. A preliminary comparison shows that GIA-ICE-4Gb is closest to *Paulson et al.*'s model among the four models above, while the former is 2.9 mm/yr greater than the latter on an area-weighted average within the GLB.

[63] We have repeated Figure 6 after the GIA correction with GIA-ICE-4Ga. The correlation coefficients are: 0.67 (300 km Gaussian), 0.67 (450 km Gaussian), 0.71 (non-isotropic), and 0.74 (the two-step method) in contrast to the values before the GIA correction: 0.60, 0.56, 0.62, 0.67 in Figure 6. The correlations for all the filtering strategies have been improved.

6. Estimation of Total and Groundwater Storage Changes

6.1. Total Water Storage Changes

[64] The GLB water system is one of the largest fresh water systems on the Earth covering an area of about 766,000 km^2. It consists of lakes, surface snow or ice, soil moisture and groundwater components. The TWS is the sum of water storage in these reservoirs.

[65] The average TWS variations in WTE for the period of 2002–2009 have been estimated with respect to the reference mean of 2005–2009 within the GLB from the monthly GRACE models by CSR and GFZ, respectively. The GRACE signal is refined using the two-step method described in section 3.2, and the TWS variations are computed using the 2D window weight method described in section 4. The GIA effect is corrected by predictions of GIA-ICE-3G, GIA-ICE-4Ga and GIA-ICE-4Gb models, respectively. The results and their standard errors are shown Figure 12. The standard errors are estimated from the calibrated standard deviations of GRACE model coefficients through error propagation.

[66] As expected, annual seasonal cycles dominate the GRACE TWS with a peak-to-peak magnitude of about 150 mm. They represent seasonal water storage variations. The variations from CSR's models are consistent with those from GFZ's in terms of standard errors except for July to October of 2004, where the latter show enormous standard errors caused by short repeat cycles of GRACE satellites for this period [see *Wagner et al.*, 2006]. The differences between the CSR and GFZ time series for common months vary from −68 mm (Sept. 2004) to 56 mm (Feb. 2003) with a RMS of 24 mm. The standard errors are smaller than 24 mm for most months indicating the commission error level of GRACE in terms of the water thickness equivalent within the GLB. The CSR GRACETWS correlates to the level of 0.79 (top), 0.82 (middle) and 0.85 (bottom) with the SMSL water storage shown in Figure 7, respectively. This suggests that the annual cycles from GRACE largely reflect the SMSL water storage variations. This observation conforms to the common understanding that the SMSL water change is the dominant water change signal in the GLB. Furthermore

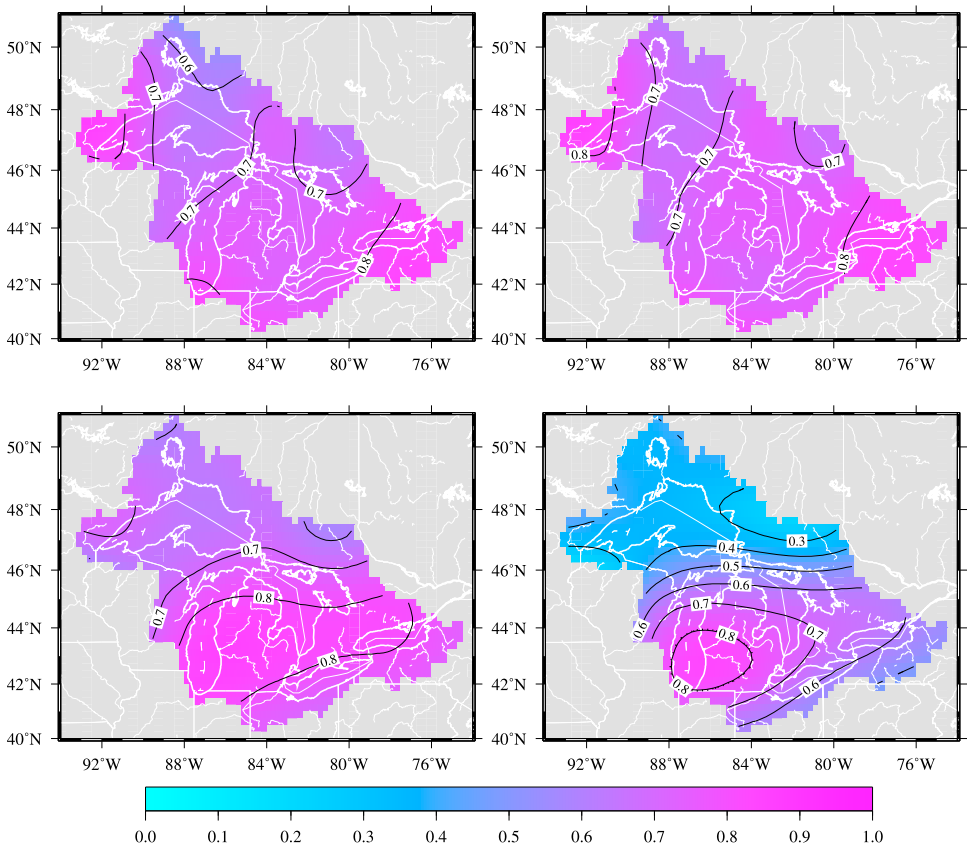

Figure 11. The correlation between the combined soil moisture, snow and lake water storage and the GRACE TWS corrected for four different GIA models. (top left) GIA-ICE-3G; the mean correlation coefficient is 0.72. (top right) GIA-ICE-4Ga; the mean correlation coefficient is 0.74. (bottom left) GIA-ICE-4Gb; the mean correlation coefficient is 0.74. (bottom right) GIA-ICE-5G; the mean correlation coefficient is 0.55.

the time series of the basin-wide TWS show higher correlation than the node-wise TWS (see Figure 11) with the respective SMSL water storage due to the basin-wide smoothing effect.

[67] In the meantime, the use of the three GIA models has resulted in different annual trends of TWS while having negligible effects on the seasonal variation of TWS. In the case of GIA-ICE-4Ga, the mean GIA correction is 16.2 mm/yr over the GLB. After this correction, the TWS demonstrate a slight decline of 8.4 $\pm$ 1.8 and 6.4 $\pm$ 1.6 mm/yr for CSR's and GFZ's models, respectively (see Table 2). These declines are 2.4 mm/yr larger if GIA-ICE-3G is used. In both cases, the decline started in 2002 and ended in 2007 followed by a visible return from 2008 to 2009. On the contrary, using GIA-ICE-4Gb gives a slight increase of 3.4 $\pm$ 1.8 and 5.2 $\pm$ 1.6 mm/yr for CSR's and GFZ's models, respectively. While the consistency of GIA-ICE-3G and GIA-ICE-4Ga indicates that they are more likely realistic, GIA-ICE-4Gb results in the highest basin-wise correlation (0.85) between the GRACE TWS and the SMSL water storage. Thus the GIA correction poses an uncertainty of about 10 mm/yr on the estimation of TWS trend.

[68] To understand the TWS trend better, its spatial variation has been studied over the GLB and its surrounding region. The trend maps without the GIA corrections are shown in Figures 13. These trend maps are estimated in two steps. First the trend SH models are derived from the filtered monthly GRACE models by the least squares estimation. Second the trend values are predicted from the trend SH models. The GIA fingerprint can be easily recognized over the northeast corner. The GIA effect reaches 40 mm in WTE at maximum. On the other hand, both maps show similar negative patterns southwest of Lake Superior, that represents the integrated GIA and water storage change. The differences of about 5 mm between them over Lake Michigan and Lake Huron are likely due to the uncertainty of processing methods used by CSR and GFZ and the different numbers of monthly models.

[69] After the GIA corrections, the TWS trend maps are derived and shown in Figure 14. The major GIA fingerprint has largely disappeared for each case. Each column of maps show similar spatial features. The northern part of the basin is losing water storage relative to the southern part during this period. However, there appear constant shifts between the maps with different GIA corrections. They suggest biases between the three GIA models in this region. The question is which GIA model contains the least bias.

[70] One way to analyze these biases is to compare the TWS changes with the SMSL water storage charges. In particular, Lake Superior covers a major part of the northern GLB and its water storage changes are equal to the TWS changes because the groundwater under the lake is considered saturated and unchanged over time. In Figure 2, the water level in Lake Superior dropped by 597 mm from June 2005 to March 2007. The NOAH025 soil moisture and snow

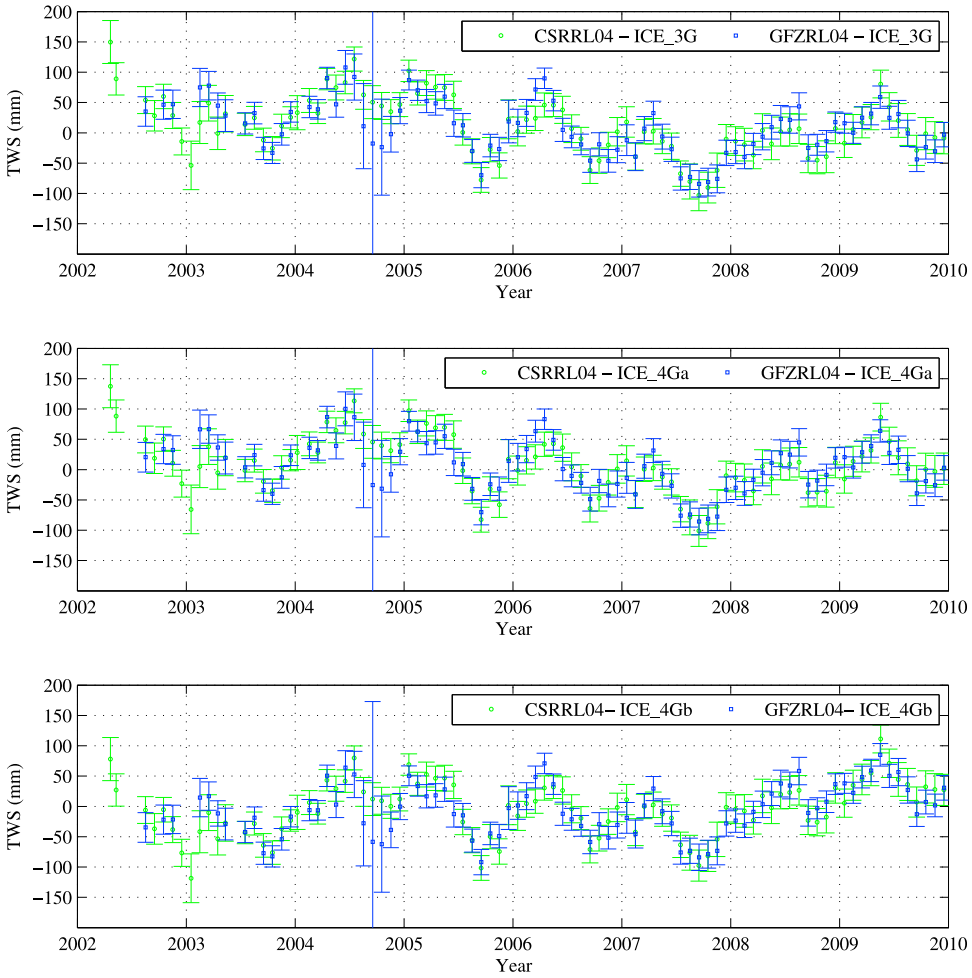

Figure 12. The mean total water storage variations in WTE within the GLB from the monthly GRACE models by CSR (green circle) and GFZ (blue square), respectively. The GIA effect has been corrected by (top) GIA-ICE-3G, (middle) GIA-ICE-4Ga, and (bottom) GIA-ICE-4Gb, respectively.

storage to the northwest of Lake Superior shows a drop of 150–200 mm for the same period. The SMSL water storage change smoothed by the GRACE-equivalent filter is compared with CSR's GRACE TWS change in Figure 15. It is evident that the GRACE-equivalent SMSL water storage change over Lake Superior is considerably smaller than the actual lake level drop. This is mainly due to the truncation and filtering of GRACE models which does not significantly affect the estimates of TWS changes within the larger GLB as shown in Figure 7. We can see the TWS results using GIA-ICE-3G and GIA-ICE-4Ga are closer to the SMSL one than the one using GIA-ICE-4Gb over Lake Superior. Among the three GRACE TWS maps, the one using GIA-ICE-4Ga shows the smallest mean bias of −0.4 mm with respect to the SMSL map, while GIA-ICE-3G and GIA-ICE-4Gb give rise to biases of −6.4 and 21.0 mm, respectively. The biases restricted to the GLB also suggest that GIA-ICE-4Gb gives a bias more than 20 mm greater than the other two GIA models between which there is a bias difference of 4 mm only. This analysis of biases suggests that GIA-ICE-3G and GIA-ICE-4Ga are more realistic. The latter is used for the GIA correction while the former is used to quantity the uncertainty of GIA corrections for groundwater estimation in the following sections.

Table 2. Annual Trend Rates, Annual Cycle Amplitudes, and Phases ϕ of GRACE (GRC) Total Mass Storage[a]

	Trend Rate		Amplitude		
Source	WTE (mm/yr)	Volume (km^3/yr)	WTE (mm)	Volume (km^3/yr)	ϕ (yr)
GRC(CSR)	7.8 ± 1.8	6.0 ± 1.4	32.4	24.8	0.70
GRC(GFZ)	9.8 ± 1.6	7.5 ± 1.2	37.4	28.6	0.74
GIA-ICE-4Ga	16.2	12.4			
TWS(CSR)	−8.4 ± 1.8	−6.4 ± 1.4	32.5	24.9	0.71
TWS(GFZ)	−6.4 ± 1.6	−4.9 ± 1.2	37.5	28.7	0.74

[a]The volume values are products of the WTE values and the area of GLB, which is 766,000 km^2.

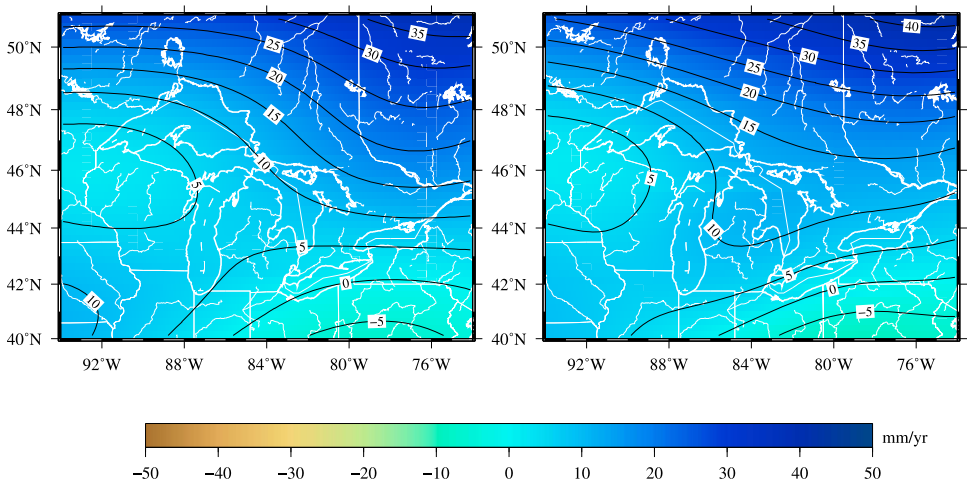

Figure 13. The annual trend of total mass storage from GRACE using (left) CSR's and (right) GFZ's models for the period of 2002 to 2009 before removing the GIA effect.

6.2. Groundwater Storage Changes

[71] Given GRACE TWS, SMSL water storage, the groundwater storage (GWS) Δh_{GW} as deviation from a mean for the same reference period as GRACE and SMSL can be estimated by the following water storage balance relation:

$$\Delta h_{GW} = \Delta h_{GRC} - \Delta h_{SMSL} \tag{20}$$

where the subscript GRC stands for GRACE, and SMSL for soil moisture, snow and lake.

[72] Note that the water storage variations in the unsaturated soil zone between the simulation depth of the land surface models and the groundwater table are not considered. Thus the estimated GWS includes the water storage variations within this zone. The difference is insignificant around lakes or swampy areas but can be larger in dry areas where the water table is deeper. A summary of the difficulties in characterizing the unsaturated region is given by http://wwwrcamnl.wr.usgs.gov/uzf/unsatflow/unsatflow.html. Essentially, water content and flow are determined by the size, shape and characteristics of soil cavities as well as by the amount of water present. The water table is at or very near the surface in much of the study area and there are no really arid regions, so the effect of including water in this zone in the GWS estimates should not be significant.

[73] The GWS variations within the GLB have been estimated from the GRACE TWS shown in Figure 12 (middle) that uses the least-biased GIA correction, and the SMSL water storage that are filtered using the GRACE-equivalent filtering coefficient f'_{nm} given by equation (15).

[74] The GRACE-equivalent filtering is to make the SMSL water storage field spectrally equivalent to the GRACE field so that the derived groundwater storage field has the same spectral content as the GRACE field. The de-striping step is not part of the GRACE-equivalent filtering, and should not be used for the SMSL field which does not contain the stripe-like error. Strictly speaking, the de-striping is not a spectral filtering step. Instead it is a correction step to SH coefficients which contain the stripe-like errors. The GRACE-equivalent filtering depends on the GRACE models. The RMS difference of the average SMSL time series by CSR's and GFZ's GRACE-equivalent filters is 4 mm in WTE for the period of 2005–2009.

[75] Three SMSL fields are created by combining the lake water storage field with the three land surface models: NOAH25, AVG10 and WGHM. The term 'field' here is defined as a set of monthly water storage grids of the same data source. Their trend and annual periodic parameters are estimated and shown in Table 3. Overall, the SMSL variations using NOAH025 and AVG10 are similar and comparable with each other as well as the GRACE TWS variations in Table 2, but significantly different from the SMSL variations using WGHM in terms of phase and amplitude. The former two do not show a significant trend while the latter indicates an increase of water storage. These results illustrate the uncertainty associated with land surface models [see, e.g., *Fan et al.*, 2011].

[76] Considering the GWS under the lakes does not change over time, the estimated GWS time series within the GLB are converted to the land GWS time series by an area factor of 1.468, as shown in Figure 16. The time series using NOAH025 agree well with the ones using AVG10 for the same set of GRACE models (CSR or GFZ), but considerably disagree with the ones using WGHM in terms of phase and amplitude (see also Table 4). There are generally similarities between the CSR and GFZ GWS series for the same SMSL field, except for the period of the short repeat cycles of GRACE in 2004. However, the differences of the GWS series between using WGHM and GLDAS models are too large to lead to an overall conclusion on the amplitude of GWS variations within the GLB. Nevertheless the agreement between NOAH025 and AVG10 is a good indicator that their resulting GWS estimates tend to be realistic as AGV10 is an average of four LSMs. Two important features are: GWS seasonal variations are two thirds smaller in amplitude, and have a phase shift of about one half year relative to the seasonal SMSL water storage variations. On the other hand, the seasonal undulations of GWS series using NOAH025 and AGV10 appear more random than the ones using WGHM.

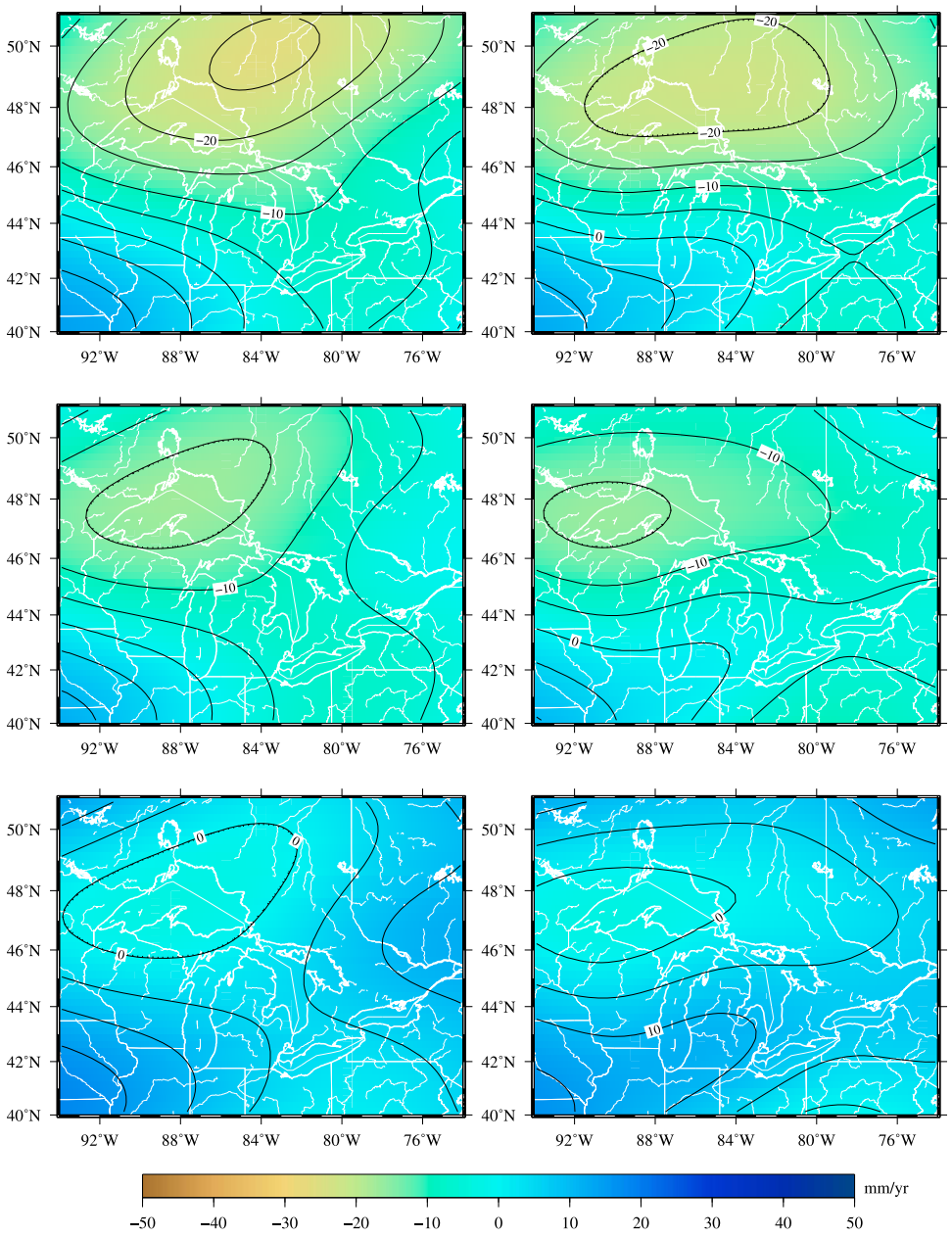

Figure 14. The annual trend of TWS from GRACE for the period of 2002 to 2009. The left column show the trends from CSR's models while the right column shows the trends from GFZ's models. The GIA effect has been corrected by (top) GIA-ICE-3G, (middle) GIA-ICE-4Ga, and (bottom) GIA-ICE-4Gb, respectively.

[77] Trend rates of the GWS time series have also been estimated using the least squares estimation, and are listed in Table 4. All the trend estimates indicate a loss of GWS within the GLB. Of them, the use of NOAH025 results in a net loss of GWS at a rate of 5.2 km^3 ± 0.9/yr for CSR's models or 3.6 km^3 ± 1.1/yr for GFZ's models. The mean rate of GWS loss is 6.0 km^3/yr for CSR's models and 4.7 km^3/yr for GFZ's models when averaging the GWS changes from the uses of the three LSMs. Since the TWS loss rate is estimated at 6.4 km^3 ± 1.4/yr for CSR's models or 4.9 km^3 ± 1.2/yr for GFZ's models, it appears that the GWS loss dominates the TWS loss in the GLB for the period of study.

[78] An important question is where the GWS loss takes place. To answer this question, trend maps of the GWS have been estimated using the GRACE-equivalent monthly SH models of SMSL fields and the GIA-corrected monthly GRACE SH models, respectively. First the GWS monthly SH models are computed by subtracting one set of the monthly SMSL SH models from one set of the GIA-corrected and filtered GRACE SH models by the two-step method. Then a trend SH model is derived from this set of GWS SH

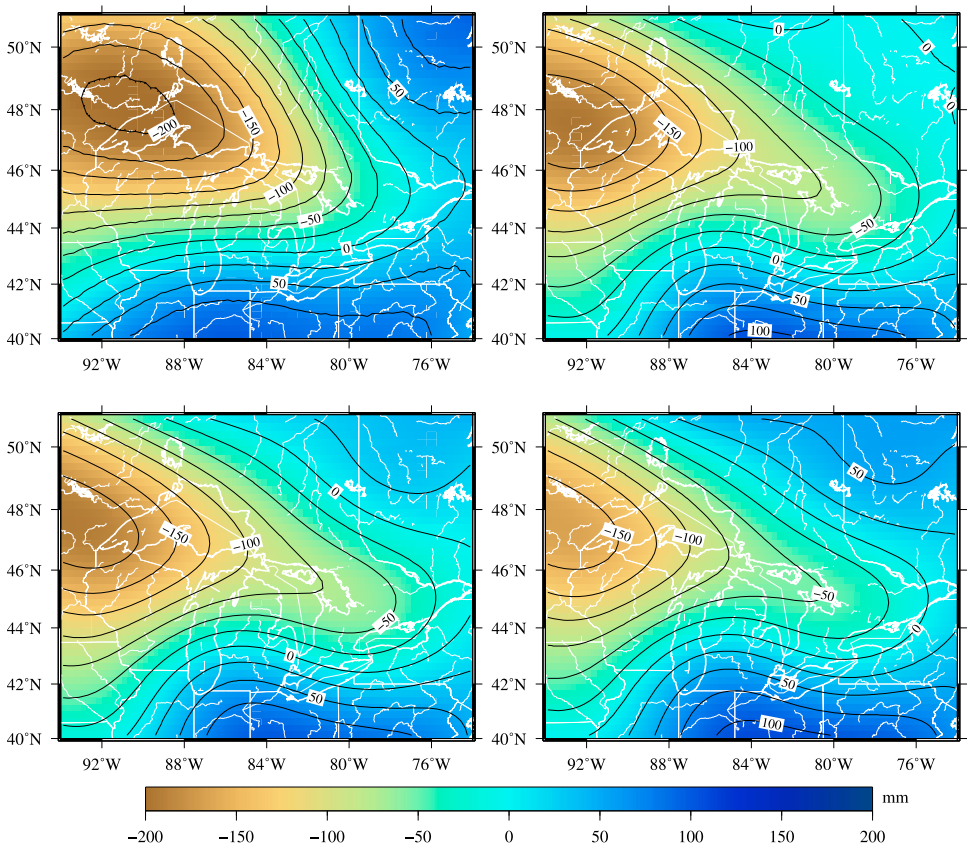

Figure 15. (top left) The GRACE-equivalent SMSL water storage change using NOAH025 and the TWS changes from CSR's GRACE between June 2005 and March 2007. The GIA effect has been corrected by (top right) GIA-ICE-3G, (bottom left) GIA-ICE-4Ga, and (bottom right) GIA-ICE-4Gb.

models by estimating the linear trend rate of each individual SH coefficient using the least squares estimation. Finally the trend map is synthesized from the trend SH model using equation (3). Since both monthly SMSL and GRACE SH models are filtered, so are the resulting trend SH model and map.

[79] Six trend maps are shown in Figure 17 for the two sets of GRACE monthly models and the three sets of SMSL monthly models. They show a similar and negative broad pattern which somewhat correlates to the shape of GLB but with differences in details. The GWS trend maps from CSR's models suggest that the southern part of the basin is losing more GWS than the northern part for the period of study, while the GWS maps from GFZ's models suggest an overall loss of GWS with less spatial variation. The differences between the left and right panels are due to the differences between CSR's and GFZ's models as well as different numbers of monthly GRACE models. The differences among three rows are caused by the three SMSL fields.

[80] To illustrate the impact of the SMSL fields, their trends are computed and displayed in Figure 18. The upper-left map is computed from the SH degree 2 to 60 of the SMSL field of NOAH025 without any filtering, while the upper-right map is computed with the GRACE-equivalent filtering

Table 3. Annual Trend Rates, Annual Cycle Amplitudes, and Phases ϕ of SMSL Water Storage Fields Which Combine Soil Moisture and Snow Fields From Different Land Surface Models With the Field of Lake Water Storage[a]

	Trend Rate		Amplitude		
Land Surface Model	WTE (mm/yr)	Volume (km^3/yr)	WTE (mm)	Volume (km^3/yr)	ϕ (yr)
NOAH025 (90 months)	0.0 ± 2.1	0.0 ± 1.6	49.0	37.6	0.68
AVG10 (90 months)	−2.3 ± 2.0	−1.8 ± 1.6	51.7	39.6	0.67
WGHM (90 months)	4.1 ± 1.2	3.2 ± 0.9	80.8	61.9	0.81
NOAH025 (84 months)	−1.9 ± 2.3	−1.4 ± 1.7	48.7	37.2	0.69
AVG10 (84 months)	−3.7 ± 2.3	−2.9 ± 1.8	51.9	39.7	0.68
WGHM (84 months)	3.6 ± 1.3	2.8 ± 1.0	79.1	60.6	0.82

[a]The SMSL fields have been smoothed by the GRACE-equivalent filters of CSR (90 months) and GFZ (84 months), respectively. The volume values are products of the WTE values and the area of GLB, which is 766,000 km^2.

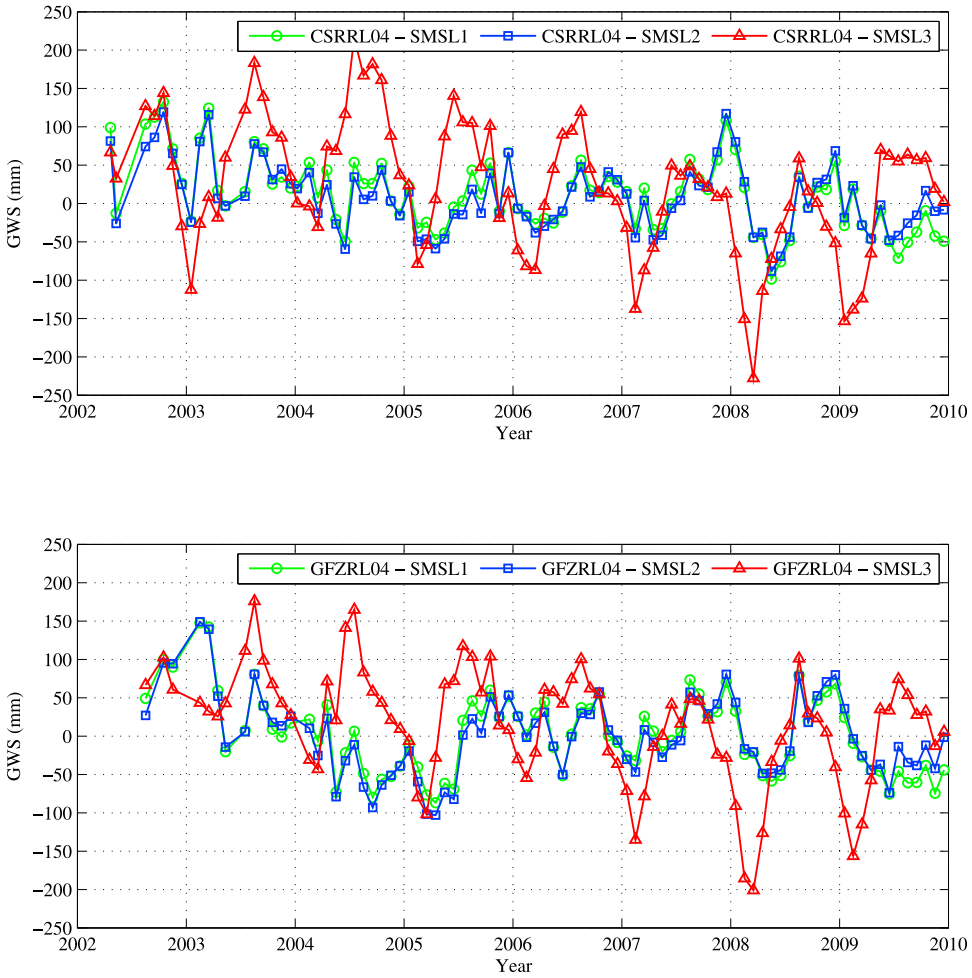

Figure 16. Estimates of the groundwater storage variations in WTE over the land of the GLB from the GRACE TWS of (top) CSR and (bottom) GFZ and the SMSL water storage. The three SMSL fields are created by combining the lake water storage field with the soil moisture and snowfields from NOAH025 (SMSL1), AVG10 (SMSL2) and WGHM (SMSL3). They are smoothed by the GRACE-equivalent filters of CSR and GFZ, respectively.

of CSR. These two trend maps show a SMSL water storage loss west of Lake Superior and a gain over the region encompassing the other four lakes. Comparing the two upper trend maps, it can be found that the GRACE-equivalent filtering smooths out detailed features. A relevant question is whether the loss of detailed features also affects the estimate of the net GWS change over the whole GLB. The average rates for both panels have been estimated at 0.4 mm/yr (unfiltered) and 0.1 mm/yr (filtered) for the whole GLB. These values suggest that the loss of detailed features does not have a significant effect on the GWS change. In addition, the GRACE-equivalent SMSL trend map of GFZ (not shown here) shows a RMS difference of 2 mm/yr from the one of CSR, that is significant when being compared with the GWS

Table 4. Annual Trend Rates, Annual Cycle Amplitudes, and Phases ϕ of Groundwater Storage (GWS) Estimated From the GRACE TWS and Three GRACE-Equivalent SMSL Fields[a]

	Trend Rate		Amplitude		
Source	WTE (mm/yr)	Volume (km^3/yr)	WTE (mm)	Volume (km^3/yr)	ϕ (yr)
CSR,NOAH025	-10.0 ± 1.8	-5.2 ± 0.9	27.6	14.4	0.18
CSR,AVG10	-6.8 ± 1.6	-3.5 ± 0.8	31.9	16.6	0.17
CSR,WGHM	-17.9 ± 2.1	-9.3 ± 1.1	91.3	47.6	0.37
GFZ,NOAH025	-6.9 ± 2.2	-3.6 ± 1.1	27.2	14.2	0.17
GFZ,AVG10	-4.4 ± 2.2	-2.3 ± 1.1	32.6	17.0	0.16
GFZ,WGHM	-15.7 ± 2.1	-8.2 ± 1.1	81.9	42.7	0.38

[a]The volume values are products of the WTE values and the land area of GLB, which is 521,840 (= 766,000 − 244,160) km^2.

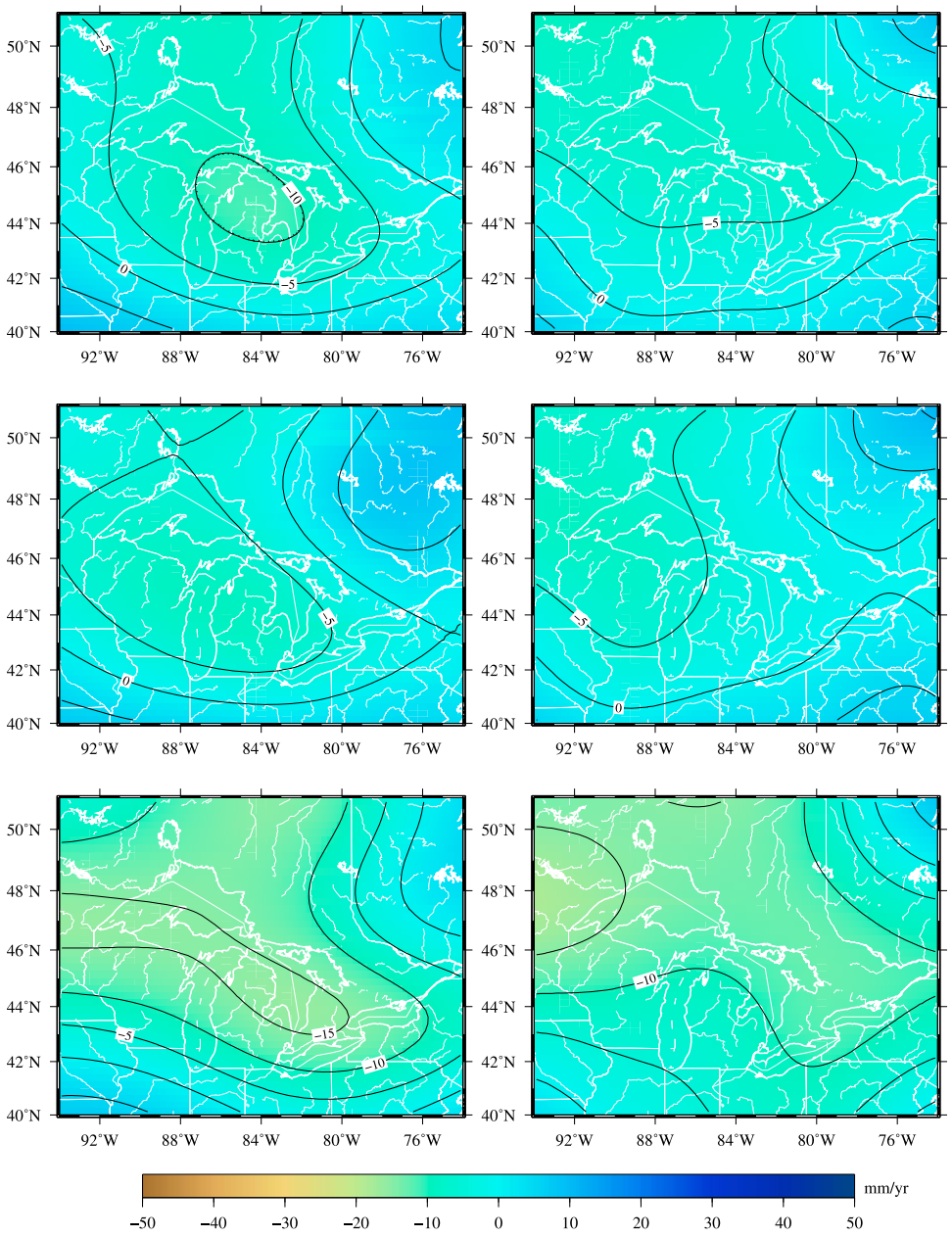

Figure 17. The annual trend of groundwater storage from GRACE for the period of 2002 to 2009. (left) Estimated from the CSR monthly models. (right) Estimated from the GFZ monthly models. Three rows corresponds to uses of (top) NOAH025, (middle) AVG10, and (bottom) WGHM for the SMSL fields.

trend in Figure 17. This difference is mainly due to the difference between the GRACE-equivalent filters of CSR and GFZ.

[81] The lower-left and lower-right maps are the GRACE-equivalent trends of the SMSL fields of AVG10 and WGHM, respectively. A common pattern is that the northern part of the basin is losing SMSL water storage, and the southern part is gaining the SMSL storage. Recall that the TWS trends in Figure 14 (middle) that are the least biased show an overall loss of TWS with stronger loss over the northern part of the GLB. It is evident that the SMSL loss contributes to the TWS loss over the northern part while the SMSL gain implies a greater loss of GWS so as to give rise to a TWS loss over the southern part. The differences among the three SMSL trends are as large as 5 mm/yr, contributing to the uncertainty of the GWS trends.

6.3. Discussions

[82] We have estimated the groundwater storage changes from GRACE. A central question then is how realistic these GRACE estimates are. Contributing errors can be classified into two categories: data errors and methodology errors. The data errors include the GRACE commission and omission errors, the GIA correction error, and the surface water (soil moisture, snow and lake) data error. Methodology errors include filtering error, the basin integration error, and nonlinearity of the least squares estimation process.

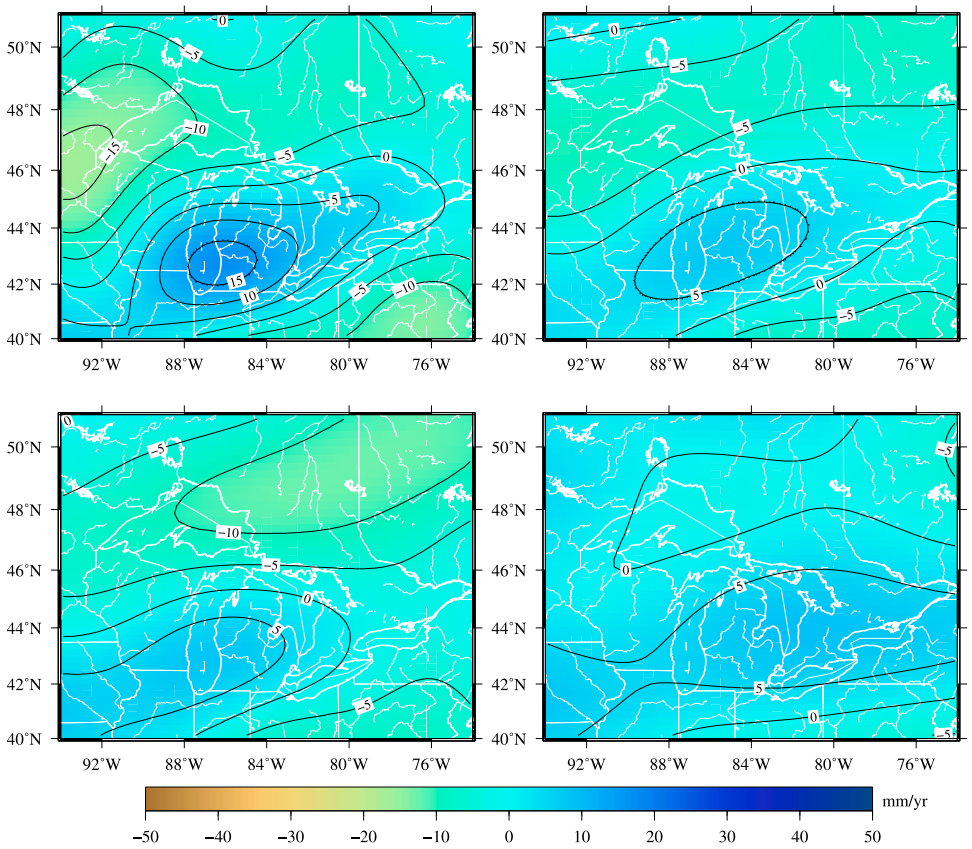

Figure 18. The annual trend of SMSL water storage for the period of 2002 to 2009. The top two maps show the SMSL trends using NOAH025 (top left) before and (top right) after applying the GRACE-equivalent filtering. The bottom two maps show the SMSL trends using (bottom left) AVG10 and (bottom right) WGHM after applying the GRACE-equivalent filtering.

[83] The GRACE commission errors are estimated at the level of 24 mm in the resulting WTE values (see Figure 12). The GRACE omission errors are at the level of 10 mm according to the discussion in section 4.2. Therefore the total GRACE errors for the TWS within the GLB are at the level of $(24^2 + 10^2)^{1/2} = 26$ mm in WTE. It should be pointed out that the relatively low omission error for the GLB is due to the large size of the GLB area, i.e., the larger the study area, the smaller the omission error for the TWS in the area. The omission error also depends on how similar water storage variations are in the surroundings of region of interest. Similar signals in the surroundings reduce the omission error due to cancellation of the leakage-in and -out errors.

[84] The GIA effect is corrected as a linear long-term trend. There is no error information associated with the GIA-ICE-4Ga model employed here. The next best model, GIA-ICE-3G, was also used to correct for the GIA effect and gives a difference of 2.4 mm/yr from GIA-ICE-4Ga. This value indicates the magnitude of uncertainty caused by the GIA model, though it is unknown how realistically this value reflects the uncertainty of the GIA correction to the GRACE trend rate for the GLB. Considering the approximately 2 mm uncertainty of the trend estimates in Table 4, the total trend uncertainty is about $(2.4^2 + 2^2)^{1/2} = 3.1$ mm/yr in WTE. It is necessary to quantify the uncertainty of the GIA model in order to use GRACE in monitoring water storage in the GLB region. It is encouraging that more GPS data are becoming available in recent years to better constrain GIA models.

[85] The SMSL data errors are a major error source for the estimates of GWS. The soil moisture and snow error is estimated at the level of 19 to 44 mm according to the comparisons of the three different sources (NOAH025, AVG10 and WGHM) of soil moisture model in section 2.2. It is at the level of 13 to 30 mm when being averaged over the whole GLB. However, this error estimate can be too pessimistic for the soil moisture model. This calls for a more comprehensive evaluation. Another source of error in the SMSL data error is from the lake level data. Multiple water level gauges have been used to compute the lake levels. The standard deviations of the monthly lake levels for each lake range from 5 to 28 mm. The area-weighted standard deviation for the total lake storage is 6 mm, which translates to 2 mm when averaged over the whole GLB. Taking the errors from all sources into consideration, the SMSL data errors are at the level of 13 to 30 mm.

[86] Considering all the uncertainties above, the GWS estimates have an uncertainty range of 29 to 40 mm, which are averages over the whole GLB. When converting them to the estimates over land only, they become 43 to 59 mm. These estimates suggest that the seasonal cycles shown in Figure 16 are not statistically significant. In other words, the

Table 5. Groundwater Storage and Fluxes in the GLB.

Sources	Volume (km^3)	Flux (km^3/yr)
In (long-term) storage	4009[a]	
Annual recharge		30 – 135[b]
Use		2.9[c], 2.08[d]
Precipitation		550 – 760[e]

[a]This volume is likely to be higher if the Canadian side is included.
[b]This is a range including the whole basin.
[c]Canadian side of the basin (the uncertainty in this estimate is high).
[d]US side of the basin.
[e]This range is equivalent to 720 mm to 1000 mm of precipitation over the basin per year.

uncertainties associated with the GWS estimates suggest that the estimated time series of GWS are not statistically reliable in terms of seasonal variations. However, it does not exclude the possibility that one of the time series is realistic because the error estimates for soil moisture and snow data are based on inter-comparisons between the soil moisture models, and can be too pessimistic. The consistence of the GWS estimates between using NOAH025 and AVG10 indicate that they likely reveal the dynamics of the GWS within the GLB, albeit the fact that all GLDAS models run with the same climate forcing weakens this argument.

[87] At this time, long-term (geological) groundwater storage, and groundwater storage changes with time (annual) in the GLB are not systematically measured on either side of the border. Most of the existing values are "estimates" obtained from modeling or from indirect measurements, such as base-flow from rivers. Well levels are the most commonly used measure of groundwater volume, but water mass depends on porosity, which can vary widely throughout the region. There are also considerable variations in water levels from well to well due to local factors such as flow rates and topography. They are not reliable data to verify the GWS time series from GRACE.

[88] A joint research project by the USGS and Environment Canada in 2008 [*Neff et al.*, 2005], estimated shallow groundwater recharge in the GLB using base-flow measurements in streams located within the basin. The main assumption in this study was that the base flow in a given stream is equal to the amount of shallow groundwater recharge to the surrounding watershed, minus losses to evapotranspiration. The recharge values estimated using this approach ranged from 40 mm to 420 mm per year, representing a range of volumes for the whole GLB between 30.6 km^3/yr to 326 km^3/yr. A more recent study based on numerical modeling, estimated a total of 134.9 km^3/yr for the US-side of the basin only [*Reeves*, 2010].

[89] On the other hand, it is well known that information on real volumes of exact groundwater use, all users combined, is extremely scarce. However, using detailed water budgets, water-supply management data, and consumptive-use coefficients, some values for groundwater use have been recently reported in the GLB. *Solley et al.* [1998] originally reported a total groundwater use on the US-side of the GLB of 2.08 km^3/year. In a more recent refined study, *Reeves* [2010] reported a total groundwater withdrawal of 1.78 km^3/year. These numbers cover the eight states of the US-side of the GLB.

[90] On the Canadian side, some values for the Province of Ontario have been reported as indirect estimates. In 2004, using a municipalities' database, *Rutherford* [2004] reported a total of 207 million m^3/year of groundwater use in the Ontario. However, this number is not consistent with a report by the National Groundwater Association [*National Groundwater Association*, 2007], who made an extensive literature review and a survey using questionnaires sent to well drillers and citizens in Ontario. Their results report a total of 2.9 km^3/year for all groundwater users combined, which is an order of magnitude higher than the value from *Rutherford* [2004]. A summary of storage and fluxes is provided in Table 5.

[91] The groundwater storage changes, reported above, in particular the annual amplitude of 14.4 km^3 and the trend -5.2 km^3/yr as estimated with CSR's GRACE models and NOAH025, seem to be reasonable for the large-scale basin covered by the Great Lakes. The GRACE trend of -5.2 km^3/yr represents between 0.7% and 0.9% of the precipitation in the GLB. Annual groundwater recharge in the GLB represents between 17% and 24% of precipitation and it can be from two times to ten times larger than the changes in groundwater storage estimated with GRACE. On the other hand, groundwater use in the basin ranges between 0.29% and 0.52% of precipitation in Ontario and between 0.20% and 0.37% in the US side. Groundwater use in the whole basin is less than 1% of precipitation.

[92] There is no evidence that climate effects or groundwater exploitation are the cause of the decreasing trend estimated with GRACE. A few groundwater levels in Southern Ontario show normal changes of recharge and discharge of groundwater on that part of the GLB.

[93] When looking at water budgets for large regions such as the GLB, the groundwater system, expressed in terms of long-term recharge provides little information regarding the response of the system to increased pumping. The pumping will initially remove water from storage and will lower the hydraulic head in the aquifer. At some time, however, the increased pumping will be balanced by either an increase in recharge to the pumped aquifer or a decrease in discharge from the pumped aquifer. This is true for local or semi-regional scales; however when a region covered in the analysis is as large as the GLB, the smaller spatial and temporal scales of the observations cannot be representative of the whole region.

[94] Given that groundwater is such a large component of streamflow in the Great Lakes Basin [*Holtschlag and Nicholas*, 1998; *Reeves*, 2010], and given the difficulty in increasing recharge to the aquifer by groundwater pumping, the dominant process in the GLB is likely to be reduction in discharge of groundwater to surface water. On the whole, water budgets in the Great Lakes basin indicate that water in storage and water flux through the system are very large compared to observed and estimated groundwater withdrawals [*Reeves*, 2010].

[95] The phase for each storage time series have been estimated in terms of a parameter equation similar to equation (6). The results are listed in Tables 2, 3, and 4. It can be seen that the SMSL amplitudes are greater than the TWS suggesting that the GWS variation has a phase shift of about half a year with respect to the SMSL to reduce the SMSL amplitude to the TWS one. The estimated GWS reaches the maximum level in September and October and the minimum in March and April, while the SMSL storage peaks are

opposite to the GWS ones. The TWS variations have a close phase to the SMSL which dominates the TWS variations.

[96] One possible explanation is that the phase shift is due to the complex mechanisms of infiltration, recharge and discharge. It is known that within the GLB, changes in groundwater levels and groundwater storage over time are important, both seasonally and annually. Depending on the depth to the water table in phreatic aquifers, very important temporal variations of water storage occur within the unsaturated zone.

[97] Infiltration into the ground helps to form the near-surface stock of water needed for evaporation and transpiration in late spring and summer. However, in cooler seasons, water infiltrates deeper into the ground, recharging the groundwater contained in the soils and rocks.

[98] When rain arrives on the ground and snowmelts, one part infiltrates and is essentially used to recharge the "soil reservoir" (soil moisture increases), from where evapotranspiration takes it back to the atmosphere (discharge). It is mainly in the cooler seasons when evaporation is lowest, that water continues downward and reaches the water table (groundwater storage increases, recharge). However, this process is complex and variable depending on the regions of the GLB and on the aquifer types (i.e., phreatic, confined, porous media, fractured media). The estimates with GRACE are averages over the whole GLB.

[99] Regarding the filtering of GRACE models, it is worth noting that the principle merit of the two-step filtering method suggested in this study is the preservation of the GRACE signal. On the one hand, due to insufficient spatio-temporal coverage of observations and limited sensitivity, GRACE cannot recover all spectral components evenly [see, e.g., *Seo et al.*, 2008]. On the other hand, the spectral power of the gravity field change is not distributed evenly at all wavelengths. Consequently, one group of the spectral components is more sensitive to the signal than the other group. The latter group typically has a low signal-noise ratio that does not allow a direct extraction of the signal. Thus, a spectral component-specific filtering should be applied for the monthly GRACE models. The improvement by the two-step method shown in Figure 6 suggests that this spectral component-specific method is effective in extracting the GRACE signal. Further, the agreement between the GRACE-equivalent-filtering results and the true values shown in Figure 7 suggests that no re-scaling is necessary for the two-step method because the Gaussian filtering is only applied for the low-signal components. However, this conclusion may not be applicable for other basins. A basin-specific analysis is required to quantify the scaling factor.

[100] The least squares estimation or fitting has been used to compute the trend of spherical coefficients in terms of equation (6). A similar parameter equation has also been used to estimate the annual trend of the TWS, SMSL storage and GWS. All results are listed in Tables 2, 3, and 4. Of them, the GWS trend is estimated after the SMSL storage has been removed from the GRACE TWS. They suggest that the estimation process for the trend is nonlinear, i.e., the trend of GW is not equal to the difference of the TWS and SMSL trends because of the consideration of GRACE coefficient errors in the least squares estimation. What is interesting is that the CSR and GFZ GWS trend rates show a slightly better agreement than their respective TWS rates when using AVG10 and WGHM. An explanation is that the TWS variations contain stronger seasonal and inter-annual cycles than the GWS variations. The removal of these cycles leads to a better estimation of the trend rate [see *Tregoning et al.*, 2009]. A longer GRACE time series may help to substantiate this point.

7. Summary

[101] Groundwater is a primary hydrological reservoir of the Great Lakes Water Basin. However, there is a knowledge gap about its storage variation and interaction with the Great Lakes due to insufficient observations. The Gravity Recovery And Climate Experiment (GRACE) satellite is a new type of remote sensor that can infer changes in terrestrial total water mass storage from its monthly gravity solutions. The objective of this study is to examine the detectability of the groundwater storage change within the Great Lakes Water Basin using GRACE. To achieve this objective, we have looked into each essential step of the process to estimate the groundwater storage variation from the GRACE monthly gravity models.

[102] First, a two-step filtering method has been developed to optimize the extraction of GRACE signal. It takes advantages of the effectiveness of the de-striping, statistical test, and non-isotropic methods while minimizing the possible signal loss associated with these methods. Numerical comparisons demonstrate that the two-step method gives significantly more details in the filtered GRACE results, and more importantly, the two-step method results in higher correlation between GRACE and simulated GRACE data.

[103] Secondly, a two-dimensional window weight function has been suggested to integrate the water storage within the Great Lakes Basin (GLB). It performs better than the exact averaging weight function by reducing ringing artifacts. A simulation analysis based on the surface water (lake water, soil moisture and snow) storage for the period of study suggests that the GRACE omission error is insignificant in estimating the total water storage variation for the GLB. It also indicates that the use of the two-step filtering method can avoid the rescaling by the Gaussian filtering.

[104] Thirdly, the total water storage (TWS) variations have been estimated for the period of 2002 to 2009. These TWS variations are derived by correcting for the significant glacial isostatic adjustment (GIA) effect on the GRACE total mass change. The TWS estimates from both CSR's and GFZ's monthly models show dominant annual cycles which highly correlate with the total surface water storage (soil moisture, snow and lake, or SMSL) variations. They show a declining trend of 6.4 ± 1.4 and 4.9 ± 1.2 km^3/yr in volume for CSR's and GFZ's models, respectively. The two derived TWS trend maps from both CSR's and GFZ's monthly models show a dominant water loss over the GLB with the maximum loss rate around Lake Superior at about 20 mm/yr in water thickness equivalent (WTE). The two maps show a difference of about 5 mm/yr over Lakes Michigan and Huron.

[105] Fourthly, the groundwater storage (GWS) variations have been estimated for the period of study from the GRACE TWS and SMSL storage in terms of the water balance relation, i.e. TWS = SMSL + GW. The average GWS variations clearly show annual cycles with an amplitude of about 30 to 90 mm depending on the LSM being

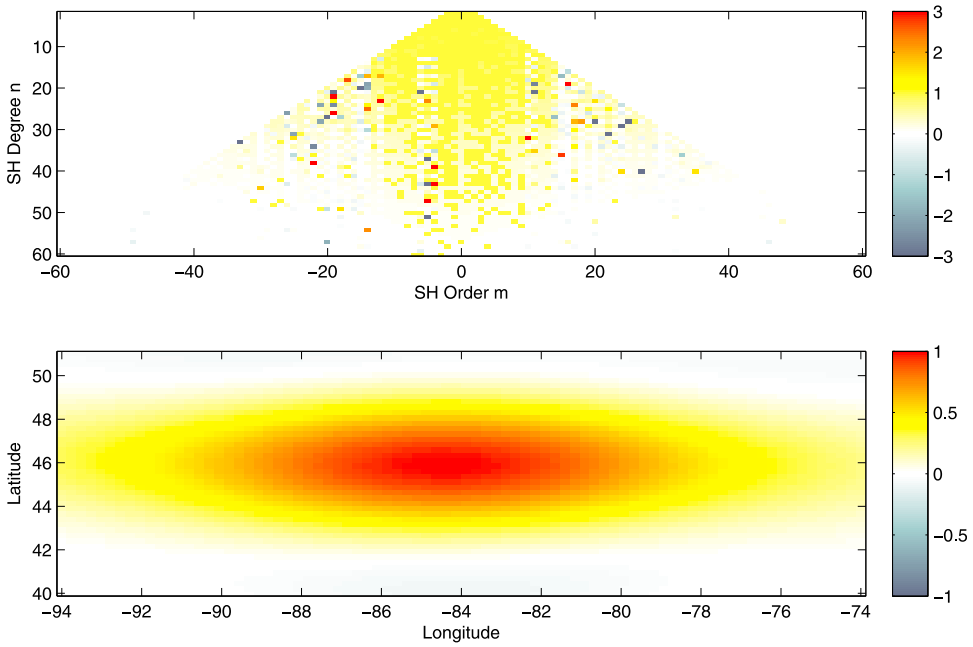

Figure A1. (top) Equivalent filter coefficients f_{nm}. Minimum and maximum values range from −3 and 3. Coefficients outside this range are displayed as the minimum and maximum values, respectively. (bottom) Equivalent filter function in spatial domain. The computational point is located at the center of map. All values are normalized by the maximum function value in the region of computation.

used. The GWS variations have a phase shift of about half a year with respect to the SMSL variations. The GWS peaks in September and October and has a minimum in March and April while the SMSL peaks are opposite. The TWS variations have a close phase to the SMSL storage variations which dominates the TWS. The least squares estimation gives the loss for GWS trend of 5.2 ± 0.9 and 3.6 ± 1.1 km^3/yr from CSR's and GFZ's GRACE models and the NOAH025 land surface model. In contrast, the TWS change loss trend is estimated at 6.4 ± 1.4 and 4.9 ± 1.2 km^3/yr, respectively. They suggest that the GWS loss dominates the TWS loss in the GLB for the period of study. Using an average GLDAS model and WGHM also gives rise to the loss trend of GWS but with different magnitudes due to the difference of the land surface model output.

[106] The GWS trend maps are derived from both CSR's and GFZ's models and three soil moisture models: NOAH025, AVG10 and WGHM. Six trend maps are dominantly negative indicating a net water mass loss. Broad patterns in these maps are similar, and somewhat correlate with the shape of the GLB. However, there exist significant differences in detail among these maps, which only give broad features of the GWS loss due to the limited resolution of GRACE and the uncertainty of LSMs which makes interpretation at a specific location unrealistic without other local hydrological information.

[107] Finally, the uncertainty of the GRACE GWS result is caused largely by the LSMs, and to a lesser extent by the GRACE commission and omission errors, and the GIA model error. This study shows that the GIA effect can be effectively removed based on the latest models over the GLB. With about 8 years of GRACE monthly models, the average GWS change trend for the GLB can be detected with a standard error of about ±2.0 mm/yr in WTE that is equivalent to 1.0 km^3/yr in water volume ignoring the GIA correction error. A longer GRACE observation allows a better determination of the long-term GWS change in the GLB.

Appendix A: Two-Step Filtering Function

[108] The two-step filter function can be constructed as

$$f(\phi,\lambda,\phi',\lambda',t_i)=\sum_{n=0}^{N}\sum_{m=-n}^{n} f_{nm}(t_i)\overline{Y}_{nm}(\phi,\lambda)\overline{Y}_{nm}(\phi',\lambda') \quad \text{(A1)}$$

[109] Evidently, it is anisotropic and two-point symmetric in space and non-stationary in time. The filtered anomalous potential is accordingly written as

$$\Delta T^{\mathrm{f}}(\phi,\lambda,t_i)=\frac{1}{4\pi}\int\int_{\Omega'} f(\phi,\lambda,\phi',\lambda',t_i)\cdot\Delta T(\phi',\lambda',t_i)\cos\phi' d\lambda' d\phi' \quad \text{(A2)}$$

or equivalently

$$\Delta T^{\mathrm{f}}(\phi,\lambda,t_i)=\frac{GM}{R}\sum_{n=0}^{N}\sum_{m=-n}^{n} f_{nm}(t_i)\Delta\bar{C}_{nm}(t_i)\overline{Y}_{nm}(\phi,\lambda) \quad \text{(A3)}$$

[110] The filter function in equation (A1) is determined by both the de-striping and statistical test steps. Figure A1 displays the two-step filter in both spectral and spatial domains for the monthly model of December 2009. Equivalent filter coefficients f_{nm} = 1 (yellow pixels in the top panel) indicate that the corresponding $\Delta\bar{C}_{nm}$ remain unchanged, i.e., there are no significant striping (or correlation) errors being detected in them, and there are significant signals (either of linear, quadratic, annual and semiannual) being detected in

them. When f_{nm} are not equal to 1, the corresponding $\Delta\bar{C}_{nm}$ are either de-striped or failed in the statistical test, or both. Of them, those $\Delta\bar{C}_{nm}$ being de-striped, but passing the statistical test are not Gaussian filtered. Otherwise they are Gaussian filtered. The second step (statistical test) can be understood as a remove-restore process, i. e., those coefficients which are signal-dominant are identified and removed first, then the rest of the coefficients are smoothed by Gaussian filter, finally the signal-dominant coefficients are restored. This approach preserves the major part of signals while depressing only the noise part. It is suitable for the irregular distribution of error across all coefficients.

[111] An evident feature is that most of the higher order coefficients ($m > 20$) are either de-striped or failed in the statistical test, while a significant part of the higher-degree (up to 60) and lower-order coefficients remains unchanged in the two-step process. It suggests that GRACE models have higher spatial resolution in the north-south direction than the west-east direction at lower and mid latitude. The filter function shown in Figure A1 (bottom) is clearly wider in the west-east direction in the Great Lakes region which has a median latitude of 45.5°N even with the consideration of meridian convergence toward the North by a factor of cos ϕ.

[112] **Acknowledgments.** We are thankful to the Associate Editor and two anonymous reviewers for their constructive comments which have helped to improve this paper greatly. Discussions with Tony Lambert have been helpful. This study has been supported by the Groundwater Geoscience Program (see http://www.nrcan.gc.ca/earth-sciences/about/current-program/groundwater-geoscience/4106), Earth Science Sector (ESS), Natural Resources Canada (NRCan) and ESF-EUROCORES programme TOPO-EUROPE. We would like to thank the German Space Operations Center (GSOC) of the German Aerospace Center (DLR) for providing continuously and nearly 100% of the raw telemetry data of the twin GRACE satellites. We would also like to thank NASA GEC DISC for GLDAS models, and Heike Hoffmann-Dobrev for WGHM model output. This is ESS contribution 20110368.

References

Braun, A., C.-Y. Kuo, C. K. Shum, P. Wu, W. van der Wal, and G. Fotopoulos (2008), Glacial isostatic adjustment at the Laurentide ice sheet margin: Models and observations in the Great Lakes region, *J. Geodyn.*, *46*, 165–173.

Chambers, D. P. (2006), Evaluation of new GRACE time-variable gravity data over the ocean, *Geophys. Res. Lett.*, *33*, L17603, doi:10.1029/2006GL027296.

Chao, B. F. (2005), On inversion for mass distribution from global (time-variable) gravity field, *J. Geodyn.*, *39*, 223–230.

Chen, J. L., C. R. Wilson, J. S. Famiglietti, and M. Rodell (2007), Attenuation effect on seasonal basin-scale water storage changes from GRACE time-variable gravity, *J. Geod.*, *81*, 237–245, doi:10.1007/s00190-006-0104-2.

Chen, J. L., C. R. Wilson, B. D. Tapley, Z. L. Yang, and G. Y. Niu (2009a), 2005 drought event in the Amazon River basin as measured by GRACE and estimated by climate models, *J. Geophys. Res.*, *114*, B05404, doi:10.1029/2008JB006056.

Chen, J. L., C. R. Wilson, D. Blankenship, and B. D. Tapley (2009b), Accelerated Antarctic ice loss from satellite gravity measurements, *Nat. Geosci.*, *2*, 859–862, doi:10.1038/NGEO694.

Cheng, M. K., and J. Ries (2007), Monthly estimates of C20 from 5 SLR satellites, *Tech. Note 05*, GRACE Proj., Cent. for Space Res., Univ. of Tex. at Austin, Austin.

Davis, J. L., M. E. Tamisiea, P. Elósegui, J. X. Mitrovica, and E. M. Hill (2008), A statistical filtering approach for Gravity Recovery and Climate Experiment (GRACE) gravity data, *J. Geophys. Res.*, *113*, B04410, doi:10.1029/2007JB005043.

Döll, P., K. Frank, and L. Bernhard (2003), A global hydrological model for deriving water availability indicators: Model running and validation, *J. Hydrol.*, *270*, 105–134.

Döll, P., H. Hoffmann-Dobrev, F. T. Portmann, S. Siebert, A. Eicker, M. Rodell, G. Strassberg, and B. R. Scanlon (2012), Impact of water withdrawals from groundwater and surface water on continental water storage variations, *J. Geodyn.*, doi:10.1016/j.jog.2011.05.001, in press.

Famiglietti, J. S., M. Lo, S. L. Ho, J. Bethune, K. J. Anderson, T. H. Syed, S. C. Swenson, C. R. de Linage, and M. Rodell (2011), Satellites measure recent rates of groundwater depletion in California's central valley, *Geophys. Res. Lett.*, *38*, L03403, doi:10.1029/2010GL046442.

Fan, Y., and H. M. van den Dool (2004), Climate Prediction Center global monthly soil moisture data set at 0.5 resolution for 1948 to present, *J. Geophys. Res.*, *109*, D10102, doi:10.1029/2003JD004345.

Fan, Y., H. M. van den Dool, and W. Wu (2011), Verification and intercomparison of multimodel simulated land surface hydrological datasets over the United States, *J. Hydrometeorol.*, *12*, 531–555, doi:10.1175/2011JHM1317.1.

Georgakakos, K. P., and O. W. Baumer (1996), Measurement and utilization of on-site soil moisture data, *J. Hydrol.*, *184*, 131–152.

Grannemann, N. G., R. J. Hunt, J. R. Nicholas, T. E. Reilly, and T. C. Winter (2000), The importance of groundwater in the Great Lakes region, *U.S. Geol. Surv. Water Resour. Invest. Rep.*, *00-4008*, 19 pp.

Han, S. C., C. K. Shum, C. Jekeli, C. Y. Kuo, C. Wilson, and K. W. Seo (2005), Non-isotropic filtering of GRACE temporal gravity for geohysical signal enhancement, *Geophys. J. Int.*, *163*(1), 18–25.

Harris, F. J. (1978), On the use of windows for harmonic analysis with the discrete Fourier transform, *Proc. IEEE*, *66*(1), 51–83.

Henry, C. M., D. M. Allen, and J. Huang (2011), Groundwater storage variability and annual recharge using well-hydrograph and GRACE satellite data, *Hydrogeol. J.*, *19*, 741–755, doi:10.1007/s10040-011-0724-3.

Holtschlag, D. J., and J. R. Nicholas (1998), Indirect ground-discharge to the Great Lakes, *U.S. Geol. Surv. Open File Rep.*, *98–579*, 25 pp.

Huang, J., M. Véronneau, and A. Mainville (2008), Assessment of systematic errors in the surface gravity anomalies over North America using the GRACE gravity model, *Geophys. J. Int.*, *175*, 45–64, doi:10.1111/j.1365-246X.2008.03924.x.

Jekeli, C. (1981), Alternative methods to smooth the Earth's gravity field, *Rep. 327*, Dep. of Geod. Sci. and Surv., Ohio State Univ., Columbus.

Mainville, A., and M. R. Craymer (2005), Present-day tilting of the Great Lakes region based on water level gauges, *Geol. Soc. Am. Bull.*, *117*(7/8), 1070–1080, doi:10.1130/B25392.1.

Meredith, D. D. (1975), Temperature effects on Great Lakes water balance studies, *Water Resour. Bull.*, *11*(1), 60–68.

Moran, M. S., C. D. Peters-Lidard, J. M. Watts, and S. McElroy (2004), Estimating soil moisture at the watershed scale with satellite-based radar and land surface models, *Can. J. Remote Sens.*, *30*(5), 805–826.

National Groundwater Association (2007), *Groundwater Use in Ontario: Fact Sheet*, Westerville, Ohio.

Neff, B. P., A. R. Piggott, and R. A. Sheets (2005), Estimation of shallow groundwater recharge in the Great Lakes Basin, *U.S. Geol. Surv. Sci. Invest. Rep.*, *2005-5284*, 20 pp.

Paulson, A., S. Zhong, and J. Wahr (2007), Inference of mantle viscosity from GRACE and relative sea level data, *Geophys. J. Int.*, *171*, 497–508, doi:10.1111/j.1365-246X.2007.03556.x.

Peltier, W. R. (1994), Ice Age paleotopography, *Science*, *265*, 195–201.

Peltier, W. R. (2004), Global glacial isostasy and the surface of the ice-age Earth: The ICE-5G (VM2) model and GRACE, *Annu. Rev. Earth Planet. Sci.*, *32*, 111–149.

Reeves, H. W. (2010), Water availability and use pilot: A multiscale assessment in the U.S. Great Lakes Basin, *U.S. Geol. Surv. Prof. Pap.*, *1778*, 105 pp.

Rivera, A. (2008), Groundwater sustainable development in Canada—Emerging issues, *Geosci. Canada*, *35*(2), 73–87.

Rodell, M., et al. (2004), The global land data assimilation system, *Bull. Am. Meteorol. Soc.*, *85*(3), 381–394.

Rodell, M., J. Chen, H. Kato, J. S. Famiglietti, J. Nigro, and C. R. Wilson (2007), Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE, *Hydrogeol. J.*, *15*, 159–166.

Rodell, M., I. Velicogna, and J. S. Famiglietti (2009), Satellite-based estimates of groundwater depletion in India, *Nature*, *460*, 999–1002, doi:10.1038/nature08238.

Rutherford, S. (2004), Groundwater use in Canada, report, West Coast Environ. Law, Vancouver, B. C., Canada.

Schwab, D. J., G. A. Leshkevich, and G. C. Muhr (1999), Automated mapping of surface water temperature in the Great Lakes, *J. Great Lakes Res.*, *25*(3), 468–481.

Sella, G. F., S. Stein, T. H. Dixon, M. Craymer, T. S. James, S. Mazzotti, and R. K. Dokka (2007), Observation of glacial isostatic adjustment in "stable" North America with GPS, *Geophys. Res. Lett.*, *34*, L02306, doi:10.1029/2006GL027081.

Seo, K. W., C. R. Wilson, J. L. Chen, and D. E. Waliser (2008), GRACE's spatial aliasing error, *Geophys. J. Int.*, *172*, 41–48, doi:10.1111/j.1365-246X.2007.03611.x.

Solley, W. B., R. R. Pierce, and H. A. Perlman (1998), Estimated use of water in the United States in 1995, *U.S. Geol. Surv. Circ.*, *1200*, 71 pp.

Strassberg, G., B. R. Scanlon, and D. Chambers (2009), Evaluation of groundwater storage monitoring with the GRACE satellite: Case study of the High Plains aquifer, central United States, *Water Resour. Res.*, *45*, W05410, doi:10.1029/2008WR006892.

Swenson, S., and J. Wahr (2002), Methods for inferring regional surface-mass anomalies from Gravity Recovery and Climate Experiment (GRACE) measurements of time-variable gravity, *J. Geophys. Res.*, *107*(B9), 2193, doi:10.1029/2001JB000576.

Swenson, S., and J. Wahr (2006), Post-processing removal of correlated errors in GRACE data, *Geophys. Res. Lett.*, *33*, L08402, doi:10.1029/2005GL025285.

Swenson, S., J. S. Famiglietti, J. Basara, and J. Wahr (2008), Estimating profile soil moisture and groundwater variations using GRACE and Oklahoma Mesonet soil moisture data, *Water Resour. Res.*, *44*, W01413, doi:10.1029/2007WR006057.

Tapley, B. D., S. Bettadpur, M. Watkins, and C. Reigber (2004), The gravity recovery and climate experiment: Mission overview and early results, *Geophys. Res. Lett.*, *31*, L09607, doi:10.1029/2004GL019920.

Tiwari, V. M., J. Wahr, and S. C. Swenson (2009), Dwindling groundwater resources in northern India, from satellite gravity observations, *Geophys. Res. Lett.*, *36*, L18401, doi:10.1029/2009GL039401.

Tregoning, P., G. Ramillien, H. McQueen, and D. Zwartz (2009), Glacial isostatic adjustment and nonstationary signals observed by GRACE, *J. Geophys. Res.*, *114*, B06406, doi:10.1029/2008JB006161.

Tushingham, A. M., and W. R. Peltier (1991), ICE-3G: A new global model of late Pleistocene deglaciation based upon geophysical predictions of post-glacial relative sea level change, *J. Geophys. Res.*, *96*(B3), 4497–4523.

van der Wal, W. (2009), Contributions of space gravimetry to postglacial rebound modeling with different rheologies, PhD thesis, Univ. of Calgary, Calgary, Alberta, Canada.

van der Wal, W., A. Braun, P. Wu, and M. G. Sideris (2009), Prediction of decadal slope changes in Canada by glacial isostatic adjustment modelling, *Can. J. Earth Sci.*, *46*, 587–595.

Velicogna, I., and J. Wahr (2006), Acceleration of Greenland ice mass loss in spring 2004, *Nature*, *443*(7109), 329–331.

Wagner, C., D. McAdoo, J. Klokocnik, and J. Kostelecky (2006), Degradation of geopotential recovery from short repeat-cycle orbits: application to GRACE monthly fields, *J. Geod.*, *80*, 94–103.

Wahr, J., M. Molenaar, and F. Bryan (1998), Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE, *J. Geophys. Res.*, *103*(B12), 30,205–30,229.

Werth, S., A. Güntner, R. Schmidt, and J. Kusche (2009), Evaluation of GRACE filter tools from a hydrological perspective, *Geophys. J. Int.*, *179*, 1499–1515, doi:10.1111/j.1365-246X.2009.04355.x.

Yeh, P. J.-F., S. C. Swenson, J. S. Famiglietti, and M. Rodell (2006), Remote sensing of groundwater storage changes in Illinois using the Gravity Recovery and Climate Experiment (GRACE), *Water Resour. Res.*, *42*, W12203, doi:10.1029/2006WR005374.