



GEOProcessing 2021

The Thirteenth International Conference on Advanced Geographic Information
Systems, Applications, and Services

ISBN: 978-1-61208-871-6

July 18 – 22, 2021

Nice, France

GEOProcessing 2021 Editors

Claus-Peter Rückemann, Westfälische Wilhelms-Universität Münster (WWU) /
DIMF / Leibniz Universität Hannover, Germany

Yerach Doytsher, Technion - Israel Institute of Technology, Haifa, Israel

GEOProcessing 2021

Forward

The Thirteenth International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2021), held in Nice, France, July 18 - 22, 2021, addressed the aspects of managing geographical information and web services.

The goal of the GEOProcessing 2021 conference was to bring together researchers from the academia and practitioners from the industry in order to address fundamentals of advances in geographic information systems and the new applications related to them using the Web Services. Such systems can be used for assessment, modeling and prognosis of emergencies

GEOProcessing 2021 provided a forum where researchers were able to present recent research results and new research problems and directions related to them. The topics covered aspects from fundamentals to more specialized topics such as 2D & 3D information visualization, web services and geospatial systems, geoinformation processing, and spatial data infrastructure.

We take this opportunity to thank all the members of the GEOProcessing 2021 Technical Program Committee as well as the numerous reviewers. The creation of such a broad and high-quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to the GEOProcessing 2021. We truly believe that, thanks to all these efforts, the final conference program consists of top quality contributions.

This event could also not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the GEOProcessing 2021 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that GEOProcessing 2021 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in geographic information research.

GEOProcessing 2021 Chairs

GEOProcessing 2021 General Chair

Claus-Peter Rückemann

Westfälische Wilhelms-Universität Münster (WWU) / DIMF / Leibniz Universität Hannover
Germany

GEOProcessing Steering Committee

Thomas Ritz, FH Aachen, Germany

Yerach Doytsher, Technion - Israel Institute of Technology, Haifa, Israel

Francisco Javier Nogueras-Iso, Universidad de Zaragoza, Spain

Alexey Cheptsov, High Performance Computing Center Stuttgart, Germany

Roger Tilley, Sandia National Laboratories, USA

Jui-Hsin (Larry) Lai, Ping-An Technology - Research Lab, USA

GEOProcessing 2021 Advisory Committee

Monica De Martino, Consiglio Nazionale delle Ricerche - Genova, Italy

Jianhong Cecilia Xia, Curtin University, Australia

GEOProcessing 2021 Publicity Chair

Alvaro Liebana, Universitat Politecnica de Valencia, Spain

Mar Parra, Universitat Politecnica de Valencia, Spain

GEOProcessing 2021

COMMITTEE

GEOProcessing 2021 General Chair

Claus-Peter Rückemann, Westfälische Wilhelms-Universität Münster (WWU) / DIMF / Leibniz Universität Hannover

GEOProcessing Steering Committee

Thomas Ritz, FH Aachen, Germany
Yerach Doytsher, Technion - Israel Institute of Technology, Haifa, Israel
Francisco Javier Nogueras-Iso, Universidad de Zaragoza, Spain
Alexey Cheptsov, High Performance Computing Center Stuttgart, Germany
Roger Tilley, Sandia National Laboratories, USA
Jui-Hsin (Larry) Lai, Ping-An Technology - Research Lab, USA

GEOProcessing 2021 Advisory Committee

Monica De Martino, Consiglio Nazionale delle Ricerche - Genova, Italy
Jianhong Cecilia Xia, Curtin University, Australia

GEOProcessing 2021 Publicity Chair

Alvaro Liebana, Universitat Politecnica de Valencia, Spain
Mar Parra, Universitat Politecnica de Valencia, Spain

GEOProcessing 2021 Technical Program Committee

Alia I. Abdelmoty, Cardiff University, Wales, UK
Danial Aghajarian, Georgia State University, USA
Nuhcan Akçit, Middle East Technical University, Turkey
Zaher Al Aghbari, University of Sharjah, UAE
Md Mahbub Alam, Dalhousie University, Canada
Heba Aly, University of Maryland, College Park, USA
Francisco Javier Ariza López, Escuela Politécnica de Jaén - Universidad de Jaén, Spain
Thierry Badard, Centre de Recherche en Géomatique (CRG) | Université Laval, Canada
Abderrazak Bannari, Arabian Gulf University, Bahrain
Fabian Barbato, Ort University of Uruguay, Uruguay
Melih Basaraner, Yildiz Technical University, Turkey
Peter Baumann, rasdaman GmbH Bremen / Jacobs University Bremen, Germany
Lorenzo Carnevale, University of Messina, Italy
Mete Celik, Erciyes University, Turkey
Dickson K.W. Chiu, The University of Hong Kong, Hong Kong
Alexey Cheptsov, High Performance Computing Center Stuttgart, Germany
Alexandre Corrêa da Silva, HEX Geospatial Technologies, Brazil
Mehmet Ali Çullu, Harran University, Turkey

Monica De Martino, CNR-IMATI (National research Council, Institute of applied Mathematics and Information technology), Italy
Cláudio de Souza Baptista, University of Campina Grande, Brazil
Subhadip Dey, Indian Institute of Technology Bombay, India
Yerach Doytsher, Technion - Israel Institute of Technology, Haifa, Israel
Suzana Dragicevic, Simon Fraser University, Canada
Emre Eftelioglu, Cargill Inc., USA
Süleyman Eken, Kocaeli University, Turkey
Salah Er-Raki, Université Cadi Ayyad, Morocco
Javier Estornell, Universitat Politècnica de València, Spain
Jamal Ezzahar, Université Cadi Ayyad, Morocco
Francisco R. Feito, University of Jaén, Spain
Anabella Ferral, Instituto de Altos Estudios Espaciales Mario Gulich | Centro Espacial Teófilo Tabanera - CONAE, Córdoba, Argentina
Douglas Galarus, Utah State University, USA
Erica Goto, University of California Santa Barbara (UCSB), USA
Mohd Helmy Abd Wahab, Universiti Tun Hussein Onm Malaysia, Malaysia
Arif Hidayat, Monash University, Australia / Brawijaya University, Indonesia
Masaharu Hirota, Okayama University of Science, Japan
Qunying Huang, University of Wisconsin, Madison, USA
Chih-Cheng Hung, Kennesaw State University - Marietta Campus, USA
Sergio Ilarri, University of Zaragoza, Spain
Ge-Peng Ji, Wuhan University (WHU) & Inception Institute of Artificial Intelligence (IIAI), China
Katerina Kabassi, Ionian University, Greece
Hassan A. Karimi, University of Pittsburgh, USA
Baris M. Kazar, Oracle America Inc., USA
Saïd Khabba, Université Cadi Ayyad, Marrakech, Morocco
Kyoung-Sook Kim, National Institute of Advanced Industrial Science and Technology (AIST), Tokyo, Japan
Mel Krokos, University of Portsmouth, UK
Piyush Kumar, Florida State University, USA
Jui-Hsin (Larry) Lai, Ping-An Technology - Research Lab, USA
Robert Laurini, INSA Lyon | University of Lyon, France
Dan Lee, Esri Inc., USA
Lassi Lehto, Finnish Geospatial Research Institute, Finland
Jugurta Lisboa-Filho, Federal University of Viçosa, Brazil
Ying Lu, DiDi Research America, Mountain View, USA
Dandan Ma, Northwestern Polytechnic University, China
Ahmed Mahmood, Google, USA
Ali Mansourian, Lund University, Sweden
Jesús Martí Gavilá, Research Institute for Integrated Management of Coastal Areas (IGIC) | Universitat Politècnica de València, Spain
Sara Migliorini, University of Verona, Italy
Sobhan Moosavi, The Ohio State University, USA
Tathagata Mukherjee, The University of Alabama in Huntsville, USA
Beniamino Murgante, University of Basilicata, Italy
Ahmed Mustafa, The New School University, New York, USA
Aldo Napoli, MINES ParisTech - CRC, France
Maurizio Napolitano, Fondazione Bruno Kessler, Trento, Italy

Javier Nogueras-Iso, University of Zaragoza, Spain
Alexey Noskov, Philipps University of Marburg, Germany
Xiao Pan, Shijiazhuang Tiedao University, China
Shray Pathak, School of Geographic Sciences | East China Normal University, Shanghai, China
Kostas Patroumpas, Athena Research Center, Greece
Davod Poreh, Università degli Studi di Napoli "Federico II", Italy
Satish Puri, Marquette University, Wisconsin, USA
Kuldeep Purohit, Michigan State University, USA
Thomas Ritz, FH Aachen, Germany
Ricardo Rodrigues Ciferri, Federal University of São Carlos (UFSCar), Brazil
Armanda Rodrigues, Universidade NOVA de Lisboa | NOVA LINCS, Portugal
Claus-Peter Rückemann, Westfälische Wilhelms-Universität Münster (WWU) / DIMF / Leibniz Universität Hannover, Germany, Germany
Ibrahim Sabek, University of Minnesota, USA
André Sabino, Universidade Autónoma de Lisboa, Portugal
Markus Schneider, University of Florida, USA
Raja Sengupta, McGill University, Montreal, Canada
Shih-Lung Shaw, University of Tennessee, Knoxville, USA
Yosio E. Shimabukuro, Brazilian Institute for Space Research - INPE, Brazil
Spiros Skiadopoulos, University of the Peloponnese, Greece
Dimitris Skoutas, Athena Research Center, Greece
Francesco Soldovieri, Istituto per il Rilevamento Elettromagnetico dell'Ambiente - Consiglio Nazionale delle Ricerche (CNR), Italy
Katia Stankov, University of British Columbia, Canada
Payam Tabrizian, IDEO, San Francisco, USA
Ergin Tari, Istanbul Technical University, Turkey
Brittany Terese Fasy, Montana State University, USA
Roger Tilley, Sandia National Laboratories, USA
Goce Trajcevski, Iowa State University, USA
Linh Truong-Hong, Delft University of Technology, Netherlands
Taketoshi Ushiyama, Kyushu University, Japan
Marlène Villanova-Oliver, Univ. Grenoble Alpes - Grenoble Informatics Lab, France
Massimo Villari, University of Messina, Italy
Tin Vu, University of California, Riverside, USA
Hong Wei, University of Maryland, College Park, USA
John P. Wilson, University of Southern California, USA
Jianhong Cecilia Xia, Curtin University, Australia
Ningchuan Xiao, The Ohio State University, USA
Xiaojun Yang, Florida State University, USA
Zhaoming Yin, Google LLC, USA
Qiangqiang Yuan, School of Geodesy and Geomatics | Wuhan University, China
F. Javier Zarazaga-Soria, University of Zaragoza, Spain
Shenglin Zhao, Tencent, Shenzhen, China
Qiang Zhu, University of Michigan - Dearborn, USA

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission to reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Cross-border and Cross-domain Integration of Content in a European Geospatially Enabled Ecosystem <i>Lassi Lehto, Jaakko Kahkonen, Jari Reini, Timo Aarnio, and Roope Tervo</i>	1
Deep Reinforcement Learning for Spatial Motion Planning in 3D Environments <i>Oren Gal and Yerach Doytsher</i>	6
Image Classification Methods Assessment for Identification of Small-Scale Agriculture in Brazilian Amazon <i>Flavia Domingos Pacheco, Maira Ramalho Matias, Gabriel Maximo Silva, Anielli Rosane Souza, Yosio Edemir Shimabukuro, and Maria Isabel Sobral Escada</i>	12
Parametrising a Model of Clinopyroxene/Melt Partition Coefficients for Sodium to Higher Upper Mantle Pressures <i>Julia Marleen Schmidt and Lena Noack</i>	20
Methodology of Chorological and Coherent Conceptual Knowledge Contextualisation: Approaches for Multi-disciplinary Contexts in Prehistory and Archaeology <i>Claus-Peter Ruckemann</i>	28

Cross-border and Cross-domain Integration of Content in a European Geospatially Enabled Ecosystem

Lassi Lehto, Jaakko Kähkönen
 Department of Geoinformatics
 Finnish Geospatial Research Institute
 Masala, Finland
 e-mail: lassi.lehto@nls.fi, jaakko.kahkonen@nls.fi

Jari Reini^a, Timo Aarnio^a, Roope Tervo^b
 National Land Survey of Finland^a
 Finnish Meteorological Institute^b
 Helsinki, Finland
 e-mail: jari.reini@nls.fi, timo.aarnio@nls.fi,
 roope.tervo@fmi.fi

Abstract—The Geospatially Enabled Ecosystem for Europe (GeoE3) project has commenced with the main goal to develop use case-oriented, cross-border and cross-domain geospatial services conforming to the second-generation interface standards. The five-nation project focusses on cross-domain integration of content by applying dynamic, service level methods for joining meteorological and statistical data with geospatial features. Various technologies will be tested for providing seamless, cross-border access to geospatial resources in context of the new Open Geospatial Consortium’s family of service interface standards. 3D aspects of geospatial data will also be considered, both from content encoding and service interface point of view. The project aims at supporting renewable energy- and smart city-related applications.

Keywords—*geospatial ecosystem; cross-border; cross-domain; OGC API; solar energy; smart city.*

I. INTRODUCTION

In an interoperable computing environment, every participating actor should gain benefits. This is the main guiding principle in the spatial data ecosystems currently being developed. A major action, Geospatially Enabled Ecosystem for Europe (GeoE3), has commenced to establish an ecosystem of use case-oriented geospatial services in five European countries [1]. The action is part of the Connecting Europe Facility (CEF) programme advancing European transport, energy and digital infrastructure development [2]. The project runs for three years and involves five national mapping and cadastral agencies, together with a few other governmental organizations and private companies. The main goals of the project include adaptation of modern geospatial APIs (Application Programming Interfaces) to establish cross-border services for renewable energy and urbanization related use cases, use of dynamic service-level mechanisms for cross-domain content integration, and promotion of the developed approaches through extensive online innovation and education programme.

GeoE3 consortium membership includes the national mapping or cadastral agency from Finland, Norway, Estonia, The Netherlands and Spain. Furthermore, the consortium includes the national meteorological and statistical agency of Finland: Finnish Meteorological Institute and Statistics Finland, respectively. Private companies include Spatineo

from Finland and Aventi Intelligent Communication from Norway. Open Geospatial Consortium (OGC) Europe participates in the project as a representative of the standardization community. National Land Survey of Finland coordinates the project that will be finalized by Sep 2023 [3].

The main use cases identified by GeoE3 include renewable energy applications, specifically in the context of buildings construction and use. In particular, this involves solar energy potential, based on solar panels on rooftops. Wind energy and heating/cooling facilities of the buildings are considered too. The other use cases deal with traffic applications, in particular efficient use of electric cars, and with various urbanization challenges. The last one is considering urban expansion efficiency employing the United Nation’s Sustainable Development Goal (SDG) indicator 11.3.1 ‘Ratio of land consumption rate to population growth rate’ [4].

The paper is organized as follows. In Section II, the envisioned service architecture of the GeoE3 platform is described. Section III discusses various mechanisms for content integration. In Section IV, some novel ideas concerning OAPIF functionality are discussed. Section V presents the initial considerations of the project on dealing with 3D geodata. The paper ends with conclusions in Section VI.

II. SERVICE ARCHITECTURE

The GeoE3 project aims at establishing a set of use case - oriented services that enable content integration both across national borders and domain boundaries. The idea is to set up an integration layer on top of national services. The integration layer would provide modern service interfaces to the client side and accommodate various categories of source services on country level. The GeoE3 services are to be based on the OGC API family of second-generation, internationally standardized service interfaces [5]. These include services like OGC API Features [6], OGC API Coverages and OGC API Records. Internally the GeoE3 platform will apply at least OGC API Processes [7] for introducing analysis-based content layers to the clients and Table Joining Service (TJS) [8] for content integration across domain boundaries. OGC API Environment Data Retrieval (EDR) [9] interface will be used for requesting meteorological data, particularly for supporting transport-related applications.

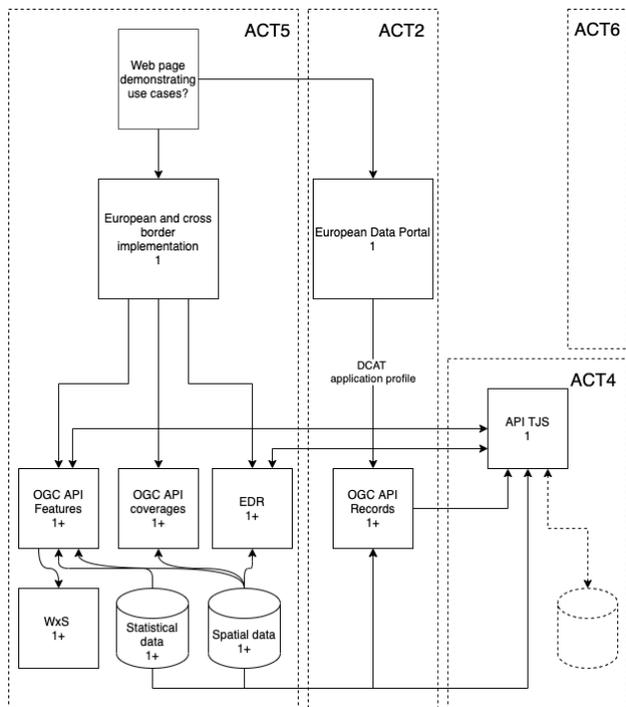


Figure 1. The preliminary GeoE3 service architecture. Dashed boxes indicate division of responsibility among GeoE3 subgroups (Activities).

The illustration in Figure 1 shows the envisaged GeoE3 service architecture. The WxS box in the left bottom corner represents all the country level data resources that will mostly be accessed through service access points conforming to the OGC standard interfaces, like WFS, WCS and the OGC API family of services. The databases at the bottom emphasize the GeoE3 goal of integrating spatial data with statistical and other tabular data sets. All the data resources will be made available via modern interface standards and be integrated in a cross-border manner by the GeoE3 integration platform. Metadata records of all GeoE3 services will be stored into the European Data Portal (EDP) and also be available from the OGC API Records service interface maintained on the GeoE3 platform. An implementation of the TJS standard is an essential component of the GeoE3 service architecture. It will be responsible for joining tabular data sets with geospatial features. The resulting combined resource is then made accessible to client applications via OGC API Features services.

III. CONTENT INTEGRATION

A. Content Integration with OAPIF

An essential component on the GeoE3 integration layer is OGC API Features (OAPIF) service interface. The successor for the OGC’s Web Feature Service (WFS) interface specification, OAPIF follows the principles of the so-called REST (Representational State Transfer) service architecture.

The output of the service consists of individual geospatial features, organized as a GeoJSON -encoded feature collection [10], geometries expressed in WGS84 Coordinate Reference System (CRS). The OAPIF service instance offers a set of data collections, accessible from the given service address paths. This provides a natural setting for organizing individual countries’ data sets as collections inside a single OAPIF instance. For instance, the buildings data sets of Finland, Norway and Spain could be organized as follows.

- /geo3/buildings/collections/buildings_FI/items
- /geo3/buildings/collections/buildings_NO/items
- /geo3/buildings/collections/buildings_ES/items

The OAPIF implementation adopted in the GeoE3 platform is the OAPIF reference implementation, called pygeoapi [11], a Python library supporting OAPIF and partially supporting OGC API Coverages [12], OGC API Tiles and OGC API Processes. pygeoapi provides a flexible plugin architecture for content provision and formatting. The off-the-shelf library supports, via OpenGIS Simple Features Reference Implementation (OGR) [13] and rasterio [14] provider plugins, a vast array of vector and raster source formats and services. For example, WFS is supported as source service. OGR provider also supports OAPIF service access, thus enabling cascading OAPIF service approaches. Specific project-related needs can be accommodated by developing a tailored plugin. As the plugin architecture is based on Python language, all the Python-based geospatial libraries can be easily utilized. The set of already available preliminary GeoE3 components include plugins for WFS versions 1.1.0 and 2.0.0 with basic schema transformation capabilities, and a plugin for WCS version 1.0.0.

B. Content Integration with OGC API Processes

OGC API Processes is a new resource-based process definition standard. The standard defines how the client program can start the process and how the process inputs are provided, and where the results of the process are stored.

Computational tasks are wrapped inside the process and the results are provided to the client application. Processes can be for example geometry buffering or routing from point A to point B. The process can be either synchronous, giving the result immediately, or asynchronous, in which case the service informs the client, when the results are ready and where they can be found.

The GeoE3 project develops a process, in which the footprint of a building is taken as a starting point and the 3D geometry of the building is analysed using the Digital Surface Model (DSM). In addition, a process is being developed to calculate the average amount of solar energy in a building area. This dynamically created information can then be integrated back to the original features.

The OGC API Processes interface standard follows the same principles as the other OGC API standards (e.g., OAPIF). The path structure of the GeoE3 OGC API Processes services might be as follows.

- The service landing page (/geoe3)
- Conformance classes (/geoe3/conformance) describing what service can do
- The list of processes in the service (/geoe3/processes)
- Process to create 3D buildings from the building footprint (/geoe3/processes/3d-building-creator)
- Solar energy potential calculator (/geoe3/processes/solarenergy-potential)

C. Content Integration with Table Joining Service

OGC Table Joining Service (TJS) provides a method for joining spatial and tabular data. The standard is currently being revised by the OGC and the new draft standard is going to be used in GeoE3 [15]. It allows integration of statistical (tabular) data to the platform.

To carry out the join, the service needs the spatial and tabular data and a common identifier between them. Several columns of the tabular data can be joined at once. Both the spatial and tabular data can either be configured to be available from the server or they can be uploaded on demand. Spatial data can thus be delivered via OGC Features APIs and tabular data, e.g., via PX Web API.

The goal of the project is to evaluate the fitness for purpose of TJS in the use cases identified in the project. A TJS instance will be set up and configured, along with a browser-based end user interface. While TJS can be demonstrated and used with a user interface, it can be more valuable when used in the background as a part of an automated process. The obvious benefit of joining data is that spatial data need not be stored in several places.

The output of a TJS has several options. The combined data can be downloaded for further use in several spatial formats, such as GeoJSON. Furthermore, results from the joining operation can be calculated from the data. This includes information about successful joins, unsuccessful joins and extra data rows. In the project this derived data will be visualized in the user interface as a report for the user.

IV. OAPIF RETHOUGHT

Some of the OAPIF concepts have been rethought in the development of the preliminary GeoE3 platform. For instance, the HTML-formatted feature browsing in the OAPIF development is understood as a tabular presentation of feature properties, browsed through the feature collection page by page. When there are potentially millions of features in the national databases, and when the order of features presented is arbitrary, the browsing process becomes unfeasible. The feature browsing in the GeoE3 OAPIF is returned back to the traditional spatially organized, map-based browsing. A plain background map is first shown to the user. The individual vector features are requested and presented only when the user zooms in deep enough.

Another new concept introduced in the GeoE3 project is the idea of treating the HTML-formatted visualization of an individual spatial feature as an application-specific dashboard

of the feature [16]. Into this dashboard, a wide set of feature properties and visualizations can be collected by the GeoE3 integration platform. For example, a sun energy-related building dashboard could contain all the relevant building attributes - like the area, volume and heating information, and the DSM from the building's area - useful for solar energy analyses. Furthermore, climate-related information could be integrated to the same dashboard on the GeoE3 platform by accessing appropriate values from a meteorological service. The dashboard could be seen as a shop window for the available content, and could later be downloaded into a digital analysis process as the corresponding JSON-formatted data set. An example of a prototypical building dashboard for renewable energy related applications is shown in Figure 2. OAPIF browsing by map view is shown on the left. On the right, a detail window of the selected feature (The Finnish Parliament) is shown, together with the DSM rendered inside the building footprint, and average speed of wind in the area shown below.

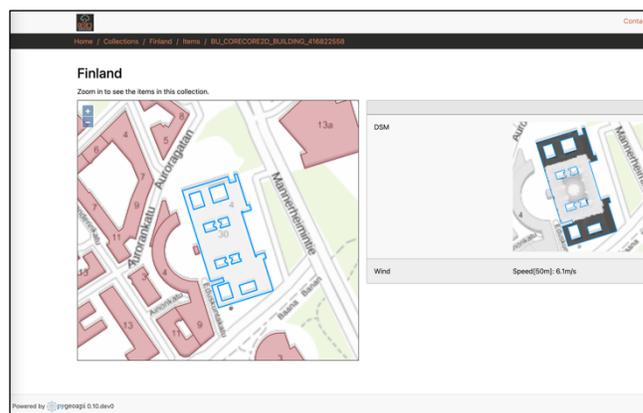


Figure 2. Preliminary OAPIF html output as a building's renewable energy dashboard from the GeoE3 platform.

V. 3D CONSIDERATIONS

The GeoE3 project introduces the idea of HTML-formatted application-specific dashboard, especially in the context of building construction and use with renewable energy applications, such as solar energy potential. While computing the energy potential, many data sources and service interfaces can be utilized, depending on the computing tasks in hand. Using gridded Digital Elevation Models (DEMs), DSMs and building footprints, one can automatically compute a 2D raster layer that indicates rooftops suitable for solar energy production. This raster layer can then be portrayed together with the DSM and the orthophoto for 3D visual exploration.

In the exemplary visualization shown in Figure 3, locations suitable for solar panels are presented with yellow colour. The visualization is a result of a parameterized on-the-fly process, in which the DMS slope and aspect values are used, together with the building footprints and the height difference between the DSM and the DEM, to compute the areas well exposed to sun energy. Figure 3 shows a three.js-based [17] Web application that requests the analysis result

layer and the DSM from OGC-compatible service interfaces and drapes the raster on top of the DSM. Corresponding visualization could also be integrated as part of the GeoE3's HTML-based feature dashboard.



Figure 3. Visual exploration of potential solar panel locations on rooftops. Analysis is based on the use of DSM slope and aspect values, height differences between the DSM and the DEM, and building footprints.

One of the goals of the GeoE3 project is to make use of 3D vector data, provided that project partners are able to produce 3D data during the project's lifetime. For instance, the project coordinator, NLS Finland, aims at making 3D building models available by the end of 2021. The production process is based on the use of the building footprints and DSM, and aims at constructing LoD2 category building geometries. A widely accepted, production-level service interface specification for 3D models does not exist yet. A possible workaround for this situation is to store the 3D geometries into a 3DCityDB database and export them as a CityGML, CityJSON or gltf -formatted file on demand.

At this moment, CityJSON format is the most interesting proposal for the 3D vector representation. CityJSON is a JSON-based subset of the OGC's CityGML [18] data model and is designed for storage and transfer of 3D city models. CityJSON is not yet standardized by the OGC, but is currently being considered as an official OGC Community Standard. Current version of the CityJSON standard is available at the CityJSON development site [19].

Freely available browser-based applications for visualization of CityGML does not yet exist. However, JSON processing is widely supported in browsers, and CityJSON can naturally benefit from this too [20]. There are development initiatives for providing an easy-to-use CityJSON-supporting software component for Web applications. On the left side in Figure 4, an example is presented of browser-based visualization displaying NLS Finland buildings in CityJSON format utilizing a software component called Ninja [21]. On the right side in Figure 4, the same features are presented using the 3D map view of the QGIS application [22]. GeoE3 is also interested in utilizing 3D Tiles, an OGC Community Standard for visualizing 3D geospatial content in desktop, Web and mobile applications [23].

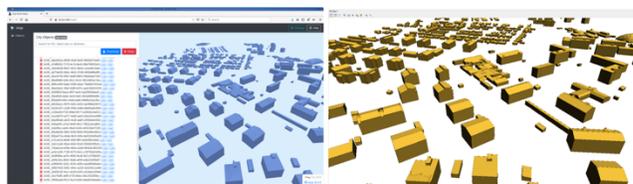


Figure 4. Visualization of 3D vector buildings in browser using Ninja (on the left) and in the 3D map view of the QGIS application (on the right).

VI. CONCLUSIONS AND FUTURE WORK

GeoE3 project has commenced with the main goal of developing a geospatially enabled ecosystem of use case-oriented services conforming to modern, second-generation interface standards. The project will tackle the challenge of cross-border and cross-domain integration of content employing a flexible, state-of-the-art development platform, based on OGC API Python library called pygeoapi. The plugin architecture of pygeoapi facilitates development of tailored components for connecting the platform to various different country level source services.

The overall service architecture of the GeoE3 project can be seen as a three-tier architecture with country level services forming the bottom data layer. The GeoE3 integration platform with advanced capabilities for dynamic content integration and use case-specific adaptation of the resulting data sets works as the middle processing layer. Finally, end user applications form the third architecture layer.

An application-specific output from an OAPIF service can be provided in the form of an html-formatted dashboard of the feature in focus. Based on the user needs, an appropriate set of information items can be collected and put together by the GeoE3 platform, and the resulting data set be formatted in html as a shop window for potential users. Finally, the same data set could be requested as a JSON-encoded package, to be fed into an external analysis process.

The project will continue developing the described functionalities over the coming two and half years and aims at wide adoption of the services by actors working in the identified use case domains.

ACKNOWLEDGMENT

The GeoE3 project is co-financed by the Connecting Europe Facility (CEF) of the European Union with the grant agreement number INEA/CEF/ICT/A2019/2063390.

REFERENCES

- [1] GeoE3, Geospatially Enabled Ecosystem for Europe, GeoE3 Home Page. [Online]. Available from: <https://geoe3.eu>, 2021, [retrieved: June, 2021]
- [2] CEF, Connecting Europe Facility, Home Page. [Online]. Available from: <https://ec.europa.eu/inea/en/connecting-europe-facility>, 2021, [retrieved: June, 2021]
- [3] NLS, National Land Survey of Finland Home Page. [Online]. Available from: <https://www.maanmittauslaitos.fi/en>, 2021, [retrieved: June, 2021]
- [4] United Nations, SDG Indicators, Goal 11. [Online]. Available from:

- <https://unstats.un.org/sdgs/metadata/?Text=&Goal=&Target=11.3>, 2021, [retrieved: June, 2021]
- [5] OGC, OGC API Home Page. [Online]. Available from: <https://ogcapi.org>, 2021, [retrieved: June, 2021]
- [6] OGC, OGC API – Features Home Page. [Online]. Available from: <https://ogcapi.org/features/>, 2021, [retrieved: June, 2021]
- [7] OGC, OGC API – Processes Home Page. [Online]. Available from: <https://ogcapi.org/processes/>, 2021, [retrieved: June, 2021]
- [8] OGC, Table Joining Service (JTS) Home Page. [Online]. Available from: <https://www.ogc.org/standards/tjs>, 2010, [retrieved: June, 2021]
- [9] OGC, OGC API – Environmental Data Retrieval (EDR) Home Page. [Online]. Available from <https://ogcapi.org/edr/>, 2021, [retrieved: June, 2021]
- [10] GeoJSON, GeoJSON Home Page. [Online]. Available from: <https://geojson.org>, 2021, [retrieved: June, 2021]
- [11] T. Kralidis, pygeoapi Home Page. [Online]. Available from: <https://pygeoapi.io>, 2021, [retrieved: June, 2021]
- [12] OGC, OGC API – Coverages Home Page. [Online]. Available from: <https://ogcapi.org/coverages/>, 2021, [retrieved: June, 2021]
- [13] GDAL, GDAL/OGR Home Page. [Online]. Available from: <https://gdal.org>, 2021, [retrieved: June, 2021]
- [14] Mapbox, rasterio Home Page. [Online]. Available from: <https://rasterio.readthedocs.io/en/latest/>, 2021, [retrieved: June, 2021]
- [15] OGC, OGC API Table Joining Service, Draft. [Online]. Available from: <https://github.com/opengeospatial/tjs/>, 2019, [retrieved: June, 2021]
- [16] C. Jing, M. Du, S. Li, and S. Liu, Geospatial Dashboards for Monitoring Smart City Performance, *Sustainability*, 2019, vol. 11, no. 20, pp. 1-23, [Online]. Available from: <https://www.mdpi.com/2071-1050/11/20/5648>, [retrieved: June, 2021]
- [17] three.js, three.js Home Page. [Online]. Available from: <https://threejs.org>, 2021, [retrieved: June, 2021]
- [18] OGC, CityGML Home Page. [Online]. Available from: <https://www.ogc.org/standards/citygml>, 2012, [retrieved: June, 2021]
- [19] CityJSON, CityJSON Home Page. [Online]. Available from: <https://www.cityjson.org>, 2021, [retrieved: June, 2021]
- [20] J. Virtanen, et al., “Near real-time semantic view analysis of 3D city models in web browser”, *ISPRS International Journal of Geo-Information*, 2021, vol. 10, no. 3, pp. 1-23, [Online]. Available from: <https://www.mdpi.com/2220-9964/10/3/138>, [retrieved: June, 2021]
- [21] Vitalis, S., et al. CITYJSON + WEB = NINJA, *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, VI-4/W1-2020, pp. 167–173, <https://doi.org/10.5194/isprs-annals-VI-4-W1-2020-167-2020>, 2020.
- [22] QGIS, QGIS Home Page. [Online]. Available from: <https://www.qgis.org/en/site/>, 2021, [retrieved: June, 2021]
- [23] OGC, 3D Tiles Home Page. [Online]. Available from: <https://www.ogc.org/standards/3DTiles>, 2018, [retrieved: June, 2021]

Deep Reinforcement Learning for Spatial Motion Planning in 3D Environments

Oren Gal and Yerach Doytsher

Mapping and Geo-information Engineering
Technion - Israel Institute of Technology
Haifa, Israel

e-mails: {orengal@alumni.technion.ac.il, doytsher@technion.ac.il}

Abstract—In this paper, we present a spatial motion planner in 3D environments based on Deep Reinforcement Learning (DRL) algorithms. We tackle 3D motion planning problem by using Deep Reinforcement Learning (DRL) approach, which learns agent's and environment constraints. Spatial analysis focus on visibility analysis in 3D setting an optimal motion primitive considering agent's dynamic model based on fast and exact visibility analysis for each motion primitives. Based on optimized reward function, consisting of generated 3D visibility analysis and obstacle avoidance trajectories, we introduce DRL formulation which learns the value function of the planner and generates an optimal spatial visibility trajectory. We demonstrate our planner in simulations for Unmanned Aerial Vehicles (UAV) in 3D urban environments. Our spatial analysis is based on a fast and exact spatial visibility analysis of the 3D visibility problem from a viewpoint in 3D urban environments. We present DRL architecture generating the most visible trajectory in a known 3D urban environment model, as time-optimal one with obstacle avoidance capability.

Keywords—Deep Reinforcement Learning; Visibility; 3D; Spatial analysis; Motion Planning.

I. INTRODUCTION

Spatial clustering in urban environments is a new spatial field from trajectory planning aspects [1]. The motion and trajectory planning fields have been extensively studied over the last two decades [2][4][6]. The main effort has focused on finding a collision-free path in static or dynamic environments, i.e., in moving or static obstacles, using roadmap, cell decomposition, and potential field methods [11].

The efficient computation of visible surfaces and volumes in 3D environments is not a trivial task. The visibility problem has been extensively studied over the last twenty years, due to the importance of visibility in GIS and Geomatics, computer graphics and computer vision, and robotics. Accurate visibility computation in 3D environments is a very complicated task demanding a high computational effort, which could hardly have been done in a very short time using traditional well-known visibility methods.

In this paper, we present, unique spatial trajectory planning method based on DRL algorithm based on exact

visibility analysis in urban environment. The generated trajectories are based on visibility motion primitives as part of the planned trajectory, which takes into account exact 3D visible volumes analysis clustering in urban environments.

The proposed planner includes obstacle avoidance capabilities, satisfying dynamics' and kinematics' agent model constraints in 3D environments, using Velocity Obstacles (VO) in 3D for Unmanned Aerial Vehicle (UAV) model.

In the following sections, we first introduce the DRL algorithm and method and our extension for a spatial analysis case, such as 3D visibility. Later on, we present the our planner, using VO method and planner model. In the last part of the paper, with planner simulation using DRL method.

II. PROBLEM STATEMENT

We consider the basic visibility problem in a 3D urban environment, consisting of 3D buildings modeled as 3D cubic parameterization $\sum_{i=1}^N C_i(x, y, z = \frac{h_{\max}}{h_{\min}})$, and viewpoint

$V(x_0, y_0, z_0)$.

Given:

- Parameterizations of N objects $\sum_{i=1}^N C_i(x, y, z = \frac{h_{\max}}{h_{\min}})$ describing a 3D urban environment model

Computes:

- *Trajectory*, which consist of optimal set of all visible points, i.e., most visible points of $\sum_{i=1}^N C_i(x, y, z = \frac{h_{\max}}{h_{\min}})$, from starting point q_s to the goal, q_g , without collision.

This problem seems to be solved by conventional geometric methods, but as mentioned before, it demands a long computation time. We introduce a fast and efficient computation solution for a schematic structure of an urban environment that demonstrates our method based on DRL.

On the first part, we present DRL algorithm, formulated to our planning problem, and the visibility analysis along with obstacles avoidance planner.

III. DEEP REINFORCEMENT LEARNING (DRL)

In most Deep Reinforcement Learning (DRL) systems, the state is basically agent's observation of the environment. At any given state, the agent chooses its action according to a policy. Hence, a policy is a road map for the agent, which determines the action to take at each state. Once the agent takes an action, the environment returns the new state and the immediate reward. Then, the agent uses this information, together with the discount factor to update its internal understanding of the environment, which, in our case, is accomplished by updating a value function. Most methods are using the use well-known simple and efficient greedy exploration method maximizing Q-value.

In case of velocity planning space as part of spatial analysis planning, each possible action is a possible velocity in the next time step, that also represent a viewpoint. The Q-value function is based on greedy search velocity, with greedy local search method. Based on that, TD and SARSA methods for DRL can be used, generating visible trajectory in 3D urban environment.

A. Markov Decision Processes (MDP)

The standard Reinforcement Learning set-up can be described as a MDP as can be seen in Figure 1, consisting of:

- **A finite set of states** S , comprising all possible representations of the environment.
- **A finite set of actions** A , containing all possible actions available to the agent at any given time.
- **A reward function** $R = \psi(s_t, a_t, s_{t+1})$, determining the immediate reward of performing an action at from a state s_t , resulting in s_{t+1} .
- **A transition model** $T(s_t, a_t, s_{t+1}) = p(s_{t+1} | s_t, a_t)$, describing the probability of transition between states s_t and s_{t+1} when performing an action a_t .

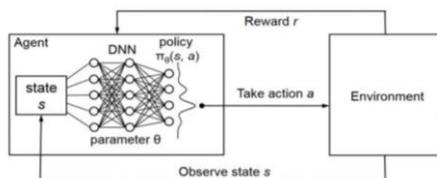


Figure 1. standard Reinforcement Learning Methodology [20].

B. Temporal Difference Learning

Temporal-difference learning (or TD) interpolates ideas from Dynamic Programming (DP) and Monte Carlo methods. TD algorithms can learn directly from raw experiences without any model of the environment.

Whether in Monte Carlo methods, an episode needs to reach completion to update a value function, Temporal-difference learning can learn (update) the value function

within each experience (or step). The price paid for being able to regularly change the value function is the need to update estimations based on other learnt estimations (recalling DP ideas). Whereas in DP a model of the environment's dynamic is needed, both Monte Carlo and TD approaches are more suitable for uncertain and unpredictable tasks.

Since TD learns from every transition (state, reward, action, next state, next reward) there is no need to ignore/discount some episodes as in Monte Carlo algorithms.

C. Spatial Planning Using DRL

In this section, we present DRL approach based on the proposed spatial planning method. The spatial planner seeks to obtain the trajectory T^* that based on visibility motion primitives set as part of the planned trajectory, which takes into account exact 3D visible volumes analysis clustering in urban environments, based on optimizing value function f along T .

The generated trajectories are then represented by a set of discrete configuration points $T = \{x_1, x_2, \dots, x_N\}$. Without loss of generality, we can assume that the value function for each point can be expressed as a linear combination of a set of sub-value functions, that will be called features $c(x) = \sum c_j f_j(x)$. The cost of path T is then the sum of the cost for all points in the path. Particularly, in the Velocity Obstacles as will be presented later on, the value is the sum of the sub-values of moving between pairs of states in the path:

$$\begin{aligned} c(\zeta) &= \sum_{i=1}^{N-1} c(x_i, x_{i+1}) = \sum_{i=1}^{N-1} \frac{c(x_i) + c(x_{i+1})}{2} \|x_{i+1} - x_i\| \\ &= \omega^T \sum_{i=1}^{N-1} \frac{f(x_i) + f(x_{i+1})}{2} \|x_{i+1} - x_i\| = \omega^T f(\zeta) \end{aligned} \quad (2)$$

Based on number of demonstration trajectories D , $D = \{\zeta_1, \zeta_2, \dots, \zeta_D\}$, by using DRL, weights ω can be set for learning from demonstrations and setting similar planning behavior. As was shown by [23][24], this similarity is achieved when the expected value of the features for the trajectories generated by the planner is the same as the expected value of the features for the given demonstrated trajectories:

$$\mathbb{E}(f(\zeta)) = \frac{1}{D} \sum_{i=1}^D f(\zeta_i) \quad (3)$$

Applying the Maximum Entropy Principle [25] to the DRL problem leads to the following form for the probability density for the trajectories returned by the demonstrator:

$$p(\zeta | \omega) = \frac{1}{Z(\omega)} e^{-\omega^T f(\zeta)} \quad (4)$$

where $Z(\omega)$ is a normalization function that does not depend on ζ . One way to determine ω is maximizing the (log-) likelihood of the demonstrated trajectories under the previous model:

$$L(D|\omega) = -D \log(Z(\omega)) + \sum_{i=1}^D (-w^T f(\zeta_i)) \quad (5)$$

The gradient of the previous log-likelihood with respect to ω is given by [23]:

$$\nabla \mathcal{L} = \frac{\partial \mathcal{L}(D|\omega)}{\partial \omega} = \mathbb{E}(f(\zeta)) - \frac{1}{D} \sum_{i=1}^D f(\zeta_i) \quad (6)$$

As mentioned in [23], this gradient in equation (6) can be intuitively explained. If the value of one of the features for the trajectories returned by the planner are higher from the value in the demonstrated trajectories, the corresponding weight should be increased to increase the value of those trajectories.

The main problem with the computation of the previous gradient is that it requires to compute the expected value of the features $\mathbb{E}(f(\zeta))$ for the generative distribution (4).

We suggest setting large amount of D cases, setting the relative w values for our planner characters.

TABLE I. DRL PLANNER PSEUDO CODE

```

DRL Planner
Setting Trajectory S Examples D, D= T*.init (xinit);
Calculate function features Weight, w
fD ← AverageFeatureCount(D);
w ← random_init();
Repeat
    for each T* do
        for VelocityObstacles_repetitions do
            ζi ← getVOstarPath(T*,ω)
            f(ζi) ← calculeFeatureCounts(ζi)
        end for
        fvo(T*) ← ∑i=1VO_repetitions f(ζi)/VO_repetitions
    end for
    fvo ← (∑i=1S fvo)/s
    ∇L ← fvo - fD
    w ← UpdatedWeighths (∇L)
Until convergence
Return w
    
```

IV. UAV MODEL

We introduce an Unmanned Aerial Vehicle (UAV) model, based on the well-known simple car and Dubins airplane [26]. Dubins airplane [27] model extends Dubins car model with continuous change of altitude without reverse gear, avoiding sudden altitude speed rate variation. Our UAV model includes kinematic and dynamic constraints which ignore pitch and roll rotation or winds disturbances.

A. Kinematic Constraints

We use a simple UAV model with four dimensions, each configuration is $q = (x, y, z, \theta)$, when x, y, z are the coordinates of the origin, and θ is the orientation, in x-y plane relative to x-axis, as can be seen in Figure 2 for a simple car-like model.

The steering angle is denoted as ϕ . The distance between front and rear axles is equal to 1. The kinematic equations of a simple UAV model can be written as:

$$\begin{aligned} \dot{x} &= u_s \cos \theta, \\ \dot{y} &= u_s \sin \theta, \\ \dot{z} &= u_z, \\ \dot{\theta} &= u_s \tan u_\phi \end{aligned} \quad (7)$$

Where u_s is the speed parallel to x-y plane, climb rate (speed parallel to z-axis) is u_z and the control on steering angle u_ϕ . We denote the control vector as $u = (u_s, u_z, u_\phi)$. Each of the controllers is bounded, $u_\phi \in [-\phi^{\max}, \phi^{\max}]$ where $\phi^{\max} < \pi/2$, the speed $u_s \in [u_s^{\min}, u_s^{\max}]$ and climb rate $u_z \in [-u_z^{\max}, u_z^{\max}]$. $u_s^{\min} > 0$, so UAV cannot stop.

B. Dynamic Constraints

The UAV model has to take into account the dynamic constraints, preventing instantaneous changes (increase or decrease) of the control vector $u = (u_s, u_z, u_\phi)$.

UAV model also includes dynamic constraints, $\dot{u}_s \in [-a_s, a_s]$, $\dot{u}_z \in [-a_z, a_z]$ and $\dot{u}_\phi \in [-a_\phi, a_\phi]$.

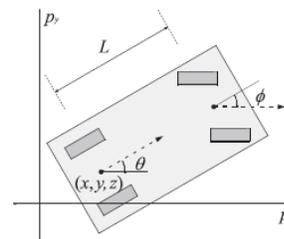


Figure 2. The Simple Car Model. The z-axis can be changed for a Simple -Airplane (Source [26])

V. DEEP REINFORCEMENT LEARNING (DRL) PLANNER

Our planner, as described in Table 1, based on DRL method, generate visible sequence of optimal-visible waypoints as a candidate trajectory. We extend previous planners which takes into account kinematic and dynamic constraints [29,30] and present a local planner for UAV with these constraints, which for the first time generates fast and exact visible trajectories based on analytic solution. The fast and

efficient visibility analysis of our method presented above, allows us to generate the most visible trajectory from a start state to the goal state in 3D urban environments, and demonstrates our capability, which can be extended to real performances in the future. We assume knowledge of the 3D urban environment model and use the well-known Velocity Obstacles (VO) method to avoid collision with buildings presented as static obstacles.

For obstacle avoidance capability, at each time step, the planner computes the next eighth Attainable Velocities (AV). The safe nodes not colliding with buildings, i.e., nodes outside Velocity Obstacles [28], are explored. The planner computes the cost for these safe nodes and chooses the node with the lowest cost. Trajectory can be characterized by the most visible roofs only, surfaces only, or another combination of these kinds of visibility types. We repeat this procedure while generating the most visible trajectory.

A. Velocity Obstacles

The VO [28] is a well-known method for obstacle avoidance in static and dynamic environments, used in our planner to prevent collision between UAV and the buildings (as static obstacles), as part of the trajectory planning method.

The VO represents the set of all colliding velocities of the UAV with each of the neighboring obstacles, in our case static obstacles as can be seen in Figure 3 and Figure 4.

Based on the dynamic and kinematic constraints, UAVs velocities at the next time step are limited. At each time step during the trajectory planning, we map the Attainable Velocities (AV), the velocities set at the next time step $t + \tau$, which generate the optimal trajectory, as is well-known from Dubins theory [27].

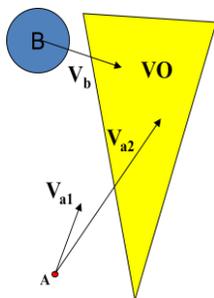


Figure 3. Linear Velocity Obstacles

We denote the allowable controls as $u = (u_s, u_z, u_\phi)$ as U , where $V \in U$.

We denote the set of dynamic constraints bounding control's rate of change as $\dot{u} = (\dot{u}_s, \dot{u}_z, \dot{u}_\phi) \in U'$.

Considering the extremals controllers as part of the motion primitives of the trajectory cannot ensure time-optimal trajectory for Dubin's airplane model [27], but is

still a suitable heuristic based on time-optimal trajectories of Dubin - car and point mass models.

We calculate the next time step's feasible velocities $\tilde{U}(t + \tau)$, between $(t, t + \tau)$:

$$\tilde{U}(t + \tau) = U \cap \{u \mid u = u(t) \oplus \tau \cdot U'\} \quad (14)$$

Integrating $\tilde{U}(t + \tau)$ with UAV model yields the next eight possible nodes for the following combinations:

$$\tilde{U}(t + \tau) = \begin{pmatrix} \tilde{U}_s(t + \tau) \\ \tilde{U}_z(t + \tau) \\ \tilde{U}_\phi(t + \tau) \end{pmatrix} = \begin{pmatrix} u_s^{\min} u_s(t) + a_s \tau \\ -u_s^{\max} \tan \phi^{\max} u_s(t) \tan u_\phi(t) + u_s^{\max} \tan a_\phi \\ u_s^{\max} u_s(t) - a_s \tau \end{pmatrix} \quad (15)$$

At each time step, we explore the next eight AV at the next time step as part of our tree search.

Each node (q, q) , where $q = (x, y, z, \theta)$, consist of the current UAVs position and velocity at the current time step. At each state, the planner computes the set of Admissible Velocities (AV), $\tilde{U}(t + \tau)$, from the current UAV velocity, $U(t)$, as shown in Figure 4. We ensure the safety of nodes by computing a set of Velocity Obstacles (VO).

In Figure 4, nodes inside VO, marked in red, are inadmissible. Nodes out of VO are further evaluated; safe nodes are colored in blue. The safe node with the lowest cost, which is the next most visible node, is explored in the next time step. This is repeated while generating the most visible trajectory.

Admissible velocities profile is similar to a trunked cake slice, as seen in Figure 4, due to the Dubins airplane model with one time step integration ahead. Simple models admissible velocities, such as point mass, create rectangular profile [28].

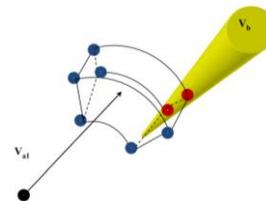


Figure 4. Tree Search Method. Admissible Velocities marked in Blue and Red Circles; Nodes inside VO (marked Red) are Inadmissible; Nodes outside VO, Colored in Blue with Lowest Cost, are Explored

B. Cost Function

Our search is guided by minimum invisible parts from viewpoint V to the 3D urban environment model. The cost function for each node is a combination of IRV and ISV, with different weights as functions of the required task.

The cost function is computed for each safe node $(q, q) \notin VO$, i.e., node outside VO, considering UAV

position at the next time step $(x(t + \tau), y(t + \tau), z(t + \tau))$ as viewpoint:

$$w(q(t + \tau)) = \alpha \cdot ISV(q(t + \tau)) + \beta \cdot IRV(q(t + \tau)) \quad (16)$$

Where α, β are coefficients, effecting the trajectory character. The cost function $w(q(t + \tau))$ produces the total sum of invisible parts from the viewpoint to the 3D urban environment, meaning that the velocity at the next time step with the minimum cost function value is the most visible node in our local search.

C. Planner Neural Network

In our DRL model, we are using fully-connected layers, consisting of: the state space of 37 dimensions, two hidden layers (64 nodes each), an output of four actions. Our network structure can be seen in Figure 5.

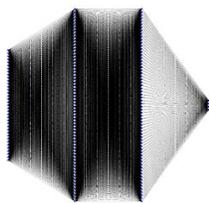


Figure 5. DRL planner network model based on fully-connected layers

D. Simulation Results

We have implemented the presented algorithm and tested some urban environments. We computed the visible trajectories using our DRL planner, as described above. We used the proposed UAV model with several types of trajectories consisting of roof and surfaces visibility, based on the introduced visibility computation method. Obstacle avoidance capability tested by VO method.

The initial parameters values are: $u_s(t = 0) = 10$ [m/s], $u_z \theta(t = 0) = 5$ [deg]. UAV dynamic and kinematic constraints are $\phi^{\max} = \pi/4$, $u_z^{\max} = 0.3$ [m/s]. $u_s^{\min} = 1$ [m/s], $u_s^{\max} = 15$ [m/s].

In the following simulations, Figures 6 till Figure 10, the start and goal points are marked, in number of scenarios with various start's and goal's points location.

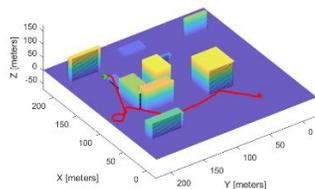


Figure 6. Trajectory Planning in Urban Environment Using DRL. Start and Goal Points with Scenario Demonstration.

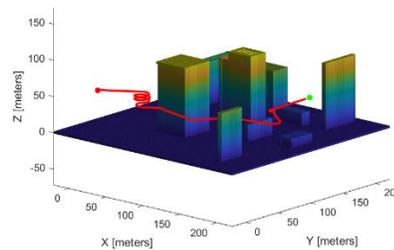


Figure 7. Trajectory Planning in Urban Environment Using DRL. Setting other Start and Goal Points with Scenario Demonstration.

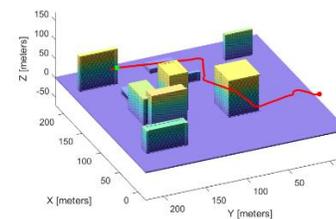


Figure 8. Trajectory Planning in Urban Environment Using DRL. Setting other Start and Goal Points with Scenario Demonstration.

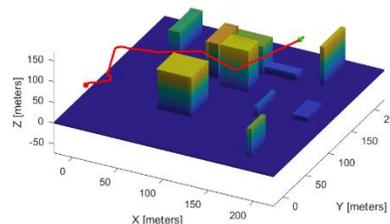


Figure 9. Trajectory Planning in Urban Environment Using DRL. Setting other Start and Goal Points with Scenario Demonstration.

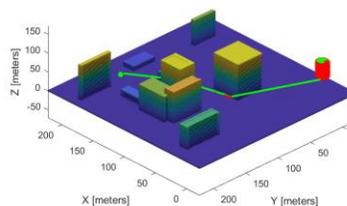


Figure 10. Trajectory Planning in Urban Environment Using DRL. Setting other Start and Goal Points with Scenario Demonstration.

VI. CONCLUSIONS

In this paper, we present a spatial motion planner in 3D environments based on Deep Reinforcement Learning (DRL) algorithms. We tackled 3D motion planning problem by using Deep Reinforcement Learning (DRL) approach which learns agent's and environment constraints.

Spatial analysis focus on visibility analysis in 3D setting an optimal motion primitive considering agent's dynamic model based on fast and exact visibility analysis for each

motion primitives. Based on optimized reward function, consist of generated 3D visibility analysis and obstacle avoidance trajectories, we introduced DRL formulation which learns the value function of the planner and generates an optimal spatial visibility trajectory.

We presented DRL architecture generating the most visible trajectory in a known 3D urban environment model, as time-optimal one with obstacle avoidance capability.

REFERENCES

[1] O. Gal and Y. Doytsher, "Spatial Visibility Clustering Analysis In Urban Environments Based on Pedestrians' Mobility Datasets," The Sixth International Conference on Advanced Geographic Information Systems, Applications, and Services, pp. 38-44, 2014.

[2] J. Bellingham, A. Richards, and J. How, "Receding Horizon Control of Autonomous Aerial Vehicles," in Proceedings of the IEEE American Control Conference, Anchorage, AK, pp. 3741-3746, 2002.

[3] A. Borgers and H. Timmermans, "A model of pedestrian route choice and demand for retail facilities within inner-city shopping areas," Geographical Analysis, vol. 18, No. 2, pp. 115-128, 1996.

[4] S. A. Bortoff, "Path planning for UAVs," In Proc. of the American Control Conference, Chicago, IL, pp. 364-368, 2000.

[5] B. J. Capozzi and J. Vagners, "Navigating Annoying Environments Through Evolution," Proceedings of the 40th IEEE Conference on Decision and Control, University of Washington, Orlando, FL, 2001.

[6] H. Chitsaz and S. M. LaValle, "Time-optimal paths for a Dubins airplane," in Proc. IEEE Conf. Decision. and Control., USA, pp. 2379-2384, 2007.

[7] B. Donald, P. Xavier, J. Canny, and J. Reif, "Kinodynamic Motion Planning," Journal of the Association for Computing Machinery, pp. 1048-1066, 1993.

[8] Y. Doytsher and B. Shmutter, "Digital Elevation Model of Dead Ground," Symposium on Mapping and Geographic Information Systems (Commission IV of the International Society for Photogrammetry and Remote Sensing), Athens, Georgia, USA, 1994.

[9] W. Fox, D. Burgard, and S. Thrun, "The dynamic window approach to collision avoidance," IEEE Robotics and Automation Magazine, vol. 4, pp. 23-33, 1997.

[10] O. Gal and Y. Doytsher, "Fast and Accurate Visibility Computation in a 3D Urban Environment," in Proc. of the Fourth International Conference on Advanced Geographic Information Systems, Applications, and Services, Valencia, Spain, pp. 105-110, 2012 [accessed February 2014].

[11] O. Gal and Y. Doytsher, "Fast and Efficient Visible Trajectories Planning for Dubins UAV model in 3D Built-up Environments," Robotica, FirstView, Article pp. 1-21 Cambridge University Press 2013 DOI: <http://dx.doi.org/10.1017/S0263574713000787>, [accessed February 2014].

[12] M. Haklay, D. O'Sullivan, and M.T. Goodwin, "So go down town: simulating pedestrian movement in town centres," Environment and Planning B: Planning & Design, vol. 28, no. 3, pp. 343-359, 2001.

[13] S. Karaman and E. Frazzoli, "Sampling-based algorithms for optimal motion planning," Int. J. Robot. Res., vol. 30, no. 7, pp. 846-894, 2011.

[14] N.Y. Ko and R. Simmons, "The lane-curvature method for local obstacle avoidance," In International Conference on Intelligence Robots and Systems, 1998.

[15] S. M. LaValle, "Rapidly-exploring random trees: A new tool for path planning," TR 98-11, Computer Science Dept., Iowa State University, 1998.

[16] S. M. LaValle and J. Kuffner. "Randomized kinodynamic planning," In Proc. IEEE Int. Conf. on Robotics and Automation, Detroit, MI, pp. 473-479, 1999.

[17] L.R. Lewis, "Rapid Motion Planning and Autonomous Obstacle Avoidance for Unmanned Vehicles," Master's Thesis, Naval Postgraduate School, Monterey, CA, December 2006.

[18] C. W. Lum, R. T. Rysdyk, and A. Pongpunwattana, "Occupancy Based Map Searching Using Heterogeneous Teams of Autonomous Vehicles," Proceedings of the 2006 Guidance, Navigation, and Control Conference, Autonomous Flight Systems Laboratory, Keystone, CO, August 2006.

[19] S. Okazaki and S. Matsushita, "A study of simulation model for pedestrian movement with evacuation and queuing," Proceedings of the International Conference on Engineering for Crowd Safety, London, UK, pp. 17-18, March 1993.

[20] P. Abbeel and P. Ng, "Apprenticeship learning via inverse reinforcement learning," in Proceedings of the twenty-first international conference on Machine learning, ICML '04, ACM, New York, NY, USA, <http://doi.acm.org/10.1145/1015330.1015430>, 2004.

[21] M. Kuderer, S. Gulati, and W. Burgard, "Learning driving styles for autonomous vehicles from demonstration," in Proceedings of the IEEE International Conference on Robotics & Automation (ICRA), Seattle, USA. vol. 134, 2015

[22] B. Ziebart, A. Maas, J. Bagnell, and A. Dey, "Maximum entropy inverse reinforcement learning," in Proc. of the National Conference on Artificial Intelligence (AAAI), 2008.

[23] Kuderer, M., Gulati, S., Burgard, W. (2015) "Learning driving styles for autonomous vehicles from demonstration", In: Proceedings of the IEEE International Conference on Robotics & Automation (ICRA), Seattle, USA. vol. 134.

[24] Abbeel, P., Ng, A.Y., (2004) "Apprenticeship learning via inverse reinforcement learning" In: Proceedings of the twenty-first international conference on Machine learning, ICML '04, ACM, New York, NY, USA, <http://doi.acm.org/10.1145/1015330.1015430>.

[25] Ziebart, B., Maas, A., Bagnell, J., Dey, A., (2008) "Maximum entropy inverse reinforcement learning", In: Proc. of the National Conference on Artificial Intelligence (AAAI).

[26] S. M. LaValle, Planning Algorithms. Cambridge, U.K.: Cambridge Univ. Pr., (2006).

[27] H. Chitsaz and S. M. LaValle, Time-optimal paths for a Dubins airplane, in Proc. IEEE Conf. Decision. and Control., USA, pp. 2379-2384, (2007).

[28] P. Fiorini and Z. Shiller, Motion planning in dynamic environments using velocity obstacles, Int. J. Robot. Res.17, pp. 760-772, (1998).

[29] G. Elber, R. Sayegh, G. Barequet and R. Martin. Two-Dimensional Visibility Charts for Continuous Curves, in Proc. Shape Modeling, MIT, Boston, USA, pp. 206-215, (2005).

[30] S. A. Bortoff, Path planning for UAVs. In Proc. of the American Control Conference, Chicago, IL, pp: 364-368 ,(2000).

Image Classification Methods Assessment for Identification of Small-Scale Agriculture in Brazilian Amazon

Flávia Domingos Pacheco, Maíra Ramalho Matias, Gabriel Máximo da Silva,
Anielli Rosane de Souza, Yosio Edemir Shimabukuro, Maria Isabel Sobral Escada

Earth Observation and Geoinformatics Division – DIOTG

Brazilian National Institute for Space Research

São José dos Campos, Brazil

e-mail: {flavia.pacheco; maira.matias; gabriel.maximo; anielli.souza; yosio.shimabukuro; isabel.escada}@inpe.br

Abstract—This paper aims to test different methods for image classification focusing on small-scale agriculture in the region of Mocajuba and Cameté, municipalities in the Northeast of Pará state, Brazil. It is an important land use class, always ignored by Land-Use and Land-Cover monitoring systems because of its small size and variable spectral signature. We used an image from the PlanetScope Surface Reflectance Mosaics (Analysis Ready) with spatial resolution of 4.77 meters and 4 spectral bands (red, green, blue and infra-red). After proceeding with a multiresolution segmentation to identify image objects, two object-oriented classification algorithms were tested: Adapted Nearest-neighbor and C5.0 Decision trees algorithms. We selected 122 random points using the images available on Google Earth Pro as reference to assess the accuracy of classifications. Afterwards, confusion matrices were generated. Both methods showed similar overall accuracy and kappa value. However, C5.0 Decision trees reached a higher producer's accuracy to small-scale agriculture (75%) in comparison to Adapted Nearest-neighbor (65%). The average size of the small-scale agriculture segments estimated was less than 1 ha in both maps, showing the need to carry out studies on scales of greater detail, preferably with images of high spatial resolution to represent these systems properly. In this study, C5.0 Decision trees had the best result, representing the most suitable method for mapping small-scale agriculture in Brazilian Amazon.

Keywords - digital image processing; segmentation; land use; land cover; smallholders; planetscope.

I. INTRODUCTION

Small-scale agriculture is a rather important land use activity that takes place in small properties, providing income for families and food for local population [1]. Although we have few definitions about it, these producers are somehow invisible to public policy and they have difficulty in accessing credits for their production [2].

Despite of its importance, small-scale agriculture is not always properly represented in the maps produced by Land-Use and Land-Cover (LULC) monitoring systems of Brazilian Amazon, such as TerraClass [3] and MapBiomass [4]. This class is included in generic mixed classes of LULC, that comprises undistinguishable features due to their small size. Besides using medium spatial resolution images, the mapping scale of these monitoring systems makes this land use undistinguishable, so they are included in mixed classes that embraces agriculture, pasture, secondary vegetation, etc.

Identifying small-scale agriculture with satellite images is a challenge and demands new approaches. Only by knowing about the presence of this type of agriculture system and its spatial pattern, cartographic representation can be complete and small-scale agriculture can take its part in policies, since it makes possible to understand how this system works, who are the agents involved on it and so on [5]. Also, it is important to highlight that a large part of the food products that supply cities come from family farming included in small-scale agriculture. Therefore, we reiterate the importance of studying small-scale agriculture in the Amazon as a way of demonstrating quantitative and qualitative data of this important land use to subsidize public policies for the regional economy.

Classifying an image requires high analyst experience and good background knowledge about the region of classification. Moreover, combining diverse parameters in supervised classification can improve the methods and provide good results [6]. Therefore, it is important to address the specific characteristics regarding small-scale agriculture in Brazilian Amazon. Especially for Amazon, scenes from optical sensors face problems with high proportion of cloud cover, which is worsened by low temporal resolution of some types of images [7]. Also, there is a significant confusion between small-scale agriculture with pasture and secondary vegetation, once they show similar spectral responses [8]. However, those LULC classes vary in shape and size [5].

To overcome these issues, different methods have been used to collect meaningful and useful information from image processing, GIS and modeling [9]. To consider shape information, the analyst may select Object-Based Image Analysis (OBIA) methods for image classification [10][11]. These methods address spectral and shape attributes using image segmentation. In addition, an alternative for high spatial resolution images to reduce cloud cover is to make a mosaic of images with the best pixels found in a certain period.

Overall, considering the gap of information and challenges, there is still a lot to do when identifying small-scale agriculture using remote sensing techniques. To assess how different methods perform, this work proposes to classify small-scale agriculture in an area that includes part of two municipalities in the Baixo Tocantins region (Brazilian Amazon), testing different image classification

algorithms based on OBIA and high spatial resolution image, the PlanetScope product of Surface Reflectance Mosaic.

We organized this paper in four sessions. In Section 2, we provide the state of the art regarding small-scale agriculture mapping, we discuss related work, existing solutions and limitations. In Section 3, we describe the PlanetScope scene, discussing its characteristics, as well as the study area. We explain all the methods applied, from image segmentation to classification and accuracy assessment. In Section 4, we show and discuss the results for each classification method, comparing their performance mainly related to small-scale agriculture class. In Section 5, we give our final remarks, highlighting pros and cons of the tested methods and future work opportunities for this matter.

II. STATE OF THE ART

In Brazilian Amazon several studies on agriculture have been carried out with remote sensing techniques. However, most of them addresses large-scale agriculture. Few studies can be found related to small-scale agriculture. On the other hand, there are plenty of techniques that can be tested for mapping this land use class. Therefore, the main contribution of this study relies on testing and evaluating techniques capable of detecting this type of agriculture, which is largely invisible, despite its importance to society, environment and economy. By doing so, we explore the challenges, potentials and constraints of mapping small-scale agriculture in Brazilian Amazon.

When searching for techniques, we can find plenty of options of classification algorithms that can be used for agriculture classification. Some authors use traditional approaches and algorithms to identify large-scale agriculture, such as Maximum likelihood (ML) and ISODATA [12][13]. Those algorithms are pixel-wise and depend uniquely on spectral response of the targets. However, small-scale agriculture is often composed of more than one crop types at different stages of growth, which results frequently in a spectral mixture at the pixel level.

By using ML classification, a study carried out in Ethiopia could not discriminate crop types not even when the crop classes were groupings. To improve the result, the authors applied a second approach by using a neural network for sub-pixel classification of the image [13]. Another study carried out in Brazilian Amazon used ML to map smallholders and reported that this algorithm is not appropriate for it, the results showed a very low accuracy (8%) for small-scale agriculture [6]. The authors also performed a classification by segmenting the image and applying Adapted Nearest Neighbor method. In this case, they were able to reach a higher accuracy (64%) for detecting small-scale agriculture.

Segmentation and object-based analysis are broadly used in many studies [10][14]-[17]. The main gain of using this technique for small-scale agriculture detection is because the segmentation allows the use of more features, such as shape, texture and so on, rather than only spectral ones. Once small-scale agriculture has specific shape and texture, and spectral mixture, an object-based analysis unfolds as a key technique [6]. For some authors, working with object is also an

advantage due to the ease of interpretation, for the features correspond to elements of landscape [15].

Lastly, it is important to mention the image spatial resolution. Although [18] do not work directly with LULC classification methods, their research is based on remote sensing data and techniques and the analysis of small-scale agriculture intensification. The authors used Landsat and MODIS imagery, which have a spatial resolution of 30 and 250 m, respectively. However, the use of higher spatial resolution images is more suitable for small farms, where cultivation takes place in areas smaller than 1ha, as in Northeast of Pará state [6]. In that sense, many authors used RapidEye imagery, which have a spatial resolution of 5 m [6][12][19]. Furthermore, the red-edge band from RapidEye imagery is adequate to discriminate the stage of vegetation, supporting small-scale agriculture mapping [20]-[22].

Overall, we can notice that frequently small-scale agriculture is not considered in LULC mapping and there are only a few researches regarding this matter. In the researches we reviewed we could commonly observe the use of the combination of different techniques: ML and neural networks [13]; multiresolution segmentation and adapted nearest neighbor [6]; and segmentation and random forest algorithm [14][15]. In some studies, we can observe that the authors adapted and tested techniques used to large-scale agriculture, but considering the unique features of small-scale agriculture in Amazon.

III. MATERIAL AND METHODS

PlanetScope product of Surface Reflectance Mosaic is a free access and analysis-ready level product. This product is a Level-1 processing data, including geometric and atmospheric correction. Atmospheric parameters are estimated from an external data source, such as MODIS [23]. The availability of Planet's monthly mosaics comprises an initiative of Norway's International Climate and Forest Initiative (NICFI), which aims to provide universal access to monitoring the tropics through high spatial resolution satellite imagery, to support efforts to stop the deforestation of the world's rainforests [23].

The scene is a monthly composition of pixels acquired during the month of September, 2020. This mosaic has spatial resolution of 4.77 m and has four spectral bands: red, green, blue and near-infrared [23]. The image covers part of Mocajuba and Cametá municipalities, in the Northeast of Pará state, Brazil (Figure 1). We selected this study area because it is a hotspot of small-scale agriculture in Pará State [10]. This region is historically occupied by many smallholders that are responsible to provide food for local and regional markets, besides self-consumption [24]. In both municipalities, the main crop of small-scale agriculture is cassava (Table 1), which is planted exclusively in farms smaller than 10 ha. Açai, black pepper and cocoa are important crops, but they are not exclusive to small-scale farms, they are also cultivated in farms with area up to 200 ha [25].

Furthermore, this region is a hotspot of secondary vegetation [26], that can be an indicator of shifting

cultivation, usually performed in Amazon by smallholders that leave part of the land to fallow and return their cultivation afterwards, reincorporating the nutrients and minerals to the soil [27]. The main river body are Tocantins River and its tributaries.

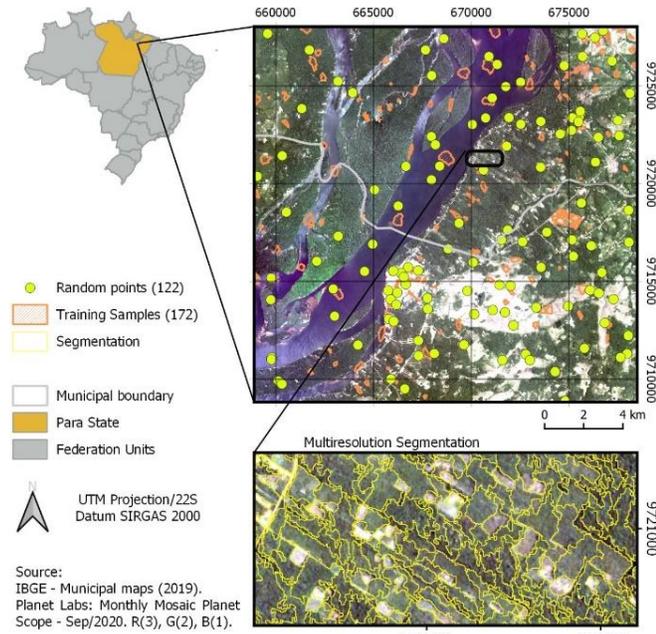


Figure 1. Study area (Mocajuba and Cametá municipalities, Pará, Brazil), training samples, random validation points, and segmentation example.

TABLE I. MAIN CROPS OF SMALL-SCALE AGRICULTURE IN CAMETÁ AND MOCAJUBA, CONSIDERING PROPERTIES UP TO 10 HA (2017)

Crop	Harvested area (ha)
Cassava	4,507
Açaí	1,676
Black pepper	413
Cocoa	402
Corn	319

Source: [25].

We used two different supervised methods: Adapted Nearest-neighbor (NN) [28] and C5.0 Decision trees, an improved version of C4.5 [29]. The main steps applied for each approach were: a) multiresolution segmentation and parameters definition; b) training sampling; c) classification; and d) assessment of accuracy. We chose LULC classes considering visual surveys and landscape descriptions in similar studies developed at the same location (Table 2).

TABLE II. LAND-USE AND LAND-COVER CLASSES

Classes	Description
Water	Water bodies: rivers, lagoons, etc.
Forest	Natural vegetation with predominance of trees
Secondary vegetation	Natural vegetation in regeneration emerged from previously deforested areas, with trees, shrubs and herbs

Urban areas	Built-up areas with population clusters: city, village and community
Pasture	Predominance of herbaceous and grassy vegetation, it may occur also sparse shrub vegetation and few arboreal individuals
Small-scale agriculture	Small agriculture lands with mainly annual crops
Others	Aggregate of land use and land cover, such as rocky outcrops, sand banks
Non observed	Clouds and cloud shadows

Source: adapted from [6].

A. Multiresolution segmentation

The multiresolution segmentation was generated by eCognition 9.0.1 [30]. This kind of segmentation creates initially an 1-pixel-sized object and merges the neighbor pixels with similar features consecutively into the object [31]. When observing similar features, the algorithm considers a combination of spectral and shape criteria. The analyst can weigh the priority criteria when creating the segments. We repeatedly tested different values for segmentation parameters, until the algorithm could create appropriate objects for small-scale agriculture. We found the values of the following parameter more adequate:

- Image layer weights = 1, 1, 1, 1. It ranges from 0 to 1. The same importance was assigned to all bands (R, G, B, NIR);
- Scale parameter = 60. It defines the size of polygons;
- Shape = 0.7. It defines the weight the shape criteria must have for segmentation. It ranges from 0 to 1, the higher its value, the lower the influence of the color. The chosen value prioritizes the shape over color;
- Compactness = 0.5. It defines the weight of compactness criteria. It ranges from 0 to 1, the higher the value, the more compact image objects may be, that is the borders are closer to the center of the segment.

B. Sampling design

This step was performed using simple random sampling method throughout the scene. The sample units were objects obtained from the multiresolution segmentation. We selected the same objects/segments as training samples for both methods to keep them uniform. In total, we selected 172 objects for all the 8 classes. We sampled 20 objects for each class described in Table 2, except *Cloud shadows*, with 12 samples. We collected fewer samples of *Cloud shadows* because this class was less representative in terms of area.

Note that we collected *Clouds* and *Cloud shadows* samples individually. After performing the classification, we merged those features into *Non observed* class.

Overall, this step involved the random selection of segments from different shapes and sizes in each class. Water sample segments presented a mean area of 12.83 ha ± 8.60 ha for these specific samples, compared to other classes (Table 3). On the other hand, small-scale agriculture presented a mean area of 0.84 ha ± 0.36 ha for these samples. As we standardized only the number of samples, this

variation in sample size might affect the final classification according to spectral mixture features of the pixels, and it may impact the accuracy. The main class for this study is small-scale agriculture and the random samples selected in this class presented similar sizes. As we see in Figure 2, this is a size pattern for this class in the Amazon for both algorithms, so the samples represented well the classification for this region.

TABLE III. DESCRIPTIVE STATISTICS FROM SAMPLING POLYGONS, IN HECTARE

Classes	μ	σ	σ^2
Water	12.83	8.60	73.87
Forest	3.86	3.25	10.56
Secondary Vegetation	6.05	2.63	6.90
Urban Area	1.44	0.94	0.89
Pasture	3.09	2.32	5.37
Small-scale agriculture	0.84	0.36	0.13
Others	2.72	2.74	7.49

Units: hectare; μ = polygon mean area; σ = standard deviation; σ^2 = variance.

C. Object-based image analysis

OBIA is an alternative to pixel-to-pixel approaches as it relies on identifying regions of the image: it uses segments to extract neighborhood, spectral and spatial features, composing the feature space [10][11]. This method has shown to be more suitable to identify small-scale agriculture and to distinguish this land use from other classes [6]. Therefore, we ran two OBIA methods using the same image, samples and accuracy points. Both methods differ when it comes to classifying an object: the first one considers the nearest neighbor in the feature space, while the latter one uses multiple decision trees to identify the proper class, as explained in the next sections.

C.1 Adapted Nearest-neighbor (NN)

This method is an adapted version of nearest neighbor, which considers not only the spectral features, but also other features related to the object [28]. The algorithm formulates a feature space considering the attributes including all segments, then it searches for the closest sample and assigns that class to the segment [30]. The analyst determines the correspondent features. To proceed with classification, we chose the features according to [6]: a) Spectral attributes: brightness, mean and standard deviation of each band, b) Object attributes: shape index. In total, we used 13 attributes.

C.2 Decision trees (C5.0)

We used the objects from multiresolution segmentation to perform a feature extraction from the original image by Geographical Data Mining Analyst (GeoDMA), an open-source plug-in available for TerraView 5.5.1 [32]. In total, the algorithm generated 103 features, considering both spectral and spatial features, e.g., mean, mode, and maximum values of each band, polygon angle, shape index,

compactness etc. After finishing the feature extraction, we collected the training samples and ran a boosting C5.0 Decision trees classification.

This algorithm generates a pre-set number of decision trees from the sample features, which is applied when classifying the segments. In total, the algorithm generated 99 trees independently. The final classification for each segment is the one that was assigned by the most of the trees [33].

Among the 103 features, the algorithm highlights the ones that were more used when classifying an object. The main features, that showed 100% usage in this classification were:

- Band 2 (Green): maximum value and band ratio;
- Band 3 (Blue): median, mode, dissimilarity and contrast;
- Band 4 (Near-infrared): mean, median and band ratio.

Band ratio is the contribution of the given band to the region. Contrast is the measure of the intensity contrast between a pixel and its southeast neighbor over the object, aka Sum of Squares Variance. Dissimilarity is the measure of how different the elements of the Gray-level co-occurrence matrix are from each other [34].

Regarding object features, the more important ones, which had ca. 60% of usage, were: Perimeter, compacity, radius and bounding box area.

D. Assessment of accuracy

To assess the accuracy of both classifications, 122 points were randomly distributed throughout the scene, representing all classes. For each class, we collected 20 random points, except for *Pasture*. Note that we had only 2 points for *Pasture* due its low scene cover.

We used those points as test samples once the ground truth was assigned by inspecting the actual land use and land cover with Google Earth Pro 2020 images. Then, we used those points to validate NN and C5.0 classifications. From that point on, confusion matrix was created for each classification and we computed the accuracy and kappa index. These information were the basis for identifying the main confusion occurring to small-scale agriculture areas and the overall performance of each method, leading to the most suitable one for small-scale agricultural mapping [35][36].

IV. RESULTS AND DISCUSSION

Classification maps are showed on Figure 2. Small-scale agriculture is more present in the upland region, even though that both riverine and upland population are acknowledged as important agents involved into this land use activity [6]. As stated before, high spatial resolution sensors are more adequate to improve classification accuracy due to the small-scale agriculture's size: our results presented mean area of $0.97 \text{ ha} \pm 0.69 \text{ ha}$ for NN and $0.70 \text{ ha} \pm 0.39 \text{ ha}$ for C5.0 Decision trees (Table 4). Considering that TerraClass maps have a minimum mapping area of 6.25 ha [6], it can not identify and map properly small-scale agriculture. That explains why this class is not explicitly visible in current LULC monitoring systems.

Descriptive statistics indicate similar classification area in both algorithms for the classes of water, forest, secondary vegetation, urban area, and others. The study area shows great forest and secondary vegetation cover in mainland and at the islands, covering ca. 60% of the scene in both maps. Water covers around 18% of the area.

As presented in confusion matrix (Table 5), both algorithms presented similar overall accuracy (NN = 75% and C5.0 Decision trees = 73%) and kappa values. On the other hand, C5.0 Decision trees algorithm found better results when mapping small-scale agriculture (75%), compared to NN (65%). This performance of Adapted Nearest-neighbor algorithm is corroborated with other studies that found around 62% of producer's accuracy for small-scale agriculture carried out in the same region of Brazilian Amazon [6].

NN may be overclassifying small-scale agriculture, representing 9.3% of the mapping area. For this classification, commission error was 19%, which indicates that a significant number of polygons were classified by mistake as small-scale agriculture, increasing the area of this class. These classification errors occurred due to confusion, especially with secondary vegetation, forest, and others.

For C5.0 Decision trees, there was no commission error for small-scale agriculture class, which represents 4.8% of the mapping area. In other words, C5.0 Decision trees is more conservative for mapping small-scale agriculture and did not included other classes in small-scale agriculture by mistake as NN did.

Both algorithms showed the same omission errors for small-scale agriculture regarding secondary vegetation (15%) and pasture (10%) classes. NN also showed omission errors for small-scale agriculture with the class others (10%).

According to the literature, similar spectral attributes may affect the classification of small-scale agriculture, once this class has similar spectral responses to other classes, such as pasture [8] and secondary vegetation [5].

Small-scale agriculture and pasture differ in terms of size and shape [5]. Although the last one showed a mean area of 1.77 ha in both maps, the only area in the scene, identified as pasture by the analysts, had actually 100 ha. Yet, this single pasture area was segmented into many smaller objects by the multiresolution segmentation, once covered by grassland in different stages, for instance, clean pasture, shrubby pasture. When segmenting an image with parameters adjusted to small-scale agriculture objects, it is necessary to resort the spatial resolution with better detail, which influences other targets segmentation. In this case, although the targets differ when it comes to object features, there was still confusion among them. Anyhow, the confusion with pasture might also have been influenced by the small number of samplings due to the lack of other pasture areas in this scene. To better represent these classes we should separate them in two classes, clean and shrubby pasture. But due to the low presence of pasture on the scene, this step was not feasible. However, we highlight the confusion and the need to separate this land use in two

distinct classes in case this methodology is applied in an area where pasture is more significant.

Regarding the confusion with secondary vegetation, once again the spectral features are the main reason for it, as well as the different stages of secondary vegetation, that poses similar stratum as agriculture. The mean size of secondary vegetation polygons was 1.71 ± 1.34 for NN and 1.58 ± 1.18 for C5.0. Even though this confusion is not opportune, it is important to address that secondary vegetation can be part of small-scale agriculture production system by forming biomass while the land is under fallow [37]. In this case, secondary vegetation poses as an asset, once its function is to ensure the land fertility [27].

Overall, the results for small-scale agriculture were adequate and despite the different accuracies, both methods showed limitations when differentiating this class from pasture and secondary vegetation.

V. CONCLUSION AND FUTURE WORK

As the first challenge faced in this study, we can highlight the successful attempted of using the same training and validation samples for the classification and evaluation steps, to promote an adequate comparison between the algorithms tested. Using the same samples in different software and algorithms is not always possible or easy. For instance, C5.0 automatically generates the confusion matrix for the classification, but does not point out the samples used in the validation process. For this reason, we collected randomly distributed points in the image to evaluate the two classifications. We can summarize our main findings as following:

- Considering all small-scale agriculture identification challenges, C5.0 Decision trees results were able to reach a higher producer's accuracy compared to NN method. Note that both methods were run relying upon different magnitude of features: C5.0 Decision tree identify automatically 103 features, however the ones that mattered the most for the classification were 33 of them, which were used at least in 50% of the decision trees. On the other hand, NN used 13 features pointed out by the analyst. So, while the latter requests that the analyst decides how many and which features are going to be inserted in the feature space, C5.0 Decision trees use as default 103 features and saves the analyst from selecting the most suitable features for small-scale agriculture.
- Therefore, C5.0 showed greater results, representing the most suitable method for mapping small-scale agriculture in the study area. Nevertheless, we recommend carrying out more studies over larger areas to identify the best attributes for classifying small-scale agriculture and overcome misclassification errors, as well as in other mixed land cover types and landscape diffuse patterns.

Regarding future work, we have the following remarks:

- We recommend investigating which features are more significant for the identification of small-scale agriculture by C5.0. We suggest a systematic removal of features at the classification level and performing a sensitive analysis;
- We believe that temporal analysis can be explored in future work. The inclusion of the temporal component coupled with machine learning and deep learning techniques may contribute for selecting other important variables for small-scale agriculture classification.

Additionally, the use of these methods may contribute for the advancement of studies linked to agricultural intensification and following practices in shifting cultivation agriculture, widely used in the context of small-scale agriculture in the Brazilian Amazon. Then, the inclusion of the time component will also be important to test whether it is relevant or not to use phenological metrics for agriculture with large species diversity that does not present regular crop cycles.

- Also, we strongly recommend testing different sampling design to test better results and perform a sensitive analysis.

TABLE IV. DESCRIPTIVE STATISTICS FROM IMAGE CLASSIFICATION ACCORDING TO LAND USE AND LAND COVER, IN HECTARE

Classes	Adapted Nearest-neighbor					C5.0 Decision trees				
	μ	σ	σ^2	Total	%	μ	σ	σ^2	Total	%
Water	5.69	5.62	31.62	7,082.61	18.66	5.92	5.75	33.01	6,885.70	18.14
Forest	1.55	1.18	1.40	11,317.54	29.82	1.62	1.26	1.58	10,477.20	27.60
Secondary vegetation	1.71	1.27	1.61	11,182.62	29.46	1.58	1.18	1.39	11,934.05	31.44
Urban Area	0.69	0.60	0.36	633.66	1.67	0.70	0.63	0.40	537.14	1.42
Pasture	1.77	1.34	1.79	718.84	1.89	1.77	1.10	1.21	2,340.14	6.17
Small-scale agriculture	0.97	0.69	0.48	3,526.97	9.29	0.70	0.39	0.15	1,837.95	4.84
Others	1.22	1.37	1.88	3,493.84	9.20	1.38	1.38	1.92	3,943.92	10.39
Total	-	-	-	37,956.08	100	-	-	-	37,956.08	100

Units: hectare; μ = polygon mean area; σ = standard deviation; σ^2 = variance.

TABLE V. CONFUSION MATRIX FOR ADAPTED NEAREST-NEIGHBOR AND C5.0 DECISION TREES ALGORITHMS

Adapted Nearest-neighbor									C5.0 Decision trees										
%		Reference							User's accuracy	%		Reference							User's accuracy
		(A)	(B)	(C)	(D)	(E)	(F)	(G)				(A)	(B)	(C)	(D)	(E)	(F)	(G)	
Classification	(A) Water	100	0	0	0	0	0	0	100	Classification	(A)	90	0	0	0	0	0	0	100
	(B) Forest	0	55	35	0	0	0	0	61		(B)	0	50	30	0	0	0	0	63
	(C) Secondary vegetation	0	40	60	0	0	15	0	52		(C)	0	45	60	0	0	15	0	50
	(D) Urban area	0	0	0	95	0	0	0	100		(D)	0	0	0	85	0	0	5	94
	(E) Pasture	0	0	0	0	50	10	10	20		(E)	0	0	10	15	100	10	20	15
	(F) Small-scale agriculture	0	5	5	0	50	65	0	81		(F)	0	0	0	0	0	75	0	100
	(G) Others	0	0	0	5	0	10	90	86		(G)	10	5	0	0	0	0	75	83
	Producer's accuracy	100	55	60	95	50	65	90			Prod. acc.	90	50	60	85	100	75	75	
Samples	20	20	20	20	2	20	20		Samples	20	20	20	20	2	20	20			
Kappa = 0,70 Overall accuracy = 75%									Kappa = 0,68 Overall accuracy = 73%										

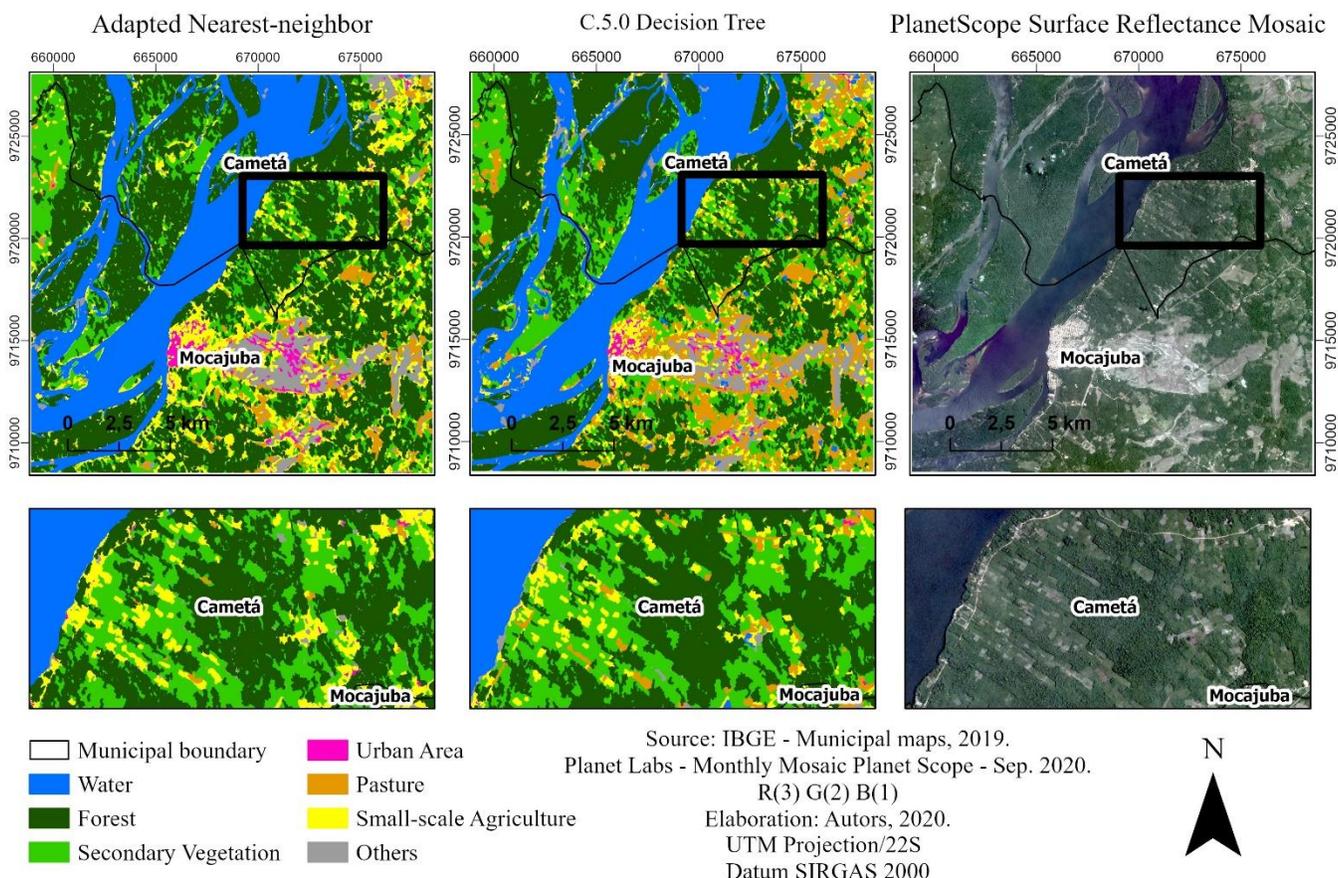


Figure 2. Land-Use and Land-Cover classification using NN and C5.0 Decision trees methods to identify small-scale agriculture.

ACKNOWLEDGMENT

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001" and in part by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). We thank the Brazilian National Institute for Space Research (INPE) for supporting the authors. Finally, we thank all anonymous reviewers for constructive suggestions on an earlier version of this paper.

REFERENCES

[1] J. Berdegué and R. Fuentealba, "Latin America: the state of smallholders in agriculture," IFAD Conference on New Directions for Smallholder Agriculture. Jan, 2011, pp. 1-37. doi: 10.1093/acprof:oso/9780199689347.003.0005.

[2] C. Caetano, "Family farming, the struggle for land in the legal Amazon and the impacts of Law 13.465/2017," "Agricultura Familiar, a luta pela terra na Amazônia Legal e os Impactos da Lei 13.465/2017," Revista de Direito Agrário e Agroambiental, vol. 4, pp. 34–55, Jun. 2018.

[3] C. Almeida et al., "High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data," Acta Amazonica, v. 46, pp. 291 – 302, Sep, 2016, doi:10.1590/1809-4392201505504.

[4] C. Souza et. al., "Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat Archive and Earth Engine," Remote Sensing, v. 12, issue 17, p. 2735, Aug. 2020, doi:10.3390/rs12172735.

[5] A. R. Souza, Economy and nature: land use and land cover patterns associated with agricultural and extractive activities in communities in the Southwest of Pará, Economia e natureza: padrões de uso e cobertura da terra associados a atividades agropecuárias e extrativistas de comunidades do Sudoeste do Pará, Brazilian National Institute for Space Research, 2016. [Online]. Available from: <http://urlib.net/8JMKD3MGP3W34P/3M8SD78>. 2021.06.13.

[6] A. R. Souza, M. I. S. Escada, R. Marujo, and A. M. Monteiro, "Cartography of the invisible: unveiling small-scale agriculture with Rapideye Images in the Baixo Tocantins Region, Pará," "Cartografia do Invisível: Revelando a Agricultura de Pequena Escala com Imagens Rapideye na Região do Baixo Tocantins, Pará," Revista do Departamento de Geografia, v. 38, pp. 137–153, Dec. 2019, doi:10.11606/rdg.v38i1.151603.

[7] M. Reis, M. I. S. Escada, S. Sant’anna, and L. Dutra, "Land use and land cover trajectory classification and analysis methods in the amazon: implications for forest regeneration studies," Revista Brasileira de Cartografia, v. 72, p. 1114, Dec. 2020, dx.doi.org/10.14393/rbcv72nespecial50anos-56535

[8] M. Rodrigues, H. Bendini, H. A. Soares, T. S. Körting, and L. Fonseca, "Remote sensing image time series metrics for distinction between pasture and croplands using the Random Forest Classifier," 2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS), Mar. 2020, pp. 149-154, ISSN: 978-1-7281-4350-7, doi: 10.1109/LAGIRS48042.2020.9165671.

[9] E. M. L. M. Novo, Remote sensing: principle and applications, Sensoriamento remoto: princípios e aplicações, 4th ed., São Paulo: Blucher, 2010.

- [10] T. Blaschke, "Object based image analysis for remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, n. 1, pp. 2–16, 2010. [10.1016/j.isprsjprs.2009.06.004](https://doi.org/10.1016/j.isprsjprs.2009.06.004).
- [11] G. J. Hay and G. Castilla, "Object-Based Image Analysis: Strengths, Weaknesses, Opportunities and Threats (SWOT)," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVI-4/C42, pp. 4–5, 2006.
- [12] O. Morariu And F. Burescu, "Agricultural land cover classification using RapidEye satellite imagery," *Journal of Young Scientist*, vol. 6, pp. 115-122, 2018, ISSN 2284-8012.
- [13] J. Delrue et al., "Crop mapping in countries with small-scale farming: a case study for West Shewa, Ethiopia," *International Journal of Remote Sensing*, vol. 34, pp. 2566-2582, Dec. 2012, <https://doi.org/10.1080/01431161.2012.747016>.
- [14] M.F.A. Vogels, S.M. Jong, G. Sterk, and E.A. Addink, "Mapping irrigated agriculture in complex landscapes using SPOT6 imagery and object-based image analysis – A case study in the Central Rift Valley, Ethiopia," *Int J Appl Earth Obs Geoinformation*, vol. 75, pp. 118-129, Mar. 2019, <https://doi.org/10.1016/j.jag.2018.07.019>.
- [15] L. P. Dutrieux, C. C. Jakovac, S. H. Latifah, and L. Kooistra, "Reconstructing land use history from Landsat time-series: Case study of a swidden agriculture system in Brazil," *International Journal of Applied Earth Observation and Geoinformation*, vol. 47, pp.112-124, Nov. 2015, <http://dx.doi.org/10.1016/j.jag.2015.11.018>.
- [16] Y. Inoue et al., "Traceability of slash-and-burn land-use history using optical satellite sensor imagery: a basis for chronosequential assessment of ecosystem carbon stock in Laos," *International Journal of Remote Sensing*, vol. 28, n. 24, Dec. 2007, doi: 10.1080/01431160701656323.
- [17] H. T. T. Nguyen, T. M. Doan, E. Tomppo, and R. E. McRoberts, "Land Use/Land Cover Mapping Using Multitemporal Sentinel-2 Imagery and Four Classification Methods—A Case Study from Dak Nong, Vietnam," *Remote sensing*, vol. 12, n. 9, pp. 1-27, doi:10.3390/rs12091367.
- [18] M. Jain, P. Mondal, R. S. DeFries, C. Small, and G. L. Galford, "Mapping cropping intensity of smallholder farms: A comparison of methods using multiple sensors," *Remote Sensing of Environment*, vol. 134, pp. 210-223, Jul. 2013, <http://dx.doi.org/10.1016/j.rse.2013.02.029>.
- [19] B. Tapsall, P. Milenov, and K. Tasdemir, "Analysis of RapidEye imagery for annual landcover mapping as an aid to European Union (eu) common agricultural policy," *ISPRS TC VII Symposium (100 Years ISPRS)*, Jul. 2010, pp. 568-573.
- [20] S. Adelabu, O. Mutanga, E. Adam, "Evaluating the impact of red-edge band from Rapideye image for classifying insect defoliation levels," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 95, pp. 34–41, Sep. 2014, <https://doi.org/10.1016/j.isprsjprs.2014.05.013>.
- [21] H. O. Kim, J. M. Yeom, "Effect of red-edge and texture features for object-based paddy rice crop classification using RapidEye multi-spectral satellite image data," *International Journal of Remote Sensing*, vol. 35, n. 19, pp. 7046–7068, Jan. 2014, <https://doi.org/10.1080/01431161.2014.965285>.
- [22] J. González-Piqueras et al., "Monitoring crop N status by using red edge-based indices," *Advances in Animal Biosciences*, vol. 8, n. 2, pp. 338–342, Jul. 2017, doi:10.1017/S2040470017000243.
- [23] Norway's International Climate and Forest Initiative. *NICFI Data Program*. 2020. [Online]. Available from: assets.planet.com/docs/NICFI_General_FAQs.pdf. 2021.06.13.
- [24] I. M. C. C. Cordeiro, L. G. T. Rangel-Vasconcelos, G. Schwartz, and F. A. Costa, Northeast of Pará: overview and sustainable use of secondary forests, Nordeste Paraense: panorama geral e uso sustentável das florestas secundárias. Belém: EDUFRA, 2017. ISBN 978-85-7295-118-0.
- [25] Brazilian National Institute of Geography and Statistics. *Agricultural census. Censo agropecuário*. [Online]. Available from: sidra.ibge.gov.br/pesquisa/censo-agropecuário/censo-agropecuário-2017. 2021.06.13.
- [26] R. Carvalho, M. Adami, S. Amaral, F. Bezerra, and A. P. Aguiar, "Changes in secondary vegetation dynamics in a context of decreasing deforestation rates in Pará, Brazilian Amazon," *Applied Geography*, v. 106, pp. 40-19, May, 2019, doi: 10.1016/j.apgeog.2019.03.001.
- [27] F. A. Costa, "Contributions of fallow lands in the Brazilian Amazon to CO2 balance, deforestation and the agrarian economy: Inequalities among competing land use trajectories," *Elementa Science of the Anthropocene*, v. 4, pp. 1-24, Oct. 2016, doi: 10.12952/journal.elementa.000133.
- [28] A. R. Webb, *Statistical Pattern Recognition*, 2nd ed., Malvern: John Wiley & Sons, 2002.
- [29] J. Quinlan, *C4.5: programs for machine learning*, San Mateo, CA: Morgan Kaufmann, 1993.
- [30] Trimble Germany GmbH. *eCognition Documentation - Classification*. 2019. [Online]. Available from: http://docs.ecognition.com/v9.5.0/eCognition_documentation/User%20Guide%20Developer/6%20About%20Classification.htm. 2021.06.13.
- [31] M. Baatz and A. Schape, *Multiresolution Segmentation: An Optimization Approach for High Quality Multi-Scale Image Segmentation*. 2000. In: J. Strobl, T. Blaschke, and G. Griesbner, Eds., *Angewandte Geographische Informations-Verarbeitung, XII*, Wichmann Verlag, Karlsruhe, Germany, 12-23.
- [32] GeoDMA 2.0.1. São José dos Campos: Brazilian National Institute for Space Research (INPE), 2011. [Online]. Available from: <http://www.dpi.inpe.br/geodma>. 2021.06.13.
- [33] L. Maurano, M. I. S. Escada, and C. Renno, "Spatial patterns of deforestation and estimation of accuracy of PRODES maps for the Brazilian Legal Amazon," "Padrões espaciais de desmatamento e a estimativa da exatidão dos mapas do PRODES para Amazônia Legal Brasileira," *Ciência Florestal*, v. 29, issue 4, pp. 1763-1775, Dec. 2019, doi:10.5902/1980509834380.
- [34] T. S. Korting, *GeoDMA: a toolbox integrating data mining with object-based and multi-temporal analysis of satellite remotely sensed imagery*, Brazilian National Institute for Space Research, 2012. [Online]. Available from: urlib.net/8JMKD3MGP7W/3CCH86S. 2021.06.13.
- [35] G. Banko, *A review of assessing the accuracy of classifications of remotely sensed data and of methods including remote sensing data in forest inventory*, International Institute for Applied Systems Analysis, 1998. [Online]. Available from: <https://www.researchgate.net/publication/23738314>. 2021.06.13.
- [36] R. G. Congalton and K. Gree, *Assessing the accuracy of remotely sensed data: principles and practices*. 2nd ed., Boca Raton: CRC Press, 2009.
- [37] C. C. Jakovac, L. P. Dutrieux, L. Siti, M. Peña-Claros, and F. Bongers, "Spatial and temporal dynamics of shifting cultivation in the middle-Amazonas river: Expansion and intensification," *PLOS ONE*, v. 12, n. 7, pp. 1–15, 2017, doi.org/10.1371/journal.pone.0181092.

Parametrising a Model of Clinopyroxene/Melt Partition Coefficients for Sodium to Higher Upper Mantle Pressures

Julia Marleen Schmidt
Institute of Geological Sciences)
Freie Universität Berlin
 Berlin, Germany
 email: julia.schmidt@fu-berlin.de

Lena Noack
Institute of Geological Sciences
Freie Universität Berlin
 Berlin, Germany
 email: lena.noack@fu-berlin.de

Abstract—Mineral/melt partition coefficients describe the redistribution of trace elements during melting processes. They are highly influenced by changing pressure, temperature, and composition. We find that for sodium, partition coefficients rise in the order of two magnitudes from 0-15 GPa along the peridotite solidus temperature. Still, because of lacking high-pressure experimental data and the limited pressure range of most partitioning models, mantle and crust evolution models normally rely on constant partition coefficients. Based on an earlier model, we created a thermodynamic model which calculates sodium partition coefficients between clinopyroxene and anhydrous silicate melts. With the results, we parametrise an equation over 0-15 GPa and various melt fractions. Our new parametrised partitioning model for sodium complements previous models and is applicable up to the mantle transition zone (4-12 GPa/1850-2360 K). Knowing the sodium partitioning behavior enables us to base further calculations for many other elements on it.

Keywords—partition coefficients; thermodynamic modelling

I. INTRODUCTION

Inside the Earth and other terrestrial planets, partial melts usually control the redistribution of trace elements from the mantle to the crust. If melt is buoyant due to a lower density than the surrounding rock, it moves upwards and transports the previously incorporated incompatible trace elements towards the surface. This leads to enriched reservoirs near the surface and depleted upper mantle areas. Expanding the redistribution concept to a certain number of elements leads to the creation of trace element patterns, which are unique geochemical markers of the events that led to the exact appearance and compositions of a rock of interest [1][2][3].

During melting, partition coefficients describe if an element prefers to be incorporated into a mineral or, if it is incompatible, into melt. Partitioning data for peridotite including the minerals olivine, orthopyroxene, and clinopyroxene show that in mantle rocks, clinopyroxene is the mineral taking in most incompatible trace elements [4][5]. This makes clinopyroxene the mineral with the largest influence on trace element redistribution.

Since partition coefficients of sodium between clinopyroxene and silicate melt are highly affected by changing pressure

and temperature, partition coefficient models should consider these parameters. Because of analytical limitations, there are only few measured partition coefficients available for high-pressure conditions. Additionally, partition coefficient models are often only applicable for rather low pressures. This is the reason why so far, partial melt simulations in some mantle evolution models either neglected partition coefficients and chose pre-defined elemental abundances in mantle melts [6], or settled for constant experimentally derived and estimated partition coefficients [7][8][9][10].

In 1937, Goldschmidt postulated that both the matching size and charge of the mineral's lattice site and element of interest determine an element's ability to partition into a mineral [11]. This leads to the assumption that not only it is possible to determine the P-T sensitive partition coefficients experimentally, but also by means of numerical modelling. Based on a model of Brice (1975) [12], Blundy et al. (1995) [13] developed a quantitative model and determined a parametrised fit function for clinopyroxene/melt partition coefficients for sodium in the range of 0-4 GPa and 1000-1800°C.

As is shown in Figure 1, partition coefficients of trace elements have a parabolic relationship in an Onuma diagram. The parabolas' curves change in broadness and shift along the x-axis depending on the bulk modulus E and the ideal lattice site radius r_0 , respectively. Because of this relationship, it is possible to calculate other trace elements' partition coefficients based on the reference coefficient D_{Na} .

The aim of this study is to implement a thermodynamic model for $D_{Na}^{cpx/melt}$ for a large range of pressures suitable for mantle melt simulations. Furthermore, we parametrise a fit function from the mentioned model that is applicable for upper mantle pressures. This will open the possibility to include partition coefficients depending on pressure and temperature into mantle evolution models and to acquire more realistic model results [14].

The rest of the paper is organized as follows. In section II, we describe the methods used for our parametrisation and thermodynamic model in the appropriate subsections. Section III

is dedicated to the results for the thermodynamic model and parametrisation respectively. Section IV discusses the results and compares them to experimental data. Section V concludes the paper, summarizes the findings, and points out what can be done in future works.

II. METHODS

A. Thermodynamic Model

To obtain the most realistic trace element partitioning results for mantle melting processes, it is useful to determine a strain-compensated partition coefficient D_0 [15]. Strain-compensation means that the redistributed element has the same charge and size as the regular element on the lattice site of the crystal [12]. Therefore, it is assumed that the radius of the element of interest r_i equals the lattice site radius r_0 . This way the partition coefficient D_i can be calculated with the following equation:

$$D_i = D_0 \exp \left(\frac{-4\pi E_{M2} N_A \left(\frac{r_0}{2} (r_i - r_0)^2 + \frac{1}{3} (r_i - r_0)^3 \right)}{RT} \right), \quad (1)$$

where E_{M2} is the bulk modulus of the M2 lattice site, N_A the Avogadro's number, R the gas constant and T the temperature. Here, the lattice site of interest is the crystallographic M2 site of clinopyroxene. Sodium is used as the strain-compensated partition coefficient D_0 because it is the 1+ charge element that has a radius r_i which is closest to the M2 lattice site radius r_0 . Therefore, we assume $D_0 = D_{Na}$.

To calculate the strain-compensated partition coefficient D_0 , we took a thermodynamic approach described by Blundy et al. (1995) [13]. For this, we used the melting curve of jadeite and linked it to the activity of jadeite's components in the melt. To calculate the partition coefficients of sodium between clinopyroxene and silicate melt, jadeite was the clinopyroxene of choice because of both, its high concentration in sodium and its ability to mix nearly ideally with diopside [16] and enstatite.

Generally, partition coefficients are given as the weight fraction ratio

$$D_i = \frac{X_i^{crystal}}{X_i^{melt}}, \quad (2)$$

with X_i being the weight percentage of the given component. However, since we obtain partition coefficients with the help of thermodynamic properties, we calculate molar partition coefficients. Based on experiments on plagioclase/fluid partitioning [17], it is broadly assumed that also for other minerals molar partition coefficients nearly equal weight percentage coefficients [18]. Thus, we can make use of Flood's equation for exchange equilibria in molten salts [19]:

$$RT \ln K_f = \sum N_i \Delta G^0, \quad (3)$$

where K_f is the molar equilibrium constant, N_i the concentration of the component and ΔG^0 the Gibbs' free energy of change. Because we use a thermodynamic approach for our calculations, N_i can be neglected and is set to one. The molar

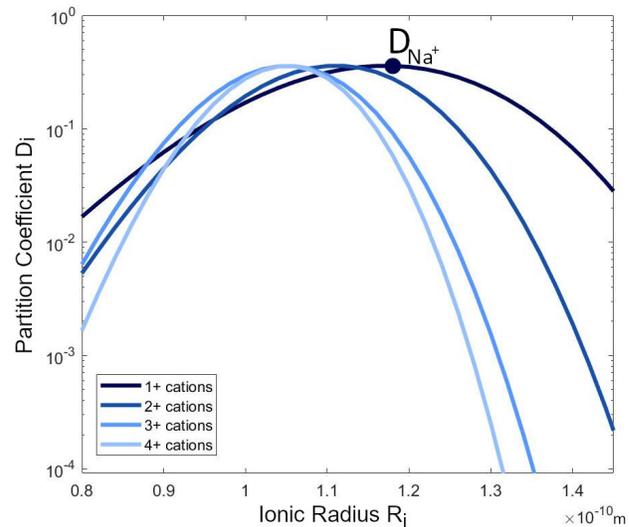


Figure 1. Partition coefficients of 1+ to 4+ charged cations in the clinopyroxene M2 lattice site. Modelled after Wood et al. (1997) [18]. Here, D_{Na} equals D_0 .

partitioning of jadeite D^{jd} can be expressed by using K_f in the expression

$$D_{jd} = \frac{1}{K_f} = \frac{a_{jd}^{cpx}}{a_{jd}^{melt}}, \quad (4)$$

where a_{jd}^{cpx} and a_{jd}^{melt} are the mole of the solid clinopyroxene and molten components, respectively. Here, we assume an ideal case where the activity equals the mole fraction of the elements of interest. Thus, we assume $a_{jd}^{cpx} = X_{Na}^{cpx}$ and $a_{jd}^{melt} = X_{Na}^{melt}$. Taking (4) and rearranging (3) leads to the following equation:

$$D_{Na}^{cpx/melt} = \exp \left(\frac{\Delta G_{f(P,T)}}{RT} \right). \quad (5)$$

Besides (5), one more equation is needed to determine D_{Na} thermodynamically. Equation (6) relates the entropy of the melting reaction to its heat capacity and melting point T_f of jadeite at the pressure $P_f=0.0001$ GPa. By including the enthalpy, we determine the Gibbs' free energy change of reaction with the following equation:

$$\begin{aligned} \Delta G_{(f(P,T))} = & \Delta H_{f(P_f,T_f)} + \int_{T_f}^T \Delta C_p dT \\ & - T (\Delta S_{f(P_f,T_f)} \\ & + \int_{T_f}^T \frac{\Delta C_p}{T} dT) + \int_{P_f}^P \Delta V_{f(T)} dP, \quad (6) \end{aligned}$$

where $\Delta H_{(f(P_f,T_f))}$ is the enthalpy, $\Delta S_{f(P_f,T_f)}$ the entropy and ΔC_p the difference between the local P,T values and P_f, T_f . $\Delta V_{f(T)}$ is the molar volume of fusion at pressure P_f extrapolated to the temperature of interest [13]. We calculate D_{Na} by inserting (6) into (5).

To be able to solve (6), we use the Maier-Kelly power function [20] to calculate the heat capacity C_p (as stated in

Table I). With the given parameters we can solve the first two integrals in (6) analytically, by evaluating the heat capacity of the solid at temperature T and the melt heat capacity at T_f . To solve the third part of the integral in (6) we use a Riemann integral, where we approximate the intervals on subintervals defined over a pressure vector of $n=1000$ steps from $P=0.0001$ to 15 GPa and assume that the volumes of melt and solid are constant in each subinterval.

$$\begin{aligned} \int_{P_f}^P \Delta V_{f(T)} dP &= \sum_{i=1}^{n-1} \int_{P_i}^{P_{i+1}} \Delta V_{f(T),i} dP \\ &\approx \sum_{i=1}^{n-1} \Delta V_{f(T),i} \int_{P_i}^{P_{i+1}} dP \\ &= \sum_{i=1}^{n-1} \Delta V_{f(T),i} (P_{i+1} - P_i) \end{aligned} \quad (7)$$

where $\Delta V_{f(T),i}$ is the volume of fusion, i.e., the volume difference between melt and solids. The calculation of V_{solid} and V_{melt} is done separately using an isothermal Birch-Murnaghan equation of state:

$$P = 1.5K_T \left(x^{\frac{7}{3}} - x^{\frac{5}{3}} \right) \left(1 + 0.75(K' - 4) \left(x^{\frac{2}{3}} - 1 \right) \right), \quad (8)$$

where $x = V^0/V$. K_T is the bulk modulus for pyroxene and K' its derivative. The volumes further depend on temperature via

$$V(T) = V \cdot \exp \left(\int_{T_0}^T \alpha(T) dT \right). \quad (9)$$

Here, we assume that the thermal expansion coefficient α is constant for melts and solids (see Table I), hence

$$\begin{aligned} \Delta V_{f(T),i} &= \\ &V_{melt}(P_i) \cdot \exp(\alpha_{melt}^0(T - T_f)) \\ &- V_{solid}(P_i) \cdot \exp(\alpha_{solid}^0(T - T_0)). \end{aligned} \quad (10)$$

T_0 is set to 298 K. The partition coefficient D_{Na} is then derived for varied pressures and temperatures by inserting ΔG into (5).

B. Fit Function Development

The thermodynamic model indicates that partition coefficients are not only dependent on pressure, but temperatures and therefore melt fractions as well. Thus, to develop a mantle melt fit function for the reference coefficient D_{Na} , we calculated partition coefficients between 0 and 15 GPa along the peridotite solidus temperature, peridotite liquidus temperature, and corresponding melt fraction F temperatures by using the thermodynamic approach described above. Melt fraction F is included because as is suggested by the batch melting equation [21], a rising melt fraction F leads to a decrease in the total amount of incompatible trace elements in the melt. Therefore, we arranged the melt fractions in $F = 0.2$

TABLE I
THERMODYNAMIC INPUT DATA FOR LIQUID AND SOLID JADEITE AND COMPARISON OF MODEL PARAMETERS.

Parameters	Pyroxene	Melt	Units
V^0	60.4	79.9	kJ GPa^{-1}
α^0	$2.81 \cdot 10^{-5}$	$6.28 \cdot 10^{-5}$	K^{-1}
K_T	141.2 ^b	14.8	GPa
$\partial K/\partial T$	-0.025	-0.0015	GPa K^{-1}
K'	4.5	4.5	
$\Delta H_{f(0.1, T_f)}$	61.1 ± 1.3		kJ mol^{-1}
$\Delta S_{f(0.1, T_f)}$	64.8 ± 0.6		J mol^{-1}
T_f	943 ± 22		K
C_p			
1	0.30113	0.28995 ^a	$\text{kJ mol}^{-1} \text{K}^{-1}$
2	$1.0143 \cdot 10^{-5}$		$\text{kJ mol}^{-1} \text{K}^{-1}$
3	-2239.3		$\text{kJ mol}^{-1} \text{K}^{-1}$
4	-2.0551		$\text{kJ mol}^{-1} \text{K}^{-1}$

Parameters as in Blundy et al. (1995) [13] at 0.0001 GPa and 298 K (unless stated otherwise).

^a At temperatures >1200 K.

* $C_p = C_{p1} + C_{p2} \cdot T + C_{p3} \cdot T^{-2} + C_{p4} \cdot T^{-0.5}$ (T in kelvins)

increments while neglecting the extraction of the melt, by using

$$F(P, T) = \frac{T - T(P)_{sol}}{T(P)_{liq} - T(P)_{sol}}. \quad (11)$$

The solidus T_{sol} and liquidus temperatures T_{liq} equations which we used [22] are third-order fits to experimental data [23][24]:

$$\begin{aligned} T_{sol,ini} &= \\ &1409K + 134.2 \frac{K}{GPa} \cdot P \\ &- 6.581 \frac{K}{GPa^2} \cdot P^2 + 0.1054 \frac{K}{GPa^3} \cdot P^3, \end{aligned} \quad (12)$$

$$\begin{aligned} T_{liq,ini} &= \\ &2035K + 57.46 \frac{K}{GPa} \cdot P \\ &- 3.487 \frac{K}{GPa^2} \cdot P^2 + 0.0769 \frac{K}{GPa^3} \cdot P^3. \end{aligned} \quad (13)$$

For pressures from 0-15 GPa, the solidus temperatures range from 1409-2297 K, while the liquidus temperatures go from 2035-2372 K. The pressure step size for each temperature profile is 0.1 GPa, which corresponds to 151 D_{Na} -P-T data points for each melt temperature profile. With the least square function, these data points were fitted to a function with parameters a, b, c, d, e, and f (14) in Python 3. Because the function has the same form as the fit function of Blundy et al. (1995) [13], their fit parameters are compared in Table II.

$$\begin{aligned} D_{Na}(T[\text{K}], P[\text{GPa}]) &= \\ &\exp \left(\frac{a + b \cdot P - c \cdot P^2}{T} - d + e \cdot P - f \cdot P^2 \right). \end{aligned} \quad (14)$$

By including the melt fraction P-T profiles, we make sure that the resulting model function will satisfy a broader range of P-T conditions and is not only valid for modelling partial melting in the mantle (where temperatures are close to the solidus), but also for crystallisation of melt (where temperatures are close to the liquidus). In contrast to our model, partition coefficients for crystallising liquid were often determined by taking only the liquidus temperatures into account [25][26][27].

For both, the thermodynamic model and the fit function development, replication data is accessible in the TRR170-DB [28].

III. RESULTS

A. Thermodynamic Model Along Melting Temperature Profiles

In a pressure range between 0 and 15 GPa and a temperature range between 1400 and 2400 K, the thermodynamic model produces increasing partition coefficients from low P/high T (0 GPa/2400 K) to high P/low T (10-12 GPa/1400 K) conditions (Figure. 2). The increase is in the order of four magnitudes and coincides with experimental observations [29][30]. At higher pressure above 12 GPa and low temperatures, the model starts to invert the trend and the partition coefficients decrease with increasing pressure. However, for our fit function development we have used the P-T space between the solidus and liquidus, which is not affected by this inversion. Along the melting temperature profiles, increasing pressure always leads to increasing partition coefficients, while increasing temperatures cause the partition coefficients to decrease (Figure 2).

In Table II, we compare the thermodynamic model results with experimental literature data. It is notable that most of the experimental data fit well to the model results. On average, the coefficients calculated by the thermodynamic model deviate from the experimental data by 26%. The best fitting value deviates only 2.9% from the experiments at 0.0001 GPa and 1526 K, while the worst fit deviates by 46.9% at 0.0001 GPa and 1524 K. The implications of this variance will be discussed in section IV.

B. Parametrisation of Thermodynamic Model Results

As described in section II-B, we developed a scaled equation along the mantle peridotite solidus, liquidus, and corresponding melt fraction temperatures in between. For this, we used the least square function to fit the thermodynamic model results presented in Figure 2 to the following equation:

$$D_{Na}(T[K],P[GPa]) = \exp\left(\frac{2183 + 2517P - 157P^2}{T} - 4.575 - 0.5149P + 0.0475P^2\right). \quad (15)$$

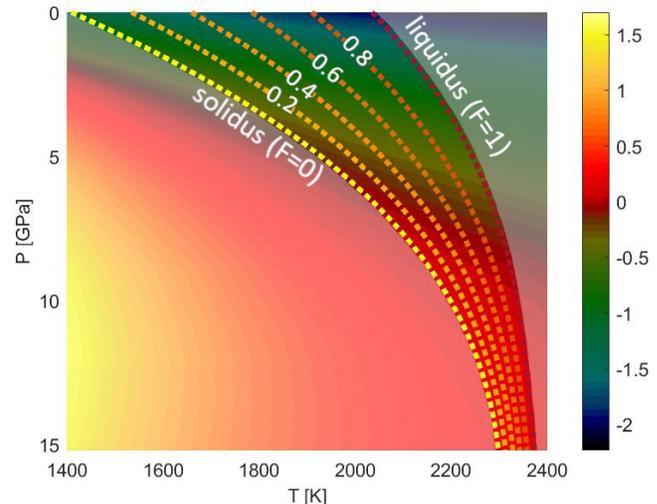


Figure 2. Sodium in clinopyroxene partition coefficients in P-T space, calculated thermodynamically with the methods of Blundy et al (1995) [13]. Solidus 12, liquidus (13) and intermediate melt fraction outlines are calculated from de Smet (1999) [22].

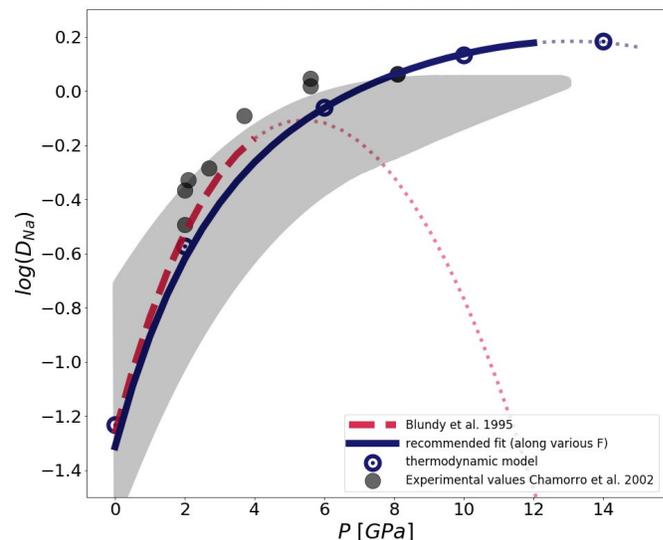


Figure 3. D_{Na} fits along the solidus from Blundy et al. (1995) [13] (red, dashed line) and this study (blue, solid line) compared to experimental data [30] (shaded grey area and grey dots) for comparison. Thin dotted lines are extrapolations of the fit functions beyond range of validity.

With the new resulting scaling law (15), it is now possible to calculate partition coefficients for sodium at varied temperature and pressure conditions.

Figure 3 illustrates how the rising D_{Na} values correlate with rising pressure along the solidus temperature. In contrast to the fit function of Blundy et al. (1995) [13], our resulting function produces steadily rising values up to 13.2 GPa at the respective peridotite solidus temperatures. After this point, the calculated values are starting to decrease. In contrast to our study, the model values of Blundy et al. (1995) [13]

TABLE II
COMPARISON OF EXPERIMENTAL DATA FOR CLINOPYROXENE/MELT PARTITION COEFFICIENTS WITH THE THERMODYNAMIC DATA

P [GPa]	T [K]	D_{Na} (experimental data)	Ref.	therm. model (this study)	Fit values Blundy et al. (1995) [13]	by Fit (this study)
0.0001	1370	0.050(4)	[13]	0.0696	0.0670	0.0507*
0.0001	1449	0.042(4)	[13]	0.0496	0.0444	0.0465*
0.0001	1524	0.070(19)	[13]	0.0372	0.0312	0.0432*
0.0001	1526	0.040(4)	[13]	0.0369	0.0309	0.0431*
0.0001	1526	0.038(5)	[13]	0.0369	0.0309	0.0431*
0.0001	1598	0.046(5)	[13]	0.0288	0.0228	0.0404*
1	1663	0.075(7)	[13]	0.0891	0.0794	0.0991*
2	1773	0.113(8)	[13]	0.1708	0.1680	0.1827*
2	1843	0.144(13)	[13]	0.1363	0.1247	0.1587*
3	1938	0.237(30)	[13]	0.2241	0.2170	0.2465*
6	2038	0.52(12)	[13]	0.7789	0.6507*	0.7774
1.2	1588	0.225(5)	[31]	0.1497	0.1494	0.1366*
1.2	1458	0.221(5)	[31]	0.2671	0.3039	0.1807*
1.6	1643	0.283(4)	[31]	0.1866	0.1927	0.1750*

Parameter	Fit [13] 0-4 GPa	Fit (this study) 4-12 GPa
a	10367	2183
b	2100	2517
c	165	-157
d	-10.27	-4.575
e	0.358	-0.5149

D_{Na} is the experimentally determined weight partition coefficient.

*Note that these are extrapolated values (beyond valid P-T range).

Parameters to be inserted into (14).

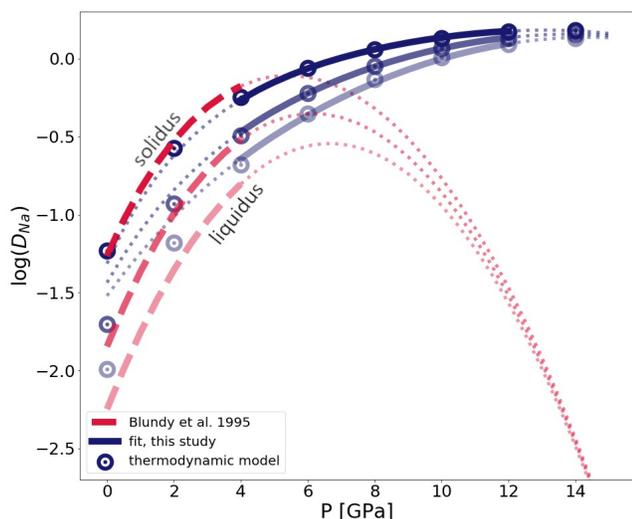


Figure 4. Partition coefficients depending on melt fractions. The line in between liquidus and solidus indicates a melt fraction $F=0.5$. Thin dotted lines are extrapolations of the fit functions beyond range of validity.

start to decline after approximately 5 GPa along the solidus temperature. However, Blundy et al. (1995) [13] themselves state that their scaling law is not to be used over a pressure of 4 GPa.

In Figure 4, the decrease of partition coefficients from the

solidus towards the liquidus becomes visible. This indicates that not only pressure, but also temperature and therefore melt fractions have an impact on the redistribution. It should be noted that at low pressures and high melt fractions, our fit seems to divert more from the thermodynamic model than the previous fit [13]. Consequences of this diversion on the applicability of our scaling law (15) are explored in section IV.

IV. DISCUSSION

Partition coefficients are highly controlled by pressure, temperature, and composition. Since it is not possible for models to include all compositional interactions between a mineral and melt, comparing them to experimental data is sometimes difficult. This is especially true for higher pressures, where partition coefficient data is, to this date, often lacking. However, the existing experimental data for lower pressures already indicate if we can expect realistic results from our calculations. As is shown in Table II, our thermodynamic model results fit well to the experimental results [13][31]. Best matches were achieved for experimental data at 0.0001 GPa/1526 K, 2 GPa/1843 K and 3 GPa/1938 K [13], where the model data deviates from the experimental values by only 2.9% and for the latter two by 5.4%, respectively. Interestingly, with 46.9%, the second largest diversion from the experimental data is at 0.0001 GPa/1524 K, which is at the same pressure and only 2 K below a very well fitting value (Table II). However, the experimentally derived partition coefficients vary more than what could be solely explained by changing P-T conditions. Deviating experimental and thermodynamic model results can happen

because of several effects. For example, analytical error margins have to be considered. Furthermore, the composition of melt as well as solid clinopyroxene can be much more diverse in experiments than it is for thermodynamic model calculations, which does not take these changes into account.

It should be noted that the fit parameters in (14) differ from the parametrisation in Blundy et al. (1995) [13], which is due to the larger parameter space applied in our model. For a comparison of the two scaling laws, we refer to Figure 3 and Figure 4. A direct comparison of the resulting parameters for the scaling law of our and the previous study [13] is given in Table II.

Figure 3 illustrates that the partition coefficients calculated with our scaling law (15) over solidus temperatures is close to the experimental sodium partitioning data. This data is indicated by a shaded grey area and, for D_{Na} in the $Ab_{80}Di_{20}$ system, by darker grey dots [30]. A direct comparison with appropriate temperatures in the range of the experimental data [30] shows that here, the thermodynamic model is within an error of on average 32%.

As is shown in Figure 3 and Figure 4, our and the previous fit [13] do not completely agree with each other. Blundy et al. (1995) [13] parametrised their thermodynamically calculated results over a peridotite solidus of McKenzie and Bickle (1988) [32], while we took the peridotite solidus equation of de Smet (1999) [22]. This could have produced a small shift in the D_{Na} -P-T field and, therefore, in the resulting fit function. Also, a different parametrisation method could have been used by the previous study [13], which might have produced slightly different fits. Finally, the previous study [13] concentrated more on low pressure partitioning, whereas we have tried to parametrise the function up to the 410 km discontinuity. The inclusion of higher pressures could also be the reason why at pressures below 4 GPa (especially at higher melt fractions) our scaling law (15) fits less well to the thermodynamic model values than the older model [13].

Between 4 and 12 GPa, our scaling law (15) fits well to the thermodynamic data over all melt fraction values from solidus to liquidus. However, at high melt fractions below 4 GPa, larger diversions from the original fit function [13] appear (Figure 4). The reason for this lies in the nature of the least square parametrisation used in Python 3, which allows the fit function to divert from the thermodynamic model at low pressures for the expense of being applicable to higher pressures and to a wide range of melt fractions. If we would only parametrise the model over the solidus, the shift would disappear at the expense of the model being applicable to any other P-T conditions. Thus, the included P-T conditions for the varying melt fractions ensure that the model is useful over a wide range of pressures and temperatures, but limits the applicability of the scaling law for higher melt fractions to a range between 4 and 12 GPa.

In the upper mantle, buoyant melt can occur up to the depth where it becomes gravitationally stable and a density inversion of melt and solid surrounding material takes place (the so-called density crossover). In other words, melts

formed at higher pressures may not be able to rise to the surface [33][34]. However, this is only true for upper mantle melts and melts rising upwards in lower regions of the Earth can not be ruled out [35]. Inside Earth, the density crossover exists at approximately 11-12 GPa at 2000°C [36], and in Mars between 7 GPa [36] and 7.5 GPa [37]. To include pressure and temperature dependent partition coefficients into a mantle evolution model, it would often suffice to be able to calculate them up to the density crossover. Overlapping with the density crossover, in the Earth's 410 km mantle discontinuity (i.e., pressures of approximately 12-15 GPa), phase changes occur [38] and pyroxene slowly starts to dissolve into a pyrope-rich garnet to form majorite [39]. This and the density crossover indicate that a parametrisation up to approximately 12-15 GPa is sufficient.

As the experimental data [30][29] suggests, the partition coefficients of sodium in clinopyroxene increase with temperature and pressure before they remain constant. At solidus P-T conditions, our fit function curve slowly starts to flatten and starts to fall at approximately 13 GPa (Figure 3). This coincides with the thermodynamic model, where (along the solidus) the partition coefficients start to decrease at 12.5 GPa. Combining these findings with the occurrence of a density crossover and transition zone at approximately 12 GPa, we suggest to not use our scaling law (15) above this pressure.

As is discussed above, the scaling law works well for melting P-T conditions between the peridotite solidus and liquidus between 4 and 12 GPa. Thus, it can be considered as a useful expansion of the previous scaling law for D_{Na} [13]. Because of the broad P-T range, the model should not only be useful for mantle melting, but also for models which crystallise melt, as is the case in a magma ocean. However, as recent studies have suggested, the solidus and liquidus temperature may change heavily depending on if material is melting or crystallising [40]. If this is the case, our model could lie outside of the P-T range between solidus and liquidus for crystallising liquids and would have to be extrapolated. Therefore, one has to be careful with the usage of the fit.

Like the thermodynamic model, phase transitions and density crossover behavior all depend on pressure. Thus, even for planets with a radius or mass different from Earth, the fit should be applicable if the mantle composition is comparable to Earth. Additionally, the composition of the terrestrial planet has to be taken into account. If there is no or if there are only very minor portions of clinopyroxene in the planets upper mantle, our partition coefficient calculations for clinopyroxene cannot be used.

V. CONCLUSION AND FUTURE WORK

With the new high-pressure scaling law (15), it is now possible to include partition coefficient models depending on pressure and temperature into mantle convection models for the entire pressure range over which upper mantle melts are

buoyant. The newly developed fit function can be used to calculate clinopyroxene/melt redistribution behavior of sodium starting from 4 GPa up to the mantle transition zone of the Earth. This is in contrast to the scaling law by Blundy et al. (1995) [13], which can be applied from 0 to 4 GPa. Our new scaling law can be used as the basis for calculating D_i in (1), with $D_0 = D_{Na}$ for 1+ charge elements. Possible approaches to model partition coefficients based on this scaling law for the charges 2+ to 4+ are described in Wood and Blundy (2014) [41] and are based on adjusted calculations for the mineral's lattice site radius r_0 and the bulk modulus E .

Compared to the existing experimental data, our scaling law (15) allows for a good approximation of clinopyroxene partition coefficients of trace elements between solid and melt. This enables us to do self-consistent calculations of local partition coefficients for variable pressures and temperatures. Because we have parametrised our model over a wide range of P-T conditions and melt fractions between the peridotite solidus and liquidus, our model can be applied for any $D_{Na}^{cpa/melt}$ calculation between 1850-2360 K and 4-12 GPa.

By combining our scaling law with the older scaling law [13], we will be able to calculate the redistribution behavior of trace elements in terrestrial planets in much more detail. Our partition coefficient calculations for clinopyroxene should be applicable as long as clinopyroxene is present in the planet's upper mantle in sufficient abundance.

Future works could not only include the application of our new scaling law (15) in numerical simulations, but also further investigations on partitioning behavior in mantle material as well. For instance, adding an orthopyroxene/melt trace element partitioning model to a mantle evolution model would provide an even more detailed tool to study on the trace element redistribution from mantle to crust if used alongside the clinopyroxene/melt partitioning model.

ACKNOWLEDGMENT

We would like to thank Jon Blundy and Timm John for fruitful discussions on partition coefficient modelling. This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 263649064 – TRR 170. This is TRR 170 Publication No. 130.

REFERENCES

- [1] H. Neumann, J. Mead, and C. Vitaliano "Trace element variation during fractional crystallization as calculated from the distribution law," *Geochimica et Cosmochimica Acta*, vol. 6 (2-3), pp. 90-99, February 1954.
- [2] P. W. Gast "Trace element fractionation and the origin of tholeiitic and alka-line magma types," *Geochimica et Cosmochimica Acta*, vol. 32 (10), pp. 1057-1086, October 1968.
- [3] B. B. Jensen "Patterns of trace element partitioning," *Geochimica et Cosmochimica Acta*, vol. 37 (10), pp. 2227-2242, October 1973.
- [4] K. T. Johnson, H. J. Dick, and N. Shimizu "Melting in the oceanic upper mantle: an ion microprobe study of diopsides in abyssal peridotites," *Journal of Geophysical Research: Solid Earth*, 95 (B3), 2661-2678, March 1990.
- [5] T. Skulski, W. Minarik, and E. B. Watson, E. B. "High-pressure experimental trace-element partitioning between clinopyroxene and basaltic melts," *Chemical Geology*, vol. 117 (1-4), pp. 127-147, November 1994.
- [6] U. Walzer, and R. Hendel "Mantle convection and evolution with growing continents" *Journal of Geophysical Research: Solid Earth*, vol. 113 (B9), September 2008.
- [7] A. Morschhauser, M. Grott, and D. Breuer "Crustal recycling, mantle dehydration, and the thermal evolution of Mars," *Icarus*, vol. 212 (2), pp. 541-558, April 2011.
- [8] T. Ruedas, P. J. Tackley, and S. C. Solomon "Thermal and compositional evolution of the martian mantle: Effects of phase transitions and melting," *Physics of the Earth and Planetary Interiors*, vol. 216 , pp. 32-58, March 2013.
- [9] A.-C. Plesa, and D. Breuer "Partial melting in one-plate planets: Implications for thermo-chemical and atmospheric evolution," *Planetary and Space Science*, vol-405 98 , pp. 50-65, August 2014.
- [10] R. E. Jones, P. E. van Keken, E. H. Hauri, J. M. Tucker, J. Vervoort, and C. J. Ballentine "Origins of the terrestrial Hf-Nd mantle array: evidence from a combined geodynamical-geochemical approach," *Earth and Planetary Science Letters*, vol. 518 , pp. 26-39, July 2019.
- [11] V. M. Goldschmidt "The principles of distribution of chemical elements in minerals and rocks," *Journal of the Chemical Society (Resumed)*, pp. 655-673, March 1937.
- [12] J. Brice "Some thermodynamic aspects of the growth of strained crystals," *Journal of Crystal Growth*, vol. 28 (2), pp. 249-253, March 1975.
- [13] J. Blundy, T. Falloon, B. Wood, and J. Dalton "Sodium partitioning between clinopyroxene and silicate melts," *Journal of Geophysical Research: Solid Earth*, vol. 340 100 (B8), pp. 15501-15515, August 1995.
- [14] J. M. Schmidt and L. Noack "Modelling clinopyroxene/melt partition coefficients for higher upper mantle pressures," *EGU General Assembly 2021*, online, EGU21-12478, <https://doi.org/10.5194/egusphere-egu21-12478>, April 2021.
- [15] P. McDade, J. D. Blundy, and B. J. Wood "Trace element partitioning on the Tinaquillo Lherzolite solidus at 1.5 GPa," *Physics of the Earth and Planetary Interiors*, vol. 139 (1-2), pp. 129-147, September 2003.
- [16] B. J. Wood, and J. Nicholls "The thermodynamic properties of reciprocal solid solutions," *Contributions to Mineralogy and Petrology*, vol. 66 (4), pp. 389-400, June 1978.
- [17] J. Blundy, and B. Wood "Crystal-chemical controls on the partitioning of Sr and Ba between plagioclase feldspar, silicate melts, and hydrothermal solutions," *Geochimica et Cosmochimica Acta*, vol. 55 (1), pp. 193-209, January 1991.
- [18] B. J. Wood, and J. D. Blundy "A predictive model for rare earth element partitioning between clinopyroxene

- and anhydrous silicate melt," *Contributions to Mineralogy and Petrology*, vol. 129 (2-3), pp. 166-181, October 1997.
- [19] K. Grjothheim, C. Krohn, and J. Toguri "Thermodynamic evaluation of activities in molten mixtures of reciprocal salt systems," *Transactions of the Faraday Society*, vol. 57, pp. 1949-1957, March 1961.
- [20] C. G. Maier, and K. Kelley "An equation for the representation of high-temperature heat content data," *Journal of the American chemical society*, vol. 54 (8), pp. 3243-3246, August 1932.
- [21] J. Hertogen, and R. Gijbels "Calculation of trace element fractionation during partial melting," *Geochimica et Cosmochimica Acta*, vol. 40 (3), pp. 313-322, March 1976.
- [22] J. De Smet, A. Van Den Berg, and N. Vlaar, N. "The evolution of continental roots in numerical thermo-chemical mantle convection models including differentiation by partial melting," In: *Developments in geotectonics*. Elsevier, vol. 24, pp. 153-170, 1999.
- [23] T. Gasparik "Phase relations in the transition zone," *Journal of Geophysical Research: Solid Earth*, vol. 95 (B10), pp. 15751-15769, September 1990.
- [24] E. Takahashi "Speculations on the Archean mantle: missing link between komatiite and depleted garnet peridotite," *Journal of Geophysical Research: Solid Earth*, vol. 95 (B10), pp. 15941-15954, September 1990.
- [25] D. C. Presnall, Y.-H. Weng, C. S. Milholland, and M. J. Walter "Liquidus phase relations in the system MgO-MgSiO₃ at pressures up to 25 GPa - constraints on crystallization of a molten Hadean mantle," *Physics of the Earth and Planetary Interiors*, vol. 107 (1-3), pp. 83-95, April 1998.
- [26] M. Walter, E. Nakamura, R. Tronnes, R., and D. Frost, D. "Experimental constraints on crystallization differentiation in a deep magma ocean," *Geochimica et Cosmochimica Acta*, vol. 68 (20), pp. 4267-4284, October 2004.
- [27] D. M. Shaw "Trace element fractionation during anatexis," *Geochimica et Cosmochimica Acta*, vol. 34 (2), pp. 237-243, February 1970.
- [28] J. M. Schmidt, and L. Noack "Replication Data for Parameterizing a Model of Clinopyroxene/Melt Partition Coefficients for Sodium to Higher Upper Mantle Pressures," TRR170-DB doi.org/10.35003/QVOL83 (V01), 2021.
- [29] W. Wang, and E. Takahashi "Subsolidus and melting experiments of a K-rich basaltic composition to 27 GPa: Implication for the behavior of potassium in the mantle," *American Mineralogist*, vol. 84 (3), pp. 357-361, November 1999.
- [30] E. Chamorro, R. Brooker, J. A. Wartho, B. Wood, S. Kelley, and J. Blundy "Ar and K partitioning between clinopyroxene and silicate melt to 8 GPa," *Geochimica et Cosmochimica Acta*, vol. 66 (3), pp. 507-519, February 2002.
- [31] G. A. Gaetani, A. J. Kent, T. L. Grove, I. D. Hutcheon, and E. M. Stolper "Mineral/melt partitioning of trace elements during hydrous peridotite partial melting," *Contributions to Mineralogy and Petrology*, vol. 145 (4), pp. 391-405, May 2003.
- [32] D. Mckenzie, and M. Bickle "The volume and composition of melt generated by extension of the lithosphere," *Journal of petrology*, vol. 29 (3), pp. 625-679, June 1988.
- [33] E. Stolper, D. Walker, B. H. Hager, and J. F. Hays "Melt segregation from partially molten source regions: the importance of melt density and source region size," *Journal of Geophysical Research: Solid Earth*, vol. 86 (B7), pp. 6261-6271, July 1981.
- [34] Z. Jing, and S.-i. Karato "The density of volatile bearing melts in the Earth's deep mantle: The role of chemical composition," *Chemical Geology*, vol. 262 (1-2), pp. 100-107, May 2009.
- [35] M. J. Beuchert and H. Schmeling "A melting model for the lowermost mantle using Clapeyron slopes derived from experimental data: Consequences for the thickness of ultralow velocity zones (ULVZs)," *Geochemistry, Geophysics, Geosystems*, vol. 14 (1), pp. 197-208, January 2013.
- [36] E. Ohtani, Y. Nagata, A. Suzuki, and T. Kato "Melting relations of peridotite and the density crossover in planetary mantles," *Chemical Geology*, vol. 402 120 (3-4), pp. 207-221, March 1995.
- [37] L. T. Elkins-Tanton, E. Parmentier, and P. Hess "Magma ocean fractional crystallization and cumulate overturn in terrestrial planets: Implications for Mars," *Meteoritics & Planetary Science*, vol. 38 (12), pp. 1753-1771, January 2003.
- [38] A. E. Ringwood, A. E. "Phase transformations and their bearing on the constitution and dynamics of the mantle," *Geochimica et Cosmochimica Acta*, vol. 55 (8), pp. 2083-2110, August 1991.
- [39] M. Akaogi "Phase transitions of minerals in the transition zone and upper part of the lower mantle," *Special Papers-Geological Society Of America*, vol. 421, p. 1, 2007.
- [40] C.-E. Boukare, Y. Ricard, and G. Fiquet "Thermodynamics of the MgO-FeO-SiO₂ system up to 140 GPa: Application to the crystallization of Earth's magma ocean," *Journal of Geophysical Research: Solid Earth*, vol. 120 (9), pp. 6085-6101, August 2015.
- [41] B. J. Wood, and J. D. Blundy, J. D. (2014) "Trace element partitioning: The influences of ionic radius, cation charge, pressure, and temperature," In H. D. Holland and K. K. Turekian (Eds.), *Treatise on geochemistry* (second edition) (Second Edition ed., p. 421 - 448). Oxford: Elsevier, 2014.

Methodology of Chorological and Coherent Conceptual Knowledge Contextualisation: Approaches for Multi-disciplinary Contexts in Prehistory and Archaeology

Claus-Peter Rückemann

Westfälische Wilhelms-Universität Münster (WWU), Germany;
 Unabhängiges Deutsches Institut für Multi-disziplinäre Forschung (DIMF), Germany;
 Leibniz Universität Hannover, Germany
 Email: ruckema@uni-muenster.de

Abstract—This paper presents research on a new methodology of general chorological and coherent conceptual knowledge contextualisation across disciplines. The presented achievements and scenarios concentrate on new information science approaches for multi-disciplinary contexts in prehistory and archaeology, targeting the use of inherent aspects and creation of new insight, strategies, and perspectives. The new advanced approach can be applied to any type of knowledge, e.g., factual, conceptual, procedural, metacognitive, and structural. The integration of chorological knowledge includes integration of spatial and geospatial knowledge and features. Goal of this research are new facilities for coherent contextualisation in multi-disciplinary contexts. Here, we are focussing on coherent knowledge in contexts with prehistory and archaeology disciplines, natural sciences, and advanced geoscientific information systems. The solution also integrates knowledge from satellite data and soil diversity and respective properties. The knowledge approach allows a multi-directional utilisation of coherent conceptual knowledge. Future research concentrates on further continuing the development of components and application in multi-disciplinary scenarios of prehistory and archaeology.

Keywords—Prehistory; Natural Sciences; Chorology; Information Science; Contextualisation; Coherent Conceptual Knowledge.

I. INTRODUCTION

These days, mostly all Geographic Information Systems and even more advanced and more complex Geoscientific Information Systems (GIS) are –by themselves– not yet taking multi-disciplinary coherent knowledge and contexts into consideration. From scientific point of view, it is a questionable approach to think of and practice a distinct discipline while considering any other required scientific discipline being auxiliary, which, for further simplification, may even be reduced to ‘data delivery’, ‘technical’, and ‘procedural’ contributions.

Multi-disciplinary scenarios often require to consider a wide range of contexts with disciplines put to their level. It is the coherent knowledge of contexts, which is most relevant for new insight. Therefore, contextualisation should not be done without considering multi-disciplinary coherency and expert views from different disciplines put on a par with respective further scientific collaboration and support.

Goal of this research are new information science facilities for a coherent conceptual contextualisation. The result of this research is a new methodology of general chorological and coherent conceptual knowledge contextualisation. The new advanced approach enables a systematical, coherent contextualisation and can be applied with knowledge complements, e.g., factual, conceptual, procedural, metacognitive, and structural. The approach is knowledge-centric, in a way “knowledge-driven” but explicitly not “development-procedure-driven” or “software-driven”. From knowledge complements’ point of view, chorology is for place what chronology is for time. The

integration of chorological knowledge includes integration of spatial and geospatial knowledge and features. Here, we are focussing on scenarios of coherent multi-disciplinary knowledge in context with prehistory and archaeology, natural sciences, and advanced geoscientific information system components. The solution also integrates knowledge from satellite data and soil diversity with respective properties. Contexts in prehistory are special in a way that there are no direct historical sources and respectively no literary reference and documentation. Contextualisation is therefore a main intrinsic task in prehistory and protohistory. An approach has to conform with information science fundamentals and universal knowledge and has to enable an integration of the required components from methodologies to realisations for knowledge representations of realia and abstract contexts [1] while many facets of knowledge, including prehistory, need to be continuously acquired and reviewed [2]. There is no published approach known, which can be reasonably compared with the implemented and presented method. Therefore, there was a strong need for an advanced methodology to contextualise knowledge, e.g., from practically available knowledge resources. This paper presents the methodological fundamentals for a chorological and coherent multi-disciplinary contextualisation. The potential for finding and integrating multi-disciplinary inherent aspects and creation of new insight, strategies, and perspectives by development of components employing coherent conceptual knowledge has been a major motivation.

The rest of this paper is organised as follows. Section II presents the method of component integration and the relevant components in ongoing development. Section III shows the solution, results of practical implementation and realisation, which have been realised for prehistory and archaeology contexts. Section IV demonstrates and discusses the resulting coherent conceptual knowledge integration. Section V summarises lessons learned, conclusions, and future work.

II. COMPONENT INTEGRATION AND METHOD

Commonly used tools are not aware of features for contextualisation from multi-disciplinary components. Therefore, advanced individual workflows need proper preparation of components and workflow procedures. Preparation requires methods for deployment of respective knowledge characteristics and properties. Further, any workflow should be created being aware of the individuality of these characteristics.

Many aspects of knowledge, including meaning, can be described using knowledge complements supporting a modern definition of knowledge and subsequent component instrumentation [3] [4], e.g., considering factual, conceptual, procedural, metacognitive, and structural knowledge. Especially, conceptual knowledge can relate to any of factual, conceptual, procedural, and structural knowledge. Knowledge complements

are a means of understanding and targeting new insight, e.g., enabling advanced contextualisation, integration, analysis, synthesis, innovation, prospecting, and documentation. The approach can be summarised based on the methodological fundamentals.

- Selection and development of a coherent, multi-disciplinary reference implementation. Knowledge complements are addressing reference implementations.
- Multi-disciplinary knowledge resources and integrated components are realised with knowledge-centric focus.
- Contextualisation employing knowledge complements.
- Analysis, synthesis, documentation can employ reference implementations for new insight and development.

An approach to the multi-disciplinary nature of this research requires significant developments of coherently integration-able, fundamental context components especially

- multi-disciplinary contexts of prehistory and archaeology and respective resources,
- chorological contexts, e.g.,
- homogeneously consistent high resolution Digital Elevation Model (DEM) for land and sea bottom, and
- natural sciences Knowledge Resources (KR), e.g., soil classification resources, standardised soil reference systems, and parameters.

The following passages give a compact overview of major component implementations and development integrated with this research. A more detailed, comprehensive discussion and examples regarding the fundamentals are available with the research on methodology, contextualisation, and conceptual knowledge. Relevant pre-existing and ongoing component developments addressing knowledge with multi-disciplinary KR have been summarised [5].

a) Conceptual knowledge frameworks: A main reference implementation developed and used in practice with ongoing long-term research and applied for KR is the prehistory-protohistory and archaeology Conceptual Knowledge Reference Implementation (CKRI), including multi-disciplinary contexts of natural sciences and humanities [6].

b) Conceptual knowledge base: Conceptual knowledge base is The Universal Decimal Classification (UDC) [7], a general plan for knowledge classification, providing an analytic-synthetic and faceted classification, designed for subject description and indexing of content of information resources irrespective of the carrier, form, format, and language. UDC-based references in this publication are taken from the multi-lingual UDC summary [7] released by the UDC Consortium under a Creative Commons license [8].

c) Integration of scientific reference frameworks: The integration includes relevant scientific practices, frameworks, and standards from disciplines and contexts. Geosciences and soil science are continuously delivering updated state of the art research and insight, including geodiversity and standardisation [9] [10] as required for contextualisation. A practical reference implementation for coherent contextualisation of prehistory-protohistory and archaeology conceptual knowledge [6] is currently in development within a long-term project accompanying this research.

d) Formalisation: All integration components, for all disciplines, require an explicit and continuous formalisation [11] process in order to conform with the information science principles according to the practices in the disciplines.

e) Methodologies and workflows integration: Methodologies for creating and utilising methods include model processing, remote sensing, spatial mapping, high information densities, and visualisation. The respective contextualisation of prehistoric scenarios should each be done under individual prehistoric conditions, especially supported by standard algorithms [12], multi-dimensional criteria, spatial operations, interpolation geodesic computation [13], triangulation [14], gradient computation [15], and projection [16].

f) Prehistory Knowledge Resources: In order to overcome basic shortcomings of public ‘data collections’ the objects, entities, and respective conceptual knowledge references’ excerpts and examples are taken from The Prehistory and Archaeology Knowledge Archive (PAKA). PAKA has been in continuous development for more than three decades [17] and is released by DIMF [18].

g) Natural Sciences Knowledge Resources: Several coherent systems of major natural sciences’ context object groups from KR realisations have been implemented [5] [7] [19].

h) Inherent representation groups: The methodology can consider a wide range of representation groups for major disciplines and context object groups regarding their inherent representation and common utilisation, e.g., points, polygons, lines, Digital Elevation Model (DEM) representations, z-value representations, distance representations, area representations, raster, vector, binary, and non-binary data.

i) Scientific context parametrisation: Scientific context parametrisation of prehistoric targets can use the overall insights, e.g., from geoscientific disciplines [20] [21]. A relevant example is contextualisation with palaeolandscapes [22].

j) Structures and symbolic representation standards: The deployment of long-term universal structure and data standards is essential. Relevant examples of sustainable implementations are NetCDF [23] based standards, including advanced features, hybrid structure integration, and parallel computing support (PnetCDF) and generic multi-dimensional table data, universal source and text based structure and code representations.

Results of the practical implementations and realisations are presented in the following sections.

III. RESULTING IMPLEMENTATION AND REALISATION

A means of choice in order to achieve overall efficient realisations even for complex scenarios, integrating arbitrary knowledge, is to use the principles of Superordinate Knowledge. The core assembly elements of Superordinate Knowledge are methodology, implementation, and realisation [24].

In the following solution, scenario targets are contexts of prehistoric cemeteries and burials at the North Sea coast, in North-Rhine Westphalia, Lower Saxony, and The Netherlands. Integration targets are natural sciences and speleological contexts, caves and cave systems in North-Rhine Westphalia, Lower Saxony, and The Netherlands, soil diversity, and overall integration with chorological, symbolical, spatial context representations, e.g., place, spatial planning, auxiliary subdivisions for boundaries and spatial forms and administrative units.

A. Coherent conceptual knowledge implementation

We can select relevant references from the implemented prehistory-protohistory and archaeology CKRI [6]. The methodology allows to address any other references on a coherent information science knowledge base, e.g., geoscientific

knowledge from natural sciences KR components. Further, the reference implementation enables to address chorology on the coherent knowledge base. Universally consistent conceptual knowledge is based on UDC references for demonstration, spanning the main tables [25] shown in Table I.

TABLE I. COHERENT CONCEPTUAL KNOWLEDGE DEPLOYED FOR CONTEXTUALISATION, SELECTED UDC CODE REFERENCES (EXCERPT).

<i>Code/Sign Ref.</i>	<i>Verbal Description (EN)</i>
UDC:0	Science and Knowledge. Organization. Computer Science. Information. Documentation. Librarianship. Institutions. Publications
<i>UDC:004</i>	<i>Computer science and technology. Computing.</i>
UDC:1	Philosophy. Psychology
<i>UDC:2</i>	<i>Religion. Theology</i>
UDC:3	Social Sciences
UDC:5	Mathematics. Natural Sciences
<i>UDC:52</i>	<i>Astronomy. Astrophysics. Space research. Geodesy</i>
UDC:53	Physics
UDC:539	Physical nature of matter
UDC:54	Chemistry. Crystallography. Mineralogy
UDC:55	Earth Sciences. Geological sciences
UDC:550.3	Geophysics
UDC:551	General geology. Meteorology. Climatology. Historical geology. Stratigraphy. Palaeogeography
<i>UDC:551.44</i>	<i>Speleology. Caves. Fissures. Underground waters</i>
UDC:551.46	Physical oceanography. Submarine topography. Ocean floor
UDC:551.7	Historical geology. Stratigraphy
UDC:551.8	Palaeogeography
UDC:56	Palaeontology
UDC:6	Applied Sciences. Medicine, Technology
UDC:63	Agriculture and related sciences and techniques. Forestry. Farming. Wildlife exploitation
<i>UDC:631.4</i>	<i>Soil science. Pedology. Soil research</i>
UDC:7	The Arts. Entertainment. Sport
UDC:8	Linguistics. Literature
UDC:9	Geography. Biography. History
UDC:902	Archaeology
<i>UDC:903</i>	<i>Prehistory. Prehistoric remains, artefacts, antiquities</i>
UDC:904	Cultural remains of historical times
UDC (1/9)	Common auxiliaries of place
UDC:(1)	Place and space in general. Localization. Orientation
UDC:(2)	Physiographic designation
UDC:(20)	Ecosphere
UDC:(21)	Surface of the Earth in general. Land areas in particular. Natural zones and regions
<i>UDC:(23)</i>	<i>Above sea level. Surface relief. Above ground generally. Mountains</i>
<i>UDC:(24)</i>	<i>Below sea level. Underground. Subterranean</i>
UDC:(25)	Natural flat ground (at, above or below sea level). The ground in its natural condition, cultivated or inhabited
UDC:(26)	Oceans, seas and interconnections
UDC:(28)	Inland waters
UDC:(3/9)	Individual places of the ancient and modern world
UDC:(3)	Places of the ancient and mediaeval world
UDC:(4/9)	Countries and places of the modern world
<i>UDC:(4)</i>	<i>Europe</i>
UDC:“...”	Common auxiliaries of time.
UDC:“6”	Geological, archaeological and cultural time divisions
<i>UDC:“62”</i>	<i>Cenozoic (Cainozoic). Neozoic (70 MYBP - present)</i>

For this research, major references from both main and auxiliary tables are highlighted in italics with bluish colour.

B. Multi-disciplinary views: Prehistory and archaeology

Table II shows an excerpt of UDC:903...:2 ritual/burial object and subgroup examples, and conceptual view groups [7] for prehistory and protohistory (PAKA, [17] [18]).

TABLE II. PREHISTORY AND PROTOHISTORY RITUAL/BURIAL OBJECT AND SUBGROUP EXAMPLES, AND CONCEPTUAL VIEW GROUPS [7] (EXCERPT).

<i>Major Object Group</i>	<i>Selected Objects</i>	<i>Conceptual View Group</i>
Ritual places, burials	yes	UDC:903...:2
Cemetery	–	UDC:903...:2
Barrow	–	UDC:903...:2
round	–	UDC:903...:2
long	–	UDC:903...:2
Cist	–	UDC:903...:2
Dolmen	–	UDC:903...:2
Tomb	–	UDC:903...:2
chamber	–	UDC:903...:2
court	–	UDC:903...:2
portal	–	UDC:903...:2
rock cut	–	UDC:903...:2
wedge	–	UDC:903...:2
Pithos burial	–	UDC:903...:2
Cave	–	UDC:903...:2
Body finding	–	UDC:903...:2
Urn	–	UDC:903...:2
...	–	UDC:903...:2

For this illustrative object scenario, the excerpt does not show individual micro-groups. Besides different distributions and different origins, object context can be referred, e.g., artificial origin and natural origins as well as relevant object properties, materials, and soil contexts can be considered.

C. Resulting realisation components: Soil diversity

A suitable UDC:631.4... base soil reference system for prehistory and archaeology has been compiled along with this research. The results are available in Table III.

TABLE III. COMPILATION OF CONCEPTUAL REFERENCE SYSTEM (UDC:631.4...), IMPLEMENTED AND REALISED WRB STANDARD SOIL TYPE REFERENCE GROUPS AND SOIL TYPE SPECIFICATIONS.

<i>Soil type Reference group</i>	<i>Soil type specification Name in WRB 2006/WRB 1998</i>
Acrisol	Haplic / Ferric, Gleyic, Haplic, Humic, Plinthic
Alisol	Plinthic
Albeluvisol	Haplic / Endoeutric, Gleyic, Haplic, Histic, Stagnic, Umbric
Andosol	Aluandic / Dystric, Humic, Umbric, Mollic, Vitric
Anthrosol	Anthrosol, Plaggic
Arenosol	Albic, Haplic, Protic
Calcisol	Aridic
Chernozem	Calcic, Haplic, Gleyic, Haplic, Luvic
Cambisol	Haplic / Calcaric, Haplic / Chromic, Haplic / Dystric, Haplic / Eutric, Gleyic, Haplic, Mollic, Vertic
Fluvisol	Haplic / Calcaric, Haplic / Dystric, Haplic / Eutric, Gleyic, Haplic, Histic, Mollic, Salic, Thionic
Gleysol	Haplic / Calcaric, Haplic / Dystric, Haplic / Eutric, Haplic / Haplic, Histic, Humic, Mollic, Thionic
Gypsisol	Haplic / Aridic
Histosol	Histosol, Hemic / Dystric, Hemic / Eutric, – / Fibric, – / Gelic, – / Sapric
Kastanozem	Calcic, Haplic, Luvic
Leptosol	Haplic / Calcaric, Haplic / Dystric, Haplic / Eutric, Haplic / Haplic, Haplic / Humic, Rendzic, Lithic
Luvisol	Albic, Haplic / Arenic, Calcic, Haplic / Chromic, Haplic / Dystric, Haplic / Ferric, Gleyic, Haplic, Vertic
Phaeozem	– / Albic, Haplic / Calcaric, Gleyic, Haplic, Luvic, Haplic / Sodic
Planosol	Haplic / Dystric, Haplic / Eutric, Haplic
Podzol	Haplic / Carbic, Haplic / Entic, Gleyic, Haplic, Leptic, Placic, Haplic / Rustic, Umbric
Regosol	Haplic / Calcaric, Haplic / Dystric, Haplic / Eutric, Haplic
Solonchak	Gleyic, Haplic, Haplic / Takyric, Mollic
Solonetz	Gleyic, Haplic, Mollic
Umbrisol	Arenic, Gleyic
Vertisol	Haplic / Chromic, Haplic, Haplic / Pellic

For this research, the reference system is based on standard soil references and UDC, both enabling a systematic and coherent approach. Soil diversity groups are relevant for pre-historical and archaeological objects and contexts. Contextualised soil diversity groups are referenced in a consistent, standardised way. From this base compilation, a properties based reference system can be created for further contextualisation, parametrisation, and processing with the ongoing research on soil diversity for prehistory and archaeology. Associated information, e.g., on soil drainage, wetness, pH status, base saturation, chloride, subsoil organic material, and stiffness can be found as reference in the World Reference Base (WRB) for soil resources [26], [27] from the Food and Agriculture Organisation (FAO), United Nations. In this context, the conceptual references are referring to the respective categories, e.g., UDC:631.4...:903+“4...”.

D. Multi-disciplinary integration facets

Table IV shows the reference facets of a respective multi-disciplinary target contextualisation.

TABLE IV. REFERENCE FACETS OF A MULTI-DISCIPLINARY TARGET CONTEXTUALISATION, BASED ON CKRI, IMPLEMENTED AND REALISED USING UDC CODE REFERENCES (EXCERPT).

<i>Code/Sign Ref.</i>	<i>Verbal Description (EN)</i>
UDC:903... ...:2 ...;“62...” ...(4...DENW) ...(4...DENI) ...(4...NL)	<i>Geography. Biography. History</i> Prehistory, prehistoric remains, artefacts, antiquities referring to religion and rituals from Holocene ... in North-Rhine Westphalia, Germany ... in Lower Saxony, Germany ... in The Netherlands
UDC:551.44	<i>Earth sciences, geological sciences</i> Speleology, caves, fissures, underground waters
UDC:631.4	<i>Applied sciences, agriculture in general</i> Soil research data
UDC:52...,(23) UDC:52...,(24)	<i>Geodesy. Photogrammetry</i> Remote sensing data, above sea level Remote sensing data, below sea level
UDC:(4)	<i>Contextualisation Place</i> Europe

The contextualisation uses coherent conceptual knowledge and refers to the chorological references for consequent knowledge integration and symbolic representation.

IV. RESULTING COHERENT KNOWLEDGE INTEGRATION

Figure 1 shows a generated, resulting coherent conceptual knowledge integration sketch for the realisation based on the KR. The sketch considers the major conceptual references for illustration. Detailed research can further detail on prehistoric object groups, characteristics, and properties, topographic properties, soil properties, and many more. Therefore, the conceptual sketch view can result in levels of arbitrary numbers of different integrations of complements and associated properties as resulting from the KR, which are discussed in the following. The result integrates required KR components based on coherent conceptual knowledge and systematical chorological knowledge for multi-disciplinary contexts, e.g., arbitrary group representations, classification based representations, and geospatial representations.

Knowledge objects and contexts are provided by The Prehistory and Archaeology Knowledge Archive (PAKA) [17] [18]. The multi-disciplinary coherent contextualisation employs the base of a new soil system reference development with soil types (UDC:631.4...) of WRB standard, reference contexts, especially for UDC:903...:2,551.7+“628”..., prehistorical, protohistorical time spans and artefacts related to religion and rituals, geology, especially stratigraphy and paleogeography, quaternary, especially late glacial and Holocene. The integrated natural sciences KR further provide information on caves in the respective region. Contextualisation is enabled by the Conceptual Knowledge Reference Implementation (CKRI), including multi-disciplinary contexts of natural sciences and humanities [6]. The conceptual knowledge base is The Universal Decimal Classification (UDC) [7].

In this illustration plain Digital Chart of the World (DCW) data are used [28]. The coastline database is the Global Self-consistent Hierarchical High-resolution Geography (GSHHG) [29] [30], which was mainly compiled from the World Vector Shorelines (WVS) [31], the CIA World Data Bank II (WDBII) [32], and the Atlas of the Cryosphere (AC).

An equal area projection (Eckert IV) is advised due to the type of discipline knowledge representation. The compilation uses the World Geodetic System (WGS). The symbolic representation of the contextualisation is done via LX Professional Scientific Content-Context-Suite (LX PSCC Suite) deploying the Generic Mapping Tools (GMT) [33] for visualisation.

Concrete details of knowledge complements and components have been provided [5]. All basic technical aspects can be created on that base for individual application scenarios. As illustrated, the solution is explicitly not a database concept and the goal is explicitly not just to link different multi-disciplinary concepts. The solution allows to create individual conceptual knowledge based algorithms and to integrate with new and available spatial and temporal processing algorithms. Basic components and functions are given in the references.

The presented integration approach for chorological and coherent contextualisation provides a solid base for multi-disciplinary contexts in prehistory and archaeology. The implemented system of components, continuously in development, integrates relevant and beneficial methodologies and properties, especially

- coherent conceptual knowledge views,
- multi-disciplinary contextualisation,
- application approaches for multi-disciplinary contexts in prehistory and archaeology,
- allows systematical chorological consideration of knowledge, e.g., arbitrary group representations, classification based representations, geospatial representations,
- further development and valorisation of resources,
- integration of multi-disciplinary resources,
- choice for homogeneity of components,
- deployment of systematical procedures,
- effective and efficient integration and analysis,
- automation of workflows and procedures, and
- provides multi-lingual conceptual knowledge support.

The methodological approach also allows the multi-directional utilisation: Conceptual knowledge and resulting integration, e.g., symbolic representation, on the one hand and integration results delivering references to conceptual knowledge and new integrated knowledge contexts on the other hand.

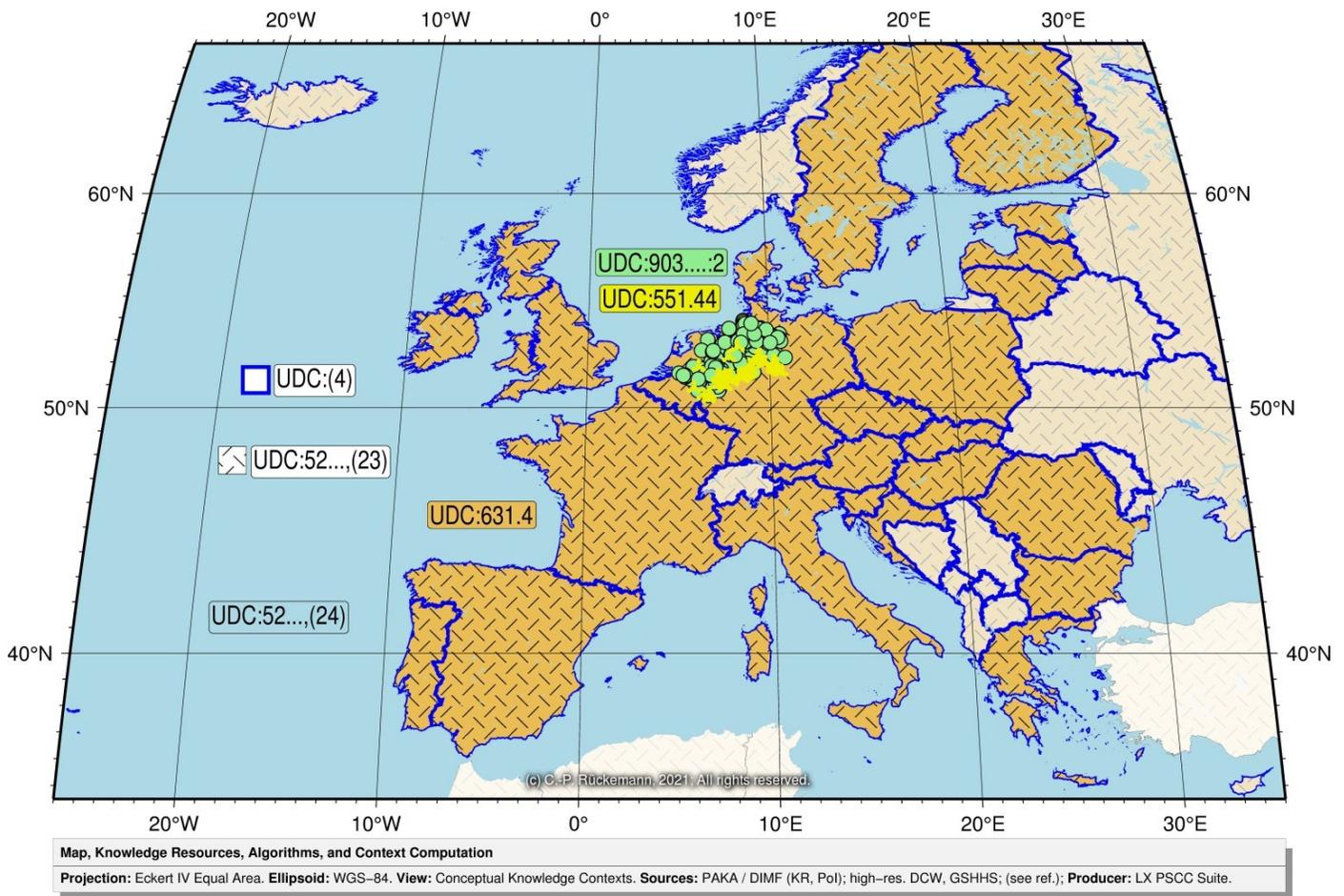


Figure 1. Resulting coherent conceptual knowledge integration sketch diagram showing knowledge resources for a prehistoric, natural sciences, and spatial contextualisation for excerpts of prehistoric cemeteries' and caves' distributions, remote sensing data, and soil properties with respective knowledge references.

As demonstrated with the integration for this research, besides coherency, general flexibility, robustness, and scalability, criteria for components employed with implementation and realisation should be evaluated carefully for being able to consider solid information science fundamentals and knowledge centricity, beyond plain technical and proprietary features.

Resulting from the methodology, the realisation integrates a wide range of relevant selection criteria, e.g., in this scenario:

- Conceptual selection (esp. prehistory, cemetery; natural sciences, caves).
- Spatial, mathematical selection.
- Regional spatial selection.
- Topographic selection (esp. above sea level).
- Contextualised selection (esp. with availability of sufficient natural sciences, soil, and other context data).

V. CONCLUSION

The practical solution for implementation and realisation of the new knowledge-based methodology showed to enable coherent conceptual knowledge for contextualisation in prehistory, archaeology, and natural sciences. The more, the approach enables to integrate multi-disciplinary contexts by a consistent system (editions) and supports multi-lingualism. Multi-disciplinary scenarios can be considered in multi-fold ways, e.g., knowledge can be documented, analysed, integrated, and selected deploying conceptual knowledge. The methodological approach also allows a multi-directional utilisation. Any KR

result can be considered starting point, intermediate result, and final result, depending on a defined task and workflow.

The general methodology provides flexibility for solid information science based methods and enables a wide range of benefits for scenarios and implementations, e.g., coherent multi-disciplinary and multi-lingual documentation and systematical knowledge based geo-spatial processing, aware of inherent knowledge spanning arbitrary disciplines. With that, geospatial scenarios are special cases of chorological contextualisation. Further recommendations from practical experiences with knowledge complements and component integration are:

- Add consistent and coherent conceptual knowledge to objects and entities of your resources and make workflows deploy conceptual knowledge.
- Choose multi-disciplinary resources with homogeneous properties, e.g., resolution and coverage.
- Use long-term standards.
- Practice scientific state-of-the-art parametrisation of respective knowledge, data, and algorithms.
- Create context dependent, suitable symbolic representations and individual methods.
- Proceed the knowledge-centric integration reasoned and levelheaded.

Future research concentrates on continuing development of components for detailed multi-disciplinary application scenarios in prehistory and archaeology and respective chorological and coherent conceptual knowledge contextualisation.

ACKNOWLEDGEMENTS

This ongoing research is supported by scientific organisations and individuals. We are grateful to the “Knowledge in Motion” (KiM) long-term project, Unabhängiges Deutsches Institut für Multi-disziplinäre Forschung (DIMF), for partially funding this research, implementation, case studies, and publication under grants D2018F1P04962, D2019F1P04998, and D2020F1P05228 and to its senior scientific members and members of the permanent commission of the science council, especially to Dr. Friedrich Hülsmann, Gottfried Wilhelm Leibniz Bibliothek (GWLb) Hannover, to Dipl.-Biol. Birgit Gersbeck-Schierholz, Leibniz Universität Hannover, to Dipl.-Ing. Martin Hofmeister, Hannover, and to Olaf Lau, Hannover, Germany, for fruitful discussion, inspiration, and practical multi-disciplinary case studies. We are grateful to Dipl.-Ing. Hans-Günther Müller, Göttingen, Germany, for providing specialised, manufactured high end computation and storage solutions. We are grateful to The Science and High Performance Supercomputing Centre (SHPC) for long-term support. / DIMF-PIID-DF98_007.

REFERENCES

- [1] C.-P. Rückemann, “From Knowledge and Meaning Towards Knowledge Pattern Matching: Processing and Developing Knowledge Objects Targeting Geoscientific Context and Georeferencing,” in Proc. of The Twelfth Internat. Conf. on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2020), November 21–25, 2020, Valencia, Spain. XPS Press, Wilmington, Delaware, USA, 2020, pp. 36–41, ISSN: 2308-393X, ISBN-13: 978-1-61208-762-7.
- [2] R. Gleser, *Zu den erkenntnistheoretischen Grundlagen der Prähistorischen Archäologie*. Leiden, 2021, 2021, (title in English: *On the Epistemological Fundamentals of Prehistorical Archaeology*), in: M. Renger, S.-M. Rothermund, S. Schreiber, and A. Veling (Eds.), *Theorie, Archäologie, Reflexion. Kontroversen und Ansätze im deutschsprachigen Diskurs*, (in print).
- [3] C.-P. Rückemann, F. Hülsmann, B. Gersbeck-Schierholz, P. Skurowski, and M. Staniszewski, *Knowledge and Computing*. Post-Summit Results, Delegates’ Summit: Best Practice and Definitions of Knowledge and Computing, Sept. 23, 2015, The Fifth Symp. on Adv. Comp. and Inf. in Natural and Applied Sciences (SACINAS), The 13th Int. Conf. of Num. Analysis and Appl. Math. (ICNAAM), Sept. 23–29, 2015, Rhodes, Greece, 2015, pp. 1–7, DOI: 10.15488/3409.
- [4] L. W. Anderson and D. R. Krathwohl, Eds., *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom’s Taxonomy of Educational Objectives*. Allyn & Bacon, Boston, MA (Pearson Education Group), USA, 2001, ISBN: 978-0801319037.
- [5] C.-P. Rückemann, “Prehistory’s and Natural Sciences’ Multi-disciplinary Contexts: Contextualisation and Context Integration Based on Universal Conceptual Knowledge,” in Proceedings of The Eleventh International Conference on Advanced Communications and Computation (INFOCOMP 2020), May 30 – June 3, 2021, Valencia, Spain. XPS Press, Wilmington, Delaware, USA, 2021, ISSN: 2308-3484, ISBN: 978-1-61208-865-5.
- [6] C.-P. Rückemann, “The Information Science Paragon: Allow Knowledge to Prevail, from Prehistory to Future – Approaches to Universality, Consistency, and Long-term Sustainability,” *The International Journal “Information Models and Analyses” (IJ IMA)*, vol. 9, no. 2, 2020, Markov, K. (ed.), ISSN: 1314-6416 (print), (to appear).
- [7] “Multilingual Universal Decimal Classification Summary,” 2012, UDC Consortium, 2012, Web resource, v. 1.1. The Hague: UDC Consortium (UDCC Publication No. 088), URL: <http://www.udcc.org/udcsummary/php/index.php> [accessed: 2021-06-12].
- [8] “Creative Commons Attribution Share Alike 3.0 license,” 2012, URL: <http://creativecommons.org/licenses/by-sa/3.0/> [accessed: 2021-06-12], (first release 2009, subsequent update 2012).
- [9] B. M. Das, *Advanced Soil Mechanics; Fifth Edition*, 5th ed. CRC Press, Taylor & Francis Group, Boca Raton, London, New York, 2019, ISBN: 978-0-8153-7913-3.
- [10] “Land Processed Distributed Active Archive Center,” 2021, LP DAAC, URL: <https://lpdaac.usgs.gov/> [accessed: 2021-06-12].
- [11] C.-P. Rückemann, R. Pavani, B. Gersbeck-Schierholz, A. Tsitsipas, L. Schubert, F. Hülsmann, O. Lau, and M. Hofmeister, *Best Practice and Definitions of Formalisation and Formalism*. Post-Summit Results, Delegates’ Summit: The Ninth Symp. on Adv. Comp. and Inf. in Natural and Applied Sciences (SACINAS), The 17th Int. Conf. of Num. Analysis and Appl. Math. (ICNAAM), Sept. 23–28, 2019, Rhodes, Greece, 2019, DOI: 10.15488/5241.
- [12] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran 90, The Art of Parallel Scientific Computing*. Cambridge University Press, 1996.
- [13] T. Vincenty, “Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations,” *Surv. Rev.*, vol. XXII, no. 176, 1975, pp. 88–93.
- [14] R. J. Renka, “Algorithm 772: STRIPACK: Delaunay Triangulation and Voronoi Diagram on the Surface of a Sphere,” *AMC Trans. Math. Software*, vol. 23, no. 3, 1997, pp. 416–434.
- [15] B. K. P. Horn, “Hill-Shading and the Reflectance Map,” *Proceedings of the IEEE*, vol. 69, no. 1, 1981, pp. 14–47.
- [16] A. J. Kent and P. Vujakovic, Eds., *The Routledge Handbook of Mapping and Cartography*. Taylor & Francis Group, London and New York, 2018, ISBN: 978-1-138-83102-5.
- [17] C.-P. Rückemann, “Information Science and Inter-disciplinary Long-term Strategies – Key to Insight, Consistency, and Sustainability: Conceptual Knowledge Reference Methodology Spanning Prehistory, Archaeology, Natural Sciences, and Humanities,” *International Tutorial, DataSys Congress 2020*, Sept. 27 – Oct. 1, 2020, Lisbon, Portugal, 2020, URL: <http://www.iaria.org/conferences2020/ProgramINFOCOMP20.html> [accessed: 2021-06-12].
- [18] “The Prehistory and Archaeology Knowledge Archive (PAKA) license,” 2021, (release 2021), Unabhängiges Deutsches Institut für Multi-disziplinäre Forschung (DIMF): All rights reserved. Rights retain to the contributing creators.
- [19] C.-P. Rückemann, “The Impact of Information Science Accompanied Structural Information on Computation of Knowledge Pattern Matching and Processing: A Prehistory, Archaeology, Natural Sciences, and Humanities Conceptual Integration Perspective,” in Proc. of The Tenth Int. Conf. on Adv. Comm. and Comp. (INFOCOMP 2020), Sept. 27 – Oct. 1, 2020, Lisbon, Portugal, 2020, ISSN: 2308-3484, ISBN: 978-1-61208-807-5, URL: http://www.thinkmind.org/index.php?view=article&articleid=infocomp_2020_1_10_60015 [accessed: 2021-06-12].
- [20] R. Brinkmann, *Abriß der Geologie, Bd. 2, Historische Geologie*, neubearbeitet von Karl Krömmelbein, 12th/13th ed. Ferdinand Enke Verlag, Stuttgart, 1986, ISBN: 3-432-80603-5, (title in English: *Epitome of Geology, Vol. 2, Historical Geology*).
- [21] P. Woldstedt and K. Duphorn, *Norddeutschland und angrenzende Gebiete im Eiszeitalter*, 3rd ed., 1974, (title in English: *Northern Germany and Adjacent Regions in the Ice Age*).
- [22] G. N. Bailey, J. Harff, and D. Sakellariou, Eds., *Under the Sea: Archaeology and Palaeolandscapes of the Continental Shelf*, ser. Coastal Research Library. Springer International Publishing AG 2017, 2017, vol. 20, ISSN: 2211-0577, ISBN: 978-3-319-53158-8, DOI: 10.1007/978-3-319-53160-1.
- [23] “Network Common Data Form (NetCDF),” 2021, DOI: 10.5065/D6H70CW6, URL: <http://www.unidata.ucar.edu/software/netcdf/> [accessed: 2021-06-12].
- [24] C.-P. Rückemann, “Principles of Superordinate Knowledge: Separation of Methodology, Implementation, and Realisation,” in *The Eighth Symp. on Adv. Comp. and Inf. in Natural and Applied Sciences (SACINAS)*, Proceedings of The 16th Int. Conf. of Num. Analysis and Appl. Math. (ICNAAM), Sept. 13–18, 2018, Rhodes, Greece, AIP CP, Vol. 2116. AIP Press, American Institute of Physics, Melville, New York, USA, 2019, ISBN: 978-0-7354-1854-7, DOI: 10.1063/1.5114033.
- [25] “UDC Summary Linked Data, Main Tables,” 2020, Universal Decimal Classification (UDC), UDC Consortium, URL: <https://udcdata.info/078887> [accessed: 2021-06-12].
- [26] “World reference base for soil resources,” 1998, World Soil Resources Report No. 84, ISSS-ISRIC-FAO, Rome.
- [27] “World reference base for soil resources,” 2006, World Soil Resources Report No. 103, ISSS-ISRIC-FAO, Rome.

- [28] P. Wessel, "The Digital Chart of the World for GMT 5 or later," 2018, URL: <ftp://ftp.soest.hawaii.edu/gmt> [accessed: 2021-06-12], URL: <http://www.soest.hawaii.edu/pwessel/dcw/> [accessed: 2021-06-12].
- [29] P. Wessel, "Global Self-consistent Hierarchical High-resolution Geography," 2016, URL: <ftp://ftp.soest.hawaii.edu/gmt> [accessed: 2021-06-12], URL: <http://www.soest.hawaii.edu/pwessel/gshhg/> [accessed: 2021-06-12].
- [30] P. Wessel and W. H. F. Smith, "A global, self-consistent, hierarchical, high-resolution shoreline database," *J. Geophys. Res.*, vol. 101, no. B4, 1996, pp. 8741–8743.
- [31] E. A. Soluri and V. A. Woodson, "World Vector Shoreline," *Int. Hydrograph. Rev.*, vol. LXVII, no. 1, 1990, pp. 27–35.
- [32] A. J. Gorny, "World Data Bank II General User Guide," 1977.
- [33] P. Wessel, W. H. F. Smith, R. Scharroo, J. Luis, and F. Wobbe, "The Generic Mapping Tools (GMT)," 2020, URL: <http://www.generic-mapping-tools.org/> [accessed: 2021-06-12], URL: <http://gmt.soest.hawaii.edu/> [accessed: 2021-06-12].