GeoERA RESOURCE GROUNDWATER

Resources of groundwater, harmonized at Cross-Border and Pan-European Scale

Deliverable 2.1

First draft of Data Management Plan

Authors and affiliation:

Klaus Hinsby (GEUS)
Hans Peter Broers (TNO)
Margarita Sanabria (IGME)

E-mail of lead author:
khi@geus.dk

Version: no. 1

This report is part of a project that has received funding by the European Union's Horizon 2020 research and innovation programme under grant agreement number 731166.

Deliverable Data		
Deliverable number	D2.1	
Dissemination level	Pubic	
Deliverable name	First draft of Data Management Plan	
Work package	WP2 Dissemination, communication and linkage to the GeoERA Information Platform	
Lead WP/Deliverable beneficiary	GEUS	
Deliverable status		
Submitted (Author(s))	30/01/2018	Klaus Hinsby
Verified (WP leader)	30/01/2018	Klaus Hinsby
Approved (Coordinator)	30/01/2018	Hans Peter Broers

SUMMARY

This report describes the first draft Data Management Plan (DMP) for the RESOURCE project. The DMP is meant to clarify what data will be produced in the RESOURCE project in order to facilitate the interaction between RESOURCE and the Geo Information Platform (GIP) project of GeoERA. In a later stage (M16) the prioritized information products of RESOURCE are identified and the corresponding data requirements are specified as input for the EGDI (The Information Platform for the GeoERA projects). This draft management plan helps to ensure the initial communication with the GIP Project.

LIST OF ABBREVIATIONS & ACRONYMS

DMP	Data Management Plan
EGDI	European Geological Data Infrastructure
FAIR	Findable, Accessible, Interoperable, Reusable
GSO	Geological Survey Organisation
JRC	The Joint Research Centre of the European Commission, ISPRA
RESOURCE	RESOURCEs of groundwater, harmonized at Cross-Border and Pan-European Scale

TABLE OF CONTENTS

1 EXECUTIVE SUMMARY

The present document is deliverable D2.1 “First draft of Data Management Plan” of the RESOURCE project: “RESOURCEs of groundwater, harmonized at Cross-Border and Pan-European Scale” developed for the specific research topic: Groundwater – GW3 Harmonization of groundwater resources information at cross-border to Pan-European scale.

The RESOURCE project aims at demonstrating the potentials of the harmonization of information about Europe’s groundwater resources through cross-border demonstrations projects, through harmonized characterization approaches for Karst and Chalk aquifers and through a first information product at Pan-European scale where available data is compiled and integrated to produce a map of the fresh groundwater resources of Europe. The set of deliverables of the RESOURCE project will provide good practices in providing harmonized data and information across borders for assessments of the 3D structure of aquifers, the water volumes available, and the water fluxes and water quality of the resource. Harmonization of such hydrogeological information is a prerequisite for any transboundary groundwater management. A range of regional and national stakeholders will be involved in the work in order to ensure both interaction with authorities that manage and protect groundwater resources and end-users. The RESOURCE project maximizes the dissemination of the results and provides stakeholders and end-users with easy-access tools through the cooperation with the GeoERA Information Platform Project, jointly prioritizing the information products that are most beneficial for society. The information products to be delivered will serve as a first prototype example of information to be accessible within a Geological Service for Europe.

The DMP is meant to clarify what data will be produced in the RESOURCE project in order to facilitate the interaction between RESOURCE and the GeoERA Information Platform project (GIP-P). In a later stage (M16) the prioritized information products of RESOURCE are identified and the corresponding data requirements are specified as input for the EGDI. This draft management plan helps to ensure the initial communication with the GIP-P.

2 INTRODUCTION

The overall aim of the RESOURCE project is to demonstrate the potentials of the harmonization of information about Europe's groundwater resources through cross-border demonstrations projects, through harmonized characterization approaches for Karst and Chalk aquifers and through a first information product at Pan-European scale where available data is compiled and integrated to produce a map of the fresh groundwater resources of Europe. The Pan-EU mapping approach and the cross-border demonstration projects under RESOURCE will lead to example products of a harmonized pan-European assessment of the 3D structure of aquifers, the volumes of water involved and its quality.

Main results from the RESOURCE project will be integrated in the European Geological Data Infrastructure (EGDI), through the GeoERA Information Platform project. The GeoERA project continues developing EGDI, as the most comprehensive information platform for dissemination of up-to-date, comprehensive, reliable and harmonised subsurface data for Europe within and beyond GeoERA. This DMP is meant to clarify what data will be produced in the RESOURCE project in order to facilitate the interaction between RESOURCE and the GeoInformation Platform (GIP) project of GeoERA. In a later stage (M16) the prioritized information products of RESOURCE are identified and the corresponding data requirements are specified as input for the GeoERA Information Platform Portal (GIP). This draft management plan helps to ensure the initial communication with the GIP Project.

The goal of the DMP entails the need for good documentation and implementation of standards, a common thesaurus, infrastructure, privacy settings and interoperability of data formats making it simple to share, use and upgrade data.

.

3 DATA MANAGEMENT PLAN

3.1 Data Summary

Within the RESOURCE project no new data will be collected by field sampling campaigns, rather existing data will be harmonized in 3 of the 4 work packages in which data are used. The main new data to be produced during the RESOURCE project is the Pan European groundwater Resources Map, which is a grid-based product. This grid- based pan-EU product is the main deliverable that will be shared through the EGDI. The TRANSFLUX and H3O-PLUS work packages will deal with data at a cross-border scale and the data are mainly described in georeferenced reports/deliverables that will be accessible through EGDI. However, the cross-border project databases will not become part of the EGDI data flow. These selected cross-border pilots cover different aspects of RESOURCE and can represent role models and state-of-the art examples for other European cross-border regions. Selected interpreted results from the projects in the form of illustrations, graphics or pdf-maps might be shared through EGDI showing important aspects that are applicable in a wider than cross-border context, and presumably links to the project databases will be made available through EGDI. All deliverables of these two work packages will be available through EGDI, preferably on a geo-referenced basis, as pdfs and as location maps as shapefile. This will allow to find these documents through his metadata in order to be displayed in a geographical viewer or via spatial selections, if this functionality will be developed under het GIP-P. The CHAKA work package is meant to make a methodological contribution to the assessment of karst and chalk areas. The deliverables (pdfs) will be made available through EGDI, and the location of the pilot sites (shapefiles) will be geo-referenced in EGDI to make the results easily available to the general public. The RESOURCE project will not deliver any raw data, such as borehole descriptions, groundwater measurements etc.

The RESOURCE project will generate 3 types of data:

1. Data on cross-border groundwater systems, that will be made available in reports and regional databases available to regional stakeholders (WP's H3O-+ and TRANSFLUX) and in formal GeoERA deliverable pdf reports accessible through EGDI and the RESOURCE (GeoERA) website.
2. Conceptual hydrogeological models and an integrated assessment framework for Karst & Chalk pilot areas (CHAKA work package) which will be made available in pdf reports and for which the GeoERA deliverable reports are accessible through EGDI http://geoera.eu/projects/resource/.
3. A pan-EU information product describing the volumes and depths of European fresh groundwater, which will be available and accessible through the EGDI platform for users that have been granted access.

This first draft of the DMP is mainly focusing on this third data type, as this is where the main interaction with the GIP project is.

The pan-EU map is a based on a 2.5 D approach at a 10 x 10 km grid scale. The basis of this grid is an ESRI shapefile that is supported with an attribute table that is populated in excel by the individual surveys (see section 4 for details), which will be converted to a PostgreSQL database by the GIP-P and distributed through EGDI. The Pan European groundwater resources map is expected to be developed in 2.5D, with the option to extract information on depth and volumes from the grids, and with the option to create cross-sections based on the grid. We anticipate that webgis functionality needed to provide the maps and cross-sections visualisation is developed by the GeoERA Information Platform Project supporting the provision of the results to the public. If considered necessary for this purpose, webservices such as WMS, WFS, WCS, Downloads services and metadata catalogue will be set up for

selected data to demonstrate possibilities of ensuring the findability, accessibility, interoperability and reusability of the compiled data according to the FAIR principles.

3.2 Fair data

The pan-EU map is based on the official INSPIRE shapefile grid over Europe (see Figure 1 in Section 4; https://www.eea.europa.eu/data-and-maps/data/eea-reference-grids-2). The coordinate system of this grid is ETRS89 / LAEA Europe (EPSG:3035) and we anticipate that these grids will be supported by the GIP-P as multiple project request that. The attribute table is a CSV file where each row indicates a grid cell. The cell-ids correspond with the existing 10x10 km INSPIRE shapefile grid over Europe. INSPIRE Codelists and vocabularies will be used on record level whenever possible, with additional codes where necessary.

The grid data of the pan-EU map, its metadata and the documentation generated along the project will be stored in the EGDI central database(s) which is intended to be sustained by the GSOs after the end of the GeoERA. This will ensure accessibility.

By default, all the data produced by RESOURCE that are stored and published via EGDI. i.e. the data related to the pan-EU map, are made available according to the FAIR principles. We anticipate that the grid will see abrupt changes at some borders as the 3D hydrogeology has seldomly been harmonized over borders within Europe. However, the fixed attributes and harmonized criteria developed under RESOURCE will guarantee that the subtraction of data from the country databases will proceed in an harmonized and reproducible way. Restrictions may apply to some data, however, and an embargo period will most probably be defined for some data to ensure that the involved partners can publish on the data before the data are made completely open to all stakeholders.

The pan-EU data produced by RESOURCE will follow common good scientific standards, practices and protocols, and it does not require special software to upload or access the data on EGDI. However, developed web services such as WFS, which are also produced and used for access to data requires some end user skills in order to be able to efficiently work with the data. Generally, the type of data that are involved (shapefiles, grids, csv files) only requires standard and widely used common GIS software to visualize the data.
We anticipate that the GIP project needs to develop custom-made software that enables the use of the grid data, with the option to visualize certain parameters as 2D map view or as cross-sections to be visualized online.

The provided pan-EU RESOURCE data are interoperable because the data from regional and National Surveys is stored in a common database according to common standards (see section 4 for details).
RESOURCE will use established European and international standards for the storage, exchange and dissemination of project data. INSPIRE (the European Directive on Infrastructure for Spatial Information) compliance will be used wherever possible. Where this is not possible, Commission for the Management and Application of Geoscience Information (CGI) standards will be used. In addition, RESOURCE and all other groundwater projects will build on and further develop the thesaurus and groundwater research classification of the European Inventory of Groundwater Research (KINDRA/EIGR).

The pan-EU data produced by RESOURCE is intended to be useable by third partied after GeoERA ends, and in principle be reusable forever or as long as the EGDI exist. The data is however not static and will

be improved and updated during and beyond GeoERA. Data may be improved by third parties and later updated in EGDI by GeoERA partners after they've quality assured by GeoERA partners.

3.3 Allocation of resources

One of the aims of RESOURCE consists in the integration of the produced pan-EU grid data into EGDI which is managed by the GIP-P. The data repository shall be stored after the project to meet the requirements of good scientific practice. A strategy for storage of the files after the project is being developed by the GIP-P and will be included in the DMP later. The costs for the maintenance of the central EGDI platform is under consideration (discussions with EGS).

3.4 Data security

Data security for the pan-EU grid data that will be stored in the central EGDI database is currently secured through the fact that it is currently operated by the French geological survey (BRGM) and included in their operational procedures.

3.5 Other issues

As the RESOURCE project is not producing or collecting basic data such as borehole descriptions or localized parameters, but instead focuses on the aggregation of data at a larger scale, we are not governed by national or institutional rules on data management. We do not foresee to use 3rd party data at this moment.

3.6 Updating the DMP

This is the draft data management plan for RESOURCE. It is anticipated that an update of the plan will be provided under deliverable D2.2 "Definition of prioritized information products for the GeoERA Information Platform as input for the GIP meeting on Groundwater" which is scheduled for M16 of the RESOURCE project.

4 DESCRIPTION OF THE TEMPLATE FOR THE PAN-EU GROUNDWATER MAP

4.1 Overview

Although EU member states generally have a comprehensive overview of the groundwater resources in their homeland, a coherent overview of all fresh groundwater over Europe is not available for policy development and evaluation. The main GIP product of the RESOURCE is the Pan-EU Groundwater Resources Map. The idea is to produce a first information product at pan-European scale where available data is compiled and integrated to produce a map of the fresh groundwater resources of Europe.

Here, we provide the first draft description of the template that is foreseen to create the map. Thee template will be used by all participating surveys for collection of required data on the average depths of aquifers and aquitards and the depth of the active fresh water groundwater system at a 10x10 km resolution.

The template is a CSV file where each row indicates a grid cell. The cell-id's correspond with an existing 10x10 km INSPIRE shapefile grid over Europe (see Figure 2 and https://www.eea.europa.eu/data-and-maps/data/eea-reference-grids-2). The coordinate system of this grid is ETRS89, with the Lambert Azimuthal Equal Area projection (LAEA) Each participant receives their own template (excel file, see figure 4) and grid (shapefile, see figure 3) which covers their country, which will be joined later to cover Europe. The grid for each participating survey is distributed along the participants. These county grids extend beyond the borders, each participant logically only fills in cells which are within their border. Cells that cover multiple countries are filled in by each nation in their own template, any large differences can be compared later.

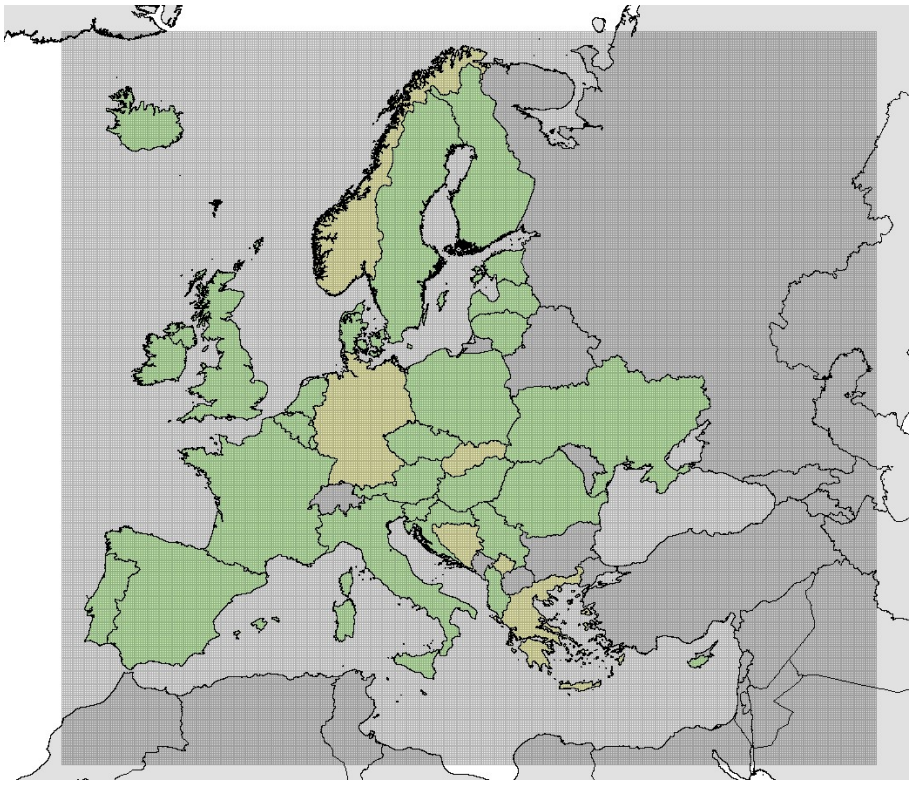

Figure 1 INSPIRE 10x10 km grid

4.2 Populating the grid with attributes

The collection of data follows the grids that were distributed. Populating the grids mean populating the attribute table that was delivered with the grid. The populating of the grids is now in its testing phase. Some important information

Grid type

- ETRS89; Official EU grid INSPIRE
- Projection: LAEA (Lambert Azimuthal Equal Area)
- Will be delivered to you with the template!
- Fixed ID's, delivered to the participants

Grid size

- Decision: scale = 10 x 10 km
- Countries that can easily deliver at higher resolution test 1 x 1, 2 x 2 and 5 x 5 km grids
- Countries that have difficulty at 10 x 10 km deliver 20 x 20 km, but on a grid that fits the 10 x 10 to make it fit with the rest
- All data is averaged for the grid cell
    - But option to indicate if aquifer extends over only a % of the grid = extension
        - % of grid cell occupied by the aquifer

Main features of the grids

- Top and bottom of the most important aquifers and aquitards

- Multiple layers
- Up to the lower boundary of the active fresh groundwater flow system
- Water level depth (stationary for agreed period, see later)

Criteria for the delineation of the active fresh groundwater flow system

Four options are offered:

- Hydrogeological boundary, depth of aquitard (H)
- Salinity boundary
  - Based on chloride (1000 mg/l chloride) (C)
  - Based on TDS (1500 mg/l TDS, still useful for irrigation) (T)
  - Based on EC (2500 µS/cm, EU drinking water standard) (E)

Water level depth

- Stationary for a fixed period (average over 2000-2010)
  - Minimal water level as measure for Karst
- But with label: water table dynamically changing (Y/N)
  - Reason indicated with symbol
    - S = almost stationary
    - R= recovery after mining
    - P = pumping increase
- Unsaturated zone depth = Altitude minus water level depth
  - Indicate lithology of unsaturated zone

Hydraulic Parameters:

- Porosity
- Conductivity, very rough k_h, K_v for each layer (m day^{-1})
- Confined/unconfined

Palaeowater

- Label of each layer based Expert judgement
  - Definition >10.000 year
  - Models that indicate these time scales
  - Identify grid cells where it occurs

Criterium for delineating an aquitard that is significant at EU scale, new proposal:

- Criterium based on vertical hydraulical resistance = thickness of the aquitard divided by the vertical hydraulic conductivity (in days)
  - C = Thickness / estimated K_v (in days)
- If C > 10000 days: take this aquitard into account as separate layer
  - Example clay layer with K_v = 0.001 m/day and thickness 10 m yields this 10.000 days

Thermal water

The presence of thermal aquifers can be indicated in the template. Similar to the fresh-salt interface, there are different thresholds for each participant (e.g. Hungary = 30 °C, Romania = 20 °C, UK 15 °C). For the thermal aquifers we suggest that each country uses their own definition and temperature, which must be indicated in the supporting information tab. If there are no thermal aquifers within the depth range that is considered for the total fresh water volume, this can be left empty.

4.3 The template for the attributes

Specifics of the attributes of the grid, given as separate worksheets in an excel spreadsheet. Worksheets include: Main Information, Layer 1, Layer 2 etc., Supplementary Info. The tabs are explained below.

4.3.1 Tab Main information

The first tab includes information that is of importance to the entire 10x10 km grid cell (see figure 4). The columns are explained below.

B. cell-id

Unique identifier for each grid cell. The id's are labelled as col_row with column numbers increasing from south to north and row numbers increasing from west to east. Please do not change anything in this column, these will be used to link the individual county grids to the Europe-wide grid.

C. Altitude_surface_level

This column contains the average altitude over the grid cell in meters relative to the EU height reference level EVRF2007 (European Vertical Reference Frame). All other depths in the template are relative to surface level.

D. GW_level

Average depth to the groundwater level (depth of unsaturated zone) in meters below the surface level. If the groundwater level is highly dynamic, please indicate in the next column which level is used (e.g. during the meeting in Vienna it was mentioned that some countries with karst systems use the minimum groundwater level instead of the average). As a period the average over 2000-2010 will be used.

E. Label_dynamic

Label which describes the dynamics of the groundwater level. There are several options which can be indicated by the following letters:

- S = Static;
- K = Karst systems with seasonal groundwater level fluctuation;
- R = Recovery after mining;
- P = Groundwater depletion by pumping;

F. GW_level_amplitude

If the groundwater level is dynamic, the amplitude can be given here. This is indicated by the difference between average lowest groundwater level and average highest groundwater level over a 10 year period

G. Unsat_lithology

Simplified lithology of the unsaturated zone

- Unconsolidated sand
- Unconsolidated clay
- Peat
- Fractured limestone
- … etc
- (list to be completed together)

H. Total_depth_active_layers

This indicates the maximum depth of the layers that are defined are of importance considering the fresh water volume. This could either be the top of layer with a very high vertical resistance (hydrogeological base) or the interface between fresh and saline groundwater if this falls above the hydrogeological base. The boundary that is used can be indicated in the next column.

I. Label_maximum_depth_active_layers

In this column you can indicate what is used as a label for the maximum depth of the active layers. This can be indicated by the following letters:

- H = hydrogeological boundary (e.g. depth of aquitard)
- C = based on chloride concentration (1000 mg/l Cl)
- T = Based on TDS concentration (1500 mg/l TDS)
- E = Based on EC (2500 µS/cm, EU drinking water standard)

4.3.2 Tabs Layer 1 – Layer 10

These tabs contain the information for each defined layer. The columns are the same for each layer. In this template there is space for 10 layers. The columns are explained below.

C. L1_top

The height of the top of the layer in meters below surface level

D. L1_bottom

The height of the bottom of the layer in meters below surface level.

E & F. L1_aquifer & L1_aquitard

Labels indicating if the layer is an aquifer or aquitard, indicated by Y (yes) or N (no).

G. L1_lithology

Description of the simplified lithology of the layer:

- Sand
- Clay
- Peat
- Limestone
- Etc… (to be completed together)

H. L1_extent
Percentage of the cell that is covered by the layer. For instance, this can be used to indicate the presence of eskers and tunnel valleys.

I. L1_Confidence_label_delineation

Label which describes the confidence level of which the above information about layer depths and extent is based on.

- EJ = Expert Judgement
- MOD = Based on subsurface model
- BH = Based on boreholes
- (etc. More labels needed?)

J. L1_Porosity

Porosity indicated with two decimals

K. L1_kh

Horizontal conductivity in m/d

L. L1_KV

Vertical conductivity in m/d

M. L1_Confidence label_hydraulic_parameters
Label which indicates the confidence level of which the hydraulic parameters are based on.

- EJ = Expert Judgement
- MOD = Based on subsurface model
- BH = Based on boreholes
- (etc.. More labels needed?)

N. L1_Paleo

Label indicating if the groundwater in the layer is Paleogenic (older than 10.000 years) or not (Y or N).

O. L1_Artesian

Label indicating if the aquifer is confined/unconfined/Confined artesian (C, U, CA).

P. L1_Thermal

Label indicating if the aquifer is (natural) thermal or not (y or n). Thermal aquifers are defined with different temperatures across Europe, we suggest that each country uses their own thresholds (as with the fresh/salt interface). Please note in supporting information which boundary is used.

4.3.3 Tab Supplementary information

This worksheet can be used to add extra information that is not yet part of the template,but is considered important for good use of the data (free format per grid cell).