

ТРУДЫ МЕЖДУНАРОДНОЙ КОНФЕРЕНЦИИ
Итоги Электронного Геофизического Года
3–6 июня 2009 • Переславль-Залесский, Россия

Доступ к мировой сети наземных геомагнитных данных и виртуальные геомагнитные обсерватории

В. Г. Петров¹ и В. О. Папаташвили²

Получено 22 октября 2009; принято 28 октября 2009; опубликовано 30 ноября 2009.

Традиционное понятие виртуальной обсерватории предполагает создание web-сервиса, который, получив от клиента запрос на данные, находит их, загружает и передает пользователю. Достоинством такого подхода является простота клиентской части и легкость поддержания сервиса, а недостатком – увеличение сетевого трафика и ограниченность самого http-протокола, не позволяющего реализовать эффективные методы обработки и представления данных. Ранее были предложены два варианта доступа к геомагнитным данным – виртуальная геомагнитная обсерватория с web-интерфейсом (web-ВГМО) и автономная ВГМО. Web-ВГМО использовала стандартный подход с загрузкой Java-апплета, автономная ВГМО полностью работала на клиентской машине. Опыт использования web-ВГМО показал, что она имеет две проблемы: 1) не все наземные геомагнитные данные могут быть перераспределены через web-ВГМО; 2) часто апплеты, использующие свои собственные протоколы, не работают в связи с ограничениями, вводимыми для обеспечения безопасности сети. Существующая сейчас автономная ВГМО написана на Fortran и Java, требует адаптации для каждой системы и перекомпиляции при любой модификации, сложна в поддержке и развитии. В работе описывается новая версия автономной ВГМО (персональная ВГМО), основанная на открытом языке Python. Python-программы могут выполняться на всех аппаратных и программных платформах без всяких изменений. Сетевой основой ПВГМО может являться web-сервер программного обеспечения ПВГМО, который позволяет всем загружать программы, поддерживает систему обновления программ и ссылок на источники данных. На основе такого сервера может возникнуть типичная для систем с открытым исходным кодом система коллективной разработки ПО, когда новые методы и алгоритмы, разработанные пользователями для своих задач, загружаются на сервер для общего пользования и постепенно включаются в состав ПВГМО. **КЛЮЧЕВЫЕ СЛОВА:** геомагнитные данные; виртуальные обсерватории; магнитные обсерватории.

Ссылка: Петров, В. Г. и В. О. Папаташвили (2009), Доступ к мировой сети наземных геомагнитных данных и виртуальные геомагнитные обсерватории, *Росс. ж. наук о Земле*, 11, RE1008, doi:10.2205/2009ES000355.

Концепция виртуальных обсерваторий зародилась на базе астрономических и солнечных данных и фактически является средством получения данных различных наблюдений за удаленным объектом на своем рабочем

¹Институт земного магнетизма, ионосферы и распространения радиоволн им. Н. В. Пушкова, Троицк, Московская обл., Россия

²Мичиганский Университет, Мичиган, США

месте, т.е. позволяет как бы иметь солнечную или астрономическую обсерваторию в собственном распоряжении. В этих дисциплинах расположение измерительных приборов (обсерватории) большого значения не имеет. В геомагнетизме ситуация несколько иная, обсерватория проводит измерения там, где расположена, и термин “виртуальная магнитная обсерватория”, по идее, должен обозначать способ получения данных о магнитном поле в некоем виртуальном месте, как например в работе [Pilipenko et al., 2002]. Однако термин “виртуальная обсерватория”, понимаемый как средство доступа к данным

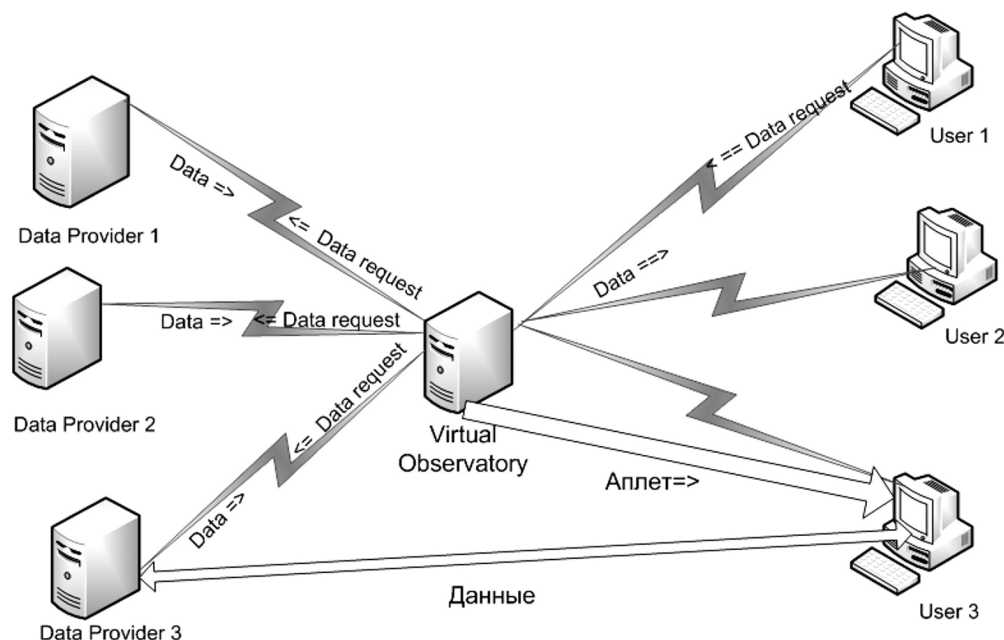


Рис. 1. Типичная схема виртуальной обсерватории.

любой магнитной обсерватории, применяется и в геомагнетизме.

Традиционное понятие виртуальной обсерватории предполагает создание web-сервиса, который, получив от клиента запрос на данные, обращается к известным серверу источникам исходных данных, загружает их, формирует требуемый ответ и возвращает его пользователю (Рис. 1). Достоинством такого подхода является простота клиентской части (обычно достаточно стандартного web-браузера) и легкость поддержания сервиса (в случае необходимости обновляется только программное обеспечение сервера). Недостатком является увеличение сетевого трафика (данные вначале загружаются на сервер, затем передаются клиенту), и ограниченность самого http-протокола, не позволяющего реализовать эффективные методы обработки и представления данных. Иногда данные из источников копируются на сервер заранее, и при запросе сразу формируется требуемый ответ, но такой подход не является виртуальной обсерваторией, а скорее представляет обычную базу данных с дистанционным доступом.

Широкому развитию виртуальных обсерваторий, понимаемых как средство доступа к данным любой магнитной обсерватории, препятствуют несколько обстоятельств. Во-первых, доступ ко многим данным ограничен и требует предварительного обращения к авторам данных с запросом конкретного интервала, и даже если доступ открыт, то правила использования данных запрещают предоставление этих данных третьим лицам, т.е. эти данные не могут быть представлены через портал виртуальной обсерватории. Во-вторых, http-протокол, исполь-

зуемый в браузерах, имеет ограниченные возможности. Эта проблема решается загрузкой на компьютер пользователя специальных программ – апплетов, которые устанавливают прямое соединение с сервером и используют свой собственный протокол обмена. Однако, в связи с ростом хакерских атак в сетях Интернета, сетевые администраторы всеми силами блокируют такие возможности. В третьих, такой подход не позволяет редактировать данные, хотя данные, особенно предварительные, могут иметь сбои и пропуски, которые должны быть исправлены перед построением графиков или анализом. И, наконец, при таком подходе трудно организовать работу с разнородными данными, находящимися на различных серверах. В этом случае пользователю все равно приходится загружать их на свой компьютер и самому конвертировать и объединять данные.

В связи с активным развитием Интернета за последние годы наблюдается быстрый рост объема геомагнитных данных, как организованных в специальные базы данных, так и данных индивидуальных обсерваторий и проектов, доступных через Интернет. В качестве примера таких баз можно привести базы данных Мировых центров данных, Интермагнет (включает примерно 100 обсерваторий), IMAGE (30 магнитометров), CARISMA (20 магнитометров), MAGDAS (до 50 магнитометров). Программное обеспечение таких центров обычно позволяет пользователю за выбранный интервал времени построить графики и/или загрузить файлы на свой компьютер. Если пользователю нужны данные из нескольких источников или интеграция магнитных данных с другими типами данных, то проблему преобразования форматов,

Табл. 1. Основные источники цифровых геомагнитных данных

Data source	Reference	N of Obs.	Access auto	Access manual	Redistrib
WDC Edinburgh	http://www.wdc.bgs.ac.uk/catalog/master.html ftp://ftp.nmh.ac.uk/pub/wdc/obsdata/1minval	110–120	Yes	Yes	Yes
WDC Kioto	http://swdcwww.kugi.kyoto-u.ac.jp/caplot/index.html	10–15	Yes	Yes	Yes
WDC Moscow	http://www.wdcb.ru/stp/data/geo_min.val/	5–10	Yes	Yes	Yes
NGDC NOAA's National Geophys. Data	http://www.ngdc.noaa.gov/geomag/data.shtml	30–40	No	Yes	Yes
NGDC NOAA's Nat. Geophys. Data Center	http://www.ngdc.noaa.gov/geomag/data.shtml	30–40	No	Yes	No
Intermagnet Definitive Data	http://ottawa.intermagnet.org/apps/dl_data_def_e.php	95–105	No(Yes?)	Yes	No
Intermagnet Preliminary Data	http://www.ngdc.noaa.gov/geomag/data.shtml	80–85	No(Yes?)	Yes	No
IMAGE Internation. Monitor for Auroral Geomag. Effects GM-210	http://www.geo.fmi.fi/image/reqform/dataform_month.html http://stdb2.stelab.nagoya-u.ac.jp/mm210/1-min_data	25–30 15–20	No Yes	Yes Yes	No No
MACCS Magnetom. Array for Cusp and Cleft Study	http://space.augsburg.edu/space/MaccsHome.html	2–5	Yes	Yes	No
SAMNET The UK Sub-Auroral Magnetom. Network	http://www.dcs.lancs.ac.uk/iono/samnet/	10–15	No	Yes	No
GIMA Geophysical Institute Magnetometer	http://magnet.gi.alaska.edu/access_data_back.html	8–12	No	Yes	No
PURAES Project for Upgrading Russian AE Stations	http://magnet.gi.alaska.edu/onemin_data.html	6	No	Yes	No
SAMBA, 1998–2004	http://samba.atmos.ucla.edu/map.html	11	No	Yes	No
AUTUMN Athabasca University THEMIS UCLA Magnetometer Network	http://autumn.athabascau.ca/index_e.php	10–12	No	Yes	No
PENGUIN (AGO)	http://www.sos.siena.edu/antarctic/PENGUIn_Program/	4–6	No	Yes	No
Canadian data	http://geomag.nrcan.gc.ca/common_apps/dl-eng.php	10–15	No	Yes	No
MIST@VT Greenland/Antarctic Geomagnetic Data	http://mist.nianet.org/webff/magdb_mist.html	1–5	Yes	Yes	No
MAGDAS Magnetic Data Greenland magnetometers	http://magdas.serc.kyushu-u.ac.jp/ http://dmiweb.dmi.dk/fsweb/projects/chain/ email: jwf@dmi.dk	30–37 15–20	No No	No No	No No
SPIDR	http://clust1.wdcb.ru/spidr/index.jsp	?	No	Yes	No
Syowa	http://polaris.nipr.ac.jp/~aurora/syowa.magne/magne.main.html	1	No	Yes	No

объединения данных, построения общих графиков пользователю все равно приходится решать самому.

В Табл. 1 представлены основные базы геомагнитных данных по состоянию на середину 2009 г. В этой таблице Access auto означает, что данные можно получить через автоматически сгенерированный скрипт, Access manual – необходимо ручное заполнение форм на

сервере, Redistribution – возможна или нет передача данных третьим лицам.

Данные одной и той же обсерватории могут присутствовать в нескольких проектах одновременно, и максимальное количество доступных обсерваторий значительно меньше суммы количества обсерваторий в Табл. 1, но все равно могут быть доступны данные 150–170 отдель-

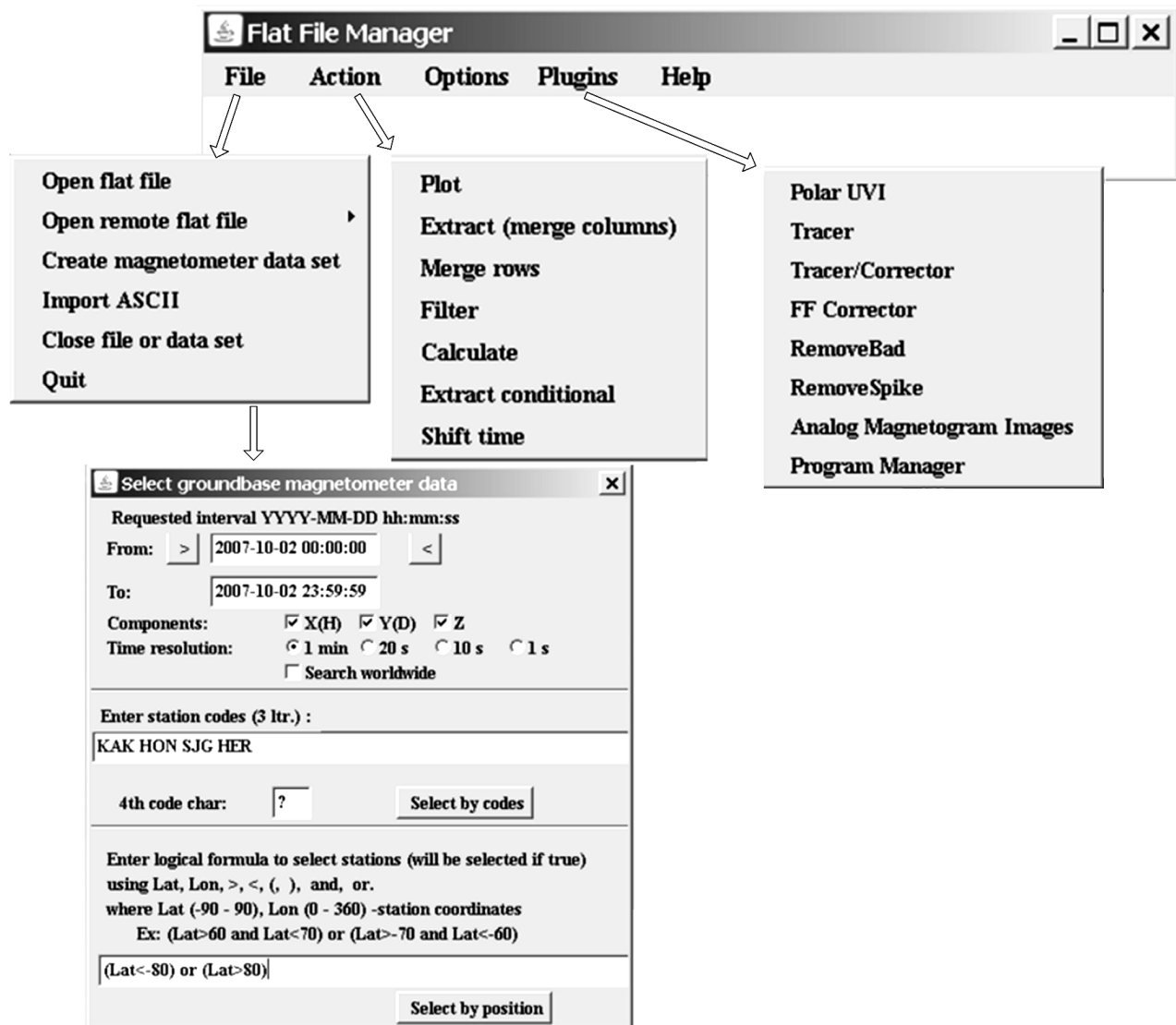


Рис. 2. Основные возможности Flat File Manager.

ных магнитных обсерваторий. Очевидно, что собрать данные такого количества обсерваторий за требуемый интервал вручную очень сложно, и необходима разработка специальных систем для этого. Примером такой системы является виртуальная геомагнитная обсерватория, предложенная в [Papitashvili et al., 2003]. Дальнейшее развитие этот подход получил в [Papitashvili et al., 2006], где было предложено два варианта доступа к данным, основанных фактически на одном и том же коде – web-ВГМО и автономная ВГМО. Web-ВГМО использовала стандартный подход с загрузкой Java-апплета, автономная ВГМО полностью работала на клиентской машине. Сейчас имеются версии автономной ВГМО для Windows и MAC OS X. Конечно, возможна компиляция программ и для других операционных систем, но это требует доступа к таким системам и, возможно, некоторой доработки, и может оказаться довольно трудоемким.

Опыт использования web-ВГМО показал, что она име-

ет две проблемы. Первая: в настоящее время имеется 20–25 источников наземных геомагнитных данных (не считая индивидуальных обсерваторий), однако только около 1/3 из них доступны для анонимного доступа и, соответственно, могут быть доступны через web-ВГМО. Вторая: в связи с ограничениями, вводимыми сетевыми администраторами для обеспечения безопасности сети, во многих случаях апплеты, использующие свои собственные протоколы, не работают (блокируются системами сетевой безопасности).

В геомагнетизме используются много форматов данных – WDC (2 разных формата), Intermagnet (2 формата), MAGA2000, MAGA2002. Данные индивидуальных проектов, как правило, используют свои собственные форматы – всего 15–20 форматов. При объединении данных необходимо выбрать какой-то один общий формат. ВГМО основана на концепции т.н. простых файлов – Flat File (FF) ([Smith and Clauer, 1984]) и названа Flat File

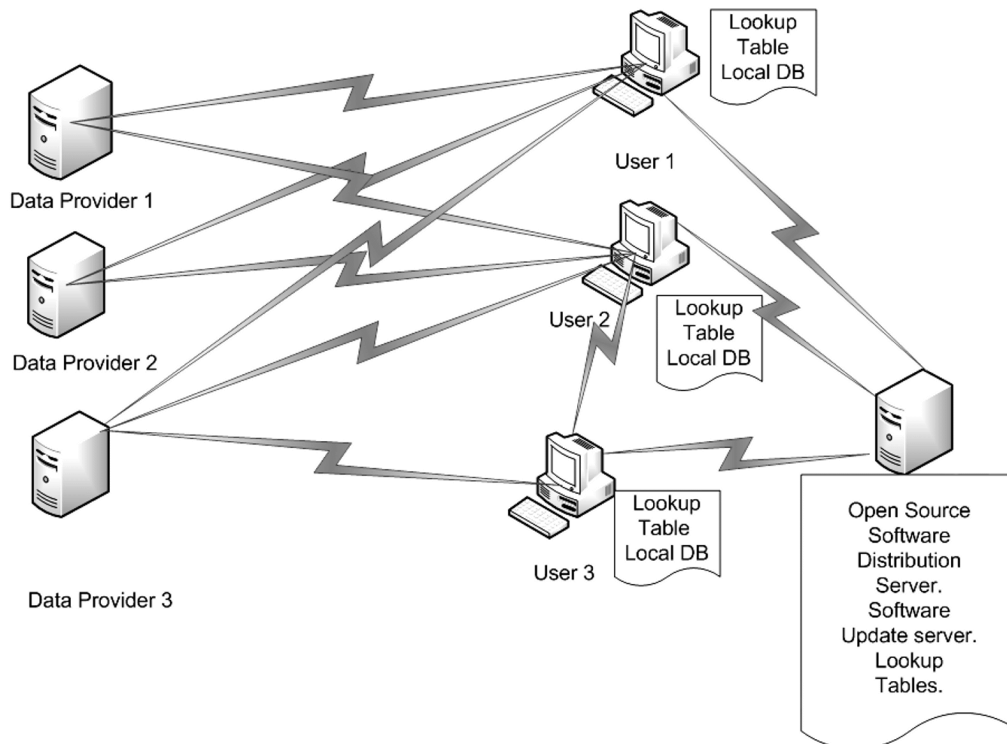


Рис. 3. Схема персональной геомагнитной виртуальной обсерватории.

Manager. Flat File – это система, в которой каждый набор данных состоит из двух файлов: 1) двоичный файл собственно данных, где каждая запись содержит одно поле времени и фиксированное количество полей данных, относящихся к этому времени; фактически, это двоичный аналог двумерной таблицы; 2) текстовый файл заголовка, описывающий эти данные (метаданные). Такая структура сочетает достоинства обоих представлений – компактность, быстрый ввод/вывод и прямой доступ в случае двоичных данных; наглядность и редактируемость в случае текстовых данных.

На Рис. 2 показаны основные меню автономной ВГМО. Flat File Manager позволяет загрузить из сети и преобразовать в FF магнитные данные или любые ASCII файлы. Затем система позволяет открыть одновременно до трех наборов данных, возможно с разным временным разрешением, и выполнять с ними большинство типовых операций с временными рядами, в том числе и очистку данных от плохих значений. При создании выборки наземных геомагнитных данных имеется возможность указать, следует ли ограничиться только локальной базой, или произвести поиск и загрузку новых данных из Интернета.

На современном этапе, когда объемы геомагнитных данных еще не очень велики (10–100 Гб), эффективнее выбирать требуемые данные и копировать их на компьютер пользователя в специально организованную базу данных. Однако источники исходных данных имеют разную

структуру самих баз и используют разные форматы хранения данных, поэтому необходима разработка отдельной управляющей СУБД, которая может формировать запросы к исходным базам и предоставлять пользователю выборку по всем оригинальным источникам данных. Такой подход имеет как свои достоинства, так и недостатки. Достоинствами является быстрый доступ к данным при последующих обращениях, большие возможности для разработки алгоритмов, возможность выполнять редактирование данных и встраивать в систему свои собственные методы обработки. Недостатками – сложность, а иногда и невозможность автоматического обновления данных в базе при обновлении данных у источника. Если сервер данных разрешает доступ непосредственно к файлам данных, то можно, сопоставив даты создания файла у источника и в базе пользователя, обновить данные в случае необходимости. Однако некоторые серверы данных создают файлы только по запросу, и выяснить дату обновления исходных файлов в базе сервера невозможно в принципе.

Существующая сейчас автономная ВГМО написана на Fortran и Java (в некоторых ОС приходится использовать и C) и требует перекомпиляции при любой модификации, сложна в поддержке и развитии. Предполагается создать новую версию автономной ВГМО – персональную виртуальную геомагнитную магнитную обсерваторию (ПВГМО), основанную на языке Python, который является наиболее подходящим средством для раз-

работки универсальной СУБД для работы с геомагнитными (и любыми другими временными рядами) данными (схема на Рис. 3). Python – это свободно распространяемый язык высокого уровня (аналог Matlab и IDL), он существует для всех платформ и операционных систем. В ОС, основанных на Unix, Python является частью ОС, для MS Windows он легко устанавливается или даже запускается без установки (Portable Application). При отказе от использования специальных системозависимых библиотек Python-программы полностью системонезависимы: один и тот же код выполняется везде. В свободном доступе имеется огромное количество готовых свободно распространяемых библиотек, а тексты программ легко читаемы и, соответственно, их легко поддерживать и модифицировать. Краткий обзор языка Python можно найти по адресу: <http://www.helloworld.ru/texts/comp/lang/python/python2/index.htm>

Прототипом ПВГМО может являться автономная ВГМО, часть процедур которой уже переписана на Python. Сетевой основой ПВГМО может являться web-сервер программного обеспечения ПВГМО, который позволяет всем загружать программы, поддерживает систему обновления программ и ссылок на источники данных. На основе такого сервера может возникнуть типичная для систем с открытым исходным кодом система коллективной разработки ПО, когда новые методы и алгоритмы, разработанные пользователями для своих задач,

загружаются на сервер для общего пользования и постепенно включаются в состав ПВГМО.

Литература

- Smith, A. Q., C. R. Clauer (1984), FLATDBMS: A flexible, sourceindependent data management system for scientific data, *STAR Laboratory Report D106-1984-1*, 25, Stanford University, Stanford.
- Pilipenko, V. A., O. V. Kozyreva, M. J. Engebretson, J. F. Watermann, K. Yumoto, S. I. Solovyev (2002), Virtual magnetograms – new tool for the monitoring of the solar wind magnetosphere coupling, *Proc. of the 4th Intern. Conf. "Problems of Geocosmos", St. Petersburg, Petrodvorets, 03–08 June 2002*, 79–82, Petrodvorets, St. Petersburg.
- Papitashvili, V. O., A. B. Saxena, V. G. Petrov, C. R. Clauer (2003), VGMO.NET – Realization and testing of a Virtual Global Magnetic Observatory (invited talk), *Session GAV. 03 "The geospace environment in near-real time: science and technology", IUGG/IAGA General Assembly, Sapporo, Japan, June 30–July 11, 2003*, IUGG, Sapporo, Japan.
- Papitashvili, V., V. Petrov, A. Saxena, N. Papitashvili (2006), Virtual Global Magnetic Observatory network: VGMO.NET, *Earth Planets Space*, 58, 765.

В. О. Папиташвили, Мичиганский Университет, Мичиган, США

В. Г. Петров, Институт земного магнетизма, ионосферы и распространения радиоволн им. Н. В. Пушкова, 42190, Троицк, Московская обл., Россия. (vpetrov@izmiran.ru)