

IMPLEMENTATION OF THE DISTRIBUTED PARALLEL PROGRAM FOR GEOID HEIGHTS COMPUTATION USING MPI AND OPENMP

Seongkyu Lee^a, Jinsoo Kim^b, Yonghwa Jung^a, Jisun Choi^a, Chuluong Choi^{a,*}

^a Dept. of Spatial Information Engineering, Pukyong National University, 45, Yongso-ro, Nam-Gu, Busan, South Korea
- geoslegend@gmail.com, invitation7@nate.com, sun33023@hotmail.com, cuchoi@pknu.ac.kr

^b ZEN21, 2nd Floor, RNC building, 981-1 Bangbae 3-dong, Seocho-gu, Seoul 137-848, South Korea -
pknu9680@gmail.com

Commission IV, WG IV/5

KEY WORDS: Distributed parallel computing, Geoid heights computation, MPI, OpenMP, Diskless-based PC Cluster

ABSTRACT:

Much research have been carried out using optimization algorithms for developing high-performance program, under the parallel computing environment with the evolution of the computer hardware technology such as dual-core processor and so on. Then, the studies by the parallel computing in geodesy and surveying fields are not so many. The present study aims to reduce running time for the geoid heights computation and carrying out least-squares collocation to improve its accuracy using distributed parallel technology. A distributed parallel program was developed in which a multi-core CPU-based PC cluster was adopted using MPI and OpenMP library. Geoid heights were calculated by the spherical harmonic analysis using the earth geopotential model of the National Geospatial-Intelligence Agency(2008). The geoid heights around the Korean Peninsula were calculated and tested in diskless-based PC cluster environment. As results, for the computing geoid heights by a earth geopotential model, the distributed parallel program was confirmed more effective to reduce the computational time compared to the sequential program.

1. INTRODUCTION

These days, global navigation satellite systems (GNSS) have been widely used and geoid is a very important element which makes it possible to determine orthometric height using GNSS positioning. Moreover, many countries have continued to make efforts to develop their own high-precision geoid model (Chandler and Merry, 2010; Kuroishi et al., 2002; Roman et al. 2009; Toth et al. 2000)

Since global geopotential models (GGMs) well expresses long wavelength component of the Earth's gravity field (Daho et al. 2008; Krynski and Lyszkowicz 2006), they not only provide basis for the gravity field to develop high-precision geoid model but also have important meanings as reference surface for calculating a local geoid (Bae et al., 2011; Dawod et al., 2010). In addition, countries which are difficult to develop geoid models have been using GGMs for rough calculation of geoid heights (or undulations) and gravity anomaly through spherical harmonic analysis (SHA) (Lee et al., 2008)

Geopotential data that is recently observed by GRACE satellite has contributed to more precise GGMs development and especially these (geopotential data) have been greatly contributed to development of Earth gravitational model 2008 (EGM2008). EGM2008 is complete to degree and order 2,159 and includes additional spherical harmonic coefficients extending to degree 2,190 and order 2,159 (Pavlis et al., 2008; Pavlis et al., 2012).

Moreover, because GGMs-derived geoid heights are referred to a global vertical datum and these are fitted with a local vertical

datum by GPS/levelling-derived geoid heights, the accuracy of these (geoid heights) could be enhanced. For this purpose, least-squares collocation (LSC) method is widely used (Kotsakis et al., 2008; Lee et al., 2008).

As mentioned above, although SHA and LSC are mostly used to evaluate accuracy and suitability of ultra-high degree geopotential models for certain local areas or countries, they all contain time-consuming problem. In geoscience simulations, to solve this problem, resources are distributed parallel computation is getting popularity to reduce data processing time. This paper is intended to carry out LSC to fit ultra-high degree SHA by EGM2008 using distributed parallel computation on PC cluster and present its performance evaluation.

2. DATA

2.1 EGM2008 geopotential model

EGM2008 is a spherical harmonic model of the Earth's gravitational potential, developed by U.S. National Geospatial-Intelligence Agency (NGA) (Lemoine, 1996; Pavlis et al., 2008; Pavlis et al., 2012).

EGM2008 is the first-ever global model that is capable of resolving the Earth's gravity field beyond spherical harmonic degree 2000. EGM96 that is a widely used GGM has spatial resolution of 30 arc minutes whereas EGM2008 is a high-resolution GGM of Earth's gravity field that allows computation of geoid heights down to a resolution of 5 arc minutes (Dawod et al., 2010; Pavlis et al. 2008).

* Corresponding author.

Especially, unlike conventional GGMs, EGM2008 contains gravity data of 15' resolution for Asian regions and also gravity data calculated from SRTM terrestrial data. Therefore, it is known to be 3-6 more accurate than a previous model, EGM96 (Pavlis et al., 2008; Pavlis et al., 2012).

2.2 GPS/levelling data

GPS/levelling data is a very important factor not only to evaluate accuracy of the developed gravimetric geoid models but also fit gravimetric geoid models to local vertical datum. And accuracy for GGMs is evaluated primarily by cross-validation between geoid heights calculated from SHA and geometric geoid heights calculated from GPS/levelling data. This paper, in order to evaluate the accuracy of EGM2008-derived geoid heights and to fit the EGM2008-derived geoid heights to local vertical datum, used a total of 1,175 points of the GPS/levelling data which is acquired from unified control points (UCPs) installed with 10 km spacing in South Korea (Figure 1).

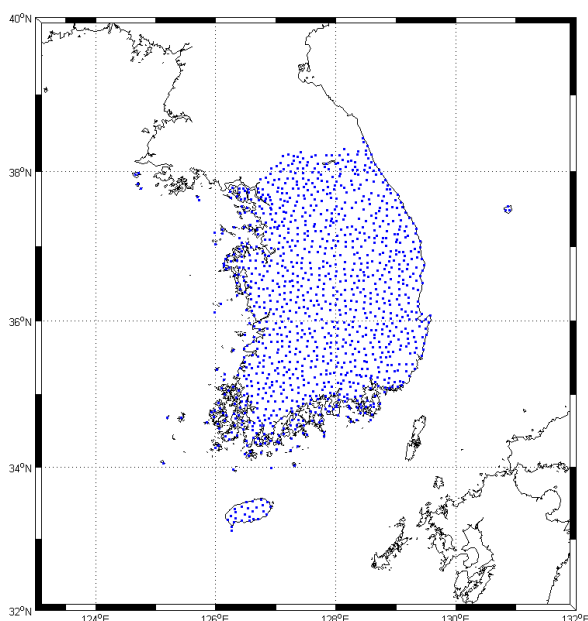


Figure 1. Distribution of a total of 1175 points of the GPS/levelling data (blue square)

3. METHODS

3.1 Spherical harmonic analysis

SHA is a process of decomposing a function on a sphere into components of various wavelengths using surface spherical harmonic as base functions. Newtonian gravitational potential V satisfies Poisson's equation and gravitational potential satisfies Laplace's equation in free space. Solution of Laplace's equation can be expressed by harmonic functions series. And disturbing potential is calculated and then it substitutes Bruns's formula and finally geoid heights N is calculated. Further details on this spherical harmonic analysis are given in Rapp and Pavlis (1990), Heiskanen and Moritz (1996), and Kenyon and Pavlis (1996).

3.2 Least-squares collocation

Least-squares collocation in Gravity field addresses correlation between signal and noise given by gravity data, and using this relation predicts signal where observation is not carried out. The LSC represents one of the major theoretical and practical foundations of modern physical geodesy (Tscherning, 1986; Tsaoussi, 1989).

In this paper, we used LSC to fit the EGM2008-derived geoid heights with geometric geoid heights calculated from GPS/levelling data. The most important rule to apply LSC is that data needs to be centered before it is collocated (Moritz 1980). In other words, trend has to be removed from the raw data such that mean of the data would be equal to zero (Forsberg et al., 2003). Procedure to remove this trend can be obtained by applying various trend models to raw data.

The introduction of the GRAVSOFIT package programmed by Forsberg et al. (2003) was a turning point in LSC application and adoption (Darbeheshti, 2009). In this paper, in order to fit geoid surface using LSC, GEOIP and GEOGRID in the GRAVSOFIT written by Fortran 77 language are used and program of GCOMB is ported by C++ language and then used. Further details on this spherical harmonic analysis are given in Forsberg et al. (2003) and Darbeheshti (2009).

3.3 MPI and OpenMP programming

Generally, parallel programming is classified as distributed programming using the message-passing interface (MPI) and multi-core CPU-based parallel programming using the open multi-processing (OpenMP).

MPI is a standardization for message passing and currently the *de facto* standard and the parallelization method for developing the high-performance computing (HPC) applications on distributed memory architecture (Message Passing Interface Forum, 1994; Wittwer, 2006; Chorley and Walker, 2010). The message-passing programming model is based on the abstraction of a parallel computer with a distributed address space where each processor has a local memory to which it has exclusive access (Rauber and Rünger, 2010). However, in the case of the communication of high volume data in the low-performance distributed computing environment which consists of low-speed network environment, MPI application's performance can be reduced.

OpenMP is a standardization and an application programming interface (API) that supports multi-platform shared memory multi-processing programming. It consists of a set of compiler directives, library routines, and environment variables that influence run-time behavior (Chapman et al, 2007; Refianti et al., 2011). Many of the major compiler vendors (e.g. Microsoft, Intel, HP) support OpenMP technology. And one of advantages of the OpenMP technology is to allow relatively fast development of parallel applications through the global access of application memory address space (Jin et al., 2011). Especially, it is available in multiple operating system environments. However, the OpenMP technology does not support the distributed computing environment.

In order to overcome this problem, researches to solve time-consuming problem through hybrid MPI and OpenMP approach which combined MPI and OpenMP have been recently proceeding.

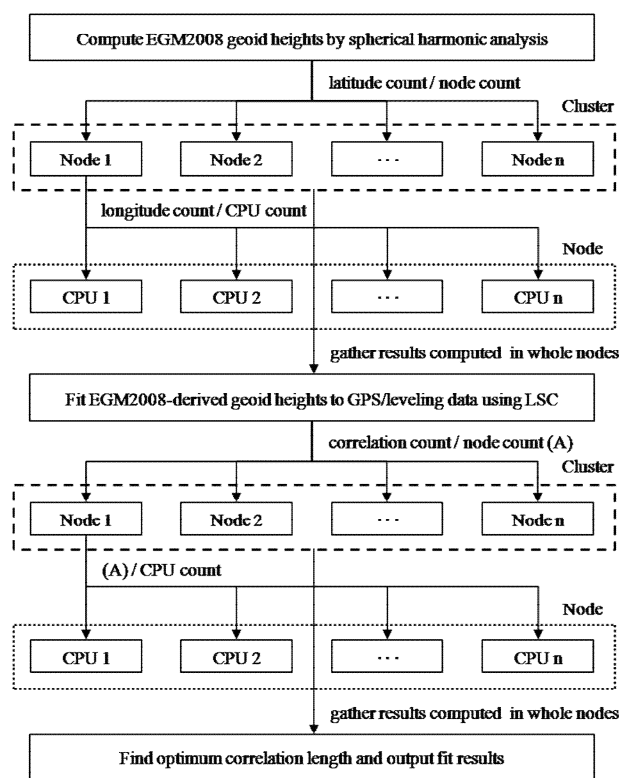


Figure 2. Parallelization model flow chart of SHA and LSC

4. PARALLELIZATION MODELING AND IMPLEMENTATION

4.1 Parallelization modelling

This research performed parallelization modeling by applying hybrid approach to analyze accuracy of geoid heights that is finally determined by fitting the EGM2008-derived geoid heights to geometric geoid heights calculated from GPS/levelling data and determine optimum correlation length. The parallelization modeling consists of SHA and LSC (Figure 2). And a task of the parallelization model is separated into one cluster master and cluster nodes. The cluster master performs to allocate tasks and gather results in each node. And the cluster nodes perform tasks of allocated regions.

First, the master node is to allocate EGM2008 coefficients, working scopes to compute EGM2008-derived geoid heights to each cluster node. Here, generally, working scopes used for distributed parallel processing are divided with grid shapes but their associated Legendre function values are same in same latitude (Xiao and lu, 2007). Therefore, the working scopes can be divided based on latitude. Next, the master node performs to gather EGM2008-derived geoid heights and then distributes the EGM2008-derived geoid heights and working scopes for LSC tasks to each cluster node. Next, each cluster node carries out LSC fitting using the EGM2008-derived geoid heights, geometric geoid heights calculated from GPS/levelling data and correlation length and explores correlation length which has minimum standard deviation at corresponding node. Finally, the cluster master performs to gather optimum correlation length in each node and then chooses final optimum correlation length. And hybrid MPI and OpenMP application to carry out SHA and LSC is ported so that it can be executed on Linux platform

written with Fortran 77 language (Forsberg et al., 2003; Rapp, 1982; Yecai, 1994). These are implemented by C++ using MPICH2 and OpenMP library and also compiled by GNU C Compiler.

4.2 Computing platform

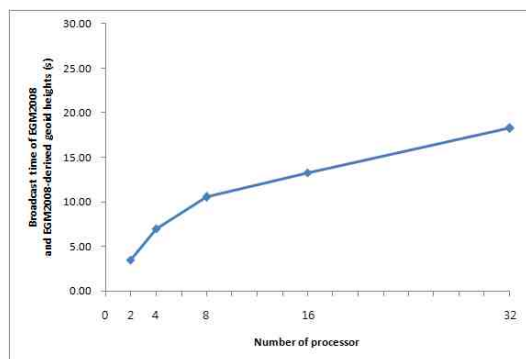
Diskless-based PC cluster system is implemented for distributed parallel computing using hybrid MPI and OpenMP approach proposed in this paper (Table 1). This cluster system consists of 1 cluster master and 16 cluster nodes, and its network environments is made of 100MB bps switch hub. Operating system software for cluster master used Community ENTERprise Operating System (CentOS) 6.2 x86 64bit, with GCC 4.4.6 compiler used for compilation of code and MPI library used MPICH2 1.2.1 (<http://www.mcs.anl.gov/mpi/mpich2>). Cluster nodes are structured to be driven by diskless method using Perceus 1.6.1 (<http://www.perceus.org/>) which is a program package made by Infiscale for driving diskless cluster while each node of cluster system is structured in a way that OS image stored at cluster master is connected by PXE network booting method.

Table 1. Summary of the diskless-based PC cluster system

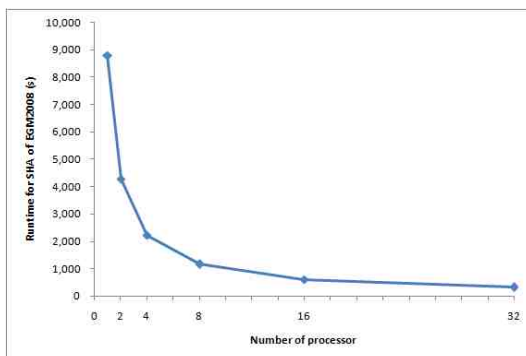
	Cluster Master			Cluster node		
No. of nodes	1			16		
Processor Model	Intel® 2120	Core™ i3-		Intel® Pentium® CPU E2200	Dual	
# of Cores	2			2		
# of Threads	4			2		
Processor speed	3.3 GHz			2.2 GHz		
Memory	4 GB			2 GB		
Ethernet card	Realtek gigabit ethernet card					

5. RESULTS AND DISCUSSION

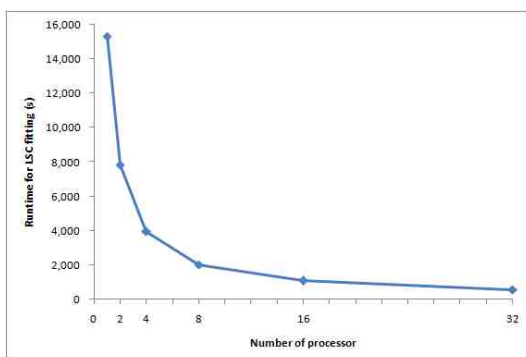
Experiment of the parallelization modeling in this paper is carried out in the diskless cluster system consisting of low-performance computers by hybrid MPI and OpenMP application. And this research carried out SHA of EGM2008 for around the Korean peninsula using parallelization modeling applying hybrid approach developed in this research. SHA is performed at latitude 32°N-43°N and longitude 123°E-132°E by applying maximum degree and order (2,190) of EGM2008. From the results of analysis, the EGM2008-derived geoid heights are calculated for a total of 357,601 grid points with 1 arc minute of interval for latitude and longitude respectively. And LSC was performed between 0.1 and 150 km by 0.1 km interval. And broadcast time of EGM2008 and the EGM2008-derived geoid heights runtime for SHA of EGM2008, runtime to find optimum correlation length, and total computation time are measured (Figure 3). 2 processors means 1 node and 1 processor means an environment without node i.e. test result by serial algorithm at single computer not at cluster system.



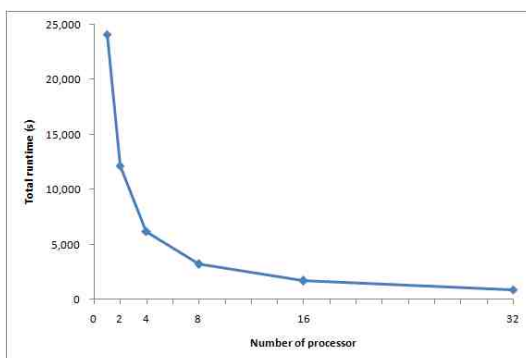
(a)



(b)



(c)



(d)

Figure 3. Running time of hybrid MPI and OpenMP application for computing of SHA and LSC. (a) broadcast time of EGM2008 and the EGM2008-derived geoid heights, (b) runtime for SHA of EGM2008, (c) runtime to find optimum correlation length using LSC fitting, and (d) total computation time.

As results, the broadcast time of EGM2008 and the EGM2008-derived geoid heights needed approximately 4 seconds at 1 node (2 processors) and approximately 18 seconds for maximum node (32 processors). The share (ratio) of broadcast time in total computation time showed almost 0%. This means that although it has almost no influence upon total computation time, it increased linearly depending on the number of nodes.

For SHA, the calculations were possible with approximately 3.6% of 1 processor. And LSC fitting was possible to be calculated with approximately 3.8% of 1 processor. The total computation time was approximately 6 hours and 40 minutes (0 node) and approximately 15 minutes (maximum node). This is similar to the result of SHA research using supercomputer, which means parallelization modeling for SHA and LSC fitting suggested by this paper is well executed in low-performance cluster system (Xiao and lu, 2007).

6. CONCLUSION

In this paper, we proposed spherical harmonic analysis used to evaluate suitability of GGMs and hybrid MP and OpenMP approach to reduce data processing time for least-squares collocation fitting.

Hybrid implementation using parallelization modeling showed more advanced result in calculation speed than serial implementation. Moreover, cluster system with low-performance not high-performance computing (HPC) like supercomputer performed SHA and LSC fitting more effectively.

7. REFERENCES

- Bae, T.S., Lee, J., Kwon, J.H. and Hong, C.K., 2011. Update of the precision geoid determination in Korea. *Geophysical Prospecting*, 60(3). pp. 555-571.
- Chandler, G. and Merry, C., 2010. The South African geoid 2010: SAGEOID10. PositionIT, pp. 29-33.
- Chapman, B., Jost, G. and Van der Pas, R., 2007. Using OpenMP. The MIT Press, Boston. pp. 1-34.
- Chorley, M.J. and Walker, D.W., 2010. Performance analysis of a hybrid MPI/OpenMP application on multi-core clusters. *Journal of Computational Science* 1, pp. 168-174.
- Daho, S.A.B., Fairhead, J.D., Zeggai, A., Ghezali, B., Derkaoui, A., Gourine, B. and Khelifa, S., 2008. New investigation on the choice of the tailored geopotential model for Algeria. *Journal of Geodynamics*, 45, pp. 154-162.
- Darbeheshti, N., 2009. Modification of the least-squares collocation method for non-stationary gravity field modelling. Ph.D. thesis, Curtin University of Technology, Perth, Western Australia
- Dawod, G.M., Mohamed, H.F. and Ismail, S.S., 2010. Evaluation and adaptation of the EGM2008 geopotential model along the Northern Nile Valley, Egypt: Case Study. *Journal of Surveying Engineering*, 136, pp. 36-40.
- Forsberg, R., Tscherning, C.C. and Knudsen, P., 2003. An overview manual of the GRAVSOF geodetic gravity field modelling programs. Kort & Matrikelstyrelsen.

- Heiskanen, W.A. and Moritz, H., 1996. Physical Geodesy. Institute of Physical Geodesy, Technical University Graz, Austria, pp. 46-125.
- Kotsakis, C., Katsambalos, K., Ampatzidis, D. and Gianniou, M., 2008. Evaluation of EGM2008 using GPS and leveling heights in Greece. In: *Gravity, Geoid and Earth Observation IAG Commission 2: Gravity Field*, Chania, Crete, Greece, 23-27 June. http://users.auth.gr/~kvek/Crete_Paper_2008.pdf (18 Jan. 2012).
- Jin, H., Jespersen, D., Mehrotra, P., Biswas, R., Huang, L. and Chapman, B., 2011. High performance computing using MPI and OpenMP on multi-core parallel systems. *Parallel Computing*, 37, pp. 562-575.
- Kenyon S.C. and Pavlis, N.K., 1996. The Development of a global gravity anomaly database used in the DMA/GSFC geopotential model. In: *International Symposium on Gravity, Geoid, and Marine Geodesy 1996 (GraGeoMar96)*, September 30-October 5, p.148.
- Kuroishi, Y., Ando, H. and Fukuda, Y., 2002. A new hybrid geoid model for Japan, GSIGEO2000. *Journal of Geodesy*, 76, pp. 428-436.
- Krynski, J. and Lyszkowicz, A., 2006. Regional quasi-geoid determination in the area of Poland. In: *The 5th FIG Regional Conference*, Accra, Ghana, pp. 1-17.
- Lee, S., Kim, J. and Kim, C., 2008. Evaluation of EGM2008 earth geopotential model using GPS/leveling data. *Journal the Korean Society for Geospatial Information System*, 16, pp. 117-126.
- Lemoine F.G., Smith, D.E., Smith, R., Kunz, L., Pavlis, E.C., Pavlis, N.K., Klosko, S.M., Chinn, D.S., Torrence, M.H., Williamson, R.G., Cox, C.M., Rachlin, K.E., Wang, Y.M., Kenyon, S.C., Salmen, R., Trimmer, R., Rapp, R.H. and Nerem, R.S., 1996. The Development of the NASA GSFC and DMA joint geopotential model. In: *International Symposium on Gravity, Geoid, and Marine Geodesy 1996 (GraGeoMar96)*, September 30-October 5, p.146.
- Message Passing Interface Forum, 1994. MPI: A message-passing interface standard, *International Journal of Supercomputer Application* 8 (3&4), pp. 159-416.
- Moritz, H. 1980. Advanced physical geodesy. Wichmann Verlag, Karlsruhe, Germany.
- Pavlis, N.K., Holmes, S.A., Kenyon, S.C. and Factor, J.K., 2008. An earth gravitational model to degree 2160: EGM2008. *the 2008 General Assembly of the European Geosciences Union*, Vienna, Austria, pp. 13-18.
- Pavlis, N.K., Holmes, S.A., Kenyon, S.C. and Factor, J.K., 2012. The development and evaluation of the earth gravitational model 2008 (EGM2008), *Journal of Geophysical Research*, doi:10.1029/2011JB008916, in press.
- Pavlis, N.K., Holmes, S.A., Kenyon, S.C. and Factor, J.K., 2008. An earth gravitational model to degree 2160: EGM2008. *the 2008 General Assembly of the European Geosciences Union*, Vienna, Austria, pp. 13-18.
- Rapp, R.H., 1982. A fortran program for the computation of gravimetric quantities from high degree spherical harmonic expansions, Reports of the depart of Geodetic science and surveying, Report No. 334, pp. 1-23.
- Rapp, R.H. and Pavlis, N.K., 1990. The development and analysis of geopotential coefficient models to spherical harmonic degree 360. *Journal of Geophysical Research*, 95(B13), pp. 21,885-21,911.
- Rauber, T. and Rünger, G., 2010. Parallel programming for multicore and cluster systems. Springer, Berlin Heidelberg, pp. 93-112.
- Refianti, R., Mutiara, A.B., Hasta and D.T., 2011. Workshare process of thread programming and MPI model on multicore architecture. *International Journal of Advanced Computer Science and Applications*, 2, pp. 99-107.
- Roman, D.R., Wang, Y.M., Saleh, J. and Li, X. 2009. National geoid height models for the United States: USGG2009 and GEOID09. *ACSM-MARLS-UCLS-WEPS Conference 2009*, Salt Lake City, UT USA.
- Toth, G.Y., Rozsa, S.Z., Andritsanos, V.D., Adam, A. and Tziavos, I.N. 2000. Towards a cm-geoid for Hungary: Recent efforts and results. *Physics and Chemistry of the Earth*, 25, pp. 47-52.
- Tscherning, C.C. 1986. Mathematical and numerical techniques in physical geodesy. *Lecture Notes in Earth Science*, 7, 1-47, Springer, Berlin/Heidelberg, Germany.
- Wittwer, T., 2006. An introduction to parallel programming. VSSD, Leegwaterstraat 42, 2628 CA Delft, The Netherlands. <http://mail.vssd.nl/hlf/9071301788.pdf> (18 Jan. 2012).
- Xiao, H. and Lu, Y., 2007. Parallel computation for spherical harmonic synthesis and analysis. *Computer & Geosciences*, 33, pp. 311-317.
- Yecai, L., 1994. Software GRAVT_GM. University of Calgary, Alberta, Canada.