Author's final version

INCREASING THE IMPACT OF OPEN DATA THROUGH VISUALISATION: CARTOGRAPHIC VISUALISATION OF OCEAN DATA

Andrea Miletić, MSc[1]

Assist. Prof. Dr. Ana Kuveždić Divjak, PhD[1]

Karlo Kević, MSc[1]

[1] University of Zagreb, Faculty of Geodesy, **Croatia**

ABSTRACT

Many national and international agencies and organizations collect and provide open access to data on parts of the Earth's surface covered by water. Data on ocean temperature and acidity, precipitation, river flows, lake, sea, and groundwater levels, and many other open data sets allow researchers to monitor and observe water at different scales and time frames, and to better understand ecological processes and human impacts on the water environment.

The main challenge in analyzing the vast amount of openly available data on seas, oceans, and other parts of the Earth's surface covered by water is their heterogeneity and the complexity of dynamic processes, which can be highly variable both spatially and temporally. Effective visualization of such data on a map can help in the effort to make mutual relationships obvious, to recognize the impacts of different actions, and model and evaluate tradeoffs among possible solutions.

This paper explores the options available to cartographers for designing various thematic representations of data related to the water-covered portions of the Earth's surface. Using a set of openly available data, we demonstrate how choice of map projection, color schema type, and color palette can influence the transfer of information related to ocean health. In this process we examine the role of cartography in designing various thematic representations of data related to parts of the Earth's surface covered by water.

Keywords: cartography, visualization, open data, ocean data

INTRODUCTION

Open data refers to data that anyone can freely access, use, modify, and share for any purpose [1]. It is well known that they can stimulate economic growth and improve citizen engagement, but recently they have also become a foundation of modern science, enabling large-scale analysis and reproducibility [2]. Opening and sharing data has therefore become an important part of scientific research and has led to the development of the open science movement. Scientists around the world are encouraged to use open data in their analyses and publish the results of their discoveries. Thanks to open data initiatives, a large amount of data, including ocean data, is now available for wider use. Ocean data are oceanographic hydrological data about tides, currents, waves, seabed topography, salinity, and other ocean properties [3]. With modern technologies (e.g., remote sensing) enabling faster and better ocean observations, more and more data on the ocean is being produced [4]. But to use all this available ocean data, open data literacy and ocean literacy are necessary. While ocean literacy is focusing on developing ocean-literaterate society and enabling science to engage with policy and society on the topics of ocean sustainability, observation, and research [5], open data literacy focuses on developing skills needed to manipulate open ocean data. These include data discovery and acquisition, data cleaning, and working with data in open data formats using common open data tools and visualizations [6]. These two concepts together have accelerated the spread of data visualizations, which often serve as a tool for a better understanding of a vast amount of geospatial ocean data.

Powerful visualization of data on a map can clarify interrelationships, show the impact of different measures, and encourage action. However, unlike official nautical charts, whose cartography has been standardized through international mediation, the cartographic framework for efficient representation of water-covered portions of the Earth's surface on bathymetric and other thematic maps intended for the general public (i.e., users who have no prior knowledge of using nautical charts and do not participate in oceanographic research) is not documented in the existing literature. The main objective of this paper is to investigate the options available to cartographers (e.g. the choice of map projection, color schema type, and color palette) in the visualization of ocean data and to make suggestions on the appropriate techniques for specific task. Using a set of openly available data on ocean health, we will demonstrate how the selection of the different cartographic techniques can affect the way information about the ocean is transmitted to the user of the map. This will result with guidelines that can be used by ocean researches to present ocean data in an effective and meaningful way.

MATERIALS & METHODS

Ocean Health Index Data

Ocean Health Index is a standardized, quantitative, transparent, and scalable measure that observes how sustainable we are managing the ocean [7]. It was introduced in 2012 as a mechanism to mitigate alterations to the ocean caused by human activities. The Index comprises of 10 goals [7] which reflect on the benefits people want and need from the ocean (e.g., food provision) and is limited to exclusive economic zones (EEZ). Each of the 10 goals is scored based on the current status and likely future status with scores ranging from 0 to 100 (0 – poor sustainable management, 100 – high sustainable management of ocean resources). Likely future status is a combination of the status of ocean health through the past – trend, and the ratio of pressures and resilience (resistance of the environment to threats - pressures). Ocean Health Index data are released yearly with the dataset for 2022 being published recently.

Cartographic techniques

The first step when designing a map is choosing the map projection. It is impossible to represent a three-dimensional surface of Earth on a two-dimensional map without distortions. Depending on the specific task, every projection has its advantages and disadvantages, and its selection depends on the peculiarities and the nature of the data that needs to be presented on the map.

In the cartographic literature, cylindrical and pseudocylindrical map projections are suggested as the appropriate choice for visualizations of ocean maps [8]. The main problem with the cylindrical projection is the unsatisfactory representation of the extreme northern and southern parts of the Pacific and Atlantic Oceans and the southern part of the Indian Ocean, but this problem is somewhat reduced by the use of pseudocylindrical projection [8]. When searching for projections used to represent the world's oceans in a continuous way, it was found that the Mollweide and Goode-Homolosin projections (both are equal area pseudocylindrical) are among the most commonly used for presenting ocean. Also worth mentioning is the Spilhaus projection, which gained great popularity in 2018 because it depicts oceans as a single body of water and makes it the main focus [9]. In addition to map projection, an important part of data visualization is choosing the right color scheme for data display. There are four main types of color schemes: binary, qualitative, diverging, and sequential [10]. The choice of the right color scheme depends on several factors; the type of data, whether there is a meaningful midpoint in the data range, the existence of extreme values, the number of data categories, the message that needs to be communicated, and more. In addition to the color scheme, the right selection of color hue can also play a big role in data communication. Colors are considered an important means of communication since they carry multiple meanings and affect attention, emotions, beliefs and behaviours. [11]

RESULTS

Depending on the map projection choice, the user can perceive the visual representation of the same data in a different way. As demonstrated in Figure 1, in the Mollweide projection OHI 2022 data distribution is represented with discontinuity of ocean water surface. In contrast, Goode Homolosine and Spilhaus (Figure 1b and 1c) have preserved the wholeness of ocean water bodies.

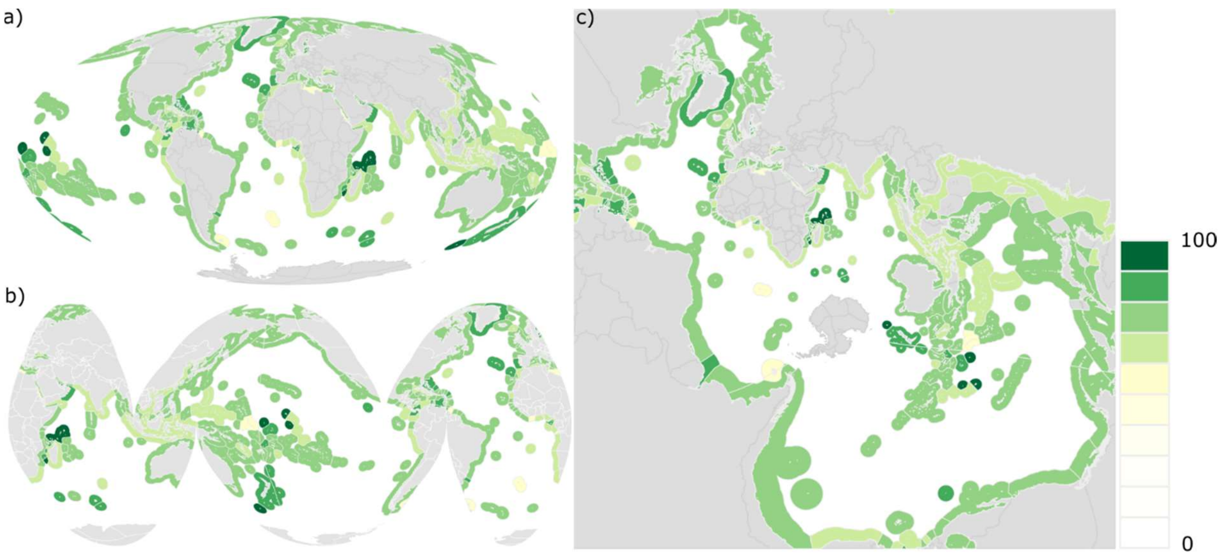

Figure 1 – Ocean Health Index 2022 visualized using different map projections. In a) *Mollweide* oceans are discontinued, while in b) *Goode Homolosine* and c) *Spilhaus* oceans have preserved the wholeness of ocean water bodies

The choice of different color schemas used in the visualization of OHI 2022 data can cause different interpretations of the same data. Sequential color schema (Figure 2a) shows ocean health ranging from white to dark green. White color is associated with lower values, while dark green relates to higher values but there is no intuitive linkage of the score with activities related to ocean management. Diverging color schema (Figure 2b) presents ocean health with colors from red to blue (from lower to higher values) but also gives the impression how the color is related to the human management of the ocean. Qualitative schema (Figure 2c) with random colors transmits no information of the rankings in score values.

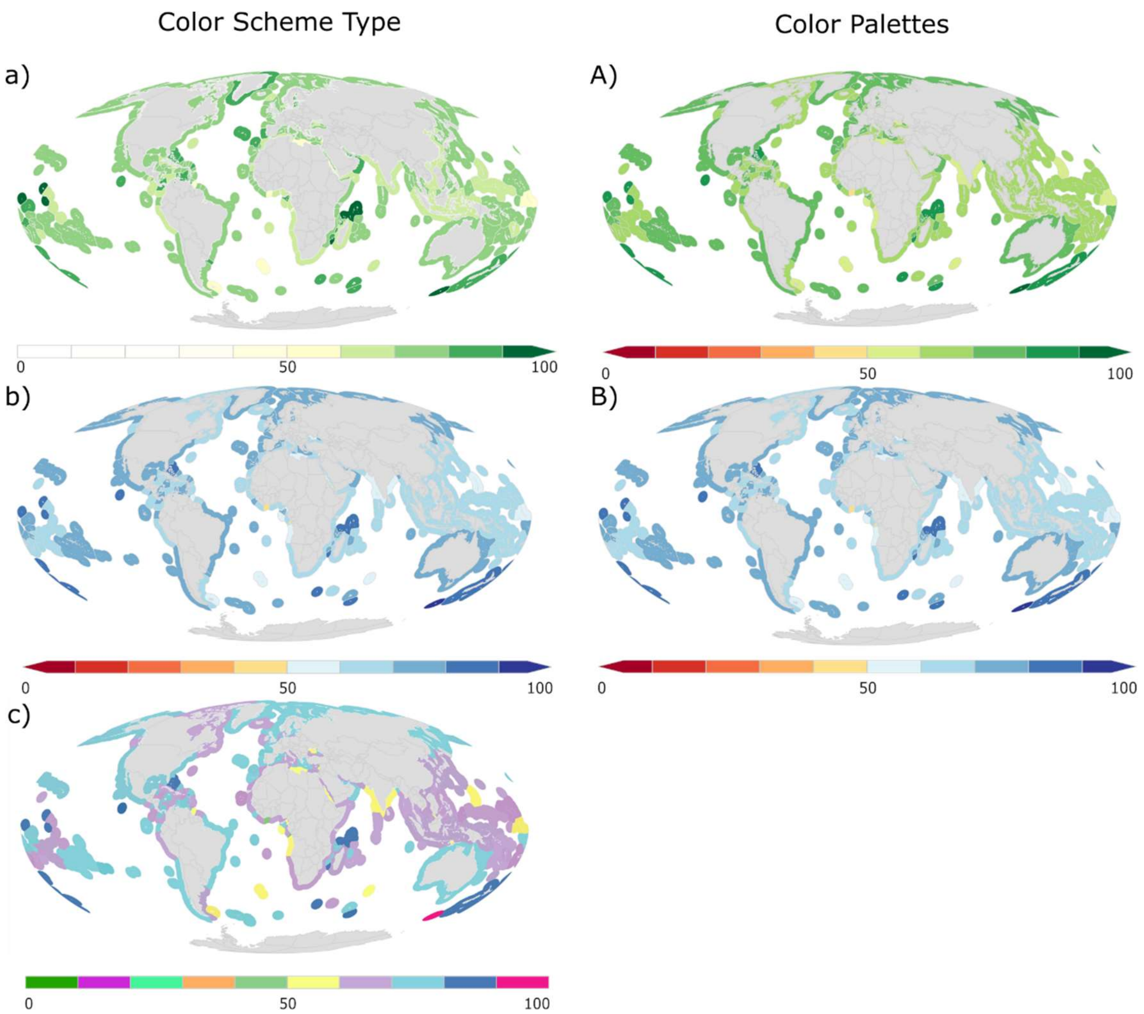

Figure 2 – Application of different color shemas (a,b,c) and color palettes (A,B) in the process of visualization of Ocean Health Index 2022 data

Color palettes also affect the way information is perceived. As seen in Figure 2A, ocean health is presented in a red-green color pallet in contrast to Figure 2B where red-blue is used. Colors are usually associated with certain anticipations so some color pallets are more suited to the presentation of ocean health data.

Finally, the Figure 3 demonstrates how on the same map, land-related data can be presented in combination with ocean data to obtain a more informative view of ocean health-related issues. Seafood production is presented using sequential schema in blue, while fish consumption is visualized using the same schema type but in a yellow color pallet.

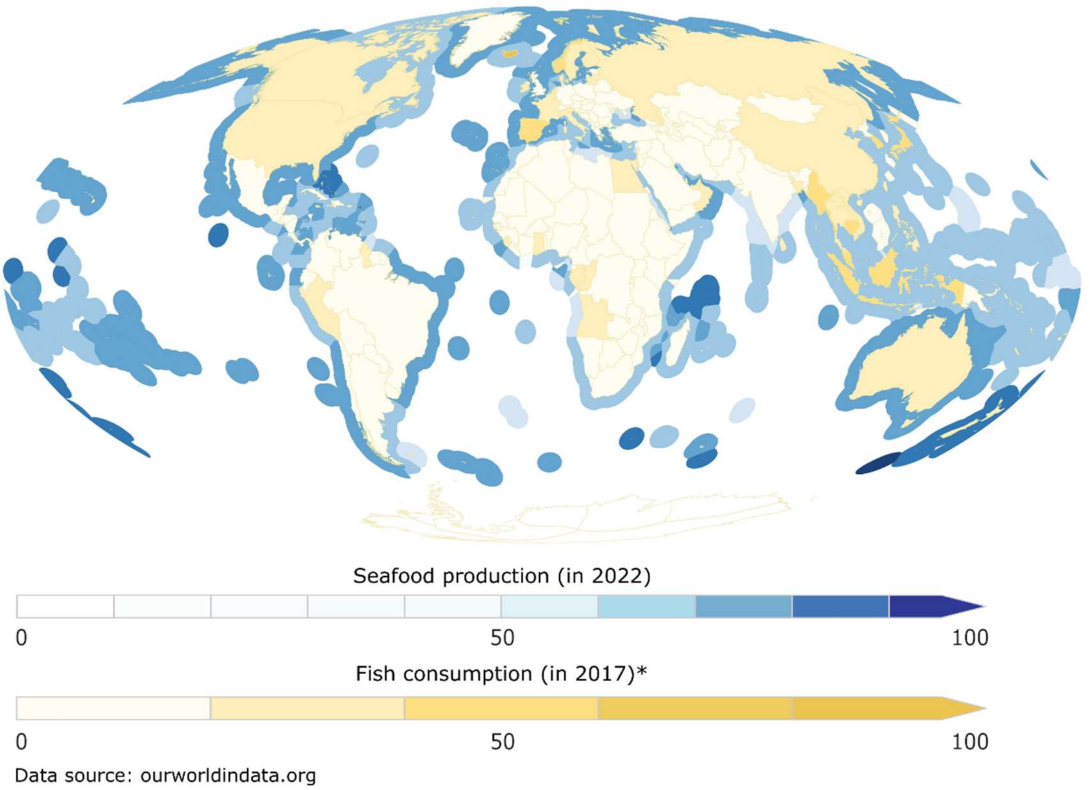

Figure 3 – Visualization of how land related data can be presented in combination with ocean related data to obtain more informative view on ocean health related issues

DISCUSSION

The presented results have demonstrated how a choice of map projection, color schema, and color palette can influence the transfer of information related to ocean health. For instance, although all proposed map projections are considered to be a suitable choice for the representation of ocean surfaces, for this particular task some of them are more suitable than others. Pseudocylindrical map projections, such as Mollweide's, are often used for small-scale world maps and since the health index is closely related to the cost (continent), these projections might be the most appropriate solution.

Ocean health index score values range between 0 and 100. Although 0 means poor sustainable management and 100 means good sustainable management of ocean resources a score of 0 might also indicate not efficient ocean management which should be interpreted as a negative human footprint. In this case, based on the nature of the underlying data, a diverging color schema would be a better choice in comparison to a sequential or qualitative scheme. Besides color schema, the choice of color pallet also plays important role in the perception of information. The green colour is usually associated with something good or healthy in contrast to red which indicates danger.

Although the red-blue colour pallet is the one that is used on the official maps of Ocean Health Index, the red-green pallet is more in line with the information transmitted.

As seen from the results (Figure 3), when more data must be visualized on a single map, color schemes and pallets must be chosen carefully. As colors evoke emotions, yellow and blue, colors associated with health and naturality, were chosen to present data on seafood production and fish consumption. Fish production is presented in blue as this color is also usually associated with water.

CONCLUSION

Opening the data has led to the broader use of ocean data for various research purposes and allowed visualization of these data by a variety of users with different data visualization skills. The scientists are encouraged to use this data as part of their research, but when visualizing the data on the map some cartographic principles need to be followed.

However, unlike official nautical charts, whose cartography has been standardized through international mediation, the choices when designing thematic maps that represent ocean health and other ocean-related thematic data rely on the personal choice of the map producer (researcher). Depending on the type of data presented on a map, the choice of cartographic techniques can impact the way information is transmitted. This process usually requires good craft when it comes to data visualization in order to explain the data in a way that is easy to understand, and does not misguide the user, which is why it is important for researchers to have appropriate ocean data and open data literacy.

As seen from the results, the choice of the most suitable cartographic techniques depends on the type and nature of underlying data. The selection of map projection, color scheme, and color pallet are important choices that can increase or decrease the impact of the map on end users. Therefore, although cartography can be appropriate means to understand the vast amount of ocean data, the use of wrong cartographic techniques can lead to misperception of the information and create confusion.

ACKNOWLEDGEMENTS

This research is part of the “Twinning Open Data Operational – TODO” project that has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement no. 857592. Work by Andrea Miletić was additionally supported by “Young Researchers’ Career Development Project – Training of Doctoral Students” (DOK-01-2020) founded by Croatian Science Foundation.

REFERENCES

[1] The Open Data Definition, https://opendefinition.org/ (5.11.2022)

[2] Tedersoo, L., Küngas, R., Oras, E., Köster, K., Eenmaa, H., Leijen, Ä., Pedaste, M., Raju, M., Astapova, A., Lukner, H., Kogermann, K., & Sepp, T. (2021): Data sharing practices and data availability upon request differ across scientific disciplines. Scientific Data, 8(1)

[3] Xie, C., Li, M., Wang, H., Dong, J. (2019): A survey on visual analysis of ocean data. Visual informatics 3 (3), 113-128

[4] Penny, S. G., Akella, S., Balmaseda, M. A., Browne, P., Carton, J. A., Chevallier, M., Counillon, F., Domingues, C., Frolov, S., Heimbach, P., Hogan, P., Hoteit, I., Iovino, D., Laloyaux, P., Martin, M. J., Masina, S., Moore, A. M., de Rosnay, P., Schepers, D., Sloyan, B. M., Storto, A., Subramanian, A., Nam, S., Vitart, F., Yang, C., Fujii, Y., Zuo, H., O'Kane, T., Sandery, P., Moore, T. and Chapman, C. C. (2019): Observational Needs for Improving Ocean and Coupled Reanalysis, S2S Prediction, and Decadal Prediction. 6:391

[5] Eparkhina, D., Pomaro, A., Koulouri, P., Banchi, E., Canu, D., Uyarra, M. and Burken, N. (2021): Ocean Literacy in European Oceanographic Agencies: EuroGOOS recommendations for the UN Decade of Ocean Science for Sustainable Development 2021-2030. Brussels. Belgium, EuroGOOS, 32pp

[6] Open Data Literacy, https://www.stateofopendata.od4d.net/chapters/issues/data-literacy.html (11.11.2022)

[7] Halpern, B. S., Longo, C., Hardy, D., McLeod, K. L., Samhouri, J.F., Katona, S.K., Kleisner, K., Lester, S. E., O'Leary, J., Ranelletti, M., Rosenberg, A. A., Scarborough, C., Selig, E. R., Best, B. D., Brumbaugh, D. R., Chapin, F. S., Crowder, L. B., Daly, K. L., Doney, S. C., Elfes, C., Fogarty, M. J., Gaines, S. D., Jacobsen, K. I., Karrer, L. B., Leslie, H. M., Neeley, E., Pauly, D., Polasky, S., Ris, B., St Martin, K., Stone, G. S., Sumaila, U. R., Zeller, D. (2012): An index to assess the health and benefits of the global ocean. Nature 448, 615-620

[8] Frančula, N. (ur.) (2020): Kartografske projekcije: Zagreb: Faculty of Geodesy

[9] The Spilhaus Projection,
https://storymaps.arcgis.com/stories/756bcae18d304a1eac140f19f4d5cb3d (11.11.2022)

[10] The Color Scheme,
https://web.natur.cuni.cz/~langhamr/lectures/vtfg1/mapinfo_2/barvy/colors.html
(21.11.2022)

[11] Wąsowicz, G., Styśko-Kunkowska, M., Grunert, K.G. (2015): The meaning of colours in nutrition labelling in the context of expert and consumer criteria of evaluating food product healthfulness. J Health Psychol; 20(6):907-20