Polarforschung 85 (2), 107–115, 2015 (erschienen 2016)

Geodatabase and WebGIS Project for Long-Term Permafrost Monitoring at the Vaskiny Dachi Research Station, Yamal, Russia

by Yury Dvornikov[1,2 *], Marina Leibman[1,3], Birgit Heim[2], Annett Bartsch[4,5,6], Antonie Haas[2], Artem Khomutov[1,3], Anatoly Gubarkov[7], Maria Mikhaylova[7], Damir Mullanurov[1], Barbara Widhalm[4,5], Tatiana Skorospekhova[8], and Irina Fedorova[8,9,10]

Abstract: The research station Vaskiny Dachi (VD) in central Yamal, Western Siberia was established in 1988. Continuous monitoring of the permafrost state is conducted since 25 years, which allows collecting a large amount of data related to permafrost state and environment of this region. To store and visualise the geospatial data, containing our knowledge of the research area and research topic, we created a geodatabase (GDB) to operatively process different types of geospatial data. The produced GDB contains so far 11 vector feature datasets and raster data in the same coordinate system. The vector data represent: 1) bathymetry; 2) social-economic objects; 3) field data; 4) geomorphology; 5) hydrography; 6) landscapes; 7) permafrost; 8) snow; 9) topography; 10) vegetation; 11) long-term measurement grids and transects (Circumpolar Active Layer Monitoring (CALM) transect, CALM measurement grid). All these feature datasets contain 60 feature classes of spatial data in total. Some of the geodata layers are directly linked to data bases of field data. The raster data contain 37 layers, including a digital elevation model with derivatives, a map of snow distribution for the key site, bathymetric maps and other maps of different scale. Moreover, the key area is a site for international research projects and the ongoing exchange of the data is supported by the VD GDB. Geographical Information System (GIS) allows collecting, storing and processing geospatial data from different sources in a wide range of types and formats. WebGIS platforms allow displaying the geospatial data for different users, giving the impression of the general processes on the certain geographic area. Also, we use the WebGIS service to publish the data and to make it available for the larger community. This paper is an overview on the permafrost studies at the VD research station, the GDB for permafrost monitoring as well as the established Yamal WebGIS project.

Zusammenfassung: Im Bereich der Forschungsstation Vaskiny Dachi (VD), in der zentralen Yamal-Region in West-Sibirien, werden seit 25 Jahren zahlreiche Messungen zur Überwachung des Zustande des Permafrost, der Vegetation, und der kryogenen Prozesse durchgeführt. Während dieser Zeit wurde eine große Menge verschiedener Informationen gesammelt, die strukturiert archiviert wurden. Für diese Aufgabe wurde eine spezielle Datenbank erstellt. Die Datenbank enthält sowohl Vektor- als auch Rasterdatensätze von Geoinformationen in einem einheitlichen Koordinatensystem. Insgesamt liegen 11 Vektordatensätze vor: 1) Bathymetrie, 2) soziale/wirtschaftliche Einrichtungen, 3) Geomorphologie, 4) Felddaten in Form von GPS-Koordinaten und Attribute, 5) Hydrographie, 6) Landschaften, 7) Permafrost, 8) Relief, 9) Schneedecke, 10) Vegetation, 11) lokale Daten (Transekte des Circumpolar Active Layer Monitoring (CALM). Diese Datensätze enthalten wiederum 60 Klassen von räumlichen Daten, wobei in einigen Fällen die Tabellen der Geodaten mit Daten von Geländemessungen verknüpft sind. Die Rasterdatensätze in der Datenbank umfassen ein Höhenmodell und verschiedene Derivate, die auf der Grundlage des Höhenmodells gemacht wurden, sowie die Verteilung der Schneedecke für Schlüsselbereiche und Rasterkarten verschiedener Maßstäbe. Des Weiteren enthält die Datenbank auch Aufnahmen von verschiedenen Fernerkundungs-Sensoren unterschiedlicher räumlicher Auflösung in einer einheitlichen Projektion. Die Struktur ermöglicht es, schnell verschiedene Daten zusammenzustellen und zu vergleichen, Untersuchungen von Abhängigkeiten zwischen Parametern durchzuführen und die Datenbank mit neuen Daten zu erweitern. Auch nutzen wir ein WebGIS, um die Geodaten zu publizieren und einem größerem Publikum zur Verfügung zu stellen. In dieser Veröffentlichung werden die Permafrostuntersuchungen, die darauf aufbauende Geodatenbank und das Yamal WebGIS Projekt in der VD Untersuchungsregion vorgestellt.

INTRODUCTION

The research station Vaskiny Dachi (VD) in central Yamal, West Siberia, Russia was established in 1988. Since then monitoring of permafrost-related and environmental parameters is on-going, which results in a large data collection related to permafrost state and environment of this region (LEIBMAN et al. 2015). Management of the collected data is important and also needed for the exchange of the data within on-going international research projects.

The long-term permafrost monitoring site VD is run by the Earth Cryosphere Institute, Siberian Branch of Russian Academy of Sciences (LEIBMAN et al. 2015) and an established site for the international Global Terrestrial Network for Permafrost (GTN-P) program (BISKABORN et al. 2015), including international CALM (Circumpolar Active Layer Monitoring, since 1993) (BROWN et al. 2000) and TSP (Thermal State of Permafrost, since 2011) (BROWN et al. 2010) programs. In 2007, VD became a part of the North Eurasian Transect within the Land Cover Land Use Change for Yamal Peninsula (LCLUC-Yamal) project with three more LCLUC-CALM grids and shallow boreholes established (LEIBMAN et al. 2012). Since 2013 also the technogenic impact due to high impact of the Bovanenkovo gas field development on the tundra environment and on permafrost state is investigated (KHOMUTOV & KHITUN 2014). Permafrost and environmental monitoring is also part of the on-going Russian-Austrian COLD Yamal (COmbining remote sensing and field studies for assessment of Landform Dynamics and permafrost state on Yamal) project that focuses on radar remote sensing

doi:10.2312/polfor.2016.007

[1] Earth Cryosphere Institute Siberian Branch Russian Academy of Sciences, Malygina 86, 625000 Tyumen, Russia;
[*] Corresponding author: <ydvornikow@gmail.com>
[2] Alfred Wegener Institute Helmholtz Center for Polar and Marine Research, Telegrafenberg A43, 14473 Potsdam, Germany.
[3] University of Tyumen, Volodarskogo 6, 625003 Tyumen, Russia.
[4] Zentralanstalt für Meteorologie und Geodynamik ZAMG, Hohe Warte 38, 1190 Vienna, Austria.
[5] Austrian Polar Research Institute, Althanstrasse 14, 1090 Vienna, Austria.in
[6] Vienna University of Technology, Gusshausstrasse 27-29, 1040 Vienna, Austria.
[7] Tuymen Industrial University, Volodarskogo 38, 625000 Tyumen, Russia.
[8] Arctic and Antarctic Research Institute AARI, Beringa 38, 199397 St. Petersburg, Russia.
[9] Institute of Earth Science, Saint-Petersburg State University, 10th line of Vasiljevsky Island 33–35, 199178 Saint-Petersburg, Russia.
[10] Kazan Federal University, Kremlyovskaya 18, 420008 Kazan, Russia.
This paper was presented as a poster presentation at the International Conference "Our Climate – Our Future: Regional perspectives on a global challenge", 6–9 October 2014 in Berlin, Germany.
Manuscript received 22 June 2015; accepted in revised form 16 March 2016.

applications for thermokarst lakes and permafrost landscape dynamics in Yamal. In 2011 first tentative water sampling for organic matter concentration in thermokarst lakes started and transformed to an ongoing Russian-German research project, POLYAR (Process of Organic matter transport to the Lakes of Yamal Region, Dvornikov et al., 2014). The main purpose of this project is to extract the environmental characteristics of lake catchments and the geochemistry of lake water bodies in Central Yamal. Geographical information systems are a useful tool for management of geodata, i.e. data with an assigned geo-referencing. Existing geodatabases in geocryology aim to collect and store historical data and ongoing monitoring data (e.g., Global Terrestrial Network for Permafrost (GTNP), Biskaborn et al. 2015), collect knowledge on permafrost (Romanovskii & Leibman 1994) as well as saving old data gathered before the "digital revolution" (e.g., CAPS, International Permafrost Association 2003). The requirements for the creation of a geodatabase in geocryology may differ in terms of its object basis (Melnikov & Minkin 1998, Minkin et al. 2001). For the geodata of the VD region we collected all available geospatial data in the GeoDataBase (GDB), undertook GIS-oriented modelling and mapping, to analyse the linkages in the "permafrost dynamics – landscape change" system.

Note that acronyms of projects and associated websites are listed in Table 7 at the end of this paper.

STUDY AREA

The Vaskiny Dachi (VD) research region is located in the western Siberian Arctic (Fig. 1). The area belongs to the continuous permafrost zone with thickness down to 500 m and is characterised by a wide distribution of tabular ground ice (Leibman 1996, Khomutov et al. 2012). The presence of tabular ground ice in connection with active layer dynamics is responsible for the cryogenic landsliding and thermal denudation in this area (Leibman & Kizyakov 2007). In 2012, a considerable activation of thermal denudation was observed in the VD region (Leibman et al. 2015) due to an extremely warm summer accompanied by active layer deepening. At least six wide exposures along the lake shore line of thermokarst lakes were documented during 2012 and 2013 field expeditions.

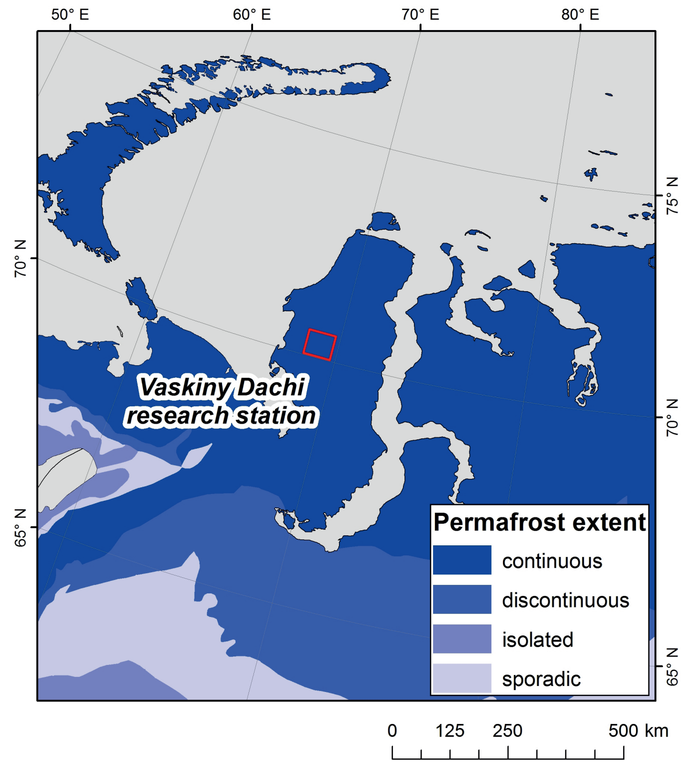

Fig. 1: Location of the region Vaskiny Dachi (VD) in central Yamal, Western Siberia. Background: Map of the permafrost extent (Brown et al. 1998).

Abb. 1: Geographische Lage der Region Vaskiny Dachi (VD) in Zentral-Jamal, West Sibirien. Hintergrund: Karte der Permafrostausdehnung (Brown et al. 1998).

DATA AND METHODS

Several datatypes are used to collect the geospatial data including:

(i) *in-situ* field observations and measurements (active layer on 4 established grids and at additional sites, ground temperature-depth profiles in 11 boreholes, geodetic measurements for the cryogenic landslide monitoring, and ground control points (GCP) collection for satellite-image geometric correction, bathymetry surveys, description of exposures, snow survey (snow depth and snow depth measurements);

(ii) geochemical and geo-optical data from lakes, soils and ground ice (results obtained in the laboratory and linked to a specific sampling point which includes ground ice samples, water samples from thermokarst lakes, soil samples);

(iii) data extracted from satellite images and aerial photographs taken in different years as well as data extracted from digital elevation models (DEM) (vegetation classes, landscape map, maps of the manifestation of cryogenic processes, maps of technogenic impact on the tundra (Khomutov & Khitun 2014), map of environmental hazards);

(iv) GIS-modelled data (snow water equivalent (SWE) map, Dvornikov et al. 2015). The sources of data for the GDB are presented in Table 1.

Field data collection

CALM and TSP data

Active layer depth is measured by a metal probe according to the CALM protocol within four grids. Such a grid is fundamental for the study of spatial and temporal variations of active layer depth and vegetation patterns. These include lithological and cryogenic controls and surface cover controls for spatial distribution. Temporal variations are controlled by climate fluctuations which are mediated by topography, lithology and surface covers (Leibman 1998).

The VD CALM grid was established in 1993 on the top and slope of a highly dissected alluvial-lacustrine-marine plain, affected by landslides, with sandy to clayey soils. Within the LCLUC program "Greening Of the Arctic" (GOA) other three grids were established in 2007 in typical conditions of

Type		Number of sites	Number of points/units	Year	Equipment
1. Field data					
ALD (Active Layer Depth) **measurements**		1 (CALM)	121	1993-	
		3 (LCLUC)	165	2007-	
Boreholes			12	2006-	HOBO
Tachymetry		3	2246	2011	TopCon© GTS-235
DGPS measurements	**Static**		7	2014	Trimble© GNSS 5700
	Kinematic		199	2014	
Bathymetry		20	>25000	2012, 2014, 2015	Humminbird© 788cxi, points tracked with 5 seconds interval
Snow survey		3	233	2013	Metal ruler, snow sampler VS-43
Water samples		29	>100	2011-2015	
2. Remote sensing data / Digital maps					
Optical data	VHR			2009	GeoEye-1 (MS)
				2010	QuickBird (Pan, MS)
				2013	GeoEye-1, WorldView-2 (Pan, MS)
	HR			2015	SPOT-5 (MS)
Radar data				2008	TerraSAR-X, ALOS PALSAR
				2010	TerraSAR-X
				2014	Terrasar-X, TanDEM-X
DEM digital elevation model				2014	TanDEM-X IDEM 12 m
Topographic maps			2	1987	1:50000
3. Laboratory data					
CDOM coloured dissolved organic matter				2011-2015	Specord 200 (Jena Analytic©)
DOC dissolved organic carbon				2015	TOC-V-CPH Shimadzu©
SPM suspended particulate matter				2014-2015	0.45 µm CA filters
MI major ions				2014-2015	Ion Chromatography

Tab. 1: List of data sources for Vaskiny Dachi GDB. ALD: active layer depth; DEM: digital elevation model; CDOM: coloured dissolved organic matter, DOC: Dissolved organic carbon, SPM: suspended particulate matter, MI: major ions.

Tab. 1: Liste der Datenquelle für Vaskiny Dachi GDB. ALD: Tiefe der Auftauschicht; DEM: Digitales Höhenmodell; CDOM: Gelbstoff, DOC: gelöster organischer Kohlenstoff, SPM: Schwebstoff, MI: Hauptionen.

IVth Coastal-marine plain (VD-1), IIIrd Alluvial-marine plain (VD-2) and IInd River terrace of Mordyakha river (VD-3). Each 50 × 50 m measurement grid covers a homogeneous terrain.

Boreholes are equipped under the CALM and TSP Programs (Tab. 2) with HOBO data loggers. The depth of boreholes ranges from 0.8 m up to 10 m. Ground temperature was first measured to the depth of 10 m in 1993 in a borehole “LGT”, which was drilled in 1988. This borehole was re-established in 2010 to the depth of only 7 m. Borehole “191-m” was drilled next to the CALM grid in 1990 to the depth 10 m and ground temperature was measured by mercury thermometers in summer only till 1996. In 2011, the borehole was replicated (new ID: “VD CALM_10”) in the same place and equipped with HOBO data loggers. Ground temperature measurements in the shallow boreholes were used to understand the process of freeze back in various landscape conditions (Leibman et al. 2015).

Bathymetry

The depth of 19 lakes was measured during 2012, 2014 and 2015 field expeditions (Tab. 3). Chartplotter Humminbird 788 cxi with internal GPS was used to collect both distances between sonar and lake bottom and position coordinates. The readings were taken along several profiles for each lake and then were filtered with 5 seconds and converted to ASCII format. Bathymetric maps were produced by interpolation.

Snow measurements

Snow survey in the study area has been carried out in March 2013. Snow depth was measured by digging a snow pit and measuring the snow layer using a metal ruler. Samples for snow density were taken with the snow sampler VS-43 designed for snow density measurements during snow surveys (Slabovich 1985). The measurements are described in detail in Dvornikov et al. (2015).

Water sampling

Water samples from lakes were collected within the POLYAR research project in order to study a number of hydrochemical parameters of the lake water: coloured dissolved organic matter (CDOM) (Dvornikov et al. 2016), dissolved organic carbon

	Site ID	X	Y	N points	established
CALM	CALM grid	68.9072	70.2836	121	1993
	VD-1 (GOA)	68.8916	70.2755	55	2007
	VD-2 (GOA)	68.8835	70.2955	55	2007
	VD-3 (GOA)	68.8413	70.3014	55	2007
TSP	**Borehole ID**			**Max depth (cm)**	**Logger installed**
	VD 1 (GOA)	68.8916	70.2755	90	27-08-2007
	VD 2 (GOA)	68.8835	70.2955	100	27-08-2007
	VD 3 (GOA)	68.8413	70.3014	100	29-08-2007
	VD CALM	68.9072	70.2836	150	25-08-2006
	VD AG19/3	68.9091	70.2826	150	29-08-2007
	VD LGT	68.8916	70.2811	700	13-09-2010
	VD CALM_10m	68.8913	70.2811	1000	25-08-2011
	VD AGG1	68.9294	70.2943	100	20-08-2011
	VD AGG2	68.9300	70.2929	80	20-08-2011
	VD Gully	68.9271	70.2912	170	03-09-2012
	VD landslide	68.9238	70.2796	140	26-08-2013
	VD Gully_2	68.9271	70.2912	200	24-08-2014

Tab. 2: List of sites for ALD (Active Layer Depth) measurements and boreholes within CALM and TSP projects. GOA: Greening of the Arctic project.

Tab. 2: Liste der Standorte von der Tiefe der Auftauschicht (ALD) sowie der Bohrlöcher von CALM und TSP-Projekten, GOA: Greening of the Arctic project.

Lake_ID	Year	Area (ha)	Y	X	N points recorded
LK-001	2012	38.30	68.8829	70.2787	790
LK-002	2012	3.26	68.9045	70.2977	448
LK-003	2012	107.72	69.0019	70.2898	1652
LK-004	2012	74.96	68.9705	70.2809	2167
LK-006	2012	3.55	68.8991	70.2878	655
LK-007	2012	38.77	68.9912	70.2672	1000
LK-010	2012	4.75	68.8642	70.3012	412
LK-012	2012	2.24	68.9216	70.2825	473
LK-013	2012	212.46	68.8843	70.2563	1156
LK-014	2015	6.55	68.8736	70.2836	1121
LK-015	2015	8.58	68.9218	70.2651	2313
LK-016	2015	10.25	68.9335	70.2668	1472
LK-017	2015	6.37	69.0221	70.2326	1195
LK-018	2015	11.79	69.0061	70.2319	1782
LK-019	2015	13.60	68.9951	70.2301	2553
LK-035	2014	66.53	68.8738	70.2155	5254
LK-037	2014	2.02	68.9938	70.2498	549
LK-038	2014	3.84	68.9974	70.2448	925
LK-039	2014	2.69	69.0024	70.2494	928

Tab. 3: List of lakes in which bathymetry survey was conducted.

Tab. 3: Liste der Seen, in denen bathymetrische Messungen ausgeführt wurden.

(DOC), suspended matter (SPM), major ions (MI) (Tab. 4). Samples were collected in calm weather conditions (to avoid mixing effects) from upper 30 cm of water column close to the shore or in the centre of lakes from boat in August-September of the 2011–2015 field expeditions. Criteria for selecting the lakes were the difference in the visual (water colour in the satellite images in "true colour" RGB composite) as well as morphometric properties of coastal area (presence of high cliffs, thermal denudation or cryogenic landslides).

Remote sensing and GIS data processing

Optical satellite sensor data and synthetic aperture radar (SAR) satellite data are used to extract additional parameters. High-spatial resolution optical satellite data were used to analyse the vegetation as well as to detect activation of cryogenic processes. Time series of TerraSAR-X data are used to extract water bodies and to assess the seasonality of the thermokarst lakes (Bartsch et al. 2012, Trofaier et al. 2013). TanDEM-X digital elevation model (DEM) was used to delineate lake catchments, for topography analysis as well as to model snow distribution in the study area.

Optical satellite data processing

The optical satellite images from high-spatial resolution sensors (GeoEye-1, QuickBird, WorldView-2) obtained from Digital Globe Foundation© (DGF) in both, very high spatial resolution panchromatic (0.5–0.6 m) and lower spatial resolution multispectral (2.0–2.4 m) bands. All the acquisitions have been taken in the summer season: QuickBird on 2010-07-30, GeoEye-1 on 2013-07-05, WorldView-2 on 2013-07-21 (Tab. 5). Pan-sharpened images for all three acquisitions were obtained, applying PANSHARP2 fusion algorithm developed by Zhang (2004) with the PCI Geomatica 2014 software (PCI Geomatics©), where the multispectral and panchromatic bands were merged to obtain the high spatial resolution together with multispectral information in one image. The pan-sharpened very high resolution images were used to manually digitise streams, areas covered by shrubs and other visual analysis, but not for the spectral processing. The ortho-rectification procedure was performed within the OrthoEngine module in PCI Geomatica. The TanDEM-X IDEM (Tab. 5) was used to correct the images for relief distortions. ATCOR (Richter 1996) ground reflectance atmospheric correction module was used to correct the image data within PCI Geomatica software.

Multispectral processing is the base for the extraction of vegetation distribution based on vegetation indices. To calculate the vegetation indices, water masking, using a near infrared

Year	CDOM	DOC	SPM	MI
2011	11	–	–	–
2012	7	–	–	–
2013	13	–	–	–
2014	21	–	16	21
2015	24	24	24	24
Total	76	24	40	45

Tab. 4: Overview on the lake sampling in 2011 to 2015. CDOM: coloured dissolved organic matter; DOC: dissolved organic carbon; SPM: suspended particulate matter; MI: major ions.

Tab. 4: Überblick über die Wasserbeprobungen von 2011 bis 2015. CDOM: Gelbstoff, DOC: gelöster organischer Kohlenstoff, SPM: Schwebstoff, MI: Hauptionen.

	Sensor/ [bands]	Type*	Acquired	Source	Spatial resolution (m)
Optical data	GeoEye-1/ [4]	PS	2009-08-15	NASA NGA License UAF	0.5
	–	MS	2013-07-05	DGF**	2
		PS			0.5
	QuickBird/ [4]	MS	2010-07-30	–	2.4
		PS			0.6
	WorldView-2/ [8]	MS	2013-07-21	–	2
		PS			0.5
Radar data	TerraSAR-X/[1]		2008-07-11, 2008-09-15	DLR PI agreement LAN1706	2.5
	–		2010-07-29, 2010-08-31	–	–
	–		2014-07-04, 2014-08-10	–	2
	ALOS Palsar/ [1]		2008-08-14 2008-09-29	JAXA PI agreement 90 and 1200	16
DEM	TanDEM-X		2013-06-19	DLR PI agreement LAN1706	12

Tab. 5: List of remote sensing data. *different types of data used; MS: multispectral; PS: pan-sharpened (PANSHARP2 model, Zhang 2004); DGF**: Digital Globe Foundation.

Tab. 5: Liste der Fernerkundungsdaten. *verschiedene Datentypen wurden genutzt; MS: multispectral, PS: pan-sharpened (PANSHARP2 Modell, Zhang 2004); DGF**: Digital Globe Foundation.

(NIR) band threshold, was performed for the further spatial analysis. NDVI index (Kriegler et al. 1969), which describes the abundance and "greenness" of vegetation, is calculated using the following equation:

$$NDVI = \frac{NIR - R}{NIR + R}$$

where NIR and R are the reflectance values in near infrared and red bands, respectively.

The relative chlorophyll absorption vegetation index CHL is calculated as follows:

$$CHL = \frac{G + NIR}{2} - R$$

where G is the reflectance value in green band.

Radar satellite data processing

TerraSAR-X time series and the TanDEM-X DEM with 12 m resolution have been available via DLR PI agreement LAN1706. The time series of water-body extents make it possible to investigate the seasonal area change of thermokarst lakes (e.g., Trofaier et al. 2013).

The TerraSAR-X data was derived as Level 1B SSC (Single-Look Slant Range Complex) Data with HH polarization. For data processing, the open source software NEST (Next ESA SAR Toolbox) was used. The data were multilooked and terrain corrected using Range Doppler Terrain Correction and the DEM product of TanDEM-X (12 m spatial resolution). The output pixel spacing was set to 2 m, resembling the source ground range pixel spacing and projected into UTM Zone 42, WGS-1984. Radiometric normalisation was applied and the resulting Sigma0 band was transferred to dB using the Software IDL. Water bodies were extracted by using a threshold method. Due to specular radar reflection on open water bodies, the resulting lower backscatter values can be used to differentiate between land and open water bodies.

ALOS PALSAR data have been available via JAXA PI agreement 90 and 1200 and were used to extract shrub contours for the study area. The data Level 1.1 HV data of two dates (14^{th} Aug. and 29 Sep. 2008) were processed within NEST. The multilooked images were ellipsoid corrected using the Geolocation-Grid method. The resulting intensity images were multiplied by a factor of 10^{12} and converted into dB. The mean of the two dates was calculated and gamma filtered. To extract the shrub layer a threshold of -25 dB was used.

TanDEM-X IDEM was used to extract catchment polygons, for terrain analysis and as a source of relief derivatives for snow redistribution modelling. Delineation of the lake catchments requires the high quality DEM. To improve the model, the raster image is converted to the point data model and re-interpolated with interpolation method in ArcGIS 10.2.2. (ESRI Inc©). After interpolation procedure the DEM is levelled, i.e., the extracted lake polygons from TerraSAR-X 2014.08.10 are used as vector polygons for assigning the specific elevation to the lake areas. Extreme sinks are filled using ArcHydro module. The optimized DEM is used to calculate the flow direction raster model in ArcHydro (Djokic et al. 2011) and then for catchment delineation. All automatically calculated catchment areas are manually corrected for the outflows from the lakes.

DEM derivatives are curvature, aspect and slope. These parameters have been used as an input data for snow modelling together with field data and the shrub layer derived from ALOS Palsar data.

VD GDB database

The established GDB for the VD region enables the data management of all the different data sets: from the long-term GTN-P (CALM) program and all the vector and raster data produced within the POLYAR project.

Geodetic measurements (DGPS and tachymetry) in the area as well as the geocoded very high resolution images provide the high accuracy of the spatial location of the data. Geospatial data in the GDB is presented in two different data models: vector and raster. The vector data in GDB is divided into several feature classes, according to specific geographic objects (e.g., topography, vegetation, permafrost, hydrography etc.). Each feature dataset is filled with a number of feature classes (points, polylines and polygons). The data of similar types are linked to each other using unique object ID.

WebGIS application

WebGIS is a useful tool for visualising and analysing the geodata of different sources and types. For a WebGIS application GIS is compiled on a server and transferred to the user web-browser in interactive web viewers. Users can perform their own GIS visualization. The Alfred Wegener Institute, Helmholtz Centre for Polar- and Marine Research (AWI) offers a WebGIS service, an application based on JavaScript. The WebGIS core components are ArcGIS for Server 10.3 and PostgreSQL databases 9.3.1 including Spatial Database Engine (SDE). The AWI WebGIS server has been used for publishing selected GIS data layers of the Vaksiny Dachi GDB as the Yamal WebGIS Project. The selected GIS data are layers, reflecting the general state of the ongoing research in the station such as lake catchments, water sampling points, active layer grids, river network, DEM raster, snow map raster etc.

RESULTS

Vaskiny Dachi Geodatabase

The compiled GDB contains 11 vector feature datasets and raster data in the WGS-84 coordinate system, projection UTM Zone 42 North. The vector data represents: 1) bathymetry; 2) social-economic objects; 3) field data; 4) geomorphology; 5) hydrography; 6) landscapes; 7) permafrost; 8) snow; 9) topography; 10) vegetation; 11) long-term measurement grids and transects (CALM transect, CALM measurement grid). All the feature datasets contain 60 feature classes of spatial data in total (Tab. 6).

Bathymetry feature dataset contains measured depths in 19 lakes (Tab. 3) as well as GIS-derived isobaths. Social-economic objects feature dataset contains only polylines characterising the off-road tracks classified according to its impact on the landscape (low, medium, high) (Khomutov & Khitun 2014). The tracks are derived from the satellite images and aerial photographs (Tab. 5). Field data feature dataset contains the points of the field monitoring sites as well as the location of samples with the laboratory measured parameters as a supplementary information in the attribute table. Geomorphology feature dataset represents polygons of different geomorphic levels obtained from DEM and topographic maps (Tab. 5). In addition, this feature dataset contains lake-catchment polygons extracted from the TanDEM-X IDEM with the large amount of supplementary information (volume of snow in form of SWE, high shrubs percentage, occurrence of cryogenic processes etc.). Each catchment polygon is linked to a specific lake polygon, using the unique object ID.

Feature datasets	Number of feature classes	Number of points	Number of polylines	Number of polygons
Bathymetry	2	61851	261	–
Social-economic objects	2	–	170	–
Field data	6	459	–	–
Geomorphology	9	–	–	761
Hydrography	9	–	360	415
Landscapes	1	–	–	47
Permafrost	9	120	–	711
Snow	5	99	–	–
Topography	10	88	673	7
Vegetation	3	–	–	16143
Local data	4	2683	–	–
Total	60	65300	1464	18084

Tab. 6: Geospatial feature datasets in the Vaskiny Dachi GDB.

Tab. 6: Räumliche Datensätze in der Vaskiny Dachi GDB.

Hydrography – digitalized thermokarst lake polygons with the large attributive information including the transferred laboratory data (CDOM, IC, SPM, DOC; Dvornikov et al. 2016), seasonal area change derived from the TerraSAR-X time series etc. Landscapes feature dataset contains only one polygonal feature class representing the landscape map for the research station (Khomutov & Leibman 2014).

Permafrost - large dataset containing the location of boreholes with a certain attributes (depth, temperature values), remote sensing derived map of the cryogenic processes, location and characteristics of observed thermocirques. Snow feature dataset represents the *in-situ* measured snow depth in snow density points and the interaction with topography and vegetation parameters. Topography feature dataset contains the results of the tacheometry survey of the area.

Vegetation – mostly remote sensing derived contours of different vegetation classes (e.g., shrubs and sedges). Local data – CALM grid and GOA grids with a multiyear ALD measurements as well as a transect data. In total, about 80000 geometry objects in vector format are in the GDB (Tab. 6). Most of the objects are *in-situ* measured depths of thermokarst lakes and vegetation contours. The specific geospatial basis of all the data allows work with the data very fast, obtaining various geographic information and geographic interaction between objects, providing the local-scale permafrost modelling. One successful application of the GIS-based local scale modelling of SWE in the research station has been already carried out in Dvornikov et al. (2015).

Raster data in the VD GDB contains: DEM and different derivatives, SWE map, bathymetric maps of several lakes, raster maps of different map scales, aerial photographs, corrected and processed satellite images (Tab. 5). The raster data are often the base for the extraction of vector objects in the GDB (e.g., vector data from vegetation, topography, hydrography). The general structure of the VD GDB (Fig. 2) is very evolving and the new data is monthly uploaded. The spatial data analysis leads to creation of new feature classes and new attribute rows.

Yamal WebGIS

The majority of created geodata was published as the Yamal WebGIS service, which serves as a tool for the data visualization (Fig. 3). WebGIS is not a flexible tool for external users in terms of data processing but it provides a visualisation of topographic and environmental variables. The WebGIS enables visualisation on which lakes have in- and outlets, where landslides and wind-blown sandy areas are in catchments, or on what is the spatial distribution of shrubs.

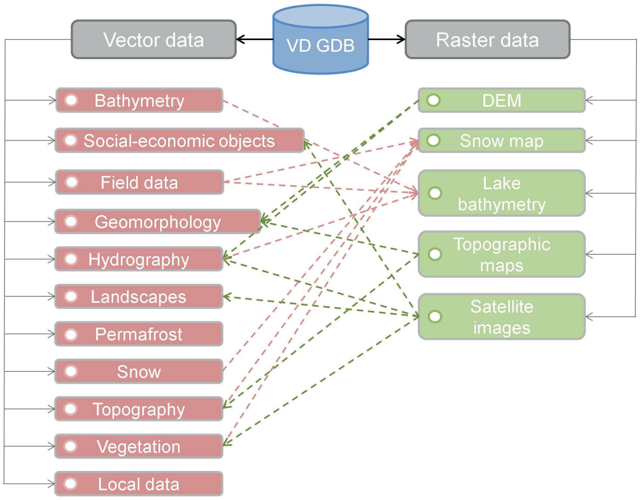

Fig. 2: Schematic structure of data layers of the Vaskiny Dachi (VD) database GDB).

Abb. 2: Schematische Struktur der Daten in der Vaskiny Dachi (VD) Datenbasis (GDB).

The Yamal WebGIS project is available and keeps the data structure form of the GDB also in the published WebGIS data sets (Fig. 4).

DISCUSSION

The created Vaskiny Dachi (VD) Geodatabase (GDB) is aimed to support the analyses of processes on this type of high-latitude tundra permafrost landscape. The GDB is adapted to fast data accumulation, processing and obtaining higher level geospatial products (e.g., SWE map, lake catchments, land cover map etc.), including spatial and temporal GIS-based modelling. The structure of GDB allows collecting the data from different sources, most of which are *in-situ* measured field parameters and remote sensing derived data. Field data is a main source for the GDB and, furthermore, is the ground truth data for validation of remote sensing products (Fig. 4). Ground truth validated environmental parameters derived from remote sensing data might be used for data extrapo-lation on the larger areas, enabling to produce new GIS layers and digital maps (Fig. 4). Both, *in-situ* measured and remote sensing data are suitable for GIS or spatial statistic based modelling. This combination is very important for the permafrost area, because the modelling of underground permafrost parameters such as ground temperature or active layer depth on the local or regional scale can only be done using landscape-based approach (e.g., STRELETSKIY et al. 2012).

Nowadays, existing databases for permafrost monitoring are predominantly oriented on global scale. For example: CAPS, which contains the information about several permafrost characteristics and supporting metadata (INTERNATIONAL PERMAFROST ASSOCIATION 2003). Another service is GTNP, collecting and disseminating the data on two major permafrost variables (ALT and permafrost temperature). National geocryological databases also exist (e.g., MINKIN et al. 2001). Many local-scale geodatabases are developed by various research groups on different objects and areas, because they significantly support the research process, providing interlinkages between environmental parameters and variables. It also

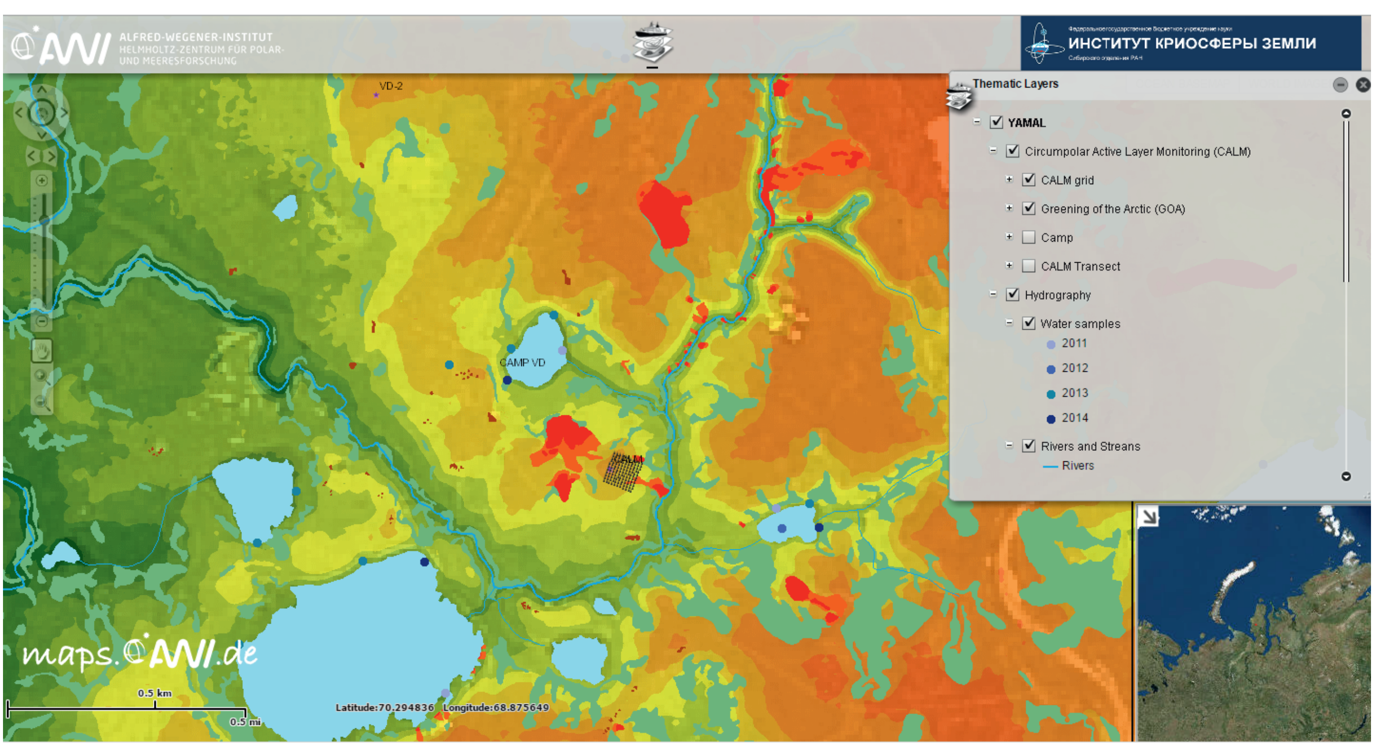

Fig. 3: Visualisation of the Vaskiny Dachi (VD) WebGIS <http://maps.awi.de/yamal/>. User-oriented interface allows to analyse the geodata. Thematic layers include field sampling points, raster surfaces and vector datasets.

Abb. 3: Visualisierung von Vaskiny Dachi (VD) WebGIS <http://maps.awi.de/yamal/>. Die benutzerorientierte Oberfläche erlaubt es, die Geodaten zu analysieren. Thematische Daten umfassen Beprobungspunkte, Rasterdaten und Vektordatensaten.

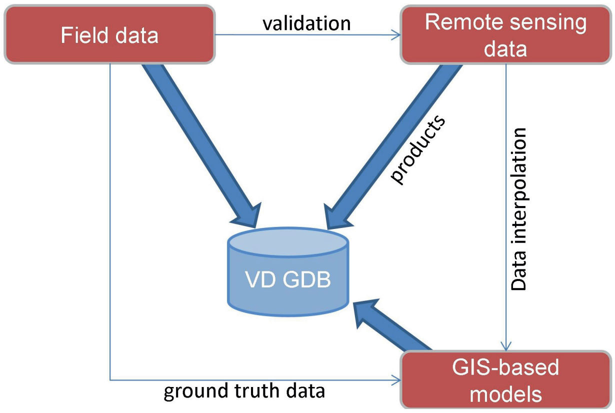

Fig. 4: Structure of data sources for the geodatabase (GDB) and its interconnection.

Abb. 4: Struktur von Datenquellen der Geodatenbasis (GDB) und deren Verbindung untereinander.

allows distributing the data within the research group members (joint research projects, international projects, PhD and master projects). The developed structure of the GDB allows fast extracting specific feature classes when it is considered as "completed" and publishing the data through world data centres such as PANGAEA. The international permafrost research covers different aspects (permafrost parameters, hydrology, geobotany, geomorphology etc.) and to establish a GDB is a very useful instrument to combine and process the knowledge of these different environmental parameters.

Both, desktop GIS and WebGIS enable visualisation. Related to the POLYAR project we visualise the location of the water sampling data related to the catchments size and properties. The Desktop GIS for example enables the combination from raster data such as the NIR band from a coarser scale resolution Landsat to optimally show the water bodies overlain by half-transparent color-coded DEM.

This raster dataset combination can be overlain by vector data of sampling points and lake catchments (Fig. 3). This combination or other possible combinations (e.g., snow depth map, vegetation indices distribution), supports the understanding of the results of the field data (different concentrations of CDOM, SPM, hydrochemistry in the lakes), and can guide the further statistical analyses. It is also possible to undertake area-based calculations to statistically derive the relationships from the range of catchment properties towards the geochemical composition in the Yamal lakes.

WebGIS is publicly available on the AWI Maps server, enabling users without technical GIS expertise to get an impression of the potential driving factors. Using the WebGIS visualisation, e.g., immediate interpretation is possible: *Which lakes have in- and outlets? Where are landslides and wind-blown sandy areas in catchments? What is the spatial distribution of shrubs?* The geomorphic spatial context is given with the GIS layer of the geomorphic terraces and the DEM (Fig. 5). The Yamal WebGIS is a very useful communication tool for the exchange between the international project partners and for the presentation of the VD study site. Recently, the developed regional scale geodatabase and WebGIS project were shown as a very useful tool for supporting big research

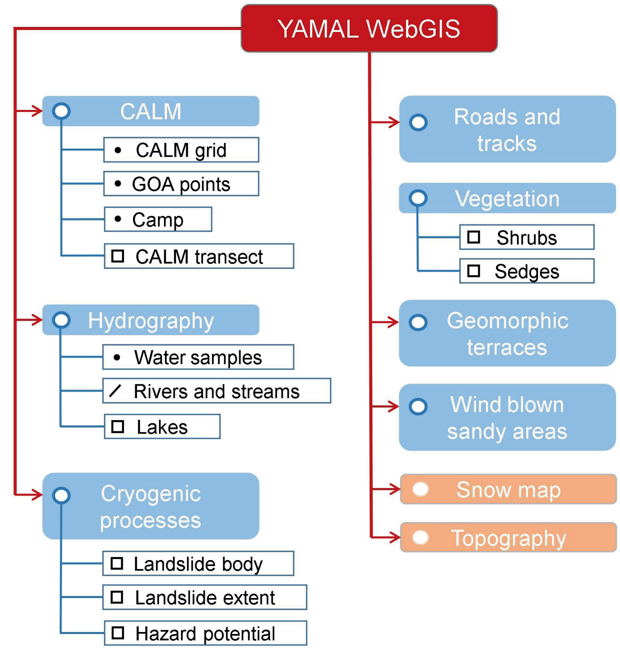

Fig. 5: Structure of the Yamal WebGIS (blue coloured boxes: vector data; orange coloured boxes: raster data).

Abb. 5: Struktur von Jamal WebGIS (blau gefärbte Felder: Vektordaten; orange gefärbte Felder: Rasterdaten).

projects, e.g., CONTINENT on Lake Baikal, and helped for obtaining new data using the existing open source datasets as well as data collected in the field (Heim et al. 2008).

SUMMARY

1) The VD GDB contains 11 vector feature datasets with more than 80000 geometry objects as well as a raster data enriched with metadata. The GDB is the base for the Yamal WebGIS published on the AWI Maps platform.
2) Field data are a main source for the GDB and is supporting the interpretation and generation of remote sensing products and GIS-based modelling.
3) Ground truth validated environmental parameters derived from remote sensing data might be used for data extrapolation on the larger areas, enabling to produce new GIS layers and digital maps.
4) Both, *in-situ* measured and remote sensing data are suitable for GIS or spatial statistic based modelling.
5) The Yamal WebGIS visualizes the study region, studied objects, and *in-situ* monitoring and sampling for a larger community. Within the POLYAR project the Yamal WebGIS visualises the lakes and lake catchments and the sampling effort of the recent years, being a useful platform to support joint international projects.

ACKNOWLEDGMENTS

This research was conducted within the framework of the Program of Fundamental Research Department of Earth Sciences No. 12 "The processes in the atmosphere and cryosphere as factors of environment changes", the RFBR grant 13-05-91001-ANF_a, Presidential grant for scientific schools

No. 5582.2012.5 and 3929.2014.5, Russian Science Foundation Grant 16-17-10203, as well as International projects CALM and TSP. Authors would like to thank the German Academic Exchange Service (DAAD), Otto Schmidt Laboratory for Polar and Marine Research (OSL) and Helmholtz graduate school for Polar and Marine Research (POLMAR) for financial support; Digital Globe Foundation and German Aerospace Center (DLR) for providing satellite images. Authors also would like to thank two anonymous reviewers and the editor of this issue for provided comments to this manuscript.

References

Bartsch, A., Trofaier, A.M., Hayman, G., Sabel, D., Schlaffer, S., Clark, D.B. & Blyth, E. (2012): Detection of open water dynamics with ENVISAT ASAR in support of land surface modelling at high latitudes.- Biogeosci. 9: 703-714. doi: 10.5194/bg-9-703-2012.

Biskaborn, B.K., Lanckman, J.-P., Lantuit, H., Elger, K., Streletskiy, D.A., Cable, W.L. & Romanovsky, V.E. (2015): The new database of the Global Terrestrial Network for Permafrost (GTN-P).- Earth Syst. Sci. Data 7: 245-259, doi: 10.5194/essd-7-245-2015.

Brown, J., Ferrians, Jr., O.J., Heginbottom, J.A. & Melnikov, E.S. (1998, revised February 2001): Circum-arctic map of permafrost and ground ice conditions.- Boulder, CO, National Snow and Ice Data Center, Digital media <http://nsidc.org/data/docs/fgdc/ggd318_map_circumarctic/>.

Brown, J., Hinkel, K.M. & Nelson, F.E. (2000): The circumpolar active layer monitoring (CALM) program: Research designs and initial results.- Polar Geogr. 24: 166-258, doi: 10.1080/10889370009377698.

Brown, J., Kholodov, A., Romanovsky, V., Yoshikawa, K., Smith, S.L., Christiansen, H.H., Vieira, G. & Noetzli, J. (2010): The thermal state of permafrost: the IPY-IPA snapshot (2007-2009).- Proc. 63rd Canadian Geotechnical Conference & 6th Canadian Permafrost Conference Geo 2010, Calgary, Canada, 12-16 Sept. 2010.

Djokic, D., Ye, Z. & Dartiguenave, C. (2011): Arc hydro tools overview.- Redland, Canada, ESRI, 1-189.

Dvornikov, Y., Bartsch, A., Khomutov, A., Heim, B., Widhalm, B., Fedorova, I., Leibman, M., Mikhaylova, M. & Skorospekhova, T. (2014): Process of organic transport in lakes of the Yamal region (POLYAR).- Proc. of Arctic Change Conference 2014, Ottawa, Canada, 6-12 December 2014.

Dvornikov, Y., Khomutov, A., Mullanurov, D., Ermokhina, K., Gubarkov, A. & Leibman, M. (2015): GIS and field data based modelling of snow water equivalent in shrub tundra.- Fennia 193: 53-65.

Dvornikov, Y., Heim, B., Roessler, S., Leibman, M., Khomutov, A. & Bartsch, A. (2016): Colored dissolved organic matter (CDOM) absorption measurements in the Vaskiny Dachu region, Central Yamal, Russia.- Inst. Earth Cryosphere Siberian Branch Russian Acad. Sci. Tuymen, doi: 10.1594/PANGAEA.860049

IPA (2003): International Permafrost Association Standing Committee on Data Information and Communication, Circumpolar Active-Layer Permafrost System (CAPS), Vers. 1. Boulder, Colorado USA: National Snow and Ice Data Center, doi: http://dx.doi.org/10.7265/N5SF2T3B

Heim, B., Klump, J., Oberhänsli, H. & Fagel, N. (2008): Assembly and concept of a web-based GIS within the paleolimnological project CONTINENT (Lake Baikal, Russia).- J. Paleolimnol. 39: 567-584, doi: 10.1007/s10933-007-9131-0.

Khomutov, A.V. & Khitun, O.V. (2014): The dynamics of vegetation cover and the depth of seasonal thawing in the typical tundra of Central Yamal under technogenic impact.- Tyumen State University Herald 4: 17-27 [In Russian].

Khomutov, A.V. & Leibman, M.O. (2014): Assessment of landslide hazards in a typical tundra of Central Yamal, Russia.- In: W. SHAN et al. (eds): Landslides in Cold Regions in the Context of Climate Change, Springer International Publishing, 271-290, doi: 10.1007/978-3-319-00867-7_20.

Khomutov, A.V., Leibman, M.O. & Andreeva, M.V. (2012): Mapping of ground ice in Central Yamal.- Tyumen State University Herald 7: 68-76 [In Russian].

Kriegler, F.J., Malila, W.A., Nalepka, R.F. & Richardson, W. (1969): Preprocessing transformations and their effects on multispectral recognition.- Proc. 6th Internat. Sympos. Remote Sensing of Environment 97-131.

Leibman, M.O. (1996): Results of chemical testing for various types of water and ice, Yamal Peninsula, Russia.- Permafrost Periglacial Processes. 7: 287-296, doi: http://dx.doi.org/10.1002%2f(SICI)1099-1530(199609)7%3a3%3c287%3a%3aAID-PPP224%3e3.0.CO%3b2-A.

Leibman M.O. (1998): Active layer depth measurements in marine saline clayey deposits of Yamal Peninsula, Russia: Procedure and interpretation of results.- In: A. LEWKOWICZ & M. ALLARD (eds): Permafrost. Proc. 7th internat. Confer. 635-639. Collection Nordicana. Centre d'Etudes Nordiques, Université Laval, Yellowknife.

Leibman, M.O. & Kizyakov, A.I. (2007): Cryogenic landslides of the Yamal and Yugorsky Peninsulas.- Earth Cryosphere Inst. SB RAS, Moscow, 1-206 [In Russian].

Leibman, M.O., Khomutov, A.V., Gubarkov, A.A., Mullanurov, D.R. & Dvornikov, Y.A. (2015): The research station rVaskiny Dachiiny Dac DacYamal, West Siberia, Russia ia ion rVaskiny Dachivor. & Dvor. & Dvorlas.- Fennia 193: 3-30, doi: 10.11143/45201.

Leibman M.O., Khomutov, A.V., Orekhov, P.T., Khitun, O.V., Epstein, H., Frost, G. & Walker, D.A. (2012): Gradient of seasonal thaw depth along the Yamal transect.- In: D. DROZDOV & V. ROMANOVSKY (eds): Proc. 10th Internat. Confer. Permafrost 2. Translations of Russian Contributions, 237-242. The Northern Publisher, Salekhard.

Melnikov, E.S. & Minkin, M.A. (1998): About the strategy of developing geoinformational systems (GIS) and data bases in geocryology.- Kriosfera Zemli 2(3): 70-76 [In Russian].

Minkin, M.A., Melnikov E.S. & Leibman, M.O. (2001): Russian national geocryological database and a strategy for its development.- In: R. PAERE & V. MELNIKOV (eds): Permafrost response on economic development, environmental security and natural resources, 591-599, Kluwer Academic Publishers, Netherlands.

Romanovskii, N.N. & Leibman, M.O. (1994): Russian Contribution to the Global Geocryological Database. Glaciological Data, Report, GD-28, WDC-A for Glaciology, Boulder, Colorado, 24-27.

Richter, R. (1996): A spatially adaptive fast atmospheric correction algorithm.- Internat. J. Remote Sensing 17: 1201-1214, doi: 10.1080/01431169608949077.

Slabovich, G.I. (ed) (1985): Instructions for hydrometeorological stations and posts.- Gidrometeoizdat, Leningrad 1-301 [In Russian].

Streletskiy, D.A., Shiklomanov, N.I. & Nelson, F.E. (2012): Spatial variability of permafrost active-layer thickness under contemporary and projected climate in Northern Alaska.- Polar Geogr. 35: 95-116, doi: 10.1080/1088937X.2012.680204.

Trofaier, A.M., Bartsch, A., Rees, G. & Leibman M. (2013): Assessment of spring floods and surface water extent over the Yamalo-Nenets Autonomous District.- Environ. Res. Lett. 8: 045026, doi: 10.1088/1748-9326/8/4/045026.

Zhang, Y. (2004): Highlight Article: Understanding Image Fusion.- Photogrammetric Engineering & Remote Sensing 70: 657-661.

Acronym	Description	Website
AWIMAPS WebGIS Service	Alfred Wegener Institute GIS maps portal	<http://maps.awi.de/awimaps/>
YAMAL WebGIS	Alfred Wegener Institute GIS maps portal for YAMAL region	<http://maps.awi.de/yamal/>
CALM	Circumpolar Active Layer Monitoring network	<http://www.gwu.edu/~calm/>
COLD Yamal	COmbining remote sensing and field studies for assessment of Landform Dynamics and permafrost state on Yamal	<http://cold.zgis.net/>
GTN-P	Global Terrestrial Network for Permafrost Database	<http://gtnpdatabase.org/>
LCLUC-Yamal	Land-cover and Land-use Changes on the Yamal Peninsula, Russia	<http://www.geobotany.uaf.edu/yamal/>
NEST	NEST ESA SAR toolbox	<https://earth.esa.int/web/nest/home>
PANGAEA	Data Publisher for Earth & Environmental Science	<http://www.pangaea.de/>
TSP	Thermal State of Permafrost	<http://ipa.arcticportal.org/activities/gtn-p/tsp/15-tsp.html>

Tab. 7: List of projects and associated websites.

Tab. 7: Liste von Projekten und zugehörigen Webseiten.