How much of the seabed is impacted by mobile fishing gear? Absolute estimates from Vessel Monitoring Systems (VMS) point data

Hans D Gerritsen1\*, Cóilín Minto2 and Colm Lordan1

1 Marine Institute, Rinville, Oranmore, Co Galway, Ireland

2 Marine and Freshwater Research Centre, Galway-Mayo Institute of Technology, Dublin Road, Galway, Ireland.

\* corresponding author: email: hans.gerritsen@marine.ie; phone: +353 91 387297; fax: +353 91 387201

Abstract

Demersal trawling impacts extensively on the seabed and the extent and frequency of this impact can be assessed using Vessel Monitoring Systems (VMS) data (positional data of fishing vessels). Existing approaches interpolate fishing tracks from consecutive VMS locations (track interpolation) and/or aggregate VMS point data in a spatial grid (point summation). Track interpolation can be quite inaccurate at the current 2-hour time interval between VMS records, leading to biased estimates. Point summation approaches currently only produce relative estimates of impact and are highly sensitive to the grid size chosen We propose an approach that provides absolute estimates of trawling impact from point data and is not sensitive to an arbitrary choice of grid cell size. The method involves applying a nested grid and estimating the swept area (area covered by fishing gear) for each VMS point. We show that the ratio of the swept area to the surface area of a cell can be related to the proportion of the seabed that was impacted by the gear a given number of times. We validate the accuracy of this swept-area ratio approach using known vessel tracks and apply the method to international VMS data in the Celtic Sea.

Keywords:

Automatic Identification System (AIS); Impact of trawling; Marine Strategy Framework Directive (MSFD); Nested grid; Swept-area ratio; Vessel Monitoring Systems (VMS)

Introduction

Demersal trawling is probably the most extensive human activity that impacts on the seabed (Eastwood et al. 2007, Foden et al. 2011). Trawling has both direct and indirect effects on benthic ecosystems and the severity and longevity and complex interactions involved are receiving increasing attention (Hiddink et al. 2006, Lambert et al. 2011). Policy developments such as the Ecosystem approach to fisheries management (FAO 2008) and the EU Marine Strategy Framework Directive (EC 2008) oblige member states to provide indicators that include the quantification of the impact of fishing on the seabed. These indicators should describe how much of the seabed is impacted by trawling and how often it is impacted.

The widespread implementation of Vessel Monitoring Systems (VMS) for surveillance purposes has, as a by-product, given scientists access to a rich dataset of fishing vessel positional data. These data have allowed major progress towards the goal of quantifying the distribution and intensity of trawling. Fishing vessels fitted with VMS transmit their position and speed at regular time intervals; in EU waters the maximum time interval between transmissions is 2 hours and since 2005 all fishing vessels >15m are required to carry VMS (EC 2003); since 2012 this has been extended to all fishing vessels >12m (EC 2009).

There are two existing approaches to estimating the area impacted by fishing gear: track interpolation and point summation methods. Track interpolation methods aim to re-construct vessel tracks between consecutive VMS points. This can be done using a straight line or a spline curve (Hintzen et al. 2010, Russo et al. 2011, and references therein). Skaar et al. (2011) found that straight line interpolations at 2-hour intervals deviated more than 3km from the real track for the majority of hauls of two Norwegian demersal trawlers. Lambert et al. (2012) found that both straight-line and spline interpolations deviated around 2km from the real track for Isle of Man scallop dredgers and trawlers. These findings show that there is considerable uncertainty in the vessel position during the 2-hour interval between VMS records. So while interpolated tracks can be used to estimate the likelihood that a location is trawled, they cannot be used to accurately estimate the number of times a location is trawled (Hintzen et al. 2010).

The other main existing approach, point summation, involves applying a grid to all VMS point locations where the vessels were deemed to be fishing (Lee et al. 2010). Each grid cell that contains VMS points is then considered to be impacted by trawling. This approach is generally applied to the observed VMS points but can also be applied to interpolated VMS points, where interpolated fishing tracks are converted back into a sequence of point locations. An important problem with the point summation approach is that the proportion of impacted cells is strongly dependent on the grid size chosen (Dinmore et al. 2003, Piet and Quirijns 2009, Hinz et al. 2012, Lambert et al. 2012, Piet and Hintzen 2012) because larger grid cells are less likely to be completely free of trawling activity than smaller cells. Related to this is the issue that generally only a part of each cell is impacted by fishing gear; for large cells with a small amount of effort this may only be a small proportion of the cell area. We are not aware of any studies that take this into account.

We modified the existing point summation approach by addressing its sensitivity to an arbitrary choice of grid cell size and by accounting for the fact that cells with fishing effort may be only partially impacted by fishing gear. We apply the proposed method to VMS data from 2011 in the Celtic Sea (south of Ireland) to illustrate the approach.

Methods

*VMS data and study area*

VMS data were available for all fishing vessels > 15m in total length inside the Irish Exclusive Economic Zone (EEZ). We used data from the most recent year available (2011) in the Celtic Sea (ICES Divisions VIIg and VIIj). We chose this region because it contains a broad range of habitats and it displays strong spatial structures in the distribution of fishing effort. Because no deepwater fishing takes place in this area, we excluded areas deeper than 800m. The study area covers an area of 90367 km2 and is shown in Figure 4. Only mobile bottom-impacting gears were included in the analysis: demersal otter trawls, beam trawls and dredges. Seines were not considered to be bottom-impacting gears for the purpose of this study because their impact is relatively minor (ICES 2000).

The VMS data itself contain no information on the gear type used. We used two sources of information to determine which gear was used: EU logbooks (EEC 1983) and the Community Fishing Fleet register (<http://ec.europa.eu/fisheries/fleet>). Logbooks document the gear used by each vessel for each day it was fishing. These data were available for all Irish vessels and for foreign vessels landing into Ireland (see Gerritsen and Lordan (2011) for further description of linking logbooks and VMS data). For all other vessels we obtained their main gear type from the Community Fishing Fleet register. Note that this may not be the actual gear used if a vessel uses more than one gear. The gear type could not be established for 7% of the VMS records; we assumed that missing gears occurred randomly and multiplied the remaining effort values by a factor of 1/(1-0.07) to account for the missing data. Demersal otter trawlers accounted for the vast majority of the effort of bottom-impacting gears in the area (90%).

Each VMS record was assigned an effort value that was equal to the time interval since the previous record (generally 2 hours). Records with time intervals larger than 4 hours were given an effort value of 4 hours. Vessels using otter trawls were assumed to be fishing if their instantaneous speed ranged between 0.5 kn and 4.5 kn. Gerritsen and Lordan (2011) found that these speed criteria correctly identify whether a vessel was fishing with 88% accuracy. We assumed fishing speeds between 0.5 kn and 5.5 kn for beam trawls and between 0.5 kn and 4.0 kn for dredges. These criteria were based on the frequency distribution of the vessel speed for these gears. Only records where the vessel was assumed to be fishing were included in the analysis.

For otter trawls, we assumed that the width of the gear that impacts the seabed was equal to the door spread. Data on door spread is not routinely recorded in the EU logbooks so we used data obtained from personal contacts within the industry and from engineering trials conducted by the Irish Sea Fisheries Board (BIM) to estimate the average door spread. For otter trawlers operating in areas where *Neprhops* are targeted, we applied a mean door spread of 60m and for all other otter trawlers we used 100m. (The locations of the *Nephrops* grounds are described by Gerritsen et al. (2012).) The width of beam trawls and dredges is generally recorded in the logbooks: the average width was 18m for beam trawls and 16.5m for dredges. The estimated mean widths of the gears are based on Irish vessels only, but they were applied to all vessels in the dataset.

*Nested grid*

Because VMS points tend to be highly clustered, the number of observations in each grid cell will vary greatly in any regular grid. This is undesirable because cells with low numbers of observations will have a low precision; while cells with high number of observations will have high precision. Therefore a coarse grid will be more appropriate in areas with few data points, while a fine grid (high spatial resolution) can be applied to areas with high numbers of observations. To achieve this, we apply a system with nested grid cells. The approach is as follows: start with a rectangular grid of a coarse size; if there are 20 or more observations in a cell then split the cell in two; repeat this procedure until no more cells can be divided or until a specified minimum cell size has been reached. The choice of a threshold of 20 observations per cell is somewhat arbitrary and reflects a compromise between precision and spatial resolution. We chose rectangles with a height-to-width ratio that was approximately equal to so the aspect ratio of rectangles of all sizes would be constant (analogous to the way an ISO A4 sheet of paper can be divided into two A5 sheets each with the same aspect ratio as the original A4 sheet). We chose a starting grid size of 0.16° latitude by 0.18° longitude (approx. 220km2) and the smallest grid size (after 11 divisions) was approx. 0.11km2.

*Measure of randomness*

It is important for our analysis to quantify whether the points within a grid cell are randomly distributed or clustered. Rijnsdorp et al. (1998) achieved this by calculating the variance of the density of points in a cell and dividing by the expected variance of the Poisson distribution. We could not apply this method because each cell had a relatively low number of observations. Instead we calculated the mean distance () of each point in the cell to its nearest neighbour (which might be in another cell). If the points are distributed at random, the expected mean nearest neighbour distance is:

(1)

where *ρ* is the density of points (Clark and Evans 1954). The ratio of the observed () and expected () mean distance is therefore expected to be 1 for randomly distributed points. A ratio of 0 indicates that the points are maximally aggregated (they are all in the same location). A ratio of higher than 1 indicates that the points are more evenly distributed than expected. Clark and Evans (1954) show that the highest possible value of the ratio is 2.1491, which would occur in an even hexagonal pattern.

*Swept-area ratio*

In order to quantify the area of the seabed that is impacted by fishing gear, we first estimate the area swept by fishing gear in each grid cell as follows:

(2)

where *i* is the VMS record (1,…n), *e* is the effort (the time interval since the previous record in hours); *v* is the instantaneous vessel speed (in km/h) and *w* is the width of the gear that impacts the seabed (in km). For each cell we can then calculate the swept-area ratio λ, which is the swept area (*SA*) divided by the cell area (*CA*).

(3)

The swept-area ratio of a cell can also be interpreted as the mean number of times the seabed in the cell was impacted by fishing gear. A swept-area ratio of one indicates that the swept area equals the cell area; however that does not mean that 100% of the cell is impacted by fishing gear (unless the tracks do not overlap at all). If we assume that the tracks are distributed randomly, we can derive an equation that closely describes the relationship between the swept-area ratio and the proportion of the cell that was impacted by gear (see online supplementary material S1 for the derivation). The proportion *p* of the cell that is impacted *k* times is given by:

(4)

It follows from equation (4) that the proportion of the cell that is impacted at least once is:

(5)

We provide R-code in the online supplementary material S2 that allows the user to apply a nested grid and estimate the area covered by fishing gear for a simulated dataset.

*Validation of the method*

The swept-area ratio approach was applied to the Celtic Sea VMS dataset but in order to validate the approach, we used an additional test dataset of known vessel tracks from Automatic Identification System (AIS) data that were received by Irish base stations. AIS data contain vessel position and speed, like VMS data, but are transmitted by VHF radio, rather than satellite. Only a small proportion of fishing vessels are equipped with AIS and spatial coverage of the data is incomplete because data are only recorded if a vessel is within VHF range of a base station. However the data have a very high temporal resolution (the most common time interval between records is 10 seconds), therefore the data can serve as a test dataset with accurately known vessel tracks.

AIS data were available to us for 157 days in 2011, mostly within 100km from the Irish coast. The dataset included 504 fishing vessels. We selected three regions with high levels of fishing effort of demersal otter trawlers: a region where monkfish, megrim, hake are mainly targeted (on the continental shelf to the south-west of Ireland) and two regions where *Nephrops* are targeted (the Aran grounds to the west of Ireland and the Irish sea to the east of Ireland). Table 1 gives some summary statistics for these regions and Figure 2d shows their locations. In each region, we re-sampled the data at 5-minute intervals and calculated the average vessel speed between records under the assumption that the vessel travelled in a straight line at a constant speed during the 5-minute interval (we found that intervals of <5 minutes resulted in imprecise speed estimates). We then identified fishing tracks as sets of records where the calculated speed remained between 0.5 kn and 4.5 kn continuously for a distance of at least 1nm. Each fishing track was given a width of 60 or 100m depending on the dominant fishery in the area by applying a spatial buffer to the track lines (60m for the *Nephrops* areas and 100m for the monkfish, megrim and hake area). Next we sampled 10 000 random locations inside the area and for each point we counted the number of track polygons that covered that point. From this we estimated the ‘true’ area that was impacted a given number of times.

We also applied the swept-area ratio approach to the test data by re-sampling the AIS fishing tracks every 2 hours to mimic a VMS dataset. We applied the nested grid cell algorithm to the re-sampled data (same settings as for the Celtic Sea VMS data) and for each grid cell we calculated the swept-area ratio and applied equation (4). In order to investigate the sensitivity of the method to the VMS time interval, we repeated the procedure with 15min and 8h intervals.

Results

*Nested grid*

Figure 1 illustrates how grid cells are divided into progressively smaller cells if there are 20 or more observations in the cell. This leads to grid cells that have similar numbers of observations. Although the VMS points in the figure are clearly clustered, the distribution of the points inside individual grid cells does not show any obvious clustering.

*Validation of the method*

We compared the impact of the fishing gear in each of the AIS areas estimated from the actual AIS tracks to the estimates from the swept-area ratio approach (Figure 2a-c). The swept-area ratio estimates agree closely to the direct estimates and the results are virtually independent of the re-sampling frequency: re-sampling every 15min only gave slightly more accurate results than re-sampling every 8 h. There appeared to be a minor, but consistent, bias: the swept-area ratio approach tends to slightly over-estimate the area that is impacted once and under-estimate the area that is impacted more than once. This is probably due to a small amount of non-randomness in the vessel tracks; they have a higher degree of overlap than expected by chance.

*Celtic Sea VMS data*

We applied the nested grid cell algorithm to the Celtic Sea VMS data which resulted in an average of 13 VMS records per cell (70% of cells had more than 10 records). The majority of cells were between 7 and 1.8km2 in size (Table 2). Figure 3 shows the distribution of the nearest neighbour distance ratio for each cell size. Cells between 7 and 1.8km2 in size had median ratios that were just below one, indicating a nearly random distribution. Large cells showed a higher degree of clustering (the mean nearest neighbour distance in these cells was smaller than expected by chance). However these cells only account for a very small proportion of the effort (Table 2). The smallest cells also showed an increasing degree of clustering. This is probably an artefact of the limited precision of the VMS positional data; in small cells the VMS points are likely to lie exactly on top of each other because there is insufficient precision in the data to distinguish their exact location. The mean nearest neighbour ratio of all cells was 0.95.

Figure 4 shows the swept-area ratio estimated from the VMS data for each grid cell in the study area. The fishing activity is highly structured and most of the study area has a swept-area ratio of more than 1. Equation (4) allows us to translate the swept-area ratios in each of the grid cells to the proportion of each cell that is impacted at least *k* times (Figure 5). By multiplying these proportions to the surface area of the cells we can estimate that 68% of the study area was impacted at least once during 2011. A considerable portion of the area (46%) was impacted at least twice and 13% of the area was impacted at least five times, particularly along the continental shelf edge and on the mud patches where *Nephrops* occur. Some of these regions were even impacted 10 times or more, although this occurred in less than 2% of the area.

Discussion

An accurate estimate of the number of times the seabed is impacted is a very important pressure indicator. These estimates are essential for managing the overall impact on benthic biomass, production and diversity (Hiddink et al. 2006). We have shown that absolute estimates can be obtained without interpolating vessel tracks.

*Grid size*

When effort values are assigned to VMS locations which are then aggregated in a grid, it is sometimes argued that the grid size should not be smaller than the maximum distance a vessel can travel between consecutive VMS points. This would prevent a situation where a vessel can travel through grid cells without registering a VMS record. For example, Fock (2008) states that for his study a 3 nm x 3 nm grid was appropriate because most interval distances were less than 3nm. In fact, there is no reason why the grid size could not be smaller than that, as long as there are a sufficient numbers of observations in the grid cells. Because the time of the first VMS transmission on any trip is effectively random, the VMS positions consist of systematic random sample along the vessel track. Therefore, the precision of the effort estimate in each cell is not directly determined by the size of the cell, but rather by the number of observations in the cell.

*Nested grid*

By applying a nested grid, maps can be constructed that show a large amount of detail in areas with high fishing effort while allowing lower spatial resolution for areas where data are more sparse. One important advantage of this approach is that within small cells, the distribution of the point data is close to random. If we assume that the tracks are therefore also randomly distributed at this scale, then we can accurately estimate the proportion of each cell that was impacted by trawl gear. For larger cells the distribution of the points tends to be more clustered than expected by chance but these cells typically account for a very small proportion of the effort (in the Celtic Sea study area only 0.5% of the effort was in cells that were larger than 28 km2). The apparent higher level of clustering of the smallest cells (Figure 3) is likely to be an artefact of rounding of the positional data.

*Gear parameters*

In this study we assumed that the trawl doors, clumps, sweeps, bridles and groundgear of otter trawls, all impact on the seabed. Eastwood et al. (2007) and Stelzenmüller et al. (2008) only accounted for the impact of the trawl doors. Either approach may be valid, depending on the question that is being asked. The actual direct impact will depend on the gear design, amount of bottom contract and on the bottom type. It is not the purpose of this study to address the impact on the benthic community, but we do provide the tools to do so in the future.

For otter trawlers there can be a large amount of variation in the door spread. Many vessels are now equipped with gear sensors and it should be possible to collect detailed information on door spread, wing spread, length and type of footrope, net type, etc. to allow a more accurate estimation of the swept area. In the current EU logbook system there is no requirement to record these parameters and as far as we are aware most fishery observer programmes do not collect this information either.

We tested the sensitivity of our method to the assumed gear with and found that it is reasonably robust: if we double the assumed width of the gear, the estimated area that is impacted at least once is only increased by 19%; halving the gear width resulted in a reduction of 25% in the estimated area.

*Validation of the method*

The availability of AIS data has allowed us to validate equation (4) by comparing known vessel tracks to the estimated area impacted using the swept-area ratio approach. We did not attempt to identify the gears used in the AIS data but we selected areas that are known to be heavily dominated by demersal otter trawls. The spatial and vessel coverage of the AIS data is incomplete but the data represent a significant amount of fishing effort from dozens of vessels (Table 1). We therefore conclude that this dataset is appropriate for validating our method. For all three AIS regions the estimated area impacted agreed very closely to the true area (Figure 2a-c). This gives us considerable confidence that the method can give absolute estimates of the area impacted by fishing gear.

*Advantages of the swept-area approach*

In contrast to previous point summation methods whose outcomes depend on an arbitrary choice of grid size (Dinmore et al. 2003, Witt and Godley 2007, Piet and Quirijns 2009, Lee et al. 2010, Hinz et al. 2012, Lambert et al. 2012), our approach can provide not just an index but absolute estimates of the area impacted and the distribution of this impact. Our method also has a number of advantages over track interpolation approaches. Interpolation methods do not perform well with 2-hour intervals, which too infrequent to accurately reconstruct the vessel track (Deng et al. 2005, Skaar et al. 2011, Lambert et al. 2012) and it is often stated that there is a need for shorter intervals between VMS records. For the swept-area approach, the time interval between VMS records has an almost negligible effect on the accuracy of the outcome. As long as the number of observations is high enough to allow for small grid cells with randomly distributed points inside these cells, then the estimates are virtually unbiased. Therefore, a decrease in the reporting interval of VMS data is not necessary for our approach, although it would be useful for more accurate classification of fishing activity (e.g. Bertrand et al. 2008, Vermard et al. 2010) on which our results are contingent.

*Conclusion*

We provide a new method to assess the frequency and spatial extent of impact by mobile bottom gears on the seabed at an appropriate spatial scale, based on VMS point data. Our method does not require the interpolation of vessel tracks and therefore only uses data that correspond to actual vessel locations. The method is not sensitive to the time interval between VMS transmissions. Furthermore, it is easy to implement and is not particularly computationally intensive. The approach does require relatively high densities of VMS data so that cell sizes can be so small that the points within each cell are not significantly clustered.

Supplementary material

Supplementary material is available at the ICESJMS online version of the paper. S1 contains the derivation of equation (4) and S2 contains example R code that allows the user to set up and plot a nested grid and calculate the areas impacted by fishing gear.

Acknowledgements

We thank the Irish Naval Service for supplying the VMS data, the Sea-Fisheries Protection Authority for supplying the logbook data and the Irish Coastguard, Department of Transport, Tourism and Sport for supplying AIS data. All data were aggregated to a level that did not compromise the confidentiality of individual vessels. We also thank two anonymous reviewers for their constructive comments.

References

Bertrand, S., Díaz, E., and Lengaigne, M. 2008. Patterns in the spatial distribution of Peruvian anchovy (*Engraulis ringens*) revealed by spatially explicit fishing data. Progress in Oceanography **79**: 379–389.

Clark, P.J., and Evans, F.C. 1954. Distance to nearest neighbor as a measure of spatial relationships in populations. Ecology **35**: 445–453.

Deng, R., Dichmont, C.M., Milton, D., Haywood, M., Vance, D., Hall, N., and Die, D. 2005. Can vessel monitoring system data also be used to study trawling intensity and population depletion? The example of Australia’s northern prawn fishery. Can. J. Fish. Aquat. Sci. **62**: 611–622.

Dinmore, T.A., Duplisea, D.E., Rackham, B.D., Maxwell, D.L., and Jennings, S. 2003. Impact of a large-scale area closure on patterns of fishing disturbance and the consequences for benthic communities. ICES Journal of Marine Science **60**: 371–380.

Eastwood, P.D., Mills, C.M., Aldridge, J.N., Houghton, C.A., and Rogers, S.I. 2007. Human activities in UK offshore waters: an assessment of direct, physical pressure on the seabed. ICES J. Mar. Sci. **64**: 453–463.

EC. 2003. Commission Regulation (EC) No.2244/2003 of 18 December 2003 laying down detailed provisions regarding satellite-based Vessel Monitoring Systems. Official Journal of the European Union **L 333**: 17–27.

EC. 2008. Directive 2008/56/EC of the European parliament and of the council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). Official Journal of the European Union **L 164/19**: 40.

EC. 2009. Council Regulation (EC) No 1224/2009 of 20 November 2009 establishing a Community control system for ensuring compliance with the rules of the common fisheries policy. Official Journal of the European Union **L 343/1**.

EEC. 1983. Commission Regulation (EEC) No 2807/83 of 22 September 1983 laying down detailed rules for recording information on Member States’ catches of fish. Official Journal of the European Union **L 276**: 1–18.

FAO. 2008. Fisheries management. 2. The ecosystem approach to fisheries. 2.1 Best practices in ecosystem modelling for informing an ecosystem approach to fisheries. *In* FAO Fisheries Technical Guidelines for Responsible Fisheries. No. 4, Suppl. 2, Add. 1. FAO, Rome. p. 78pp.

Fock, H.O. 2008. Fisheries in the context of marine spatial planning: Defining principal areas for fisheries in the German EEZ. Marine Policy **32**: 728–739.

Foden, J., Rogers, S.I., and Jones, A.P. 2011. Human pressures on UK seabed habitats: a cumulative impact assessment. Mar Ecol Prog Ser **428**: 33–47.

Gerritsen, H.D., and Lordan, C. 2011. Integrating Vessel Monitoring Systems (VMS) data with daily catch data from logbooks to explore the spatial distribution of catch and effort at high resolution. ICES J. Mar. Sci. **68**: 245–252.

Gerritsen, H.D., Lordan, C., Minto, C., and Kraak, S. 2012. Spatial patterns in the retained catch composition of Irish demersal otter trawlers: High-resolution fisheries data as a management tool. Fish Res **129-130**: 127–136.

Hiddink, J.G., Jennings, S., Kaiser, M.J., Queirós, A.M., Duplisea, D.E., and Piet, G.J. 2006. Cumulative impacts of seabed trawl disturbance on benthic biomass, production, and species richness in different habitats. Can. J. Fish. Aquat. Sci. **63**: 721–736.

Hintzen, N.T., Piet, G.J., and Brunel, T. 2010. Improved estimation of trawling tracks using cubic Hermite spline interpolation of position registration data. Fisheries Research **101**: 108–115.

Hinz, H., Murray, L.G., Lambert, G.I., Hiddink, J.G., and Kaiser, M.J. 2012. Confidentiality over fishing effort data threatens science and management progress. Fish and Fisheries. doi: doi: 10.1111/j.1467-2979.2012.00475.x.

ICES. 2000, April 10. Report of the working group on fishing technology and fish behaviour. Annex 2 - Using gear technology to understand and reduce unintended effects of fishing on the seabed and associated communities: background and potential directions.

Lambert, G.I., Jennings, S., Kaiser, M.J., Hinz, H., and Hiddink, J.G. 2011. Quantification and prediction of the impact of fishing on epifaunal communities. Marine Ecology Progress Series **430**: 71–86.

Lambert, G.I., Jennings, S., Hiddink, J.G., Hintzen, N.T., Hinz, H., Kaiser, M.J., and Murray, L.G. 2012. Implications of using alternative methods of vessel monitoring system (VMS) data analysis to describe fishing activities and impacts. ICES Journal of Marine Science **69**: 682-693.

Lee, J., South, A.B., and Jennings, S. 2010. Developing reliable, repeatable and accessible methods to provide high-resolution estimates of fishing-effort distributions from Vessel Monitoring System (VMS) data. ICES J. Mar. Sci. **67**: 1260–1271.

Piet, G.J., and Hintzen, N.T. 2012. Indicators of fishing pressure and seafloor integrity. ICES Journal of Marine Science **69**: 1850–1858.

Piet, G.J., and Quirijns, F.J. 2009. The importance of scale for fishing impact estimations. Can. J. Fish. Aquat. Sci. **66**: 829–835.

Rijnsdorp, A.D., Buys, A.M., Storbeck, F., and Visser, E.G. 1998. Micro-scale distribution of beam trawl effort in the southern North Sea between 1993 and 1996 in relation to the trawling frequency of the sea bed and the impact on benthic organisms 10.1006/jmsc.1997.0326. ICES J. Mar. Sci. **55**: 403–419.

Russo, T., Parisi, A., and Cataudella, S. 2011. New insights in interpolating ﬁshing tracks from VMS data for different métiers. Fish Res **108**: 184–194.

Skaar, K.L., Jørgensen, T., Ulvestad, B.K.H., and Engås, A. 2011. Accuracy of VMS data from Norwegian demersal stern trawlers for estimating trawled areas in the Barents Sea. ICES J. Mar. Sci. **68**: 1615–1620.

Stelzenmüller, V., Rogers, S.I., and Mills, C. 2008. Spatio-temporal patterns of fishing pressure on UK marine landscapes, and their implications for spatial planning and management. ICES J. Mar. Sci. **65**: 1081–1091.

Vermard, Y., Rivot, E., Mahévas, S., Marchal, P., and Gascuel, D. 2010. Identifying fishing trip behaviour and estimating ﬁshing effort from VMS data using Bayesian Hidden Markov Models. Ecological Modelling **221**: 1757–1769.

Witt, M.J., and Godley, B.J. 2007. A Step Towards Seascape Scale Conservation: Using Vessel Monitoring Systems (VMS) to Map Fishing Activity. PLoS ONE **2(10): e1111**. doi: doi:10.1371/journal.pone.0001111.

Tables

*Table 1. Summary statistics of the three areas with AIS data that were used to validate the swept-area method.*

|  |  |  |  |
| --- | --- | --- | --- |
| Region | Continental shelf | Aran | Irish Sea |
| Main fishery | Demersal otter trawls targeting monkfish, megrim and hake | Demersal otter trawls targeting *Nephrops* | Demersal otter trawls targeting *Nephrops* |
| Boundaries | 12°00′W - 10°30′W  51°00′N - 52°30′N | 10°30′W - 9°45′W  52°45′N - 53°15′N | 6°15′W - 5°00′W  53°30′N - 54°00′N |
| Surface area (km2) | 17 200 | 2 787 | 4 563 |
| Number of vessels | 84 | 27 | 44 |
| Fishing effort (h) | 11 503 | 4 158 | 17 291 |
| Assumed gear width (m) | 100 | 60 | 60 |

*Table 2. Summary statistics for the cells in the nested grid applied to the Celtic Sea VMS data.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Cell height (°lat) | Cell width (°lon) | Cell area (km2)\* | Number of cells | Number of vms points | Total effort (hours) |
| 0.16 | 0.18 | 220 | 11 | 85 | 128 |
| 0.08 | 0.18 | 110 | 19 | 245 | 395 |
| 0.08 | 0.09 | 55 | 68 | 739 | 1 203 |
| 0.04 | 0.09 | 28 | 244 | 3 022 | 5 652 |
| 0.04 | 0.045 | 14 | 1 035 | 14 343 | 27 319 |
| 0.02 | 0.045 | 7 | 4 051 | 53 253 | 99 311 |
| 0.02 | 0.0225 | 3.5 | 5 860 | 73 916 | 129 593 |
| 0.01 | 0.0225 | 1.8 | 4 106 | 48 501 | 77 080 |
| 0.01 | 0.01125 | 0.88 | 1 834 | 21 845 | 30 590 |
| 0.005 | 0.01125 | 0.44 | 647 | 6 945 | 9 241 |
| 0.005 | 0.005625 | 0.22 | 115 | 1 181 | 1 201 |
| 0.0025 | 0.005625 | 0.11 | 75 | 3 232 | 3 764 |

\* The cell area varies slightly with latitude so the average area is given

Figures

H:\My Documents\My Publications\VMS3\IJMS final\Figure1.tif

*Figure 1. Example of a nested grid. The points correspond to VMS records. Cells with 20 or more VMS records are recursively divided in two.*

*H:\My Documents\My Publications\VMS3\IJMS final\Figure2.tif*

*Figure 2. Validation of the swept-area ratio approach in the three AIS regions: Aran (a); Irish Sea (b) and Continental shelf (c). The solid line represents the area impacted* k *times as estimated from the actual AIS vessel tracks. The symbols represent the estimates from the swept-area ratio approach that resulted from re-sampling the AIS data at 15min, 2h and 8h intervals. The bottom-right panel (d) shows the location of the three regions; dark areas indicate fishing activity.*

*H:\My Documents\My Publications\VMS3\IJMS final\Figure3.tif*

*Figure 3. Nearest neighbour distance ratio for cells varying in size from 220km2 to 0.11km2. A ratio of one indicates a random uniform distribution.*

*H:\My Documents\My Publications\VMS3\IJMS final\Figure4.tif*

*Figure 4. The swept-area (SA) ratio (the mean number of times per year that each cell was traversed by fishing gear) in the study area during 2011. The study area is bounded by the Irish EEZ in the south-east and by the 800m depth contour in the west.*

*H:\My Documents\My Publications\VMS3\IJMS final\Figure5.tif*

*Figure 5. The proportion of each grid cell that was impacted by fishing gear at least once (a), twice (b), five times (c) and ten times (d).*

**Supplementary material S2**

Derivation of trawl coverage probabilities

Definitions

|  |  |
| --- | --- |
| Symbol | Definition |
|  | Area of a tow |
|  | Total area of the cell |
|  | Coverage probability () |
|  | Non-coverage probability () |
|  | Number of tows |
|  | Number of times covered |

Derivation

For one tow, the probability of a point within a grid cell being covered by the tow is , the probability that it is not covered by a tow is  so we have



For two independent tows, we expand the polynomial



and see that the probability of being covered twice is , the probability of being covered once is  (note  only appears once) and the probability of not being covered at all is . Below are the first few order expansions

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | successes (times covered) | | | | |
| tows | Polynomial | 0 | 1 | 2 | 3 | 4 |
| 1 |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |
| . | . | . | . | . | . | . |
| . | . | . | . | . | . | . |
| . | . | . | . | . | . | . |
|  |  |  |  |  |  |  |

A general expression for these series is



The product out front is the binomial coefficient, so the distribution of the number of times covered per  tows with success probability  is



In terms of the tow area and total area



Poisson approximation

The expected number of times an area is covered is precisely (given the assumptions) modelled using a binomial distribution as above but under the conditions of a large number of tows (as ) and a small probability of success, the binomial distribution can be well approximated using a Poisson distribution. Our derivation of the Poisson distribution as the limiting distribution of the binomial largely follows that of Mangel (2006).

The limit of the binomial distribution is required



Defining  as the constant product (swept area ratio )













This decomposes into

1. 

And

1. 

The last result follows from a definition of the exponential function. In summary, for a constant product 



The probability of a point within a grid cell being covered *k* times, directly translates to the proportion of the cell that is *k* times covered. The figure below shows for a given swept-area ratio (*λ*) the proportion of a cell that is *k* times covered.

SweptAreaRatio.tif

References

Mangel, M. (2006). The theoretical biologist’s toolbox: quantitative methods for ecology and evolutionary biology. Cambridge University Press, Cambridge, UK.

**Supplementary material S2**

Most of the data aggregation was performed using SQL server 2008. However in order to make available a self-contained, generic set of code, we have re-coded an example in R. SQL code is available from the corresponding author on request.

## R version 2.15.2

library(rgeos) # needed for readWKT

library(rgdal) # needed for spTransform

# first specify a function that returns a polygon (grid cell) for any

# given point

# gridx and gridy specify the width and height of the polygon rectangle

# the origin of each grid is (0,0)

# the output is a Well-Known Text string

poly\_fun <- function(lon, lat, gridx, gridy){

# round to the nearest rectangle mid-point

lon1 <- round((lon - gridx/2)/gridx , 0) \* gridx + gridx/2

lat1 <- round((lat - gridy/2)/gridy , 0) \* gridy + gridy/2

# create WKT sting

out <- paste('POLYGON(('

,lon1 - gridx/2 ,' ' ,lat1 - gridy/2 ,', '

,lon1 + gridx/2 ,' ' ,lat1 - gridy/2 ,', '

,lon1 + gridx/2 ,' ' ,lat1 + gridy/2 ,', '

,lon1 - gridx/2 ,' ' ,lat1 + gridy/2 ,', '

,lon1 - gridx/2 ,' ' ,lat1 - gridy/2 ,'))'

,sep='')

return(out)

}

# set the parameters

gridx <- 0.18 # width of the largest rectangle

gridy <- 0.16 # height of the largest rectangle

n <- 2000 # number of simulated datapoints

# simulate some vms data

vms <- data.frame(lon = rnorm(n, -7.5, 0.15)

,lat = rnorm(n, 51.25, 0.1)

,effort\_h = 2

,speed\_kmph = rnorm(n, 6.5, 0.5)

,trawl\_width\_km = 0.1

)

# calculate swept area

vms$swept\_area\_km2 <- vms$effort\_h \* vms$speed\_kmph \* vms$trawl\_width\_km

# generate all possible nested rectangles

# A1 represents the largest rectangles

# two A2 rectangles fit inside each A1 rectangle

# two A3 rectangles fit inside each A2 rectangle etc.

A1 <- poly\_fun(vms$lon, vms$lat, gridx, gridy)

A2 <- poly\_fun(vms$lon, vms$lat, gridx, gridy/2)

A3 <- poly\_fun(vms$lon, vms$lat, gridx/2, gridy/2)

A4 <- poly\_fun(vms$lon, vms$lat, gridx/2, gridy/4)

A5 <- poly\_fun(vms$lon, vms$lat, gridx/4, gridy/4)

A6 <- poly\_fun(vms$lon, vms$lat, gridx/4, gridy/8)

A7 <- poly\_fun(vms$lon, vms$lat, gridx/8, gridy/8)

A8 <- poly\_fun(vms$lon, vms$lat, gridx/8, gridy/16)

A9 <- poly\_fun(vms$lon, vms$lat, gridx/16, gridy/16)

# count how many points are inside each rectangle

ag1 <- aggregate(list(count = A1), list(poly = A1), length)

ag2 <- aggregate(list(count = A2), list(poly = A2), length)

ag3 <- aggregate(list(count = A3), list(poly = A3), length)

ag4 <- aggregate(list(count = A4), list(poly = A4), length)

ag5 <- aggregate(list(count = A5), list(poly = A5), length)

ag6 <- aggregate(list(count = A6), list(poly = A6), length)

ag7 <- aggregate(list(count = A7), list(poly = A7), length)

ag8 <- aggregate(list(count = A8), list(poly = A8), length)

ag9 <- aggregate(list(count = A9), list(poly = A9), length)

# match the total count for each rectangle to each vms data point

N1 <- ag1$count[match(A1,ag1$poly)]

N2 <- ag2$count[match(A2,ag2$poly)]

N3 <- ag3$count[match(A3,ag3$poly)]

N4 <- ag4$count[match(A4,ag4$poly)]

N5 <- ag5$count[match(A5,ag5$poly)]

N6 <- ag6$count[match(A6,ag6$poly)]

N7 <- ag7$count[match(A7,ag7$poly)]

N8 <- ag8$count[match(A8,ag8$poly)]

N9 <- ag9$count[match(A9,ag9$poly)]

# if an A1 rectangle has at least 20 datapoints, take A2

# if an A2 rectangle has at least 20 datapoints, take A3 etc

vms$poly <- ifelse(N1 >= 20, A2, A1)

vms$poly <- ifelse(N2 >= 20, A3, vms$poly)

vms$poly <- ifelse(N3 >= 20, A4, vms$poly)

vms$poly <- ifelse(N4 >= 20, A5, vms$poly)

vms$poly <- ifelse(N5 >= 20, A6, vms$poly)

vms$poly <- ifelse(N6 >= 20, A7, vms$poly)

vms$poly <- ifelse(N7 >= 20, A8, vms$poly)

vms$poly <- ifelse(N8 >= 20, A9, vms$poly)

# calculate the total swept area in each rectangle

vms1 <- aggregate(list(swept\_area\_km2 = vms$swept\_area\_km2)

,list(poly = vms$poly), sum)

# read Well-Known Text and return SpatialPolygons

FUN <- function(x) {

readWKT(x, p4s='+proj=longlat +ellps=WGS84 +datum=WGS84 +no\_defs')

}

Srl <- lapply(vms1$poly,FUN)

# project the data to obtain the area of each polygon

FUN <- function(x) spTransform(x

,CRS("+proj=utm +zone=29 +ellps=WGS84 +datum=WGS84 +units=m +no\_defs")

)@polygons[[1]]@Polygons[[1]]@area \* 10^-6

vms1$polygon\_area\_kmsq <- unlist(lapply(Srl,FUN))

# calculate swept\_area\_ratio

vms1$swept\_area\_ratio <- vms1$swept\_area / vms1$polygon\_area

# colours for the polygons

breaks <- seq(0, max(vms1$swept\_area\_ratio), len=10)

i <- findInterval(vms1$swept\_area\_ratio, breaks, all.inside=T)

vms1$greys <- grey(seq(1, 0, length=length(breaks) ) )[i]

# set up blank plot

plot(NA, xlim = c(-8, -7), ylim = c(51, 51.5), xlab = 'Longitude'

,ylab = 'Latitude', asp=1.6)

rect(par('usr')[1],par('usr')[3],par('usr')[2],par('usr')[4]

,density=10,col='grey')

# draw polygons

for(i in 1:length(Srl)) plot(Srl[[i]], col=vms1$greys[i], add=T)

# add the datapoints & box

points(vms$lon, vms$lat, cex=0.1, col=4)

box()

# calculate the area that is impacted exactly k times

# equation (4) from the main manuscript:

# (equivalent to dpois)

pm <- function(k, lamda) lamda ^k \* exp(-lamda) / factorial(k)

k <- 1

sum(vms1$polygon\_area \* pm(k, vms1$swept\_area\_ratio) )

k <- 2

sum(vms1$polygon\_area \* pm(k, vms1$swept\_area\_ratio) )

k <- 5

sum(vms1$polygon\_area \* pm(k, vms1$swept\_area\_ratio) )

# calculate the area that is impacted AT LEAST k times

# (we empirically established this follows the gamma distribution)

k <- 1

sum(vms1$polygon\_area \* pgamma(vms1$swept\_area\_ratio, k, 1) )

k <- 2

sum(vms1$polygon\_area \* pgamma(vms1$swept\_area\_ratio, k, 1) )

k <- 5

sum(vms1$polygon\_area \* pgamma(vms1$swept\_area\_ratio, k, 1) )