IGARSS 2021 - 2021 IEEE International Geoscience and Remote Sensing Symposium | 978-1-6654-0369-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/IGARSS47720.2021.9554208

GEOSPATIAL DATA ANALYSIS FOR GLOBAL MARITIME RISK ASSESSMENT USING THE DISCRETE GLOBAL GRID SYSTEM

Andrew Rawson[1], Zoheir Sabeur[2], Mario Brito[1]

[1] University of Southampton, UK [2] Bournemouth University, UK
A.Rawson@soton.ac.uk, zsabeur@bournemouth.ac.uk, M.P.Brito@soton.ac.uk

ABSTRACT

The effective management of the safety of navigation by coastguards is challenged by the complexity in quantifying and describing the relative risk of accidents occurrence. The discovery of patterns in observation data is reliant on the collection and analysis of significant volumes of relevant heterogenous spatial datasets. Conventional approaches of risk mapping which aggregate vessel traffic and incident data into Cartesian grids can result in misrepresentation due to inherent inadequacies in this spatial data format. In this paper, we explore how the Discrete Global Grid System (DGGS) overcomes these limitations through the development of global maps of incident rates at multiple resolutions. The results demonstrate hot spots of relative high risk across different regions and clearly show that DGGS is more suited to global analysis than conventional grids. This work contributes to a greater understanding of both the disposition of maritime risk and the advantages of adopting DGGS in supporting big data analysis.

Index Terms— Maritime Risk, Automatic Identification System, Discrete Global Grid System, Big Data.

1. INTRODUCTION

Accidents involving navigating vessels have the potential to cause significant loss of life and environmental damage. Whilst the responsibility for safe navigation ultimately lies with the vessel's master, national administrations and harbor authorities are able to manage waterways to reduce the risk of incident occurrence. Decision makers must consider the benefits of introducing mandatory traffic lanes, emergency towage vessels, vessel traffic services or whether new developments such as wind farms are a hazard to shipping. Each of these decisions sits within a context of balancing the risk of an accident against costly risk control measures.

In order to understand these hazards, and enable informed decision making, risk assessment is required. Whilst this has traditionally been achieved using expert judgement, a growing body of work in recent years has championed quantitative maritime risk analysis as a field of scientific research.

One method for conceptualizing the spatial variation in maritime risk is through computing incident rates, comparing where vessels navigate and incidents have occurred. For example, some have utilized incident rates to understand which vessels and conditions are more likely to result in an accident, while promoting targeted risk controls [1]. In addition, incident rates are often key inputs into quantitative risk models for determining the likelihood that hazardous situations might turn into incidents [2]. Studies on accident numbers alone are limited as it would be expected that more accidents would occur in busier waterways, not that they are any more hazardous.

Whilst some have analyzed global patterns of vessel traffic movements [3] or incidents [4], few have attempted to undertake global modelling of vessel risk. This is important as there is a recognized bias in research to certain locations such as the Baltic Sea [5], which might limit the transferability of models to other locations. Within this paper, we seek to firstly, explore the spatial variation in incident rates across the globe. Secondly, we demonstrate the methodological challenges with spatial data structures through which global datasets are aggregated, and the benefits of a Discrete Global Grid System (DGGS) as a means to solve them.

1.1. DGGS

One means to analyze spatial data is to aggregate the data into a system of grid cells or tessellations. This can reduce the complexity of spatial data, which is continuous with an infinite number of positions, into a finite number of elements through which traditional statistical methods can be applied. However, the spatial units are independent of the features being described and might influence any derived statistical relationships [6], sometimes summarized as the Modifiable Areal Unit Problem.

A common method of discretization is the use of Cartesian grids with fixed x-y coordinates, such as decimal degrees [3]. However, such grids are inherently flawed by attempting to represent a spherical world through square grids. This results in significant distortion of angles, areas and distances.

In order to overcome some of these challenges, we have proposed DGGS as a preferable spatial reference system that provides hierarchical tessellation of equal-area cells at a global scale [7]. This is demonstrated in Figure 1, where a hexagonal DGGS is compared to a 1-degree Cartesian Grid, showing significant distortion in cell size and shape towards the poles. At best, this would require normalization of cell areas [8] such as when measuring ship density in the Arctic [9], or at worse result in misleading analyses.

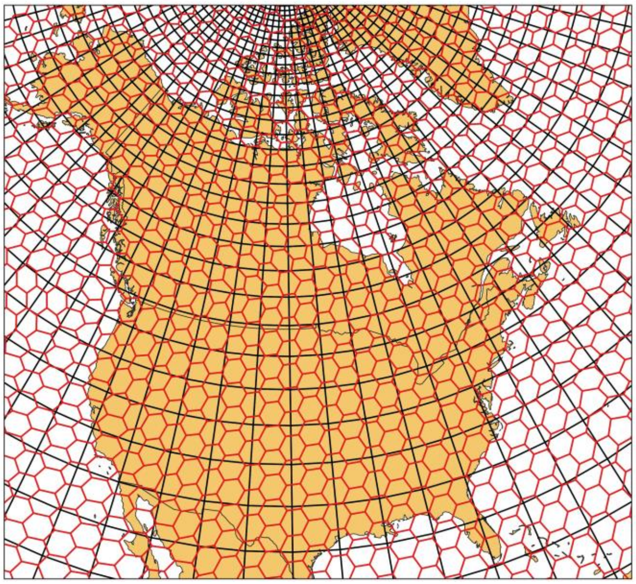

Figure 1: Comparison of 1-degree graticule and DGGS5.

2. DATA AND METHODS

2.1. Datasets

Within this assessment, two principal datasets are required: firstly, global data of vessel traffic activity and secondly, global data of maritime incidents. Vessel traffic movements are most commonly analyzed using the Automatic Identification System (AIS), a transponder system which is required on commercial vessels that transmits regular data concerning their identification and behavior. AIS data collection is challenging as transmission range is approximately line of sight, and whilst satellites can achieve a wider coverage, this has a high cost of access.

Therefore, we utilize an equivalent dataset to measure vessel traffic activity from the International Comprehensive Ocean-Atmospheric Data Set (ICOADS). This data is derived from the Voluntary Observing Ships' (VOS) scheme principally for the collection of meteorological observations from ships. Crucially, vessels report their position at six-hourly intervals which enables the mapping of vessel activity. Data from January 2010 to May 2020 was extracted, filtered to vessels only and processed into 23.5 million unique positions. It should be noted that this data is collected by only a small proportion of the global fleet, while over-representing some vessel types [10].

The incident data was obtained from the International Maritime Organization's (IMO) Global Integrated Shipping Information System (GISIS) for the years 2005-2017. A more detailed analysis of this data has been conducted by [4]. IMO's GISIS contains only a small portion of the total number of accidents which are available from commercial or national sources. However, the advantages for this study is its greater transparency and global coverage. In total, 3084 incidents were utilized form the acquired data sources, that is 237 incidents per year. Further analysis could be conducted to consider incidents by vessel type or accident type, however, we considered only the total number of accidents in this current study.

2.2. Method

In order to develop incident rates, the vessel traffic and incident locations were processed using the dggridpy python DGGS library which was developed at the University of Southampton (https://github.com/correndo/dggridpy).

DGGS can be constructed in different ways, and in this study we utilize a hexagonal DGGS with aperture 4 (ISEA4H), indicating that each change of resolution increases the number of cells by a factor of 4. Multi-resolution analysis was conducted between DGGS at resolution 2 (DGGS2) with 162 cells (area 3.2 million km^2) to DGGS7 with 164,000 cells (area 3,113 km^2). No further resolutions were analyzed due to the limited precision of the coordinates in the ICOADS dataset.

For each resolution, the total number of vessel positions and incidents in each cell were divided to derive a rate of incidents per vessel position. This enabled the mapping of incident rates and respective correlations of vessel activity to incident occurrence.

3. RESULTS AND DISCUSSION

Figure 2 shows the density of traffic, location of incidents and derived incident rates respectively. Firstly, the major global shipping routes are evident across each ocean, with the highest densities in major waterways such as the English Channel, Panama/Suez Canals and Malacca Straits, as well as key ports. Secondly, it is evident that the majority of incidents are not recorded in open ocean but rather coastal waters where vessels navigate closer together and to shore, while increasing the risk of incident. The waters in the North Sea, Baltic Sea, Mediterranean Sea and East Asia have a high number of incidents, which support previous work findings [4].

Thirdly, the incident rates at DGGS4 show a different pattern, with Europe exhibiting an incident rate which is not substantially different to other locations. The highest incident rates are where there is relatively little traffic, but an incident has occurred. For example, some hotspots are the result of incidents involving fishing vessels which might be under-represented in the ICOADS dataset.

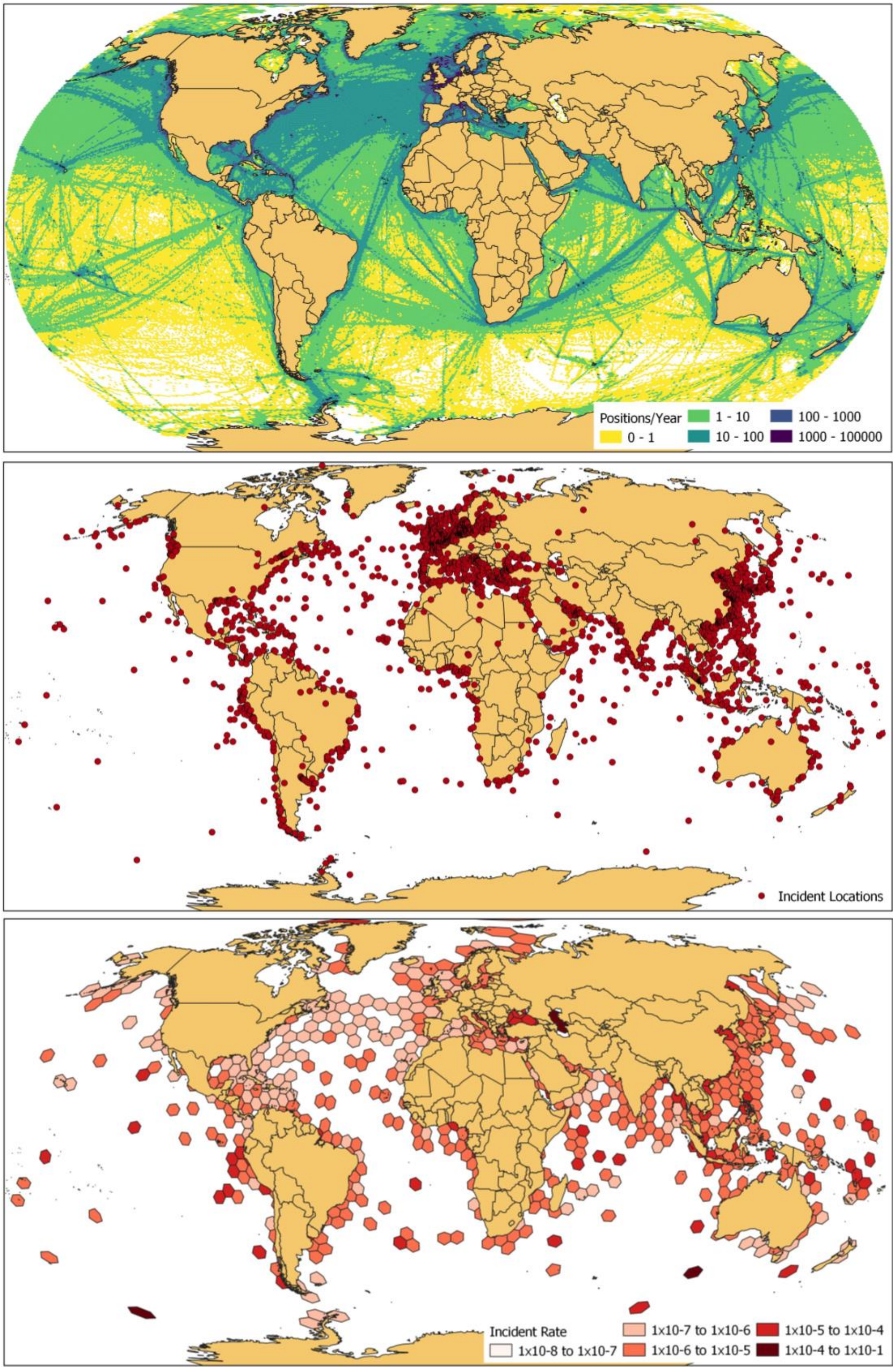

Figure 2: Top: Vessel Traffic at DGGS7, Middle: Incident Locations, Bottom: Incident Rate at DGGS4.

In addition, the number of ship positions and incidents can be compared for each cell. In general, this relationship is positive with more traffic increasing the number of incidents. However, as the resolution changes, so too does the strength of this relationship. At DGGS2, the Pearson's correlation is 0.7, 0.55 at DGGS4 and 0.24 at DGGS7. This is a widely recognized effect of the MAUP, that as the number of areal units representing data decreases, the correlation between them increases [11]. The importance of managing the scale of units must be considered in conducting aggregated spatial analysis.

A further effect of changing scale is that at finer resolutions, there is less traffic in each cell driving higher incident rates. It would be expected that there would also be fewer incidents, however, the relatively low number of incidents means that the majority of cells eventually reach only one incident, which cannot be further subdivided. At lower resolutions (DGGS2-DGGS5), the average incident rate is consistent between 3.89×10^{-5} and 5.66×10^{-5}, but as the resolution increases to DGGS6/DGGS7, the average incident rate increases to $1.82 \times 10^{-4}/2.49 \times 10^{-4}$ respectively.

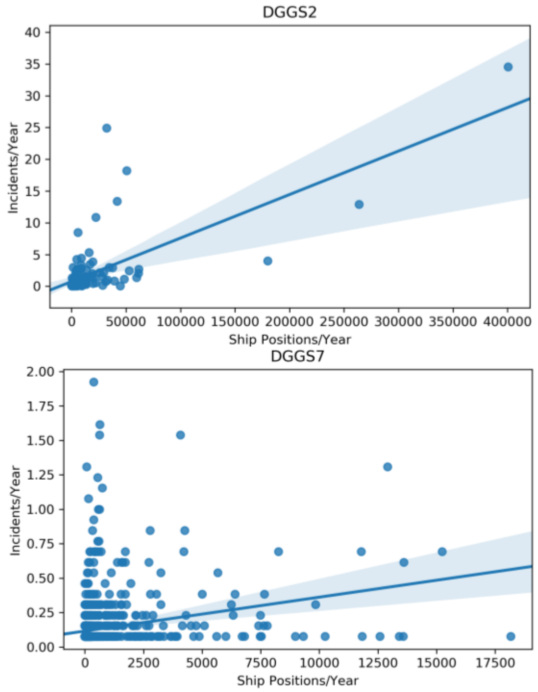

Figure 3: Incident Rates at Different Resolutions.

Finally, attention is drawn to issues of data quality in deriving incident rates. In Section 2, we have described potential issues of the use of ICOADS data that both under-represent some vessel types with limited precision. In addition, the GISIS incident data is noted to contain numerous errors in how the data is recorded, most significantly in the location. We have not sought to correct this data, nor have other analyses [4], but would inevitably result in spurious incident rates in some locations.

4. CONCLUSION

In this paper, the spatial distribution of vessel traffic and incidents have been analyzed to produce a global map of incident rates. Whilst this work can be expanded through further analysis by vessel type and incident type, it contributes to a greater understanding of where and why accidents occur, which supports decision makers in planning mitigation measures to manage navigation safety more effectively. Furthermore, we demonstrated that a DGGS is better suited to aggregating spatial data at a global scale, for overcoming issues with conventional cartesian grids. These might result in misleading conclusions being drawn concerning the evaluation of ship navigation risks.

5. REFERENCES

[1] R. Bye and P. Almklov, "Normalization of maritime accident data using AIS", *Marine Policy*, 109, 103675, 2019.

[2] P.T. Pedersen, "Collision and Grounding Mechanics", *Proceedings of WEMT 1995*, Copenhagen, Denmark, 1995.

[3] L. Wu, Y. Xu, Q. Wang, F. Wang, and Z. Xu, "Mapping Global Shipping Density from AIS Data", *Journal of Navigation*, 70, pp. 67-81, 2017.

[4] Y. Zhang, X. Sun, J. Chen and C. Cheng, "Spatial patterns and characteristics of global maritime accidents" *Reliability Engineering and System Safety*, 206, 107310, 2021.

[5] K. Kulkarni, F. Goerlandt, J. Li, O.V. Banda and P. Kujala "Preventing shipping accidents: Past, present, and future of waterway risk management with Baltic Sea focus" *Safety Science*, 129, 104798, 2020.

[6] S. Openshaw, "Optimal zoning systems for spatial interaction models" *Environment and Planning A*, 9, pp. 169-184, 1977.

[7] K. Sahr, D, White and A. Kimerling, "Geodesic Discrete Global Grid Systems" *Cartography and Geographic Information Science*, 30, pp. 121-134, 2003.

[8] S. Battersby, D. Stebe, and M. Finn, "Shapes on a Plane: evaluating the impact of projection distortion on spatial binning" *Cartography and Geographic Information Science*, 44, pp. 410-421, 2016.

[9] V. Eguiluz, J. Fernandez-Gracia, X. Irigoien and C. Duarte, "A quantitative assessment of Arctic shipping in 2010-2014" *Scientific Reports*, 6, 2016.

[10] C. Wang, J. Corbett and J. Firestone, "Improving Spatial Representation of Global Ship Emissions Inventories" *Environmental Science Technology*, 42, pp. 193-199, 2008.

[11] C. Gehlke and K. Biehl, "Certain effects of grouping upon the size of the correlation in census tract material" *Journal of American Statistical Society*, 29, pp. 169-170, 1934.