Heim, B., Klump, J., Fagel, N., Oberhänsli, H. (2008): Assembly and concept of a web-based GIS within the paleolimnological project CONTINENT (Lake Baikal, Siberia). - Journal of Paleolimnology, 39, 4, 567-584,

DOI: 10.1007/s10933-007-9131-0.

(submission to Journal of Paleolimnology)

Assembly and concept of a web-based GIS within the paleoclimate project CONTINENT (Lake Baikal, Siberia)

Heim, B.[1], Klump, J.[2], Oberhaensli, H.[2], Fagel, N.[3],

[1]Berlin Program for Equal Opportunities for Woman in Research, Berlin, Germany, [2]GeoForschungsZentrum Potsdam, Telegrafenberg, D14473, Germany; [3]UR 'Clays and Paleoclimate', Geology Department, University of Liege, Belgium

Abstract

Web-based Geographical Information Systems (GIS) are excellent tools within interdisciplinary and multi-national geoscience projects to exchange and visualize project data. The web-based GIS presented in this paper was designed for the paleolimnological project CONTINENT (Lake Baikal, Siberia) to allow the interactive handling of spatial data. The GIS data base combines project data (core positions, sample positions, thematic maps) with auxilliary useful spatial data sets that were downloaded from freely available data sources on the world wide web. The reliability of the external data was evaluated and suitable new spatial datasets were processed according to the scientific questions of the project. GIS analysis of the data was used to help answer questions such as how to quantify the sediment source provinces to assist studies on sediment provenance in Lake Baikal, and whether the visualisation of present-day vegetation distribution and surface pollen distribution supports the conclusions derived from palynologic analyses. The refined geodata are returned back to the scientific community by using online data publication portals. Data were made citeable by assigning persistent identifiers (DOI) and were published through the German National Library for Science and Technology (TIB Hannover).

Keywords: Web-based GIS, paleolimnological project, Lake Baikal, catchment analyses, open access

Introduction

The project 'High-resolution CONTINENTal paleoclimate record in Lake Baikal' (CONTINENT) was conducted during 2001-2004 on Lake Baikal, Siberia, and its catchment area. A primary objective of the project was to understand processes of eolian and fluvial input and primary production processes in the photic zone of Lake Baikal. By evaluating the transport of potential climate proxies to the deep lake bottom {Fietz, 2005 #3}, transfer functions were derived for interglacial periods applicable to longer records such as the Baikal Drilling Project (BDP) cores {Williams, 1997 #35; Appleby, 1998 #14}. The results of the project were used to reconstruct a continuous palaeoclimate record from a mid-latitude continental setting, with a centennial to decadal resolution over the last 20 kys and with a centennial time resolution for the time slice 110-130 kys B.P {Oberhaensli, 2005 #16}. The complex nature of the study called for innovative data management strategies.

Within paleoclimate projects data management strategies commonly focus on data from sediment cores and successful design patterns for data models have been established (e.g. {Diepenbroek, 2002 #1}. However, paleoclimate research in multi-disciplinary projects on large continental water bodies show that, in addition, investigations of modern hydrodynamics and biogeochemical conditions of the investigated waterbody and studies of their catchment areas need to be included. The studied area in our investigation stretches from central Siberia in the Russian Federation towards Mongolia to the South (Figure 1) where in the context of the CONTINENT project samples were collected on ten expeditions. The sampling positions are distributed over an area of approximately two million square kilometres and spread over six degrees of geographic latitude. It is evident that in a study of this scale all data have to be analysed in their spatial context and at different scales. The study also requires the incorporation of auxiliary spatial data on the topography, the geology and other relevant geodata.

The rationale to set up a Geographical Information Systems (GIS) as part of the scientific activities was that visualisation of the spatial context of the project sites and thematic background information (e.g. morphology, pedology, geology, vegetation) can help to understand and interpret research data through spatial visualisation {Butler, 2006 #15}{Lyon, 2006 #16}. However, conventional databases and GIS applications require sophisticated information, technology infrastructure and skilled users, and access is restricted locally to the institution operating the system. Accordingly, the CONTINENT Baikal-GIS has been set up as web-based GIS (doi: 10.1594/GFZ.SDDB.1201). There are significant advantages to a web-based GIS in international networked projects, such as the CONTINENT project, since it allows all participating scientists to access spatial, project related information, regardless of the local availability of a GIS.

At present, data sharing is still in its early stages, especially geospatial data are not generally available on the internet. Nevertheless, the value of data sharing is widely acknowledged (e.g. {Alexander, 2004 #20; Klump, 2006 #19}). In setting up the GIS data base we found that it was difficult to obtain geospatial data systematically from online sources because they are not treated as publications and therefore are not listed in library catalogues or similar reference works. As a result of our GIS project we published our data through the Scientific Drilling Database (http://www.scientificdrilling.org) and assigned Digital Object Identifiers (DOI) to the datasets.

Not only are geospatial data hard to find on the internet, the absence of a review system makes it difficult to judge the quality of the data. Can publicly available geospatial data be used for scientific purposes?

In this paper we present examples of GIS applications to support scientific analyses regarding research questions arising from the project: (1) the calculation of relative abundance of the lithological groups that outcrop in the tributary catchments as part of the analysis of sediment provenance, and (2) a comparison of modern pollen distribution in the lake sediments with the present-day distribution of vegetation type around the lake as part of the palynological analyses.

Lake Baikal regional setting

Lake Baikal (elevation 465 m a.s.l.) in Southern Siberia is situated at the centre of the Eurasian land mass in a string of deep rift basins which are bordered by high mountain ranges rising to 3500 m a.s.l.. The present day fluvial catchment area has three major subdivisions: the Selenga, Upper Angara and Barguzin river catchment areas (Figure 1). The Selenga River, which is the largest tributary into the lake has built up an enormous delta into the lake, separating the southern basin from the central basin of Lake Baikal. The second largest tributary, the Upper Angara River, drains the tectonic sub-basins northeast of Lake Baikal into the northern basin, while the third largest tributary, the Barguzin River, drains the tectonic Barguzin sub-basin into the central basin of Lake Baikal.

The characteristics of the catchment areas vary considerably, ranging from the boreal mountainous areas surrounding Lake Baikal towards the dryer climates of the Mongolian plains with their associated soil and vegetation cover.

Materials and Methods

Assembly of data sets for the CONTINENT Baikal-GIS

In the course of the CONTINENT project and related projects, sediment core data and a variety of limnological, geophysical, sedimentological, and botanical field data were acquired. The GFZ Lake Drilling Information System (LakeDIS) served as project data base, its data model is compatible to the data model of PANGAEA/WDC-MARE {Diepenbroek, 2002 #1}, where all data are georeferenced by coordinate pairs in decimal degrees. The experience gained from the application of the LakeDIS was later used to develop the Drilling Information System for the Antarctic Coring Expedition (ACEX) of the Integrated Ocean Drilling Program (Wallrabe-Adams et al., 2006).

The CONTINENT sediment core data, field data and additional data sets were transformed for use in a GIS data base, where all geographical objects are represented as vector elements or as raster data matrices. Vector data build a model of the real world as multipoint-layers, such as the field sampling stations, and vector lines, such as ship tracks, drainage patterns, contour lines, and polygon layers, such as geology, soil, vegetation, etc. On the other hand, digital elevation models (DEM), multispectral satellite images, and similar data, are stored as raster data, which reference parameters by grid cell location in a matrix.

All data were geographically referenced to the same map projection in the same earth coordinate system. This allows us to integrate and to visualize data from various sources into a common spatial database. To be able to calculate true lengths and areas, we chose the Universal Transversal Mercator (UTM) coordinate system with a true scale longitude at 105° W (zone 48) on the World Geodetic System (1984) reference ellipsoid (WGS84) as map datum for the CONTINENT Baikal-GIS. In this study, all data were processed using commercial GIS and image processing software (ESRI ArcGIS, ArcInfo, ArcIMS, Imagine, ENVI), but Open Source software, such as GRASS GIS and the UMN Mapserver are also suitable tools for the operations described in this paper.

The thematic layers of the CONTINENT Web-GIS can be divided into three groups:

- CONTINENT project data and data from related projects (sampling and coring positions, lacustrine sediments, bathymetric map)
- Freely available data sets from the world wide web (digital elevation model, geology, pedology, infrastructure, drainage pattern, etc.)

- Calculated higher level data (river catchments, litho-stratigraphic groups in the catchment, surface pollen distribution, distribution of plant species and plant functional groups)

The contents of the CONTINENT Web-GIS thematic layers are summarised in Tables 1a-d. As first processing step, the LakeDIS CONTINENT project data in decimal degrees and related project data, such as thematic maps and bathymetric information in different specified Cartesian coordinate systems, were transformed into UTM Cartesian coordinates. Lake Baikal and its catchment expand over three UTM zones (zones 47, 48, 49), covering zone 48 as the central UTM zone, and one third of both neighbouring zones, respectively. The error in area calculation that was generated by re-projecting zones 47 and 49 onto zone 48 was calculated for the catchment areas to be insignificant with a value of less than 1 %.{Swiercz, 2004 #30}

We searched for additional data on the world wide web and were able to obtain topographical, geo-ecological and geological, pedological and botanical thematic digital data sets. These country-related data sets from Russia, Mongolia and China were merged into single thematic map layers (Table 1b), each covering an area of 1620 km by 1760 km (including southern central Siberia, northern central Mongolia, part of north-eastern China). In some cases merging datasets required the assimilation of the factual data attributes. All new data layers were transformed on to UTM, zone 48. The process of data acquisition for selected thematic layers is described below.

The topographical digital elevation model (DEM) was calculated from the United States Geological Survey (USGS) topographic DEM ‘GTOPO30’, which offers a spatial resolution of 30 arc seconds (approx. 1 km). The GTOPO30 DEM was colour-coded as topographic background information. It was also used to calculate slope values and elevation contour lines at 100 m intervals by triangulation. It turned out that the raster-layer of slope values calculated from the GTOPO30 dataset could not resolve the rugged mountain ranges and steeply incised valleys around Lake Baikal. Owing to the low spatial resolution, the DEM only showed elevated plateaus. As a workaround, {Swiercz, 2004 #30} used the polygon layer ‘slope’ from the International Institute for Applied Systems Analysis (IIASA), showing not absolute slope values, but intervals, to extract different susceptibilities to erosion.

The catchment area outlines of Lake Baikal and its three main tributaries were digitised onto the combined layers of elevation contour lines and the high resolution drainage pattern accordingly hydrological considerations {Swiercz, 2003 #5}. Several new GIS layers were produced to estimate the proportions of parameters in the respective river catchments by clipping the parameterised layers, e.g. ‘geology’, ‘vegetation’, ‘soils’, ‘slope’, using the vector layers ‘lake catchment’ and ‘river catchments’ {Swiercz, 2004 #30}.

Implementation of the CONTINENT Web-GIS service

The web-based GIS itself has a client-server architecture. On the CONTINENT website we offer the ESRI ArcIMS client to access to the Web-GIS. This application is browser-based and uses JavaScript to provide functionalities such as zooming into and out of the map view, select thematic map layers, etc. On the server side, the ESRI ArcIMS Web Map Server receives requests to build up a thematic map from selected GIS layers, and delivers the constructed map to a client application as a raster image (JPEG). The web map service is conformable to the standards set by the Open GIS Consortium (OGC WMS).

On start-up, the CONTINENT Baikal-GIS shows the coloured DEM as standard map and as the reference map in the overview inset (Figure 2). The right-hand menu is toggled to the selection of thematic layers (e.g. geology, soils, LANDSAT satellite image mosaic, CONTINENT stations, drainage pattern, lake and river catchments, elevation contour lines). The display of map features, the associated items in the map legend, and scale bar is made scale dependent. Map features and dependent map elements are only shown at their designated scale intervals. For different map scales we integrated data layers differently, e.g. the drainage pattern shows different stream orders, and isohypes are shown at contour intervals appropriate to the map scale. The display rules of the CONTINENT Baikal-GIS are set to avoid overlapping map elements.

A feature selection (for point features, vector features, polygon features) serves as a geographical query tool to display feature attribute information of the selected layer. Attribute tables include the sample code, its coordinates in decimal degrees and UTM, date and time of sampling, attributed parameters, metric size of feature (length, area), and more. The attributed parameters may be nominals such as litho-stratigraphic groups, pollen species, and soil types, or ordinals, such as topographic elevations and bathymetrical depths, specific concentrations, relative abundances, etc., or intervals, such as ranges of slope or elevation values. The query tool makes it also possible to highlight specified features, which have been searched for by their attributes. The web-based CONTINENT Baikal-GIS does not enable the user to undertake spatial analyses through the client-server architecture, i.e. the quantitative and comparative GIS-analyses presented in this study have been carried out on the workstation-based CONTINENT Baikal-GIS.

Results

Quality of data and spatial visualisation

As a result, this study shows that it is useful to acquire substantial parts of the geodata from publicly available sources. Most data available on the world wide web are at meso-scale resolution which would be not sufficient for local studies on a scale of 1:10 000, or higher resolution. Given the size of our study area, the spatial resolution of the data incorporated in the CONTINENT Baikal-GIS (e.g., 100 m elevation contours, 1 km raster) and their thematic resolution were adequate for the research purpose. Evaluation of the quality of the downloaded data sets showed that the majority of the spatial thematic data we acquired were consistent with our investigations in the field and with the thematic content of the Baikal Atlas of {Galasy, 1993 #9} and other regional maps. The exceptions in terms of applicability and correctness were the acquired digital geological data sets.

Before we provide examples of GIS-solutions related to research questions, we would like to point out the advantage of the CONTINENT Baikal-GIS as a mapping and visualization tool. Figure 2 demonstrates its visualization capacities to produce informative maps. The background map chosen here is the colour-coded DEM {USGS, 2002 #29}. It is overlain with the polygon layers of Lake Baikal and lacustrine sediment regions {Sturm, 1998 #7}, the point layer of the CONTINENT sampling stations and the vector layers of infrastructure and elevation contour lines. The web-based GIS can be queried for specific parameters, or for codes of sampling positions. The CONTINENT Baikal-GIS allows an immediate interpretation of samples in their spatial context, i.e. their sedimentary, botanical, and geomorphologic environment, by using the supplied thematic maps and the query functionalities provided by the Web-GIS client. Also, those conventional printed maps that were available to us covered the study area only incompletely. Accordingly, the created thematic maps became a valuable orientation tool for interpreting field data.

Application of GIS in sediment provenance analyses

The sedimentological composition of lake sediments is proportionally controlled by the lithology of their source area. GIS analysis may supply the geographical distribution of lithologies in the hinterland. Therefore, we investigated the usefulness of the acquired digital data sets for provenance investigations in the CONTINENT project. Despite their richness in detail, the digital geological map data of Russia and Mongolia distributed by the USGS could not simply be merged and then used to calculate the proportions of rock types. The digital geological map of Russia is a chronostratigraphic map with petrological attributions, while the geological map data of Mongolia mainly show the chronostratigraphical units and therefore do not map the large granitoid provinces as separate entities. This difference between the map datasets resulted in a considerable underestimate of the presence of granitoids in the Selenga catchment, even though the granitoids are dominant features in this area e.g., {Jansin, 1989 #32; Kalinin, 1981 #33}. Through editing we re-attributed the granitoids in the Mongolian part of the geochronological geological map data. In addition, because hinterland lithology provides the parameters for the analysis of sediment provenance we edited and regrouped the geological map data further to produce a new polygon-layer showing the major lithological units (doi: 10.1594/GFZ.SDDB.1210).

The analysis of sediment cores from the Posolskoye site (CON01-604) in the South Basin and from the CONTINENT site (CON01-603) in the North basin showed that the clay minerals (silt-clay aggregates, mainly illite and illite-smectite) and the light and heavy mineral composition (mainly quartz and feldspars; heavy mineralogical composition dominated by hornblende) are of similar composition, despite being taken approx. 300 km apart in different deep lake basins. In contrast, cluster analysis of the mineral assemblage at the Vydrino coring site (CON01-605), also located in the South Basin near Posolskoye site (CON01-604), shows a high abundance of mica in the light mineral fraction and a different composition of the clay minerals {Fagel, 2007 #31}.

The new lithological map (Figure 3) shows that along the South East Coast of Lake Bakal, where the Vydrino site (CON01-605) is located, metamorphic rocks with mainly granulite and amphibolite facies rocks are the prevalent lithological elements. These source rocks supply the light and heavy mineral composition of non granitoid-type to the site of CON01-605.

The area and relative percentage of the main source rocks within the different river catchment areas were calculated in the main river catchment areas of the Selenga, Upper Angara, and the Barguzin rivers (Table 2), where the area calculation indicated a granitoid coverage exceeding 50 to 60 %. It is assumed that the second important lithological group, the Quaternary sediments (11 % to 31 % coverage), is composed mainly of reworked granitoides. This prevalence of granitoid source rocks strongly influences the sedimentary composition of the sediments at CON01-603 and CON01-604, even though the two positions are hundreds of kilometres apart and located in different sub-basins of Lake Baikal.

Application of GIS in palyonological analyses

Paleoclimatic reconstructions using pollen assemblages are based on the premise that there should be a direct relationship between the pollen assemblage found in lake sediment samples and the plant assemblage in the lake catchment area {Granoszewski, 2005 #4; Tarasov, 1998 #25}. The aim of this GIS application was to compare the modern pollen distribution in Lake Baikal with the present-day surrounding vegetation.

The acquired digital vegetation data from Food and Agriculture organization of the United Nations (FAO) showed the land use, i.e. the outlined boreal areas were singly classified as 'forest' and not further sub-classified. As a workaround we digitised a detailed vegetation map providing plant functional groups {Isachenko, 1990 #6}and georeferenced it to the CONTINENT Baikal-GIS map datum (UTM zone 48, WGS 84). We produced monospecific maps of relative abundance (i.e., dominant, present, absent) from the new GIS-layer and thus were able to map the relative vegetation cover in the Lake Baikal catchment area.

Pollen data from Lake Baikal sediments and surface samples on land were supplied by V. Markgraf and Y. Bezrukova {Bezrukova, 1999 #8}. The data supplied did not contain information on the taphonomy of the collected samples and therefore limit the assessment of the degree of preservation of pollen species. The pollen data were grouped to match the classification into plant functional groups used in the vegetation map of {Isachenko, 1990 #6}. The sample coordinates were converted to UTM coordinates. The georeferenced pollen data (percentages of species taxa or functional groups, respectively) were interpolated to model the pollen distribution pattern in Lake Baikal. We used the ‘Natural Nearest Neighbours’ interpolation method to prevent over-smoothing between the irregularly distributed sampling stations.

We compared the monospecific maps with the modern distribution of their pollen in the lake and its vicinity. Figures 4a–b show the distribution patterns of two selected plant species (*Pinus sylvestris* and *Larix sp.*), and their respective spatial pollen distribution. The resulting pollen and relative vegetation maps show a mismatch between the pollen distribution patterns in the lake sediments compared to the plant distribution on land. The proportion of *P. sylvestris* pollen, where present, exceeds the proportion of the recent coverage of *P. sylvestris* on land. Even though *P. sylvestris* is only moderately abundant on land, mainly in the northern Selenga and in the Barguzin catchment areas, its pollen are found widely dispersed and in high proportions. In contrast to *P. Sylvestris*, forests of *Larix sp.* are abundant in the entire lake catchment area, but its pollen are underrepresented in the pollen assemblages (Figure 4b). Also, the distribution pattern of *Larix sp.* pollen in Lake Baikal shows a distinct regional inhomogeneity. On one hand, the highest proportion of *Larix sp.* pollen is found of the Selenga delta and in Barguzin Bay, where the Barguzin River flows into Lake Baikal. On the other hand, in the Upper Angara Valley, an area with abundant *Larix sp.* forests, the relative proportion of *Larix sp.* in pollen assemblages is low, even in surface samples taken in the larch dominated light taiga.

The comparison of pollen distribution maps with the monospecific maps shows that the main mode of dispersion for the pollen of *Pinus sylvestris* was by wind transport, while the pollen of *Larix sp.* were mainly dispersed by water transport in rivers and lake surface currents. This is in agreement with previous work by other authors (e.g. {Bezrukova, 1999 #8}). Figure 5 provides information on prevailing winds and surface currents.

Our GIS-analyses showed that if maps of grouped pollen were compared with vegetation maps of plant functional groups, this disparity was not as pronounced in the maps. Grouping pollen gave a better representation of environmental conditions on land, than to use pollen data on species level {Klump, 2004 #26}.

Publishing and sharing geodata from the CONTINENT Baikal-GIS

All thematic map layers discussed in this paper have been published as data publications through the Scientific Drilling Database. Each dataset is identified persistently and globally unique by a Digital Object Identifier, which can be resolved to a Uniform Resource Locator (URL) through any server based on Handle.Net software. To facilitate discovery of the data to users not aware of their existence, all data publications in this paper are listed as electronic sources in the online catalogue of the German National Library of Science and Technology (TIB/UB Hannover). The structure of data access is displayed in Figure.6.

All data may be downloaded from the Scientific Drilling Database and used for new scientific works. The use of the data by other authors is governed by the Creative Commons licence (by-nc, 2.0, de), which allows the author to create derivative works, as long as the derivative works are non-commercial and the origin of the data is acknowledged, i.e. a proper citation is given. Thus, our data publications are treated in analogy to ‘traditional’ publications.

Discussion

Local Russian governmental GIS data bases, as well as valuable and detailed local and regional GIS data bases from Siberian and Buryatian research institutes and from international projects (e.g., (EU) INTAS Baikal mapping, (EU) Siberia-II) have yet been established for this region. The covered areas of these GIS projects mainly stretch from Lake Baikal northwestwards into the boreal regions. Within the (EU) INCO project ‘Irkutsk Regional Information System for environmental protection, IRIS', the existing GIS-layers of the Irktusk region are compiled and completed into an operative prototype for other regions of Northern Eurasia. The focus lies on the investigation of the regional environmental hazards and environmental management strategies.

This study presents a GIS data base on meso-scale whose data are related to a paleolimnological project. Through georeferencing, heterogeneous quantitative and descriptive data can be attributed to geographical objects. The data visualization – on one hand showing spatial interrelationships, on the other thematic background information – enables the interpretation of data in their spatial context and on different spatial scales. Specified data can be extracted through spatial queries and spatial analysis can be used for limnological, geological, botanical, and pedological investigations. The two presented examples show that including the spatial dimension provides a powerful tool to help understand complex limnological processes, such as sediment provenance and spatial distribution patterns of pollen.

Despite being a project-related GIS, the majority of the map layers provided by the web-GIS were created by using freely available data sets from the world wide web. There is a variety of useful available geodata sets in the world wide web that still remain underutilised. The move by academic and research institutions to create open access to knowledge has started a discussion on open access to data {Klump, 2006 #19}. However, open access to data seems unlikely before the work of data authors is acknowledged by citation of their data in the scientific literature. For data to be citeable it is necessary that they can be referred to in a persistent way. The location of internet resources, and thus their URL, may easily change, which in most cases means to the user that the data are lost {Lawrence, 2001 #24; Koehler, 2004 #23}. Therefore, a prerequisite for data access via the internet is the use of persistent identifiers, such as Digital Object Identifier (DOI), to have an address to locate the desired dataset that is reliable and available over a long time {Brase, 2004 #22; Paskin, 2005 #21}. Also, geospatial data are difficult to find on the world wide web since they are normally not listed in library catalogues or scientific databases. The project 'Publication and Citation of Scientific Data' (STD-DOI) has established a system by which data are published through accessible databases, identified by DOI and incorporated in the catalogue of the German National Library for Science and Technology (TIB Hannover). The data publications are now citeable, accessible and can be discovered through the TIB library catalogue by other users.

Intellectual property rights and fair-use have become intensely debated issues. To ensure fair use, the published CONTINENT web GIS data are licensed under a Creative Commons licence. The Creative Commons licensing system supports the idea of Open Access to scientific data, yet guards the intellectual property rights of their originators. It can be found in several localised versions that are not only translated into local languages but also into local legal systems. Creative Commons licences are legally binding and have been tested in court.

Conclusion

We show applications of GIS technology in the paleoclimate project CONTINENT at Lake Baikal, Siberia, and that it is useful to assess the spatial dimension of data in paleolimnological projects. The CONTINENT Baikal-GIS offers web-based access to various thematic data on the study area around Lake Baikal Even though the CONTINENT Baikal-GIS was intended only as a prototype, it has been operational and fully accessible to the public since 2002.

A substantial part of the CONTINENT Baikal-GIS layers has been produced from freely available spatial data sets retrieved from the world wide web. Evaluation and re-attribution of geodata lead to higher level derivative thematic map layers. This evaluation of the accuracy and validity of the available data, and the assessment of their use for specific scientific questions is considered as an important task.

The GIS data layer ‘lithology of Lake Baikal catchment’ was produced for the support of the sediment provenance analyses, because first, there were some discrepancies in the heavy and light mineral composition of the CONTINENT sediment cores. Sediment cores that were separated several hundreds of kilometres in different tectonic sub-basins of Lake Baikal showed similar sedimentologic composition in contrast to sediment cores that were relatively near-by each others. The quantification of the proportion of lithologies in the hinterland of Lake Baikal clarified the spatial derivation of source rocks.

Application of GIS in this analysis of pollen distribution in the lake supported the assumption that pollen-based transfer functions can be used for the reconstruction of the paleoenvironment of the Lake Baikal region. The disparity between mono-specific pollen and plant distribution maps showed that correction factors are needed to account for the differences in transport potential between pollen species within Lake Baikal. This disparity was not seen in maps with species grouped into plant functional groups.

A novel feature is the publication of the GIS data presented in this paper through the Scientific Drilling Database. The data publications can be referred to through Digital Object Identifiers and can thus be cited like ‘traditional’ scientific publications. The use of the data by other authors and the creation and dissemination of derived publications is governed by a Creative Commons licence. The data publications can be discovered by other scientists through their inclusion in the online catalogue of the German National Library for Science and Technology. This opens an entirely new dimension of data sharing, acquisition and dissemination, not only in paleoclimate projects.

Acknowledgements

The project ‘High Resolution CONTINENTal Paleoclimate Record in Lake Baikal, Siberia’ was supported by the European Commission under the Fifth Framework Programme and contributing to the Key Action "Global Change, Climate and Biodiversity" within the Energy, Environment and Sustainable Development (Contract EVK2-2000-00057). The project 'Publication and citation of scientific and technical data' is supported by the German Science Foundation, Div. Libraries and Information Systems (LIS4-55495(1) Hamburg BIB 46 Hhmpi 01).

The authors thank S. Swiercz, G. Dachnowski, S. Schneider, F. Kramer, A. Schulze and the technical staff of the GeoForschungsZentrum Potsdam Data Centre for their contributions to the Baikal GIS.

Appendix 1

Acronyms

DCW	Digital Chart of the World (ESRI product)
[1]DFG	Deutsche Forschungsgesellschaft (German Science Foundation); [1]project 'Chemical, isotopic, and mineralogical proxy data for aeolian input into the Lake Baikal system during the last 150 ka'
EAWAG	Swiss Federal Institute for Environmental Science and Technology, Switzerland
ESRI	Environmental Systems Research Institute
FAO	Food and Agriculture organization of the United Nations
GEOPASS	'Geochemistry, Particle Formation and Sediment Sources of Lake Baikal'
GMT	Generic Mapping Tools
IIASA	International Institute for Applied Systems Analysis
INTAS	International Association for the promotion of cooperation with scientists from the New Independent States of the former Soviet Union
RCMG	Renard Centre of Marine Geology, Ghent, Belgium
USGS	United States Geological Survey

References

Tables

Table 1a Project data, related projects

Theme	Data	Data source	Data format
CONTINENT stations CON01-1,2,3,4,5 (2001) CON02- (2002),	expedition stations (ship cruises, land and ice expeditions) in 2001, 2002 (LakeDIS attributes) doi: 10.1594/GFZ.SDDB.1203	CONTINENT team	point ordinal and nominal data, point vertical profiles, point vertical cores, vector ordinal data, point data n = 291
related projects	sediment samples (river bed, loess, etc.) in 2002, 2003 doi: 10.1594/GFZ.SDDB.1203	M. Zopperitsch, (DFG project[1])	point ordinal data, n = 64
	Baikal Drilling Project BDP cores: 1 (1993), 2 (1996), 3 (1997), 4 (1999) doi: 10.1594/GFZ.SDDB.1203	BDP (Russia, Japan, Germany, USA)	n = 4
	Baikal sediment cores doi: 10.1594/GFZ.SDDB.1203	EAWAG 1996-1998	n = 3
pollen	pollen samples (terrestrial, lacustrine) in 1999	J. Bezrukova and V. Markgraf (Bezrukova, 1999)	point ordinal and nominal data, n = 70
lacustrine sediment	lacustrine sediment map (lacustrine short cores) doi: 10.1594/GFZ.SDDB.1205	EAWAG (GEOPASS 1999)	polygon nominal data
bathymetry	new bathymetrical map of Lake Baikal doi: 10.1594/GFZ.SDDB.1204	RCMG, INTAS Project 99-1669	vector ordinal data

Table 1b Supportive digital data freely available from the world wide web

Theme	Data	Data source	Data format
Digital Elevation Model DEM	GTOPO30 DEM with 1 km x 1 km grid spacing doi: 10.1594/GFZ.SDDB.1206	USGS	raster ordinal data
geology	digital geological data sets (Russia, Mongolia, China) doi: 10.1594/GFZ.SDDB.1210	USGS	polygon nominal data
drainage pattern	GMT drainage pattern doi: 10.1594/GFZ.SDDB.1208	GMT	vector nominal data
drainage pattern	DCW drainage pattern (high stream orders) doi: 10.1594/GFZ.SDDB.1208	ESRI	vector nominal data

inland water bodies	inland water bodies doi: 10.1594/GFZ.SDDB.1208		
infrastructures	infrastructures doi: 10.1594/GFZ.SDDB.1207		
pedology physiography	IIASA digital Soil and Physiographic database for North and Central Asia doi: 10.1594/GFZ.SDDB.1210	FAO	vector, polygon, interval and nominal data
optical satellite data	MrSid Landsat Mosaic 7 4 2	NASA	raster ordinal data

Table 1c Redefined GIS data calculated from project data (Table 1a).

Theme	Data	Data integration	Data source	Data format
Chlorophyll (Chl-a) map	Averaged Chl-a concentrations	ARC GIS Nearest neighbour interpolation	CONTINENT field data, Fietz (2005)	Raster ordinal data
Pollen species map	Rel. abundance of main pollen species doi (in prep.): 10.1594/GFZ.SDDB.1211	ARC GIS Nearest neighbour interpolation	J. Bezrukova and V. Markgraf (Bezrukova, 1999)	Raster ordinal data

Table 1d Redefined GIS data calculated from the external data (Table 1b).

topography	elevation contours (m) doi: 10.1594/GFZ.SDDB.1206 Slope values (°)	ARC GIS ArcToolbox	GTOPO30 DEM USGS	vector ordinal data, Raster ordinal data
Lake Baikal catchment	lake catchment, main river catchments (Selenga, Barguzin, Upper Barguzin) doi: 10.1594/GFZ.SDDB.1208	ARC VIEW (Digitising) On layers stuck of GTOPO30 DEM, height isolines, DCW River Network	S. Swiercz (2003)	vector nominal data
precipitation	precipitation isolines (mm/a)	ARC VIEW (Digitising)	Atlas of World Water balance (UNESCO, 1979)	Vector ordinal data
vegetation	functional plant assemblages doi (in prep.): 10.1594/GFZ.SDDB.1211	ARC VIEW (Digitising)	Isachenko (1990)	polygon nominal data
geology	detailed improved and revised	ARC MAP revision of the		polygon nominal data

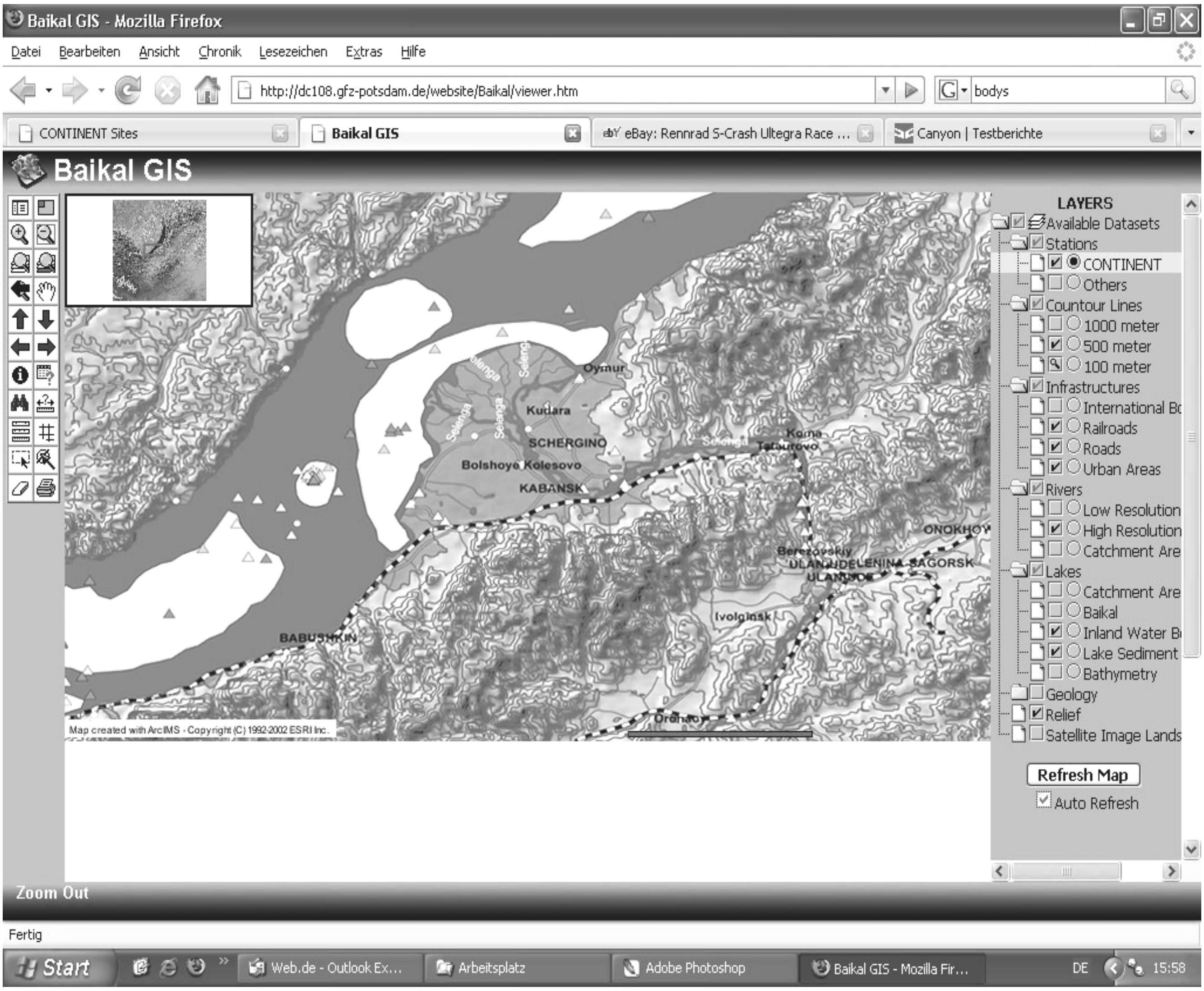
	geological map doi: 10.1594/GFZ.SDDB.1210	merged geological GIS layer		
	summarized geological map doi: 10.1594/GFZ.SDDB.1210			
lithology	lithological assemblages within the lake catchment doi: 10.1594/GFZ.SDDB.1210	ARC MAP lithological attributation	Fagel et al (2007)	polygon nominal data

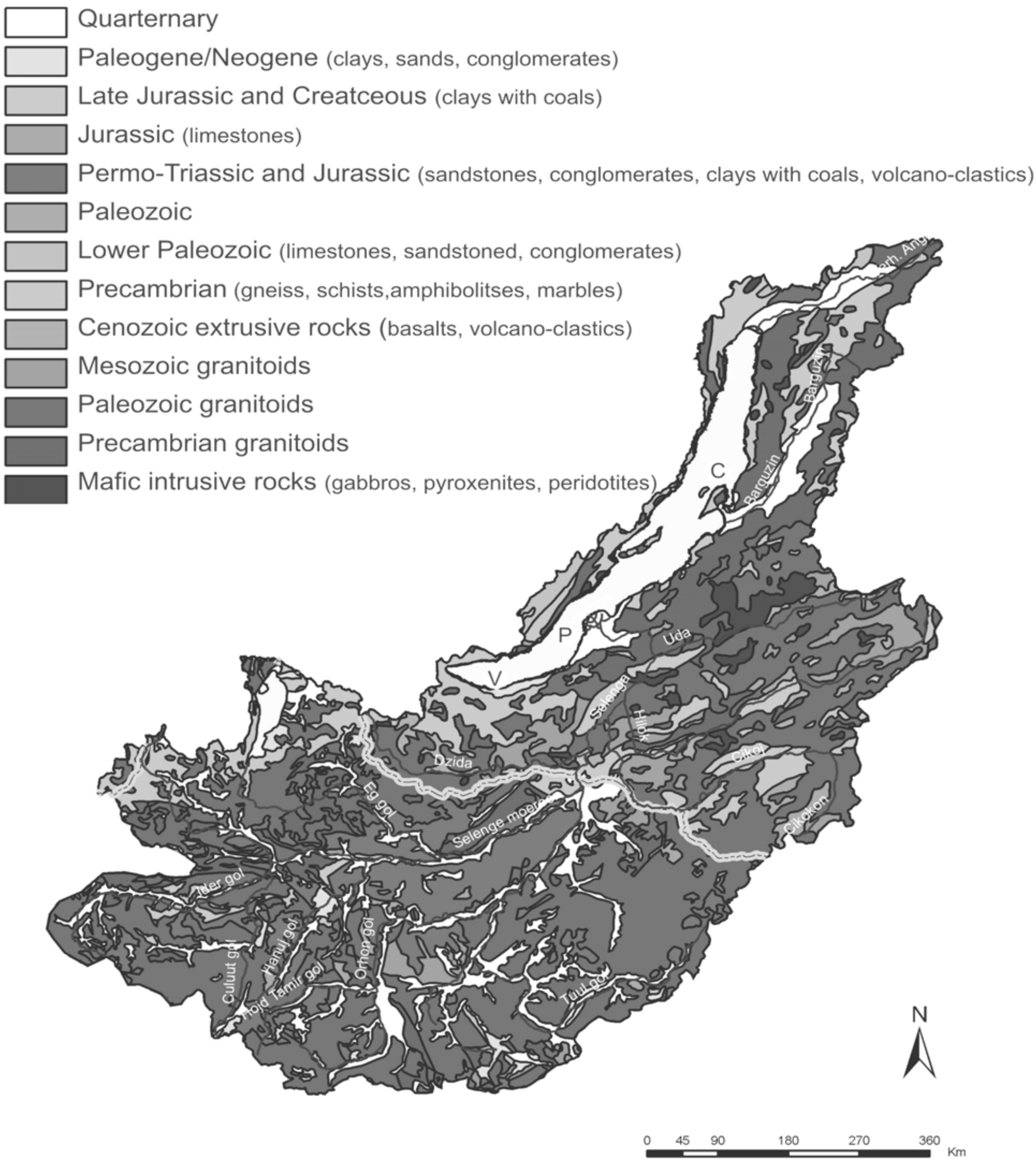
Table 2 : Relative percentages of the main stratrigraphic lithologies in the Selenga, Barguzin, and Upper Angara catchment areas**.**

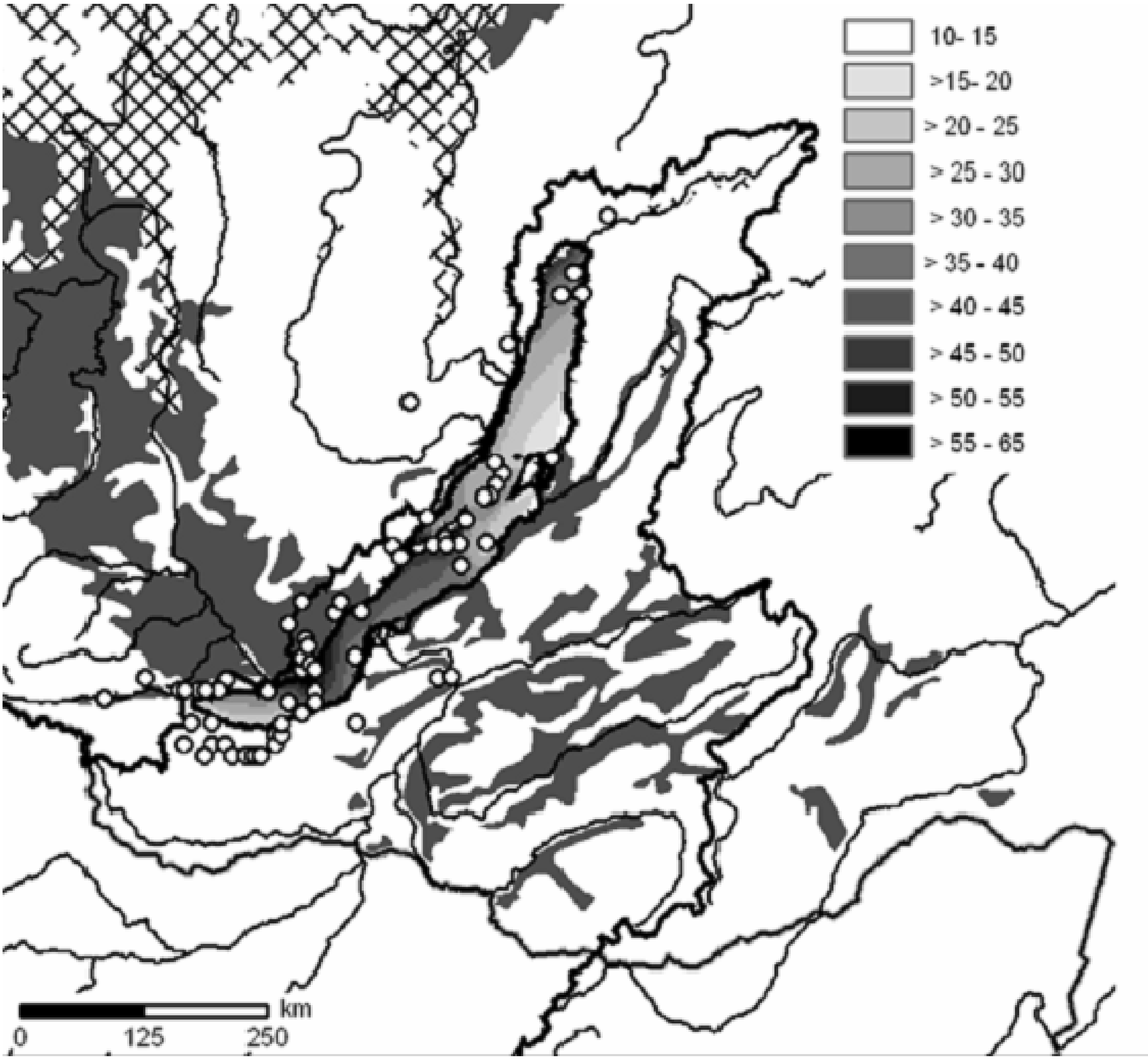
Lithology	**Selenga**	**Barguzin**	**U. Angara**
Quaternary	11 %	31 %	19 %
Paleogene/Neogene (clays, sands, conglomerates)	1 %		
Late Jurassic and Cretaceous (clays with coals)	3 %		
Permo-Triassic and Jurassic (sandstones, conglomerates, clays, volcanoclastics)	12 %		
Paleozoic (weekly to medium metamorphized sandstones, volcanoclastics, limestones)	7 %		
Precambrian (gneiss, schists, amphibolites, marbles)	7 %	16 %	24 %
mesozoic granitoids	5 %		
paleozoic granitoids	42 %	2 %	16 %
Precambrian granitoids	9 %	50 %	40%
total: granitoids	56 %	52 %	56 %
basic to ultrabasic intrusive rocks (unknown age)	2 %	1 %	1 %
extrusive rocks, Trias to Cenozoic	1 %		

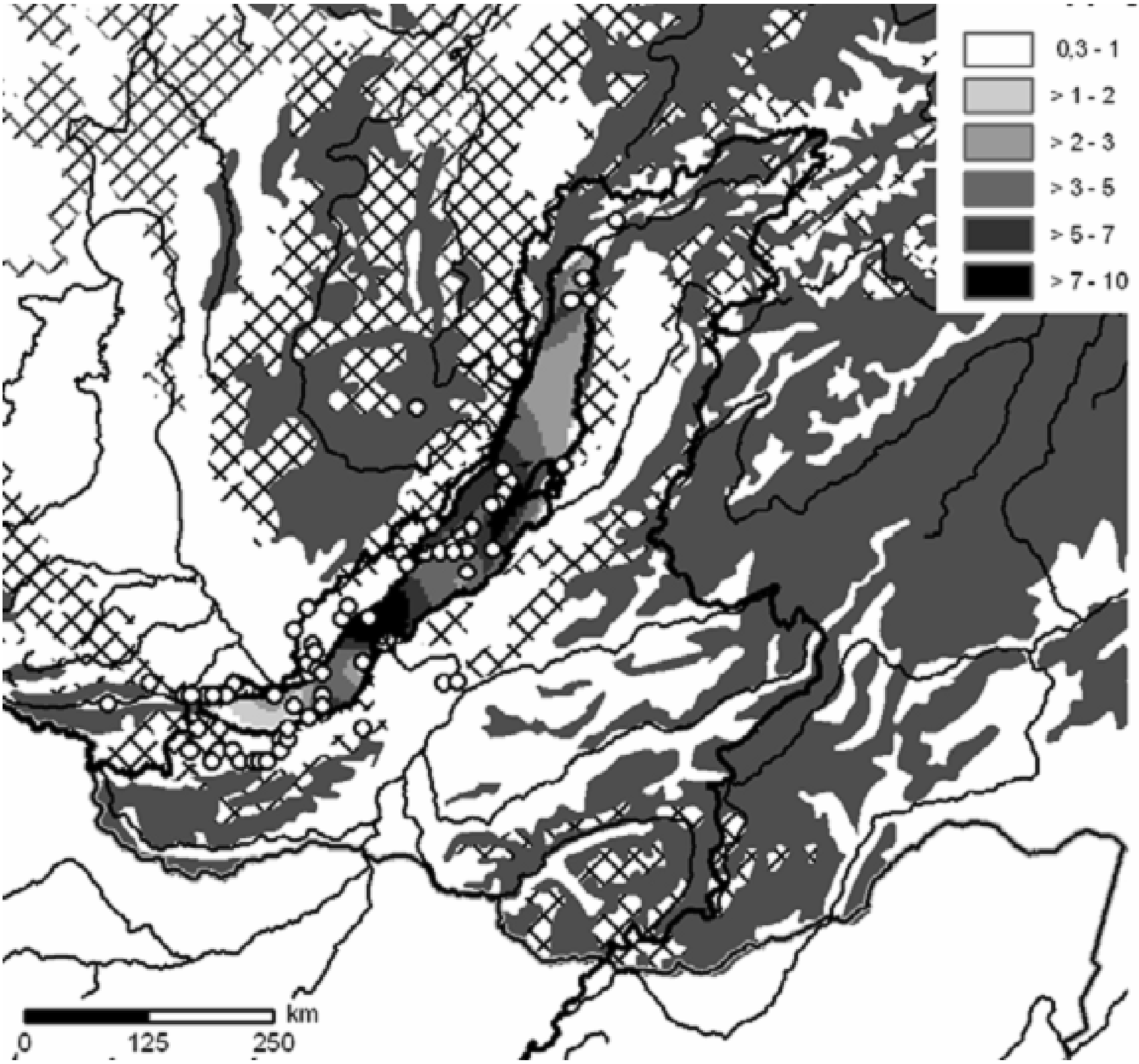
Figures

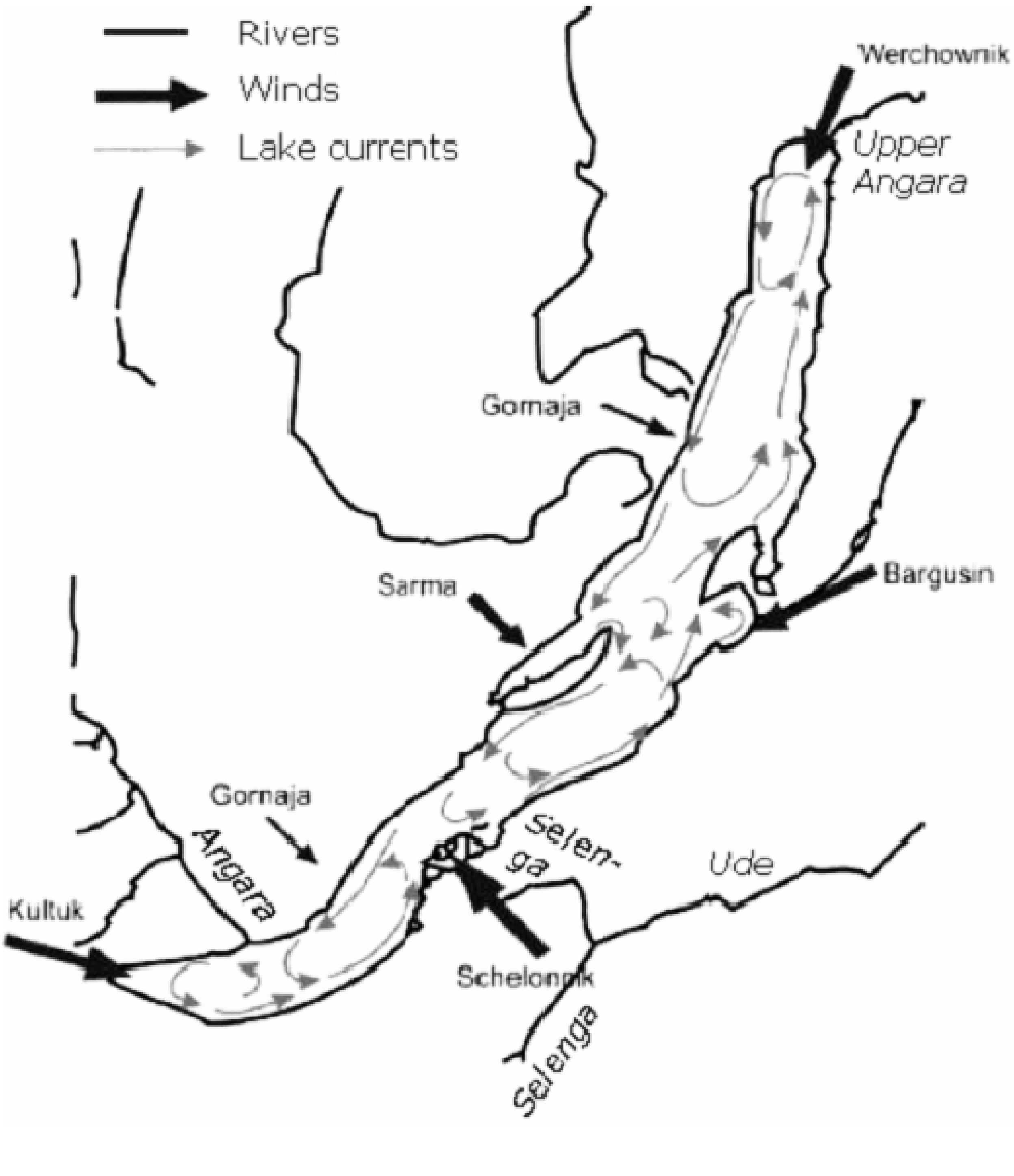

Arctic Ocean
Russia
Lake Baikal
Kazakhstan
Mongolia
Lake Balkhash
Aral Sea
Caspian Sea
Issyk Kul
China
Koko Nor
Sea of Japan
East China Sea


Baikal GIS - Mozilla Firefox
Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe
http://dc108.gfz-potsdam.de/website/Baikal/viewer.htm
bodys
CONTINENT Sites
Baikal GIS
eBay: Rennrad S-Crash Ultegra Race ...
Canyon | Testberichte
Baikal GIS
Oymur
Kudara
SCHERGINO
Bolshoye Kolesovo
KABANSK
Selenga
Koma
Tataurovo
ONOKHOY
Berezovskiy
ULAN-UDE
LENINA
ZAGORSK
Ivolginsk
BABUSHKIN
Orongoy
Map created with ArcIMS - Copyright (C) 1992-2002 ESRI Inc.
LAYERS
Available Datasets
Stations
CONTINENT
Others
Countour Lines
1000 meter
500 meter
100 meter
Infrastructures
International Bo
Railroads
Roads
Urban Areas
Rivers
Low Resolution
High Resolution
Catchment Are
Lakes
Catchment Are
Baikal
Inland Water Bo
Lake Sediment
Bathymetry
Geology
Relief
Satellite Image Lands
Refresh Map
Auto Refresh
Zoom Out
Fertig
Start
Web.de - Outlook Ex...
Arbeitsplatz
Adobe Photoshop
Baikal GIS - Mozilla Fir...
DE
15:58


Quarternary
Paleogene/Neogene (clays, sands, conglomerates)
Late Jurassic and Createeous (clays with coals)
Jurassic (limestones)
Permo-Triassic and Jurassic (sandstones, conglomerates, clays with coals, volcano-clastics)
Paleozoic
Lower Paleozoic (limestones, sandstoned, conglomerates)
Precambrian (gneiss, schists,amphibolitses, marbles)
Cenozoic extrusive rocks (basalts, volcano-clastics)
Mesozoic granitoids
Paleozoic granitoids
Precambrian granitoids
Mafic intrusive rocks (gabbros, pyroxenites, peridotites)
C
Barguzin
P
V
Uda
Selenga
Hilok
Dzida
Eg gol
Cikoi
Cikokon
Ider gol
Culuut gol
Hanui gol
Hoid Tamir gol
Orhon gol
Tuul gol
N
0 45 90 180 270 360
Km


10- 15
>15- 20
> 20 - 25
> 25 - 30
> 30 - 35
> 35 - 40
> 40 - 45
> 45 - 50
> 50 - 55
> 55 - 65
km
0
125
250


0,3 - 1
> 1 - 2
> 2 - 3
> 3 - 5
> 5 - 7
> 7 - 10
km
0
125
250


Rivers
Winds
Lake currents
Werchownik
Upper Angara
Gornaja
Sarma
Bargusin
Gornaja
Angara
Selen-ga
Ude
Kultuk
Schelonnik
Selenga

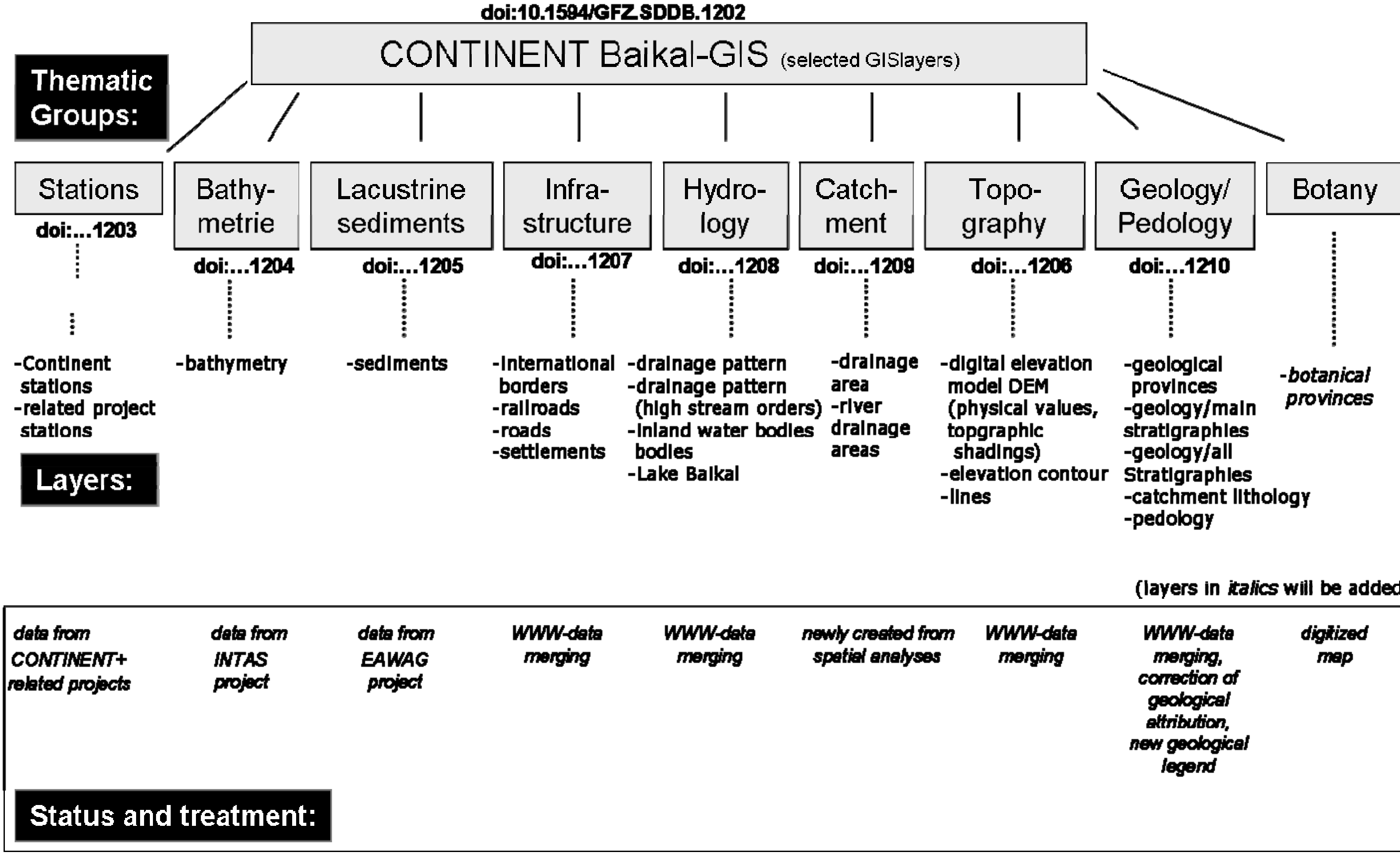

doi:10.1594/GFZ.SDDB.1202
CONTINENT Baikal-GIS (selected GISlayers)
Thematic Groups:
Stations
doi:...1203
Bathy-metrie
doi:...1204
Lacustrine sediments
doi:...1205
Infra-structure
doi:...1207
Hydro-logy
doi:...1208
Catch-ment
doi:...1209
Topo-graphy
doi:...1206
Geology/ Pedology
doi:...1210
Botany
-Continent stations
-related project stations
Layers:
-bathymetry
-sediments
-international borders
-railroads
-roads
-settlements
-drainage pattern
-drainage pattern (high stream orders)
-inland water bodies
-Lake Baikal
-drainage area
-river drainage areas
-digital elevation model DEM (physical values, topgraphic shadings)
-elevation contour
-lines
-geological provinces
-geology/main stratigraphies
-geology/all Stratigraphies
-catchment lithology
-pedology
-botanical provinces
(layers in italics will be added
data from CONTINENT+ related projects
data from INTAS project
data from EAWAG project
WWW-data merging
WWW-data merging
newly created from spatial analyses
WWW-data merging
WWW-data merging, correction of geological attribution, new geological legend
digitized map
Status and treatment:

Figure captions

Figure 1: Lake Baikal situated in the centre of the Asian continent, in southern Siberia. Its lake catchment and main tributary catchments are outlined (white lines). The square indicates the area covered by the CONTINENT Baikal-GIS data (1620 km by 1760 km).

Figure 2: Screen-shot of the CONTINENT Web GIS HTML-client, as shown in a web browser. The colour-coded digital elevation model DEM is overlaid by vector-layers of the infrastructure, elevation contour lines, drainage pattern, CONTINENT stations and lacustrine sediment regions. Here, the enlarged view provides detailed information of the Selenga delta area.

Figure 3: Thematic GIS-Layer 'lithology of catchments', showing the main litho-stratigraphic groups. The position of the CONTINENT sediment cores are indicated: C=Continent ridge (CON01-603), P=Posolskoe ridge (CON01-604), V=Vydrino shelf (CON01-605).

Figures 4a,b: GIS-Layers 'pollen distribution' (percentage in sample) overlaid with vegetation distribution (solide: dominant species, rhombs: present, blank: absent). a) distribution of *Pinus sylvestris* in the Lake Baikal catchment area, and interpolated distribution of *P. sylvestris* pollen in the lake. b) distribution of *Larix sp.* and interpolated distribution of *Larix sp.* pollen in the lake.

Figure 5: Prevailing wind systems about Lake Baikal (*Atlas of Baikal,* 1993) and the relative positions of cyclonic currents as proposed by Kozhov (1963).

Figure 6: Hybrid structure of selected data sets of the CONTINENT Baikal-GIS data as they can be withdrawn from DOIs by the scientific community.