Journal of Environmental Economics and Management
January 2021, Volume 105 Pages 102388 (16p.)
https://doi.org/10.1016/j.jeem.2020.102388
https://archimer.ifremer.fr/doc/00661/77272/

Scale-dependency in discrete choice models: A fishery application

Dépalle Maxime [1, *] , Sanchirico James N. [1, 2, 3], Thébaud Olivier [4], O'farrell Shay [2, 3], Haynie Alan C. [5], Perruso Larry [6]

[1] Department of Agricultural & Resources Economics, University of California, Davis, USA
[2] Department of Environmental Science and Policy, University of California, Davis, USA
[3] Coastal and Marine Sciences Institute, University of California, Davis, USA
[4] Ifremer, UMR 6308, AMURE, Unité d'Economie Maritime, France
[5] NOAA Alaska Fisheries Science Center, USA
[6] NOAA Southeast Fisheries Science Center, USA

* Corresponding author : Maxime Dépalle, email address : mdepalle@ucdavis.edu

Abstract :

Modeling the spatial behavior of fishers is critical in assessing fishery management policies and has been dominated by discrete choice models (DCM). Motivated by the widespread availability of micro-data on fishing vessel locations, this paper examines the complexity associated with the choice of the spatial scale in a DCM of fishing locations. Our empirical approach estimates the standard DCM at varying spatial resolutions using both simulated data and vessel monitoring system data from the Gulf of Mexico longline fishery. We assess model performance using goodness-of-fit, predictive capacity, parameter estimates, and the assessment of the fishery response to a hypothetical marine protected area. Results show that, even when the specification of the decision-making process is correct, models can be structurally biased because of the aggregation of spatial scale that neglects the value of many fishing locations. The extent of such biases can only be detected by considering various spatial aggregation levels.

Keywords : Spatial modeling, Discrete-choice model, VMS data, Fisher behavior, Monte Carlo experiments

Introduction

Understanding and anticipating fishers' spatial behavior is a critical component in the design and assessment of fishery management (Lewison et al., 2015; Smith, 2010; Valcic, 2009). Research on predicting how management measures impact the choice of fishing locations is dominated by discrete choice models (DCM), which are well-suited to the fishery setting where the resource distribution is often patchy (Sanchirico and Wilen, 1999). Although DCM provide a powerful framework for modeling spatial and temporal fishing behavior (Smith, 2010), the reliability of DCM results hinges in part on the definition of the choice set, as analyses are often limited by data availability (Huang and Smith, 2014; Smith, 2005).

Issues associated with ill-defined choice sets are neither new nor unique to fisheries economics (Manski, 1977; McFadden, 1978). Most research has focused on the issue of what to include in the consideration set – the subset of the choice set which includes only the alternatives considered by the decision-maker - and how to estimate it in a consistent manner. For example, the marketing and transportation literature has extensively investigated the heuristics of consideration sets since the 1970s (Hauser, 2014; Horowitz and Louviere, 1995; Narayana and Markin, 1975; Prato, 2009). In the 1990s, a number of important papers in recreation demand and environmental economics demonstrated the impact of choice set specification on welfare outcomes - and thus on policy analyses - and showed that the sign and magnitude of welfare biases is not necessarily systematic and can be ambiguous (Parsons et al., 2000; Parsons and Hauber, 1998; Peters et al., 1995).

Over the last decade, the spatial resolution of data capturing fishing behavior has dramatically improved with the deployment of geolocation technologies, such as vessel monitoring systems (VMS). The recent availability of vessel position data with high spatial resolution (better than 10 m accuracy) is a remarkable advance over traditional data aggregated into large statistical areas. In many fisheries, VMS transponders placed on each vessel track

vessel positions at certain specified time intervals (e.g., each hour). These new data open up possibilities to refine analyses of fisher behavior across space and time. With potentially millions of recorded locations for each vessel, VMS data allow researchers to relax some prior assumptions underlying many DCM regarding the data-generating process and spatial scale of the analysis. In particular, it allows relaxing the implicit assumption of studies constrained by coarse spatial data that the decision-making process generating the data operates at that coarse scale. VMS data allow a more refined discretization of space permitting us to investigate the implications for the performance of classical DCM of the aggregation process related to partitioning space in different manners.

Through a novel approach that combines both simulated and real data, this paper investigates the following questions. First, how do estimates of traditional DCM of spatial behavior change as the mismatch between the spatial scale of aggregation and the spatial scale of the data-generating process grows? Second, how do estimates at spatially-refined resolutions depend on the spatial distribution and heterogeneity of the observations?

Spatial aggregation impacts DCM's estimation through two main mechanisms. On the one hand, it affects the structure of the choice set. Aggregation determines what and how many choices are available in the decision-making sequence and it determines where an observed choice is located in space. Further, whether to change location in any period is a function of the opportunities available in the consideration set, which are modified by the level of spatial aggregation. On the other hand, aggregation affects the quantity and quality of the information used to estimate a DCM. Indeed, the opportunities in the different sites are assessed with the aggregated data (e.g., expected catches) and therefore the aggregation will affect the amount of information available to make those calculations, in particular by levelling out some data variability. For example, a fisher's expected catch for each site in the choice set is the result of a complex process that likely combines individual and fleet-level observations for that site over

some prior period of time.[1] The impacts of spatial aggregation on the modeling of this process, however, is not immediately clear and will likely depend on the underlying heterogeneity of the data across both space and time. Figure 1 illustrates the effects of spatial resolution where the color value of each tile (the aggregated unit of analysis) results from a weighted average of the colors associated with each individual observation. Depending on the interplay between the spatial distribution of the data, its spatial heterogeneity and the partition of space that is considered, the outcome at the aggregated spatial level can differ dramatically.

[1] In particular, the expected productivity of fishing sites has been shown to be a major driver of fishers' decisions (Eales and Wilen, 1986; Girardin et al., 2016). Yet, it is not clear which information fishers integrate to form their expectations, and how they combine information coming from different sources and of different spatial and temporal scales (Abbott and Wilen, 2011). The extent of information sharing among fishers is a rich research area (Evans and Weninger, 2014; Wilson, 1990).

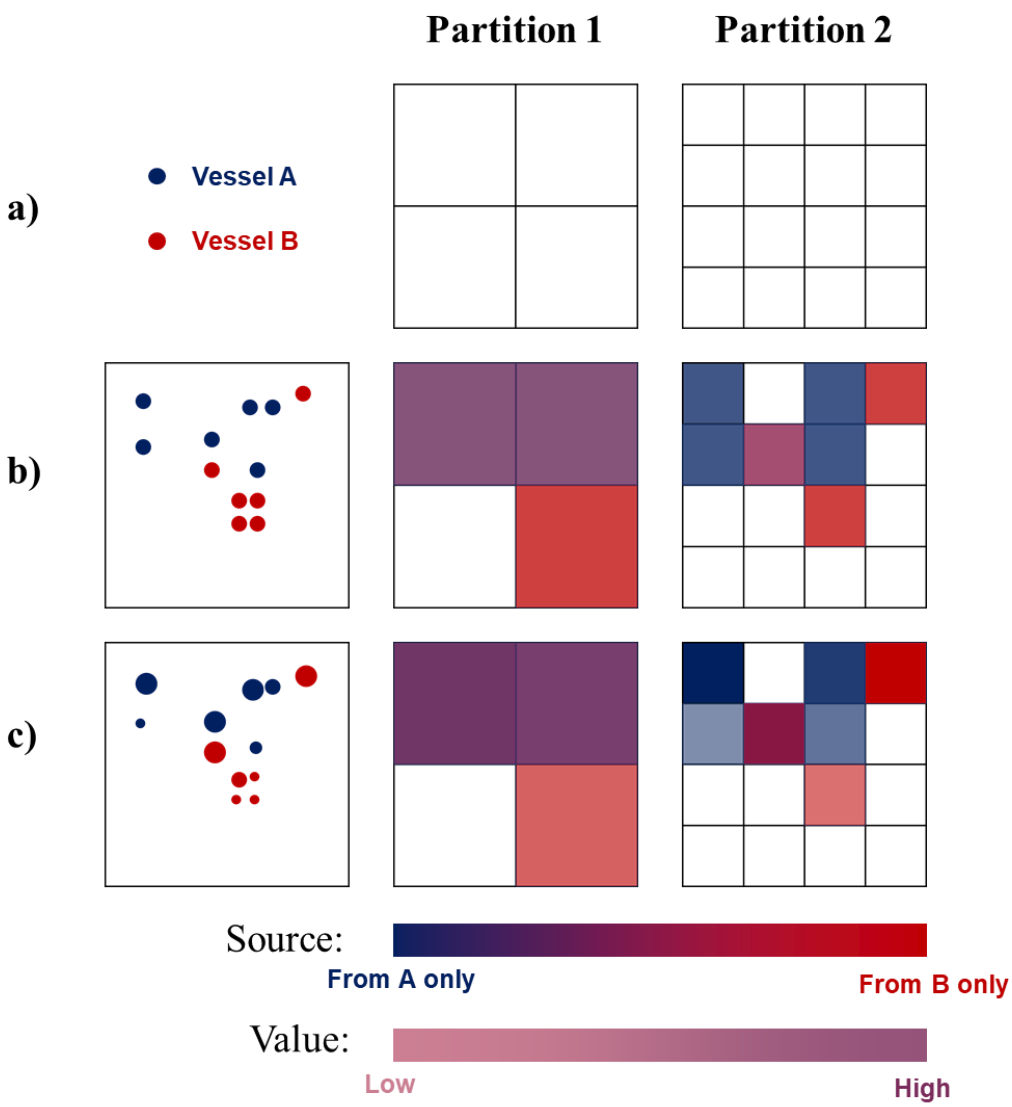


Figure 1. Conceptual figure illustrating how the partition of space in a model intrinsically shapes how information is integrated and, therefore, model's results. The panels are: a) identification of two vessels and two alternative partitions of the fishing area; b) illustration of the case where observed values (e.g., catch rates) are homogeneous in all locations; c) illustration of the case where observed values are heterogeneous across locations (dots indicate the number of observations). For instance, note the changes in the colors associated to the same area (e.g., purple in partition 1 but deep blue in partition 2) and in the share of areas with missing

information (25% in partition 1 but 56% in partition 2). Those differences highlight the information (implicitly) integrated into a spatial analysis.

We conduct a series of Monte Carlo experiments (two sets of 600 in total) mimicking data from a commercial fishery. In addition to allowing us to control for the decision-making process of fishers – thus be free from misspecification issues – the simulations enable us to control for the level of spatial heterogeneity of data by considering different spatial concentrations of observations. Each experiment consists in estimating, at varying resolutions, the true decision-making process of three datasets of observations having different spatial distributions. We consider spatial choice sets defined at both higher and lower resolutions than the resolution of the true decision-making process, and without necessarily a nested structure (see Figure C.a in the appendix). This enables us to identify more precisely the respective effects of the two aggregation mechanisms we described.

For each combination of dataset and spatial resolution, we assess how well the estimated model can recover the true parameters of the decision-making process, how well the model fits the data, and how well it can predict unobserved choices out-of-sample. We also use the model to evaluate the welfare impact of the implementation of a hypothetical Marine Protected Area (MPA) because DCM have been a popular tool to conduct the impact analysis of an MPA policy (Curtis and Hicks, 2000; Hicks and Schnier, 2010; Hynes et al., 2016; van der Lee et al., 2014). We then apply the same analytical framework to a real case study using a unique dataset on bottom longline (BLL) fishing in the Gulf of Mexico reef fish (GoMRF) fishery that includes recently available geospatial data.

Our work contributes to the literature in two ways. First, our Monte Carlo experiments provide new insight into the effects of both spatial heterogeneity and spatial aggregation on DCM. Monte Carlo experiments have been used to assess the effects of varying the choice set

generation process (Bierlaire et al., 2010; Li et al., 2015; Pramono and Oppewal, 2012; Stafford, 2018; Torres et al., 2011) and to validate the prediction capabilities of DCM (Haynie, 2005; Smith et al., 2014), but never to assess the sensitivity of the model to either the spatial heterogeneity of data or the partition of space. Second, we investigate the issue of the spatial scale of analysis in the context of VMS data utilization for DCM estimation. VMS data have been used in the context of discrete-choice modeling to refine the spatial identification of fishing sites or fishing operations (Hynes et al., 2016; Rijnsdorp et al., 2011; Russo et al., 2015) but never to explicitly address the relationship between the spatial scale of decisions and the spatial scale of analysis.

We find that even when the decision-making process is correctly specified, DCM can be biased because of the aggregation of data, and that these effects are amplified as spatial heterogeneity of a site's quality increases. In contrast, we are not able to detect any effect of the spatial irregularity of the observation sample. Second, we demonstrate how VMS data can be used to detect model bias and to assess the robustness of DCM results. Our findings suggest caution in the use of DCM for policy analysis, when the researcher is unable to verify the robustness of their results across multiple spatial resolutions (e.g., due to data limitations).

In the remainder of the paper, we first describe our methodological approach, detailing the setting of the Monte Carlo experiments and the empirical data. We then show the results of the model estimations with the simulated data and with observations from the bottom long-line (BLL) sector of the Gulf of Mexico reef fish fisheries (GoMRF). Finally, we summarize our findings, draw general conclusions about spatial models and the use of VMS data, and suggest possible future research in this area.

Methods

Identifying the effect of the spatial scale of estimation on a DCM's results is challenging because of the possible confounding effect of the misspecification of the decision-making

process. For that reason, we apply the same random utility model (RUM) - a standard framework for researchers conducting welfare analyses – to both simulated data where the data-generating process is controlled, and real empirical data where the decision-making process can only be assumed. This section presents the modeling approach applied in the simulation and empirical example.

Modeling framework

We estimate the RUM where fishers, conditional on taking a fishing trip, make a unique daily decision of where to fish according to a simple utility criterion that weights travel costs and expected rewards linearly:

$$U_{ijt}=\beta_{dist}*Dist_{ijt}+\beta_{VPUE}*E\left[VPUE_{ijt}\right]+\varepsilon_{ijt} \quad (1)$$

where i is the vessel, j is the site, and t is the day. β_{dist} and β_{VPUE} denote the marginal utilities of, respectively, the distance to a given location, $Dist_{ijt}$, and the associated expected value per unit effort ($E[VPUE_{ijt}]$) and ε_{ijt} is a random shock. The definition of what constitutes a site (i.e., the scale of space discretization) will vary across models.

Our specification is the standard workhorse model in the literature where expected revenues and travel costs are the main predictors of the choice of the fishing location (Girardin et al., 2016). A commonly used proxy for travel costs is the distance to the fishing sites, usually reduced to the centroids of the alternative and current location for computational purposes (Abbott and Wilen, 2011; Haynie and Layton, 2010). Intuitively, this variable captures that more distant fishing sites incur higher fuel costs and require more travel time.

With respect to fishers' expectations about revenues from a fishing site, we follow the literature that utilizes records of past performance for each site, aggregated at the fleet level (Girardin et al., 2015; Smith, 2005). Specifically, we assume that fishers combine both short and long-term information and weight information differently depending on what information

is available (Abbott and Wilen, 2011; Hutniczak and Münch, 2018)[2]. When we estimate the model we account for all possible combinations of information available by using dummy variables (Table I).

Table I Different cases investigated with regard to the amount of information available to vessels in the formation of their expected value per unit effort (VPUE) for the different sites in the choice set. Short-term information refers to historical VPUE records averaged over the month prior to the fishing trip ("m-1" subscript) while long-term information refers to historical VPUE records averaged over the same month of the year prior to the fishing trip ("ym-1" subscript). The superscript "ft" indicates that VPUE have been averaged across all the active vessels of the fleet.

Case	**Information available to form E[VPUE]**	β_{VPUE}*E$[\text{VPUE}_{ijt}]$ ***is equal to***
1	Both short-term **and** long-term historical VPUE are available	$\beta_{\text{VPUE}}^{Full\ info\ -\ short-term} * \overline{\text{VPUE}}^{\text{ft}}_{\text{m}-1} + \beta_{\text{VPUE}}^{Full\ info\ -\ long-term} * \overline{\text{VPUE}}^{\text{ft}}_{\text{ym}-1}$
2	Only short-term historical VPUE are available	$\beta_{\text{VPUE}}^{Short-term\ only} * \overline{\text{VPUE}}^{\text{ft}}_{\text{m}-1}$
3	O**nly** long-term historical VPUE are available	$\beta_{\text{VPUE}}^{Long-term\ only} * \overline{\text{VPUE}}^{\text{ft}}_{\text{ym}-1}$
4	N**either** short-term **or** long-term historical VPUE are available	$\beta_{\text{VPUE}}^{No\ info}$

[2] In the simulations we set the data-generating process such that when VPUE averages are available over both short- and long-term period, fishers respond primarily to the more recent information signals and weight short-term VPUE average three times more than long-term average to form their revenues expectations. When one kind of information is missing fishers are assumed to put all the weight on the available information, and when no information is available fishers' revenues expectations for the given sites are assumed to be 0 (see Appendix Section D.5).

In the empirical analysis of the GoMRF grouper fishery, we include some additional controls to those in the Monte Carlo experiments, relating to the assumed but unobserved true decision-making process. Specifically, we control for the aggregate level of effort of other BLL fishers in a given alternative the day before ($Eff.oth_{ijt}$) so as to control for the possible influence of other fishers' activity. We also include fishers' own level of effort in a given site the day before ($Eff.own_{ijt-1}$), to control for the dynamic aspect of the daily choices of the fishing site when those choices are part of the same multiple-day fishing trip:

$$U_{ijt}=\beta_{dist}*Dist_{ijt}+\beta_{VPUE}*E[VPUE_{ijt}]+\beta_{Eff.oth}*Eff.oth_{ijt-1}+\beta_{Eff.own}*Eff.own_{ijt-1}+\varepsilon_{ijt} \quad (2)$$

Indeed, the behaviors of other fishers along with fishers' past fishing patterns have been shown to influence fishers' decision-making (Girardin et al., 2016; Huang and Smith, 2014; Poos and Rijnsdorp, 2007). More details on the empirical specification and other possible representations are provided in Section E of the Appendix.

Tessellations

In the simulated and GoMRF analysis, we estimate Eq.1 or Eq. 2 with variables and alternatives defined at increasingly spatially-aggregated (i.e., coarser) resolutions (see, for example, Figure 2). For the simulations, we considered nine tessellations based on a grid, with cell sizes ranging from 1 Nautical Mile (NM – although distances are only relative in the simulations) to 30NM. We test two resolutions for the data-generating process, a 1NM scale and a 10NM scale.[3] The 1NM scale allows us to focus only on the effect of data aggregation whereas the 10NM scale allows to see the effects of both non-nested and overly spatially-refined estimation choice sets. For the analysis on the GoMRF fisheries, we consider gridded partitions of space and analyze the results of the RUM across a set of tessellations, varying the

[3] Here, we chose to focus on the issue of scale and therefore we consider only rectangle alternatives, by far the most common data collection framework.

length of the squared-shape alternatives from 0.25° ~ 15NM (1797 sites) to 3.5° ~ 210NM (15 sites). We use the 60NM-long statistical areas used by the U.S. National Marine Fisheries Service (NMFS) for logbook reporting since 2013 as a reference resolution, from which we consider both tessellations at finer and coarser scales. We also add a tessellation based on the statistical areas used by NMFS prior to 2013, which somewhat follows a longitudinal partitioning of space.

Across the tessellations and simulated and real data analyses, we measure (1) the goodness-of-fit and prediction capability of each model, (2) the normalized estimates of model parameters, and (3) the evaluation of the welfare impact of the implementation of a hypothetical Marine Protected Area[4] (MPA, shown in red in Figure 2).

Monte Carlo experiments

The simulated setting we consider mimics the case of a fishery targeting a species migrating seasonally across three different "hotspots" or, similarly, the case of three non-migrating species each located in a different hotspot with seasonal variation in catch rates. We operationalize spatial and time-varying value per unit effort for each unit of space in each decision period (t) by assuming that the productivity of a given point in space (whose coordinates are x,y) is equal to the mean of the productivity of the hotspots (recall we assume there are three hotspots) with a stochastic error of +/- 100%:

$$\mathrm{VPUE}(x,y,t) = \big(1 + u(x,y,t)\big)\left(\frac{1}{3}\sum_{h} \mathrm{VPUE}_h(x,y,t)\right) \quad (3)$$

[4] The choice of the location of the hypothetical MPA was arbitrary. Round latitudes and longitudes were chosen to define the north, south and east limits, and a round numbered bathymetric contour was used for the west limit.

with $u(x,y,t) \sim U(-1,1)$, $VPUE_h(x,y,t) = \overline{VPUE_h}(t)e^{(-\frac{(x-h_x)^2}{\sigma_x^2})}e^{(-\frac{(y-h_y)^2}{\sigma_y^2})}$ and $\overline{VPUE_h}(t) = \overline{VPUE}.\,base\left(1 + A * \cos\left(2\pi\frac{t-t_h}{T}\right)\right)$, where h represents a hotspot. Therefore, we assume that there is a hotspot (h_x, h_y)[5] with a Gaussian spatial distribution of fish abundance that oscillates through time around a base value of VPUE ($\overline{VPUE}.\,base$) with a period T and reaches its maximum at t_h[6].

The dynamics in fishing locations are driven by the seasonality and the stochasticity of the VPUEs that we generate. Specifically, fishers' experience independently and identically distributed random utility shocks when forming their decision and face different histories of VPUEs as the year progresses[7]. Figure 2a shows a map of *yearly average* VPUEs for three levels of assumed spatial heterogeneity (see Figure D.2 in the Appendix for maps at different points in time during the year). The full details for the generation of the VPUEs distributions, the specification of fishers' expectations, and the data-generating process are available in the supplemental information.

Each experiment estimates Eq. 1 across three datasets of observations with three different levels of spatial heterogeneity, with variables and choice alternatives defined at varying spatial resolutions. The three different levels of spatial heterogeneity we consider

[5] We assume 3 hotspots located in the North (2.6,2.5), Center (3.1,0) or South (1.9,-2.4).

[6] We set $\overline{VPUE}.\,base = 4$, $T = 1$ and $t_h \in \left\{\frac{2}{12}, \frac{5}{12}, \frac{8}{12}\right\}$.

[7] We assume fishers make a single choice of fishing location when they go on a trip and that they all start from the same "homeport" located at the origin (0,0). To initiate the historical records in the simulation, we use the draws from one year of simulated positions under the assumption that fishers have a perfect knowledge of the VPUE maps and we discard the subsequent first year of simulated positions.

correspond to three different spatial distributions for the VPUE. This variation in the spatial regularity of the datasets enables us to disentangle the effect of having observations with different spatial distributions from the effect of deviating from the spatial scale of the true decision making process (DMP). To capture the latter effect, we display the results according to the level of spatial aggregation that we defined as the logarithm of the ratio of the area of the aggregated alternatives to the area of the "true" alternatives (Appendix Section B). In the Monte Carlo analysis, the index equals zero when the scale matches the DMP and as the index changes (non-linearly), the aggregation moves further away from the DMP (Appendix Figure C.a). In the GoMRF analysis, the index equals zero for the 60NM-long statistical areas used by the U.S. National Marine Fisheries Service (NMFS) for logbook reporting after 2013 and negative numbers correspond to finer spatial resolutions while positive numbers correspond to larger spatial resolutions.[8]

[8] For example, a 2°x2° grid corresponds to an index of 1.4, whereas a ½°x ½° grid corresponds to an index of -1.4. Similarly, a 2.5°x2.5° grid corresponds to an index of 1.8, and a ¼°x ¼° grid corresponds to an index of -2.8 (see Figure C.b in the Appendix).

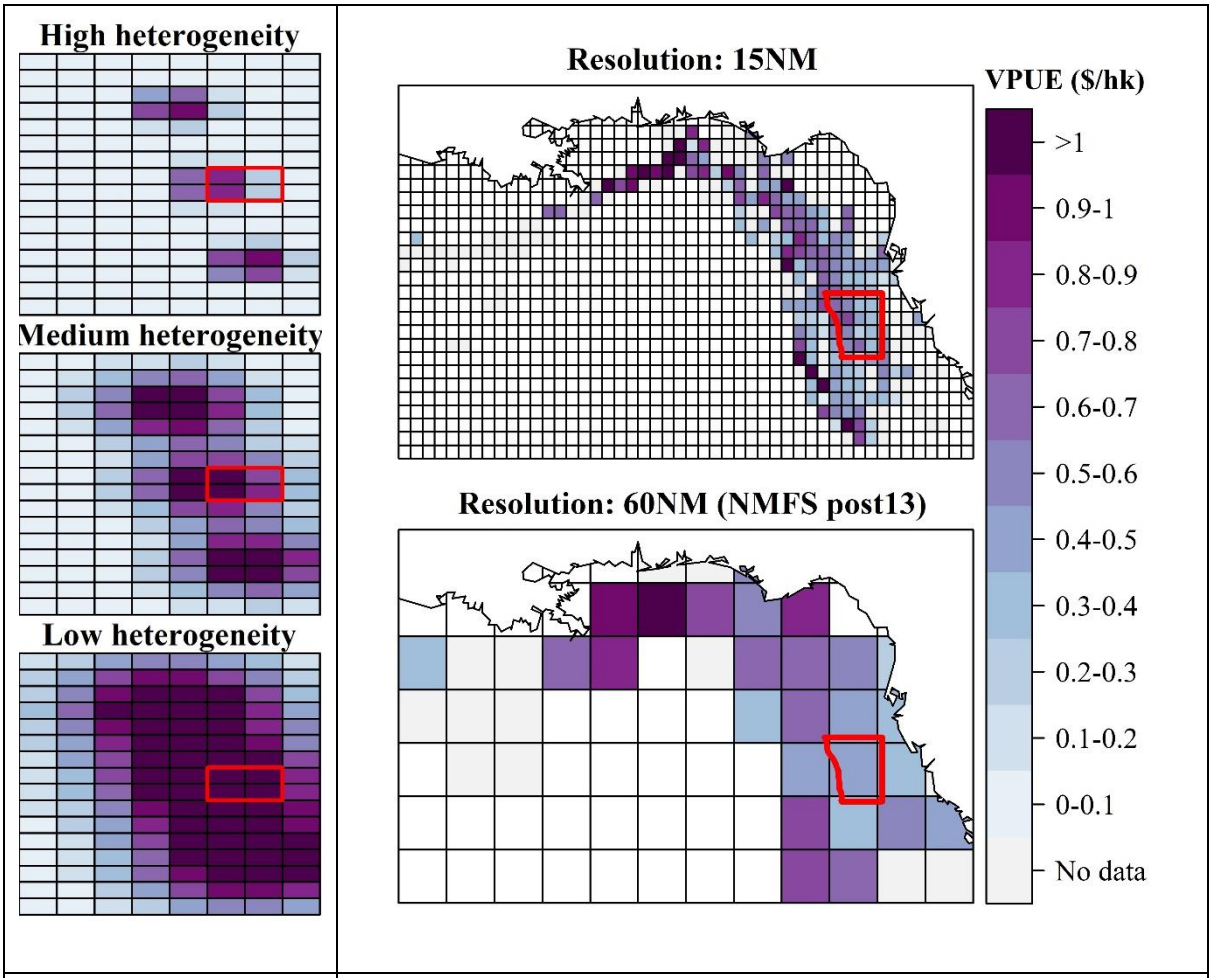


Figure 2. Yearly average of value per unit effort (VPUE, in $ per hook). Panel A shows the three spatial distributions of VPUE simulated in the Monte Carlo experiments with a 30 NM grid. Panel B shows the observed VPUE for longline fishers in the Gulf of Mexico in 2008 for 15NM or 60NM grids. The hypothetical MPAs assumed for the welfare analyses are shown in red.

Welfare impact of the implementation of a MPA

In addition to measuring the impacts of scale and heterogeneity on model parameters, we examine the implications of these results for policy advice by simulating and measuring the welfare impact of a hypothetical MPA across the different models and datasets. Specifically, we assess the distribution of welfare losses for the impacted choice occasions; that is, choice occasions where the chosen fishing location would have occurred in the MPA. For an individual

i facing a set of alternatives j the utility loss was computed as the difference between the chosen alternative without the MPA restriction and the chosen alternative when restricting the choice set (CS) to alternatives that lie outside the MPA[9]:

$$\Delta U^{i} \equiv \max_{j \in \overline{CS}} U_j^i - \max_{j \in \overline{CS}\{MPA\}} U_j^i \quad (3)$$

In the case of the Monte Carlo experiments, the "true" utility loss corresponds to the case where the choice set is the "true" set of alternatives that were considered by fishers when seeking to maximize their utility (i.e., in our case all 115,200 1x1 NM rectangles):

$$\Delta U_{true}^{i} \equiv \max_{j \in \overline{CS}} U_j^i - \max_{j \in \overline{CS}\{MPA\}} U_j^i = \sum_k \beta_k \left(\overline{X}_k^{j^{max}} - \overline{X}_k^{j^{MPA}} \right) \quad (4)$$

where j^{max} is the alternative maximizing the utility over the true choice set ($\overline{CS}$) of alternatives, j^{MPA} is the alternative maximizing the utility over the true choice set of alternatives excluding alternatives in the MPA, and $\bar{X}_k^j$ is the value of the explanatory factor X_k evaluated on the true choice set for alternative j.

In the GoMRF analysis, the true choice set of alternatives considered by fishers when forming their decisions is unknown to the researcher. In this case, a common approach is to assume that the choice set considered by fishers is the "observed" or "estimated" choice set; that is, the set of alternatives having been chosen at least once in the dataset:

$$\Delta U_{est}^{i} \equiv \max_{j \in \widehat{CS}} U_j^i - \max_{j \in \widehat{CS}\{MPA\}} U_j^i = \sum_k \widehat{\beta_k} \left(\widehat{X}_k^{\widehat{J^{max}}} - \widehat{X}_k^{\widehat{J^{MPA}}} \right) \quad (5)$$

where $\widehat{J^{max}}$ is the alternative maximizing the utility over the estimated choice set ($\widehat{CS}$) of alternatives, $\widehat{J^{MPA}}$ is the alternative maximizing the utility over the estimated choice set of

[9] Because the scale of the utility cannot be recovered, we normalize the utilities by the marginal utility of distance: $\widetilde{\Delta U^{i}} = \frac{\Delta U^{i}}{|\widehat{\beta}_{dist}|}$.

alternatives excluding alternatives in the MPA, and $\hat{X}_k^j$ is the value of the explanatory factor X_k evaluated on the estimated choice set for alternative j.

The practical assumptions made in defining the choice set in empirical applications of DCM, unfortunately, is not benign. The choice set in such a setting may be ill-specified in two ways, which we label *choice bias* and *aggregation bias*:

(1) *Choice bias (C)*: the specified choice set may not include the same set of alternatives as the one considered by fishers (i.e., $\widehat{j^{max}} \neq j^{max}$ and/or $\widehat{j^{MPA}} \neq j^{MPA}$);

(2) *Aggregation bias (A)*: the specified choice set may not account for the opportunities *within* each alternative in the same way as they were considered by fishers, and so it may be the case that $\bar{X}_k^j \neq \hat{X}_k^j$

Utilizing our Monte Carlo experiments, we disentangle the effects of these two kinds of biases along with the effect of having biased estimates of the parameters, by decomposing the error in the estimation of the utility loss:

$$\Delta U^i \equiv \Delta U^i_{est} - \Delta U^i_{true}$$

$$= \sum_k \underbrace{\widehat{\beta_k}\left(\left(\hat{X}_k^{j\widehat{max}} - \hat{X}_k^{j max}\right) - \left(\hat{X}_k^{j\widehat{MPA}} - \hat{X}_k^{j MPA}\right)\right)}_{\text{Choice bias}}$$

$$+ \underbrace{\widehat{\beta_k}\left(\left(\hat{X}_k^{j max} - \bar{X}_k^{j max}\right) - \left(\hat{X}_k^{j MPA} - \bar{X}_k^{j MPA}\right)\right)}_{\text{Aggregation bias}} + \underbrace{\left(\widehat{\beta_k} - \beta_k\right)\left(\bar{X}_k^{j max} - \bar{X}_k^{j MPA}\right)}_{\text{Parameter bias}} \quad (6)$$

Application: Bottom longline sector of the Gulf of Mexico Reef Fish fishery

We apply the same modeling and analytical framework in the Monte Carlo analysis to the BLL sector of the GoMRF fishery. The BLL fleet primarily targets grouper and tilefish (GT) species, with other minor reef fish species such as snappers and jacks. In 2012, 65 vessels that held federal GoMRF permits reported 651 GT-BLL trips and 6135 days at sea to the Southeast Coastal Fisheries Logbook Program (Coastal Logbook). These numbers are down

from 116 vessels reporting 1244 trips for 12075 days in 2008[10]. Fleet efficiency increased during the 2005-2012 period, which coincides with the GT-BLL fleet experiencing considerable consolidation. Average days at sea per trip increased from 7.5 in 2005 to an annual average of 10.0 days per trip from (2008-2012) while pounds landed (inflation-adjusted revenues) increased from 3779 pounds ($12,300) in 2008 to 6312 pounds ($21,911) in 2012. The GT-BLL fleet primarily operates off the west coast of Florida along a depth band of about 1° to 2° of latitude (Figure 2 Panel B). This corresponds to the VPUE's distribution with the lowest spatial heterogeneity in the simulated data.

The GoMRF fishery offers a well-documented, data-rich research environment. Logbooks for coastal fishing have been collected since 1993, spatial fishing data have been collected since mid-2006 by onboard observers for a subset of the vessels and trips, and VMS data for all pelagic and reef fishing trips is available since 2006. Having data through 2013, we could estimate the RUM over the whole 2007-2013 period. However, the GT-BLL sector underwent several major disruptions in 2009 and 2010[11], including the BP Deepwater Horizon oil spill, that have likely changed the dynamics of fishing decisions. Consequently, we chose to

[10] GT-BLL trips are defined as any trip reporting to the Coastal Logbook that landed at least one pound of GT species from one of five species categories (gag, red grouper, other shallow water groupers, deep water groupers, or tilefishes) covered by the GT-IFQ and used bottom longline as the primary gear on the trip. In cases of multi-gear trips, the primary gear type ("Topgear") is defined as the gear that produced a plurality of trip revenues.

[11] In 2009 a large part of the West coast of Florida was closed to longline fishing from May to October as an emergency action to reduce sea turtle bycatch, and starting 2010 an IFQ system was implemented for GT species. See Watson et al. (2018) for discussion of these management changes and their impacts on the fishery.

estimate the RUM using two different datasets, one before and one after the 2009-10 period. Using one-year lagged predictors, we choose 2008 and 2012 as our two distinct training datasets, leaving 2013 observations to evaluate out-of-sample prediction performance.

Results

Monte Carlo experiments

Overall, the results of the Monte Carlo (MC) experiments confirm the intuition that goodness-of-fit, prediction errors (out of sample), and parameter bias increase as the spatial scale used for the estimation is further away from the spatial scale of the data-generating process. They also show that the effect of having a non-nested structure with the true spatial choice set is larger than the effect of averaging information, and that data aggregation impacts more datasets with an initially high level of heterogeneity. Perhaps less evidently, our results indicate as well asymmetric effects of estimating models at overly coarse or refined resolutions: using a resolution which is twice smaller than the resolution of the DGP has a much bigger effect than using a resolution which is twice bigger. At last, they also show that, as the resolution used for estimation becomes coarser, goodness-of-fit and prediction errors tend to a threshold while the distributions of the estimated parameters become larger.

Goodness-of-fit and prediction capability

In the MC analysis, as the level of spatial aggregation increases, we observe an initial decrease of the models' McFadden pseudo-R^2 and increase in prediction errors[12] and then a leveling off (Figure 3 Panels A and B). When using overly spatially-refined resolutions,

[12] For the Monte Carlo experiments we evaluated the performance of the estimated models in terms of prediction capability by splitting each simulated dataset into a training dataset and a test dataset and we compute the rate of prediction errors of the models estimated with the training dataset on the test data.

models' goodness-of-fit and prediction accuracy also decrease, but to a much larger extent and with a sharper initial jump (Figure 3 Panels C and D). Furthermore, the decrease in the pseudo-R^2 - or increase in prediction errors - is more pronounced and the leveling off starts at lower levels of spatial aggregation when the observations are spatially concentrated; that is, as the levels of VPUE are more heterogeneous within a given area. The decrease in the pseudo-R^2 with the increase of the level of spatial aggregation means that, as the modeled resolution moves away from the resolution of the DMP, a loss of information occurs that is critical for explaining the choices. The sharper and earlier decrease observed when choices are more spatially concentrated suggests that it is the smoothing in the spatial heterogeneity of the explanatory factors that drives this loss of information[13]. The leveling off in both the goodness-of-fit and prediction errors with higher spatial aggregation is likely driven by the explanatory power of distance, which remains significant since increasing the spatial aggregation of sites maintains – and even increases – the heterogeneity of the distance variable. The same mechanisms, but in a reverse way, explain the trend observed when using overly refined resolutions. By assuming site options that have smaller spatial extents than the site options considered for decision-making, estimated models pick up more non-relevant data variations (noise). So much so that they may even end up performing worse than null models (models have negative pseudo-R^2). At last, clustering data in such a way that the spatial structure of the true choice set is not nested within the choice set used for the estimation also implies introducing non-relevant data

[13] Indeed, if choices are more spatially concentrated, it is because the variations in the VPUE levels occur at a more spatially refined scale (or, similarly, because a given variation of VPUE occurs at a smaller spatial extent), which means that when averaging those levels at a given resolution, the set of values that are pooled together are more heterogeneous than if the variations of the VPUE were occurring at a larger spatial scale.

variations which alters the capacity of DCM to properly explain the observed choices. Note, however, that these non-relevant data variations are levelled off as spatial aggregation increases, which explains why the performance of non-nested choice sets increases with the resolution (while performing less well than a nested choice set).

Parameter estimates

Given the conditional logit specification, the scale of the utility levels cannot be recovered and the β parameters of Eq. (1) or Eq. (2) – or marginal utilities (MU) - are only estimated up to a multiplicative scale. To permit comparisons across estimated models, we normalized parameter estimates by β_{dist}, the MU of distance, which amounts to comparing estimated marginal rates of substitution (MRS) with distance[14].

[14] In the simulations, we set by β_{dist} at -5 and β_{VPUE} at +20.

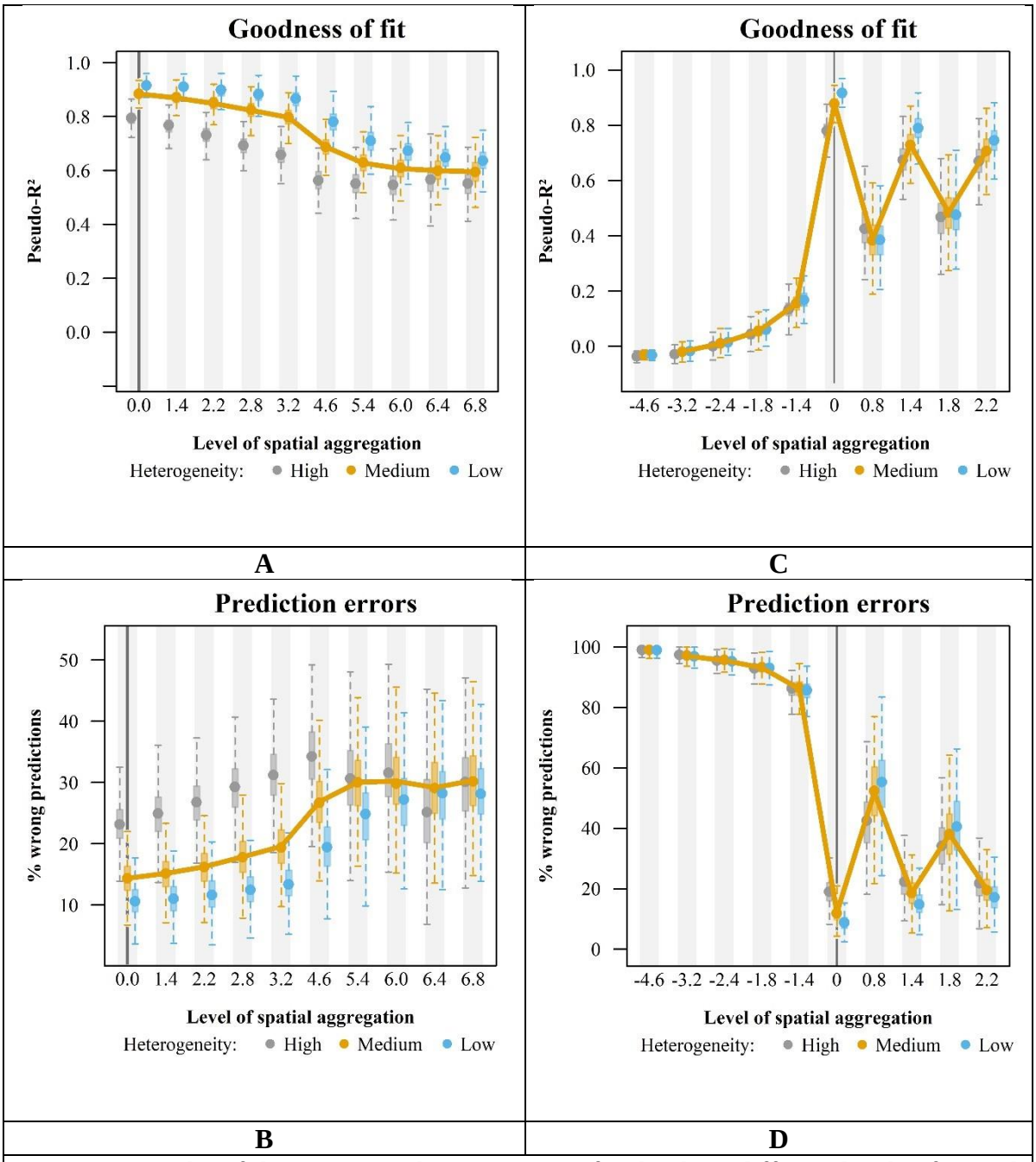
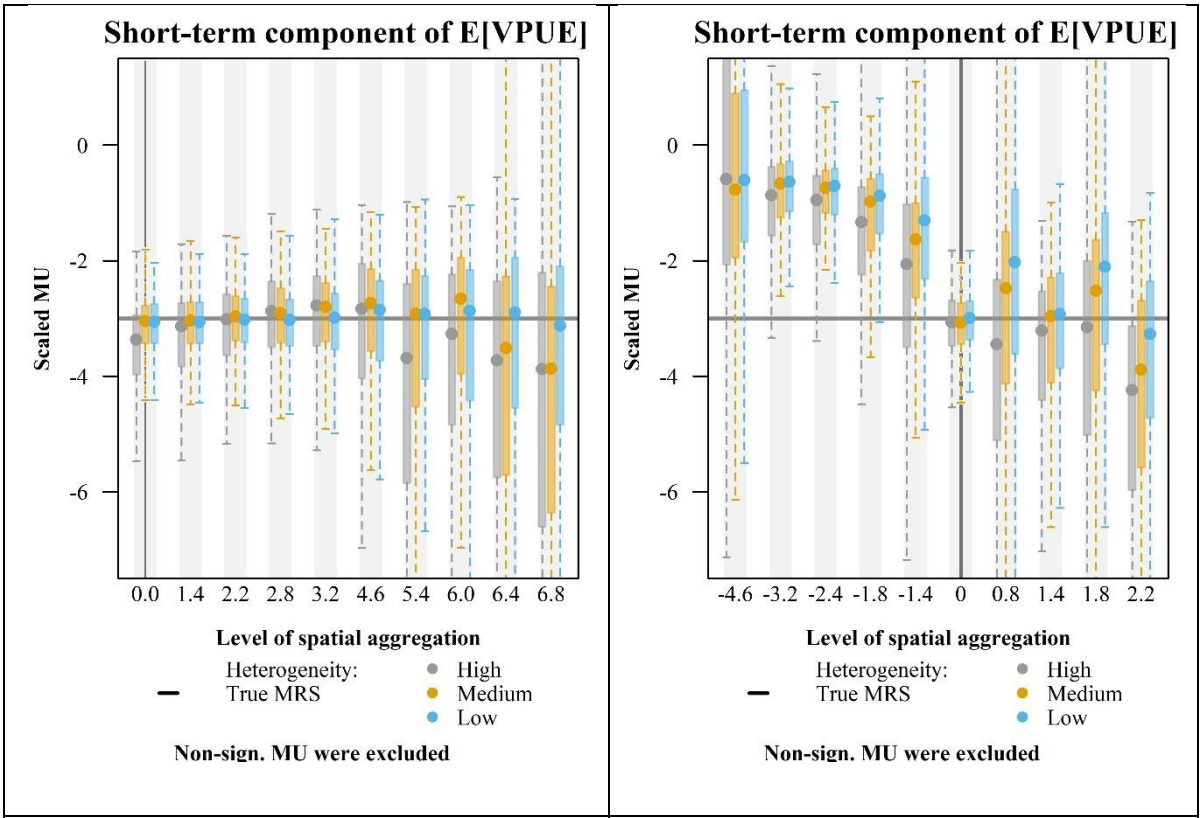


Figure 3. Results of the Monte Carlo experiments for the three different levels of spatial heterogeneity (High, Medium and Low) in terms of fitting (upper panels) and predicting (lower panels) the data, with the DGP set at on a 1x1 (left panels) or 10x10 (right panels) grid. Figures shows the median of: (A and C) model's McFadden pseudo-R^2; and (B and D) model's rate of prediction errors. Box edges are the 25th and 75th percentiles of the distribution of the losses and whiskers are located at +/- 1.5 times the interquartile ranges.



A **B**

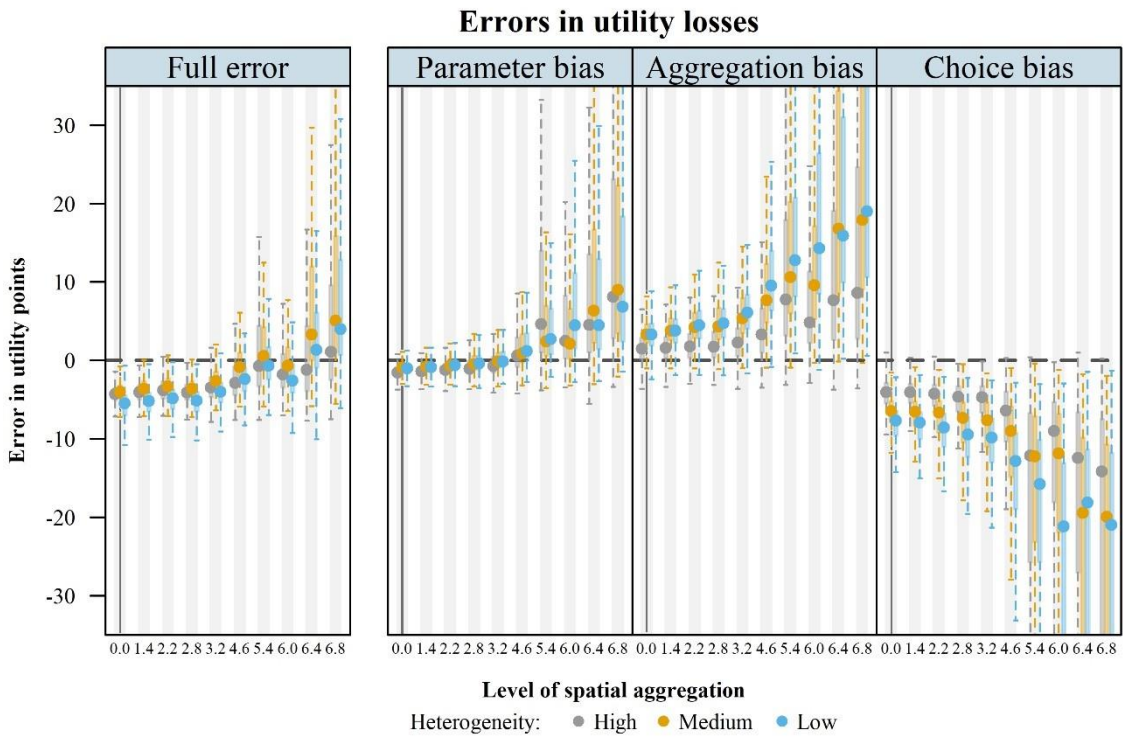
Figure 4. Results of the Monte Carlo experiments for the three different levels of spatial heterogeneity (High, Medium and Low) in terms of capacity to estimate the model's parameters, with the DGP set at on a 1x1 (Panel A) or 10x10 (Panel B) grid. Figures shows the median of the estimated marginal rate of substitution of the short-term component of the expected VPUE when no information is missing (results are similar for the other explanatory variables). Box edges are the 25th and 75th percentiles of the distribution of the losses and whiskers are located at +/- 1.5 times the interquartile ranges.

Figure 4 shows the distributions of the estimated scaled MU for the short-term component of the expected VPUE in the case where long-term information is also available. Results for the other variables are similar and are presented in table F.1 of the Appendix[15]. As the level of spatial aggregation increases, we observe that the distribution of the MU widens and, even though the MU remain not statistically different from the "true" value on average (-3 in this case), the likelihood of obtaining an estimated *significant* parameter which is far off – and possibly of the opposite sign than - the true value dramatically increases. Not surprisingly, when using overly refined spatial scales, MU are found to be less and less significantly different from 0 (Figure 4 Panel B). Regarding the effects of the spatial heterogeneity of datasets on the distribution of biases in parameter estimates, we cannot identify any particular trend.

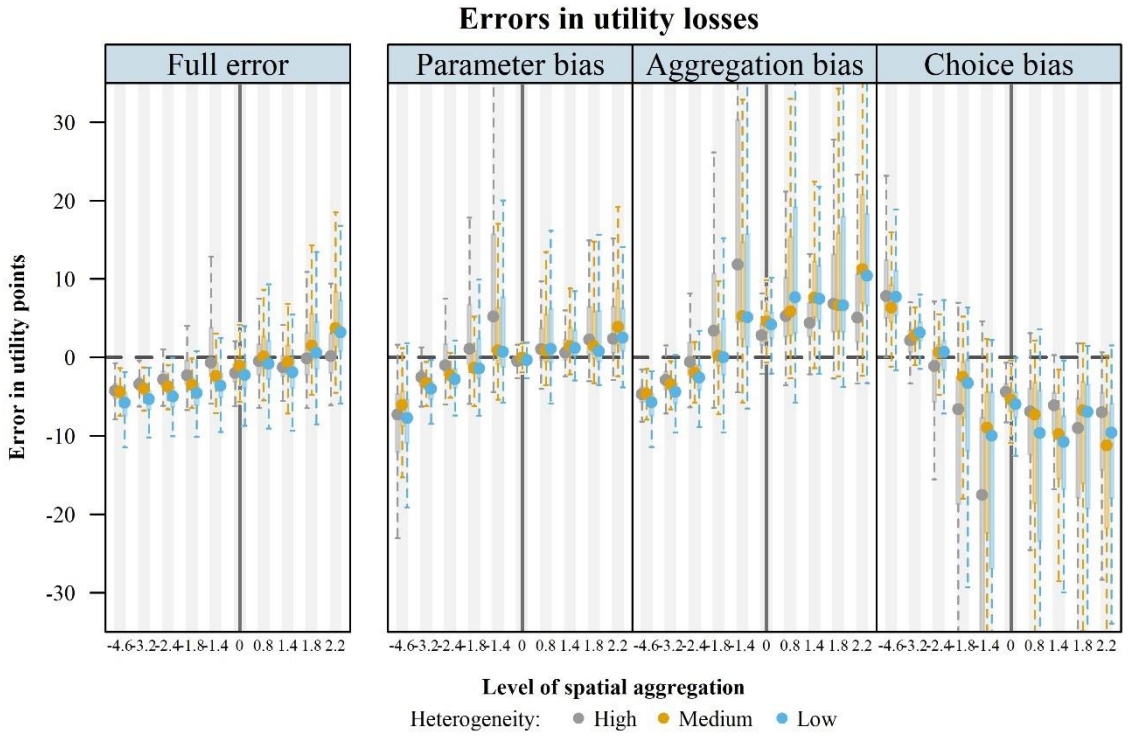
Evaluation of the welfare impact of a hypothetical MPA

Given the past use of RUM models for policy analysis, specifically the impacts of closing areas off to fishing, we investigate how the spatial resolution of the analysis interacts with welfare estimates associated with a hypothetical MPA.

[15] MRS that were not significantly different from 0 were excluded from the analysis. These cases usually occurred because the MU of distance were not significantly different from 0. They represent no more than 13% of the draws at the coarser resolution and the number of cases decline rapidly toward 0 as the resolution increases (there are zero cases for the four more spatially refined resolutions).



A



B

Figure 5. Results of the Monte Carlo experiments in terms of capacity to evaluate the welfare impact of a MPA, with the DGP set at on a 1x1 (Panel A) or 10x10 (Panel B)

grid. The figure shows the distributions of the full error in the estimated mean utility loss of fishers (left panel), along with its decomposition into the part of the error due to the bias in the estimated parameters (right panel, 1st column), due to the spatial aggregation of the explanatory variables (right panel, 2nd column) and due to the utility maximization under the incomplete choice set (right panel, 3rd column). Cases where the marginal utility of $Dist$ were found to be positive or not significant were discarded. Utility levels were computed using only significant variables (the results do no change when not doing so). Box edges are the 25th and 75th percentiles of the distribution of the losses and whiskers are located at +/- 1.5 times the interquartile ranges.

In the Monte Carlo experiments, we decompose the error in the normalized utility losses into three components. Figure 5 Panel A shows the distribution of the mean errors for increasingly aggregated nested choice sets[16]. First, we find that even when a model is estimated at the same spatial scale as the DMP, there is a systematic underestimation of the utility losses due to the restriction of the choice set to only the set of "observed" chosen alternatives (the combination of the aggregation and the choice biases). For intermediate levels of aggregation, the underestimation of the losses due to the misspecification of the choice set is compensated and eventually dominated by the overestimation due to the bias in the estimated parameters. The magnitudes of the errors are widely distributed among the levels of spatial aggregation. When looking only at truly impacted choices (i.e., excluding false positives), the estimated

[16] The effect due to improper choice set specification ("Ch bias") has been computed as the difference between the full error and the other two components. Cases where the MU of distance were positive or non-significant were excluded. Positive and significant MU of distance represent no more than 10% of cases at the most aggregated spatial scale.

utility losses are not different from the true utility losses on average but they can be as large as five times the true losses. As the level of spatial aggregation increases, the distribution of the relative errors widens even more and becomes skewed toward overestimation. On the contrary, when using more spatially-refined spatial scales we find an opposite trend, toward a reduction of the estimated welfare impact (Figure 5 Panel B). This results from the lack of significance of the estimated parameters and from the poor explanatory power of the models, which tend toward null models where the difference in utility levels of choices are directly proportional to the observed empirical distribution of choices (with a uniform distribution all options have the same utility).

Overall, the MC analysis highlights the potentially large biases of models estimated at a different spatial scale than the spatial scale of the DMP. Our analysis also reveals the great difference in the results a researcher may encounter with a single dataset when comparing spatial scales. The analysis with the data from the GoMRF longline fishery offers a typical illustration of such a situation.

Application

Results with the GoMRF longline fishery data show the same patterns as the results of the MC analysis with overly spatially-refined and non-nested estimation choice sets. Results interpretation must, however, remain cautious as, first, the true DGP is unknown and confounding effects from misspecification issues cannot be excluded ; and, then, as the metrics and parameters that are presented in this section correspond to only two samples of data (datasets from 2008 or from 2012), compared with 600 samples utilized in the MC section. Therefore, in this section the analysis of the effect of the spatial scale is carried out at the *point* level, without knowledge about the whole distribution of the observations.

Goodness-of-fit and prediction capability

We find model's goodness-of-fit to be maximal at 0.76 with the 2.5°x2.5° grid and to decrease monotonically for more refined or coarser resolutions, except for a jump at 0.54 for the 3°x3° grid (Figure 6 Panel A.). A similar pattern can be observed with model's prediction performance, with the rate of prediction errors reaching a minimum below 10% for the 2.5°x2.5° grid and increasing as the resolution used for the estimation gets finer or coarser. A jump for the 3°x3° grid can also be observed, but only for 2008 dataset (Figure 6 Panel B).[17] In this way, results for the GoMRF match those of the MC experiments where the 2.5°x2.5° grid would be close to the spatial scale used for the DMP and where the 3°x3° grid would have a spatial structure markedly mismatching the "true" choice set. Nonetheless, contrary to the results of MC experiments we do not observe a sharp decrease when halving the resolution of the supposedly "true" resolution. The fact that distance remains a meaningful predictor of spatial choices – as opposed to other predictors – even for the most refined or coarse spatial scales probably helps to explain this absence of sharp decrease of the goodness-of-fit. In addition, the large distributions of prediction accuracy we obtained in the MC experiments for the same spatial scale and the relatively limited shift in the centers of the distribution from one spatial scale to another should question the significance of the downward trends in our two

[17] To investigate, a possible effect of the spatial distribution of the observations in a more formal way than just comparing three levels of spatial heterogeneity in the Monte Carlo experiments, we looked also at possible correlations between the prediction error rates and spatial indexes capturing the distribution of the observations per alternative. Namely, the first three moments as well as Shannon diversity and equitability indexes. We found no systematic correlations, be it within or across the simulated or real case studies.

empirical datasets. All the more so as the prediction scores obtained in the applied case also demonstrate some variability as the level of spatial aggregation increases (results are not monotonic).

Parameter estimates

As mentioned previously, the spatial aggregation impacts the observations that are utilized in the calculation of expected catch (e.g., smaller areas might not have information available for calculations in a given period). To investigate the robustness of our estimation results to the availability of information both in the short-term and long-term, we estimate Eq. (2) under a number of cases, testing for different combinations of time windows and scope for the inclusion of information[18]. In each analysis, we follow the literature on how to deal with missing information. Specifically, we take the same approach as Abbott and Wilen (Abbott and Wilen, 2011) interacting the different information signals with dummies for each case of type of information available and including a dummy for when no information is available. We also estimate each model considering a mixed specification with all the estimated parameters assumed to be random. Results being almost the same between the mixed and the multinomial specification; we present only results for the latter.

Figure 6 Panel C shows the estimates for the same scaled MU as in the MC experiments, using the two GoMRF datasets for each of the 9 different spatial choice sets. Results for the tessellation based on the pre-2013 NMFS statistical areas are shown separately at the right of each panel. The MU remains close to 0 for the three lowest levels of spatial aggregation before displaying more erratic patterns at higher levels of aggregation. Such a high variability is also

[18] We consider cases where the information taken into account could be only public, or public and private; and could span only the month prior to the beginning of a fishing trip, the month and the whole year prior, or only the same month of the year prior.

observed in the other explanatory factors (Table II). In particular, none of the variables forming the expected VPUE maintain consistent signs across the different tessellations.

Table II Scaled marginal utilities (MU) of model variables estimated with the conditional logit model over the nine different tessellations with 2008 data. Numbers indicate the percentage of time the MU is significantly negative (left side) or positive (right side). Short-term information refers to the records of mean levels of VPUE across the fleet over the past 30 days. Long-term information refers to records of the mean levels of VPUE across the fleet over the 30 days prior the same date the year before.

<table>
<tr><th>Variable</th><th colspan="4">% Significant (<0|>0)</th></tr>
<tr><td rowspan="4">E[VPUE]</td><td colspan="2" rowspan="2"></td><td colspan="2">Short-term information</td></tr>
<tr><td colspan="2">Not missing</td><td>Missing</td></tr>
<tr><td rowspan="3">Long-term information</td><td rowspan="2">Not missing</td><td>Short-term:</td><td>22|33</td><td rowspan="2">0|56</td></tr>
<tr><td>Long-term:</td><td>11|67</td></tr>
<tr><td>Missing</td><td colspan="2">22|11</td><td>-</td></tr>
<tr><td>Eff.own</td><td colspan="4">11|89</td></tr>
<tr><td>Eff.oth</td><td colspan="4">100|0</td></tr>
</table>

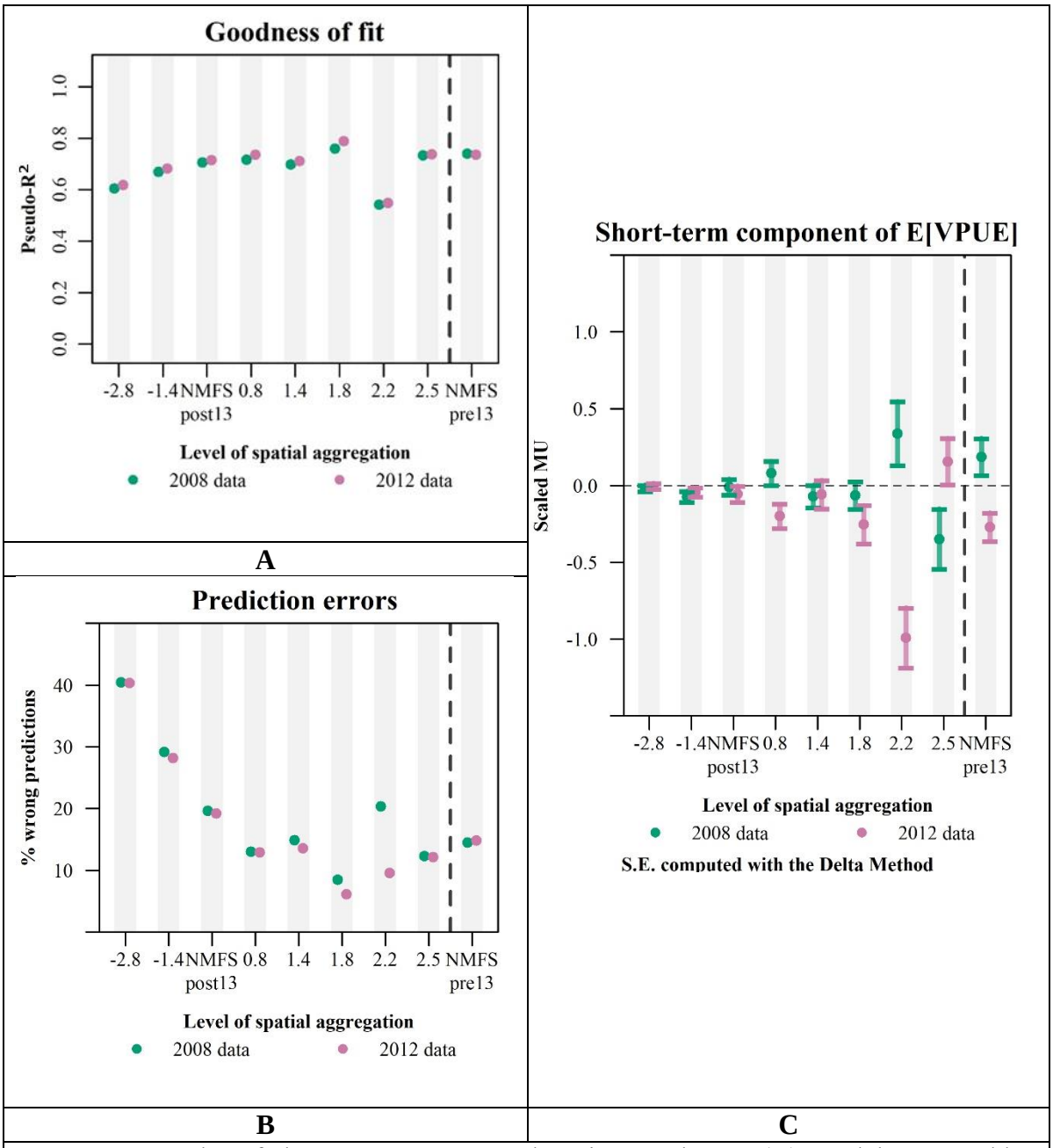


Figure 6. Results of the GoMRF case study. Figures shows: (A) model's McFadden pseudo-R^2, (B) model's rate of prediction errors; and (C) point estimates and 95% confidence intervals of the marginal utility of the short-term component of the expected VPUE when no information is missing (results are similar for the other explanatory variables).

The results of the MC experiments regarding the estimates of model parameters provide insight to the results of the case study. In both cases, we find that overly refined spatial scales lead to MU close to 0 (explanatory variables do not discriminate choices) and we observe a seemingly increase in the variability of the estimated MU with the level of spatial aggregation. However, the MC experiments reveal that even though a particular MU may be statistically significant with a specific dataset and at a given level of spatial aggregation, it may actually not be significantly different from 0 when considering other possible datasets of observations generated according to the same decision-making process. This means that a researcher, who would be limited to a single set of observations and that would operate at a single level of spatial aggregation, may well find (or search for) statistically significant MU having the opposite signs of the true MU, *even though their model is properly specified*. Therefore, when operating at too high a level of spatial aggregation, an estimated model may become inconsistent just because of the effect of spatially aggregating choices and explanatory variables.

This higher variability of the estimated MU at high levels of spatial aggregation is also reflected in the GoMRF application through the inconsistency of the estimates of model parameters and their increased invariance. When looking at the average marginal effects of each variable, we also find the same pattern of increased variability at aggregated tessellations for the VPUE variables as we did in the simulations (Appendix Figures F.1.d and F.1.e). Nonetheless, for the applied case such instability could also result from a misspecification of our model, which cannot be excluded given its simplicity.

In addition, another important caveat to these findings is that they may be due to the temporal resolution of the information used to form daily expectations of returns in a given site. Indeed, information about catches and revenues were collected at the level of the fishing trip and scaled down by days and location proportionally to the fishing activity recorded based on processed VMS data (Appendix Section A). Having had daily information about catches, for

example, may have helped capture intra-fishing trip variability which would improve the explanatory power of our model.

Estimated losses for the hypothetical MPA in the GoMRF fishery

The hypothetical MPA that we considered for the GoMRF was delineated according to the bathymetry off the west coast of Florida to mimic how spatial regulations are actually designed in this part of the Gulf.[19] To examine the robustness of the impact evaluation to the input or training datasets, we considered the implementation of the MPA for both 2008 and 2012 and looked at results using models estimated with either 2008 or 2012 data. Figure 7 shows the distribution of the normalized losses for each choice set configuration in each of the four combinations of training and input datasets. Results are robust to the choice of the input dataset and are consistent with the simulations: the mean and median of the estimated welfare losses, as well as the dispersion of the losses, increase with the level of spatial aggregation, but with two jumps for the 2°x2° and the 3°x3° grids. Once again, this pattern is similar as in the MC experiments where the 2.5°x2.5° grid would be matching the true spatial choice set used for decision-making.

An implication of these findings is that prior policy analysis of MPAs (or any spatial policy for that matter) can be confounded by the underlying spatial scale of the analysis. The

[19] Given that, depending on the spatial configuration of the fishing sites, the sites are not necessarily nested into the MPA, we slightly modified the tessellations by splitting in inside or outside fishing sites, fishing sites that were overlapping the MPA. We also estimated welfare impacts when merging together sites that were inside the MPA so that they would appear as a single alternative. In the end, we did not find large differences between the two approaches (the welfare losses with the MPA as a single alternative were just slightly lower in general).

difficulty is that this confounding is not obvious and will not be apparent to a researcher only able to consider a single spatial scale (e.g., if data are reported by management area).

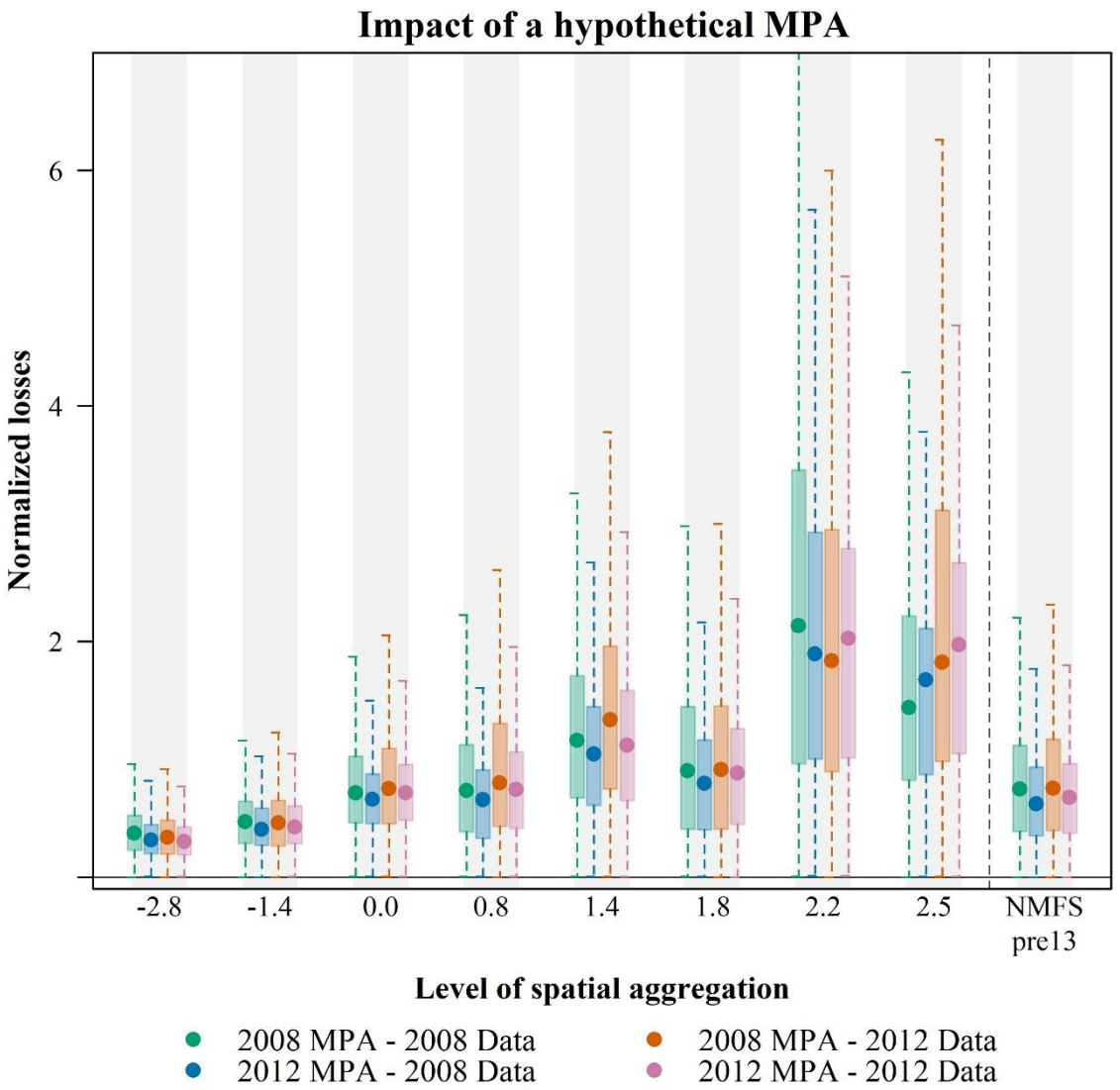


Figure 7. The estimated welfare impact of the hypothetical MPA for longline fishers in the Gulf of Mexico as a function of the level of spatial aggregation. Filled dots indicate the mean utility losses. The box edges are the 25th and 75th percentiles of the distribution of the losses and whiskers are located at +/- 1.5 times the interquartile ranges.

Conclusion

We developed an innovative empirical approach to examine the complex relationship between spatial decisions and the specification of the spatial unit of analysis in a discrete-choice model (DCM). We developed Monte Carlo simulations of fishing decisions that mimic the typical real conditions that researchers face when studying fleet dynamics. The simulations highlighted the - otherwise undetectable - biases induced by the spatial aggregation of observations. These results point to the importance for the robustness of the results of spatial models to be able to test for the consequences of alternative spatial partitioning of the choice sets.

We illustrate this using fine-scale geospatial data on the activity of commercial fishing vessels in the Gulf of Mexico to estimate a standard DCM of fishing locations over nine different spatial choice sets. We show that model outcomes can be highly sensitive to the scale of analysis, especially because of variation in the spatial heterogeneity of the values of some predictors - in particular the distance to fishing sites -, or because the data heterogeneity captured by the model mismatches the level of heterogeneity which is actually considered by decision-makers. This may be case when the attributes of choice options are computed at an overly refined resolution for instance. Therefore, the likelihood of obtaining biased results increases when using granular data to model the spatial decisions in fisheries where fishing sites cover large spatial extents.

Moving forward, our work clearly demonstrates the value for spatial modelers to use more spatially-refined data, which would allow testing the robustness of their results and gain insights into potential biases due to spatial aggregation. Whereas granular spatial data is indeed becoming more widely available, the possible increase in the reliability of results may be now limited by a lack of corresponding progress regarding the availability of temporally refined data. In the case of the modeling of fishery dynamics, the data revolution triggered by the

increase availability of vessels' spatial activity data through VMS may now turn out to be limited by the access to temporally refined information about catch. Presently, data collection frameworks across countries remain inconsistent, with the reporting of catches that can be by haul, day or even fishing trip as in our case. The lack of robustness in parameters estimates also sheds light on the need to ensure the consistency between data resolution and the scale of the decisions considered in terms of both spatial and temporal resolution.

In our analysis of the effects of varying the partitions of space considered in a spatial model, we restricted ourselves to only grid-based partitions of space and we did not consider alternative shapes for the fishing grounds. The few studies that have investigated non grid-based partitions in the DCM literature (Hynes et al., 2016; Valcic, 2009) have not examined sensitivity to the partitioning of space considered. Yet, various approaches exist to define spatial zonings in relation to the spatial distribution of observations. For instance, it would be interesting to see how randomly-generated partitions of space, such as those generated using techniques developed in the geography literature (Openshaw, 1983, 1977; Wong, 2009), compare against homogeneous grids. Certainly, more advanced geo-statistical techniques (e.g., Páez and Scott, 2005) would be worth investigating as well in the context of spatial fishery models, in particular to better handle spatial heterogeneity of data availability. On that aspect, there might be gains to considering meteorological models and other methods specifically developed to deal with highly spatially and temporally heterogeneous observations, in order to improve the predictions of spatial fishery modeling. Data re-assimilation techniques (where predicted outcomes can be looped in as training data), or model averaging (ensemble modeling) could be, in this regard, fruitful avenues to pursue.

References

Abbott, J.K., Wilen, J.E., 2011. Dissecting the tragedy: A spatial model of behavior in the commons. J. Environ. Econ. Manag. 62, 386–401. https://doi.org/10.1016/j.jeem.2011.07.001

Bierlaire, M., Hurtubia, R., Flötteröd, G., 2010. Analysis of Implicit Choice Set Generation Using a Constrained Multinomial Logit Model. Transp. Res. Rec. J. Transp. Res. Board 2175, 92–97. https://doi.org/10.3141/2175-11

Curtis, R., Hicks, R.L., 2000. The Cost of Sea Turtle Preservation: The Case of Hawaii's Pelagic Longliners. Am. J. Agric. Econ. 82, 1191–1197. https://doi.org/10.1111/0002-9092.00119

Eales, J., Wilen, J.E., 1986. An examination of fishing location choice in the pink shrimp fishery. Mar. Resour. Econ. 331–351.

Evans, K.S., Weninger, Q., 2014. Information Sharing and Cooperative Search in Fisheries. Environ. Resour. Econ. 58, 353–372. https://doi.org/10.1007/s10640-013-9701-8

Girardin, R., Hamon, K.G., Pinnegar, J., Poos, J.J., Thébaud, O., Tidd, A., Vermard, Y., Marchal, P., 2016. Thirty years of fleet dynamics modelling using discrete-choice models: What have we learned? Fish Fish. https://doi.org/10.1111/faf.12194

Girardin, R., Vermard, Y., Thébaud, O., Tidd, A., Marchal, P., 2015. Predicting fisher response to competition for space and resources in a mixed demersal fishery. Ocean Coast. Manag. 106, 124–135. https://doi.org/10.1016/j.ocecoaman.2015.01.017

Hauser, J.R., 2014. Consideration-set heuristics. J. Bus. Res. 67, 1688–1699. https://doi.org/10.1016/j.jbusres.2014.02.015

Haynie, A.C., 2005. The Expected Profit Model: A new Method to Measure the Welfare Impacts of Marine Protected Areas (Ph.D. dissertation). University of Washington, Seattle, WA.

Haynie, A.C., Layton, F.D., 2010. An expected profit model for monetizing fishing location choices. J. Environ. Econ. Manag. 59, 165–176. https://doi.org/10.1016/j.jeem.2009.11.001

Hicks, R.L., Schnier, K.E., 2010. Spatial regulations and endogenous consideration sets in fisheries. Resour. Energy Econ. 32, 117–134. https://doi.org/10.1016/j.reseneeco.2009.11.008

Horowitz, J.L., Louviere, J.J., 1995. What is the role of consideration sets in choice modeling? Int. J. Res. Mark. 12, 39–54. https://doi.org/10.1016/0167-8116(95)00004-L

Huang, L., Smith, M.D., 2014. The Dynamic Efficiency Costs of Common-Pool Resource Exploitation. Am. Econ. Rev. 104, 4071–4103. https://doi.org/10.1257/aer.104.12.4071

Hutniczak, B., Münch, A., 2018. Fishermen's location choice under spatio-temporal update of expectations. J. Choice Model. 28, 124–136. https://doi.org/10.1016/j.jocm.2018.05.002

Hynes, S., Gerritsen, H., Breen, B., Johnson, M., 2016. Discrete choice modelling of fisheries with nuanced spatial information. Mar. Policy 72, 156–165. https://doi.org/10.1016/j.marpol.2016.07.004

Lewison, R., Hobday, A.J., Maxwell, S., Hazen, E., Hartog, J.R., Dunn, D.C., Briscoe, D., Fossette, S., O'Keefe, C.E., Barnes, M., Abecassis, M., Bograd, S., Bethoney, N.D., Bailey, H., Wiley, D., Andrews, S., Hazen, L., Crowder, L.B., 2015. Dynamic Ocean Management: Identifying the Critical Ingredients of Dynamic Approaches to Ocean Resource Management. BioScience 65, 486–498. https://doi.org/10.1093/biosci/biv018

Li, L., Adamowicz, W., Swait, J., 2015. The effect of choice set misspecification on welfare measures in random utility models. Resour. Energy Econ. 42, 71–92. https://doi.org/10.1016/j.reseneeco.2015.07.001
Manski, C.F., 1977. The structure of random utility models. Theory Decis. 8, 229–254. https://doi.org/10.1007/BF00133443
McFadden, D., 1978. Modelling the Choice of Residential Location, in: Karlqvist, A., Snickars, F., Weibull, J. (Eds.), Spatial Interaction Theory and Planning Models, Studies in Regional Science and Urban Economics. North-Holland Pub. Co. : distributors for the U.S.A. and Canada, Elsevier North-Holland, Amsterdam ; New York, pp. 75–96.
Narayana, C.L., Markin, R.J., 1975. Consumer Behavior and Product Performance: An Alternative Conceptualization. J. Mark. 39, 1. https://doi.org/10.2307/1250589
O'Farrell, S., Sanchirico, J.N., Chollett, I., Cockrell, M., Murawski, S., Watson, J., Haynie, A.C., Strelcheck, A., Perruso, L., 2017. Improving detection of short-duration fishing behaviour in vessel tracks by feature engineering of training data. ICES J. Mar. Sci. 74, 1428–1436. https://doi.org/10.1093/icesjms/fsw244
Openshaw, S., 1983. The Modifiable Areal Problem, in: Classification Using Information Statistics, Concepts and Techniques in Modern Geography. Geo Books, Norwich [Norfolk].
Openshaw, S., 1977. A Geographical Solution to Scale and Aggregation Problems in Region-Building, Partitioning and Spatial Modelling. Trans. Inst. Br. Geogr. 2, 459. https://doi.org/10.2307/622300
Páez, A., Scott, D.M., 2005. Spatial statistics for urban analysis: A review of techniques with examples. GeoJournal 61, 53–67. https://doi.org/10.1007/s10708-005-0877-5
Parsons, G.R., Hauber, A.B., 1998. Spatial Boundaries and Choice Set Definition in a Random Utility Model of Recreation Demand. Land Econ. 74, 32. https://doi.org/10.2307/3147211
Parsons, G.R., Plantinga, A.J., Boyle, K.J., 2000. Narrow Choice Sets in a Random Utility Model of Recreation Demand. Land Econ. 76, 86. https://doi.org/10.2307/3147259
Peters, T., Adamowicz, W.L., Boxall, P.C., 1995. Influence of choice set considerations in modeling the benefits from improved water quality. Water Resour. Res. 31, 1781–1787.
Poos, J.-J., Rijnsdorp, A.D., 2007. An "experiment" on effort allocation of fishing vessels: the role of interference competition and area specialization. Can. J. Fish. Aquat. Sci. 64, 304–313. https://doi.org/10.1139/f06-177
Pramono, A., Oppewal, H., 2012. Accessibility and the role of the Consideration Set in Spatial Choice Modelling: A Simulation Study. J. Choice Model. 5, 46–63. https://doi.org/10.1016/S1755-5345(13)70047-2
Prato, C.G., 2009. Route choice modeling: past, present and future research directions. J. Choice Model. 2, 65–100.
Rijnsdorp, A.D., Poos, J.J., Quirijns, F.J., Grant, J., 2011. Spatial dimension and exploitation dynamics of local fishing grounds by fishers targeting several flatfish species. Can. J. Fish. Aquat. Sci. 68, 1064–1076. https://doi.org/10.1139/f2011-032
Russo, T., Pulcinella, J., Parisi, A., Martinelli, M., Belardinelli, A., Santojanni, A., Cataudella, S., Colella, S., Anderlini, L., 2015. Modelling the strategy of mid-water trawlers targeting small pelagic fish in the Adriatic Sea and its drivers. Ecol. Model. 300, 102–113. https://doi.org/10.1016/j.ecolmodel.2014.12.001
Sanchirico, J.N., Wilen, J.E., 1999. Bioeconomics of Spatial Exploitation in a Patchy Environment. J. Environ. Econ. Manag. 37, 129–150. https://doi.org/10.1006/jeem.1998.1060

Smith, M.D., 2010. Toward an econometric foundation for marine ecosystem-based management. Bull. Mar. Sci. 86, 461–477.
Smith, M.D., 2005. State dependence and heterogeneity in fishing location choice. J. Environ. Econ. Manag. 50, 319–340. https://doi.org/10.1016/j.jeem.2005.04.001
Smith, M.D., Asche, F., Bennear, L.S., Oglend, A., 2014. Spatial-dynamics of Hypoxia and Fisheries: The Case of Gulf of Mexico Brown Shrimp. Mar. Resour. Econ. 29, 111–131. https://doi.org/10.1086/676826
Stafford, T.M., 2018. Accounting for outside options in discrete choice models: An application to commercial fishing effort. J. Environ. Econ. Manag. 88, 159–179. https://doi.org/10.1016/j.jeem.2017.10.006
Torres, C., Hanley, N., Riera, A., 2011. How wrong can you be? Implications of incorrect utility function specification for welfare measurement in choice experiments. J. Environ. Econ. Manag. 62, 111–121. https://doi.org/10.1016/j.jeem.2010.11.007
Valcic, B., 2009. Spatial policy and the behavior of fishermen. Mar. Policy 33, 215–222. https://doi.org/10.1016/j.marpol.2008.06.001
van der Lee, A., Gillis, D.M., Comeau, P., Quinn, T., 2014. Comparative analysis of the spatial distribution of fishing effort contrasting ecological isodars and discrete choice models. Can. J. Fish. Aquat. Sci. 71, 141–150. https://doi.org/10.1139/cjfas-2012-0511
Watson, J.T., Haynie, A.C., Sullivan, P.J., Perruso, L., O'Farrell, S., Sanchirico, J.N., Mueter, F.J., 2018. Vessel monitoring systems (VMS) reveal an increase in fishing efficiency following regulatory changes in a demersal longline fishery. Fish. Res. 207, 85–94.
Wilson, J.A., 1990. Fishing for Knowledge. Land Econ. 66, 12. https://doi.org/10.2307/3146679
Wong, D., 2009. The Modifiable Areal Unit Problem (MAUP), in: The SAGE Handbook of Spatial Analysis. SAGE Publications, Ltd, 1 Oliver's Yard, 55 City Road, London England EC1Y 1SP United Kingdom, pp. 104–123.

Appendix

The Appendix is divided into six sections that provide further details on: A) the sources and processing of the data used for the applied case study; B) the alternative spatial indexes used to uncover possible correlations between each model's results and the spatial heterogeneity of data availability; C) the tessellations with varying levels of spatial aggregation that were considered for estimating the model; D) the full methodology followed to carry out the Monte Carlo experiments; E) the complementary estimation assumptions that have been made for estimating the model in the applied case study ; and, F) complementary results for both the Monte Carlo experiments and the applied case study.

A. Gulf of Mexico fishery Data

Data used to study the BLL fleet was provided by the National Marine Fisheries Service Southeast Fisheries Science Center (SEFSC). The SEFSC's Socioeconomic Panel (SEP) combines a variety of data sources to create a rich trip-level data set. The panel includes extensive information from the Coastal Logbook on landings disaggregated by species as well as effort data with reported variables depending on the primary gear used during the trip. In the case of BLL trips, these variables include soaking time, number of hooks per line, and the number of sets during the trip[20]. Effort data such as number of days at sea, number of crew, date of landing, and dealer identifiers are also reported. The Logbook data is further supplemented

[20] The information about soaking time was highly unreliable (fishers in some cases reported either the soaking time for the entire trip or the mean soaking time per set) and thus were discarded. Complementary analyses exploiting the observer data showed that soaking times were fairly similar across fishers - about an hour – and did not affect the level of catches of trips. Therefore, we defined and computed fishing effort for BLL as the total number of hooks having been soaked during the trip (i.e., number of hooks per line multiplied by the number of sets).

with average price data from the Accumulated Landings Service to calculate trip revenues which are also disaggregated by species. Vessel technical characteristics (e.g., vessel length) from the NMFS Southeast Regional Office Permits Office are linked to the Logbook data at the trip level. VMS data were provided by the SEFSC's Social Science Research Group.

We identified the GT-BLL fleet using the Topgear variable provided in the SEP and linked these trips to VMS using vessel identifiers along with trip start and end dates (based on reported landing date and number of days at sea). After discarding a small number of logbook entries (38 out of 4054 for the years analyzed) that had trip dates that were overlapping other entries (e.g., because of reporting mistakes or because the entries referred to a same trip), and after further subsetting GT-BLL trips to trips deriving more than 75% of the revenue from GT species, we were left with, respectively for 2008 and 2012, 816 and 420 logbook entries matched to 99,027 and 42,008 VMS observations classified as fishing, covering 362 and 350 different days of the year and representing observations from 104 and 54 vessels.

VMS pings corresponding to fishing behavior were identified using a random forest model (O'Farrell et al., 2017), trained with observer data provided by the SEFSC's Galveston Laboratory. Specifically designed and tested for this dataset, this approach makes use of observer data to devise the best classification along factors such as vessel's speed, heading and previous behavior (the estimated accuracy rate in predicting fishing activity on the training dataset is 92 %).

B. Spatial indexes used for the analysis

1. Index of spatial aggregation

To analyze the results of the simulations relative to the level of spatial aggregation used during the estimation of the DCM, we associate each tessellation to an index of spatial aggregation defined as the logarithm of the ratio of the area of the aggregated alternatives to the area of the "true" alternatives (i.e., those considered for the decision-making process):

$$I_{tess} = \ln \frac{A_{tess}}{A_0} = 2 \ln \frac{L_{tess}^{alt}}{L_0^{alt}}$$

$I_{tess} = 0$ means that the DCM was estimated at the same spatial scale as the one used during the decision-making process.

$I_{tess} = 2$ means that the DCM was estimated using alternatives that were $e^2 \approx 7.4$ times larger in terms of area, or $e^1 \approx 2.7$ times longer in terms of length, than the alternatives used during the decision-making process.

2. Indexes of spatial distribution

In addition to analyzing the effect of aggregating spatial choices made at a more refined scale, we are also interested in analyzing the effect of the spatial distribution of the observed choices. For that reason, we have considered different sizes of fishing hotspots which induced different spatial distributions of choices, with observations being more highly concentrated with smaller hotspots.

In addition to analyzing the results of the simulations individually for each of the spatial distributions, we also analyzed the results using indexes aiming at capturing the nature of the spatial distribution of choices relatively to the spatial resolution of the estimated models. For that purpose, we computed for each pair of spatial distribution/hotspot's size and tessellation, the associated Shannon entropy and equitability indexes, and the relative spatial resolution. The relative spatial resolution is defined as the logarithm of the ratio of the size of the hotspot (defined as the area including the 95% highest VPUE levels) with the size of an alternative:

$$Res_{tess}^{hs} = \ln \frac{A_{hs}}{A_{tess}} = 2 \ln \frac{1.96\sigma_{hs}\sqrt{\pi}}{L_{tess}^{alt}}$$

$Res_{tess}^{hs} = 2$ means that a given hotspot is covered by $e^2 \approx 7.4$ alternatives.

The Shannon entropy index associated with a draw d is:

$$S_d = -\sum_{i \in A_d} \hat{p}_i * \ln \hat{p}_i$$

Where, $\hat{p}_i$ is the empirical frequency of choice for the alternative i (i.e. the number of simulated fishing locations falling in alternative i over the total number of simulated fishing locations); and A_d is the whole set of alternatives considered (i.e., having $\hat{p}_i > 0$).

The more skewed the distribution of choice frequency, the smaller the corresponding entropy. If almost all of the simulated fishing locations are concentrated in only one alternative, and the other alternatives are very "rare", the entropy approaches zero.

Should the simulated fishing locations cover the alternatives in a perfectly balanced way, all the $\hat{p}_i$ would be equal, and the index would take the value $\ln N_d^A$, where N_d^A is the size of A_d (i.e., the number of alternatives considered). Therefore, a distribution of choices having a Shannon entropy index of S_d can be interpreted as "as diverse" as an even distribution of choices among $\exp(S_d)$ alternatives.

The Shannon entropy index is commonly used by ecologists as a diversity index (e.g., of species, phenotypes etc.) and as a measure of the predictability of type: the higher the index, the more diversity and the less predictable the type. This interpretation of the index can be somewhat misleading in our case since it assumes that predictions are based only on empirical frequencies and that there is not underlying model.

The Shannon equitability index is simply the Shannon entropy index normalized by its maximum value:

$$E_d = \frac{S_d}{\ln(N_d^A)}$$

C. Maps of the partitions of space considered

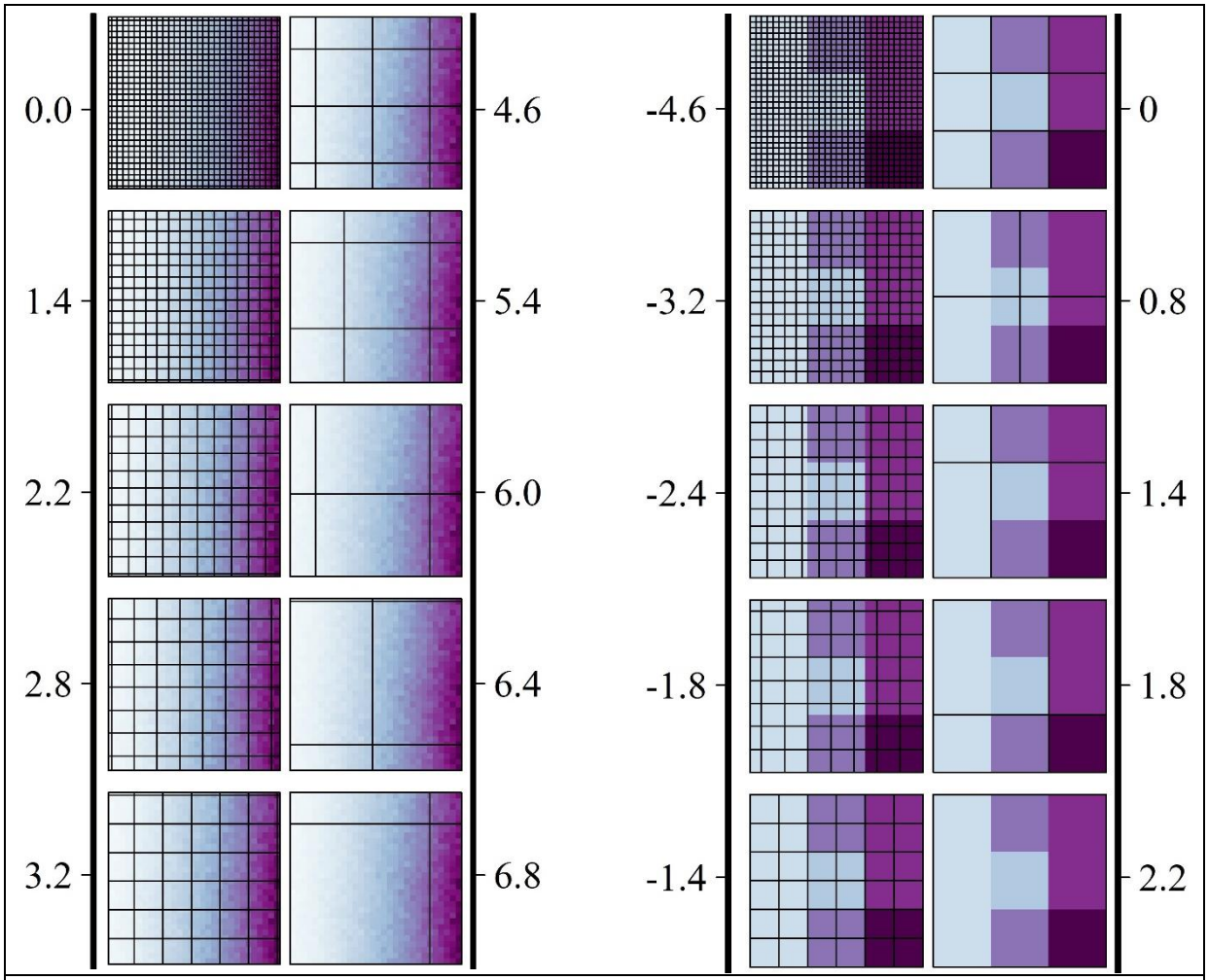


Figure C.a. Maps and indexes of spatial aggregation of the ten partitions of space considered in the Monte Carlo experiments for estimating the discrete-choice model of fishing locations, with the data-generating process set at on a 1x1 (left panels) or 10x10 (right panels) grid. The index of spatial aggregation is computed as the logarithm of the ratio of the area of the aggregated alternatives to the area of the "true" alternatives.

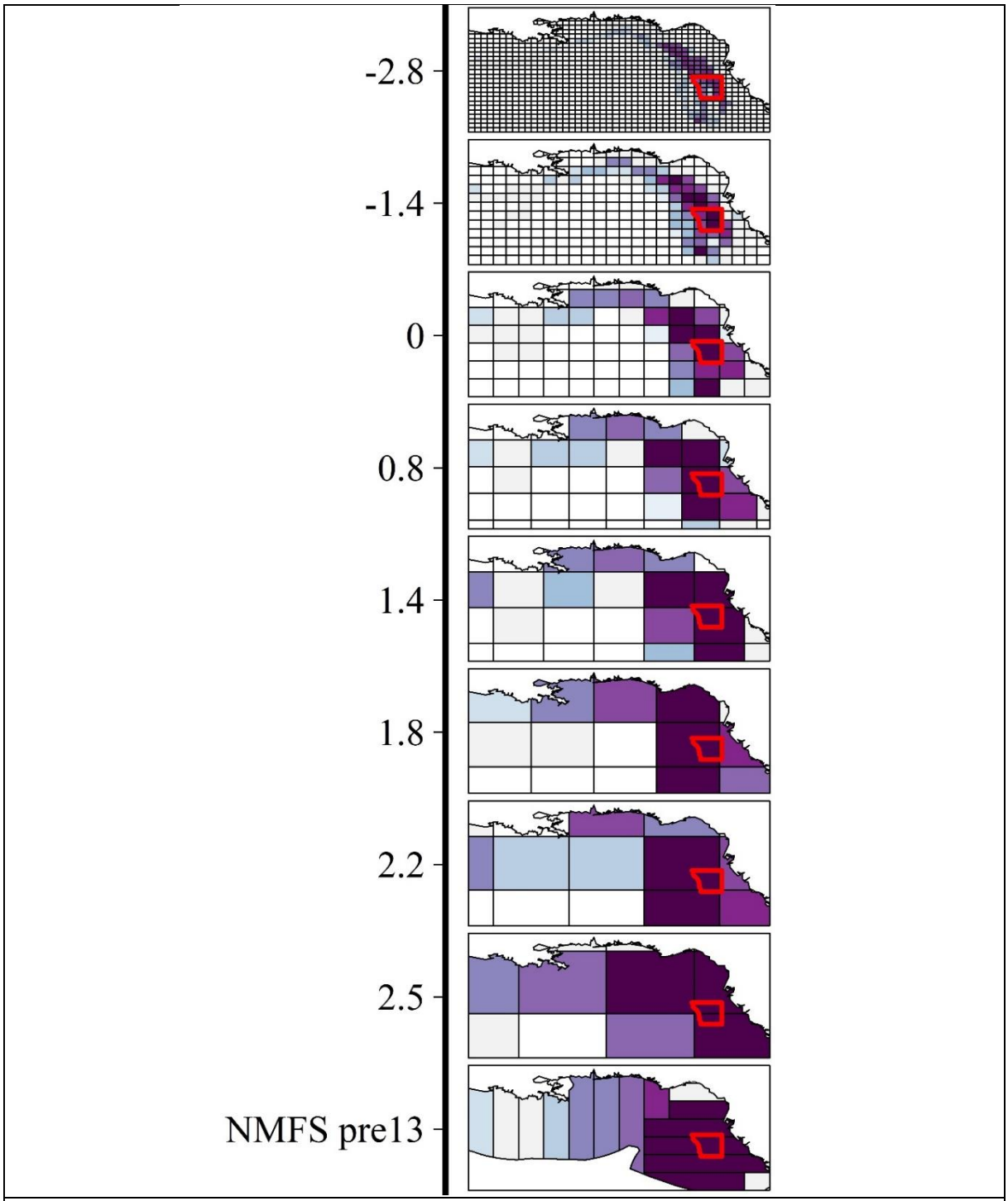


Figure C.b. Maps and indexes of spatial aggregation of the nine tessellations of the Gulf of Mexico considered for estimating the discrete-choice model of fishing locations. The index of spatial aggregation is computed as the logarithm of the ratio of the area of the aggregated alternatives to the area of the alternatives considered by the NMFS, starting 2013, for the reporting of fishing locations in logbooks.

D. Monte Carlo experiments

1. Utility function

We assume vessels' decide on their fishing location according to the following utility function:

$$U(lon, lat) = \beta_{dist} * Dist(lon, lat) + \beta_{VPUE} * VPUE(lon, lat, t)$$

With:

- $Dist(lon, lat) = C + \sqrt{(lon)^2 + (lat)^2}$ *:* We assume a "fixed-cost" of moving and that all vessels start their trip from the origin (0,0)
- $VPUE_h(lon, lat, t) = \overline{VPUE_h}(t) e^{-\frac{(lon - h_{lon})^2}{\sigma_{lon}^2}} e^{-\frac{(lat - h_{lat})^2}{\sigma_{lat}^2}}$: We assume that there is a hotspot in (h_{lon}, h_{lat}) with a Gaussian spatial distribution of fish abundance that depends on time.
    - We assume 3 hotspots located in the North (2.6, 2.5), Center (3.1, 0) or South (1.9, -2.4).
    - We consider 4 hotspot sizes taking $\sigma_{lon} = \sigma_{lat} \in \{0.4, 0.8, 1.2, 1.6\}$
- $\overline{VPUE_h}(t) = \overline{VPUE}.\,base\left(1 + A * \cos\left(2\pi \frac{t - t_h}{T}\right)\right)$:
    - The productivity of hotspots oscillates around $\overline{VPUE}.\,base$ with a period $T = 1$ and reaches their maxima at $t_h \in \left\{\frac{2}{12}, \frac{5}{12}, \frac{8}{12}\right\}$.
    - We set $\overline{VPUE}.\,base = 4$.
- The productivity of a given point in space is the mean of the productivity of each hotspot at this given point with a stochastic error of +/- 100%:
    - $VPUE(lon, lat, t) = \left(1 + u(lon, lat, t)\right)\left(\frac{1}{n_h}\sum_h VPUE_h(lon, lat, t)\right)$
    - $u(lon, lat, t) \sim U(-1,1)$

2. Expected VPUE

We assume the fishers can only form an expected VPUE at a given location based on their past observations and those of the fleet. Additional assumptions are required in this case:

a. Spatial extent of the expectations: we assume that fishers form their expectations on either a refined grid of 1x1 NM or 10x10 NM (1NM = $\frac{1}{60}$°, 115,200 alternatives for a 4°lon x 8°lat space)

b. Temporal extent of the expectations: we assume that fishers use the VPUE records of:

   i. the past 30 days

   ii. the past 30 days around the same date the year before

c. We assume the following expected VPUE (n*ot distinguishing individual records from fleet records):*

$$E\left[VPUE_{ijt}\right] = \left(\alpha_{ft,m-1} + \alpha_{ft,ym-1} * I_{NA}\left(\overline{VPUE}^{ft}_{j(ym-1)}\right)\right)\overline{VPUE}^{ft}_{j(m-1)} + \left(\alpha_{ft,ym-1} + \alpha_{ft,m-1} * I_{NA}\left(\overline{VPUE}^{ft}_{j(m-1)}\right)\right)\overline{VPUE}^{ft}_{j(ym-1)}$$

With: $\overline{VPUE}^{ft}_{j(m-1)} = \sum_{i,t\prime\in[t-30,t[} VPUE_{ijt'}$,$\overline{VPUE}^{ft}_{j(ym-1)} = \sum_{i,t\prime\in[t-370,t-350]} VPUE_{ijt'}$,

$\begin{cases} \alpha_{ft,m-1} = 0.75 \\ \alpha_{ft,ym-1} = 0.25 \end{cases}$, and I_{NA} being a dummy function valuing 1 when the argument (average of historical VPUE records) is not available.

d. To initiate the historical records we use the draws from one year of simulated positions under the assumption that fishers have a perfect knowledge of the VPUE maps and we discard the subsequent first year of simulated positions

e. When a fishing site has been chosen by a fisher, its recorded position is set by randomly picking lon and lat coordinates within the site spatial extent.

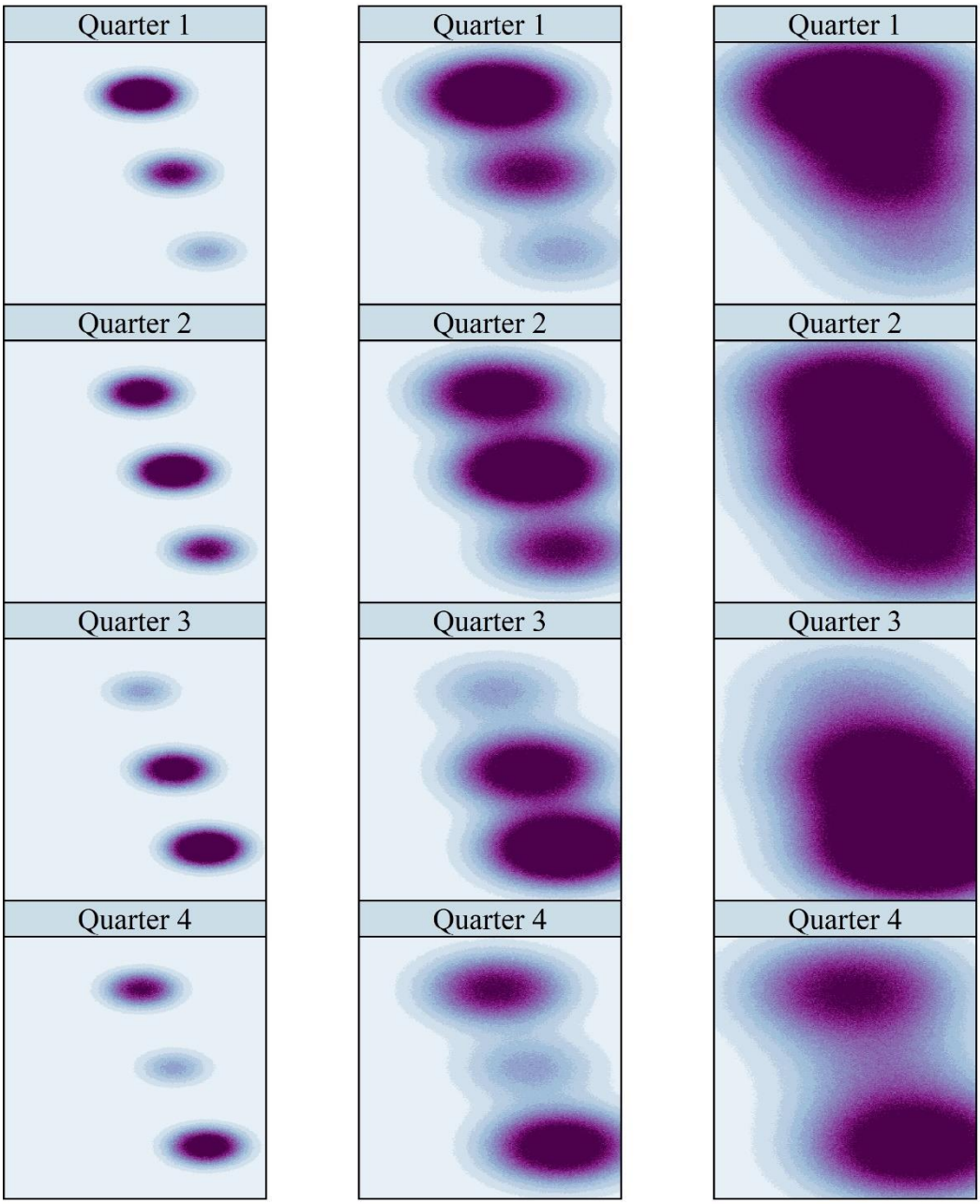


Figure D.2. Maps of simulated VPUE for four different points during the year for each of the three level of spatial heterogeneity we assume.

3. Vessels and fishing days

In the Monte Carlo analyses, we consider a fleet of identical vessels (with the same preferences) fishing during 3 periods (years) T. Vessels go out fishing all through the period but follow a probability density function proportional to the mean VPUE through the space (Figure D.3).

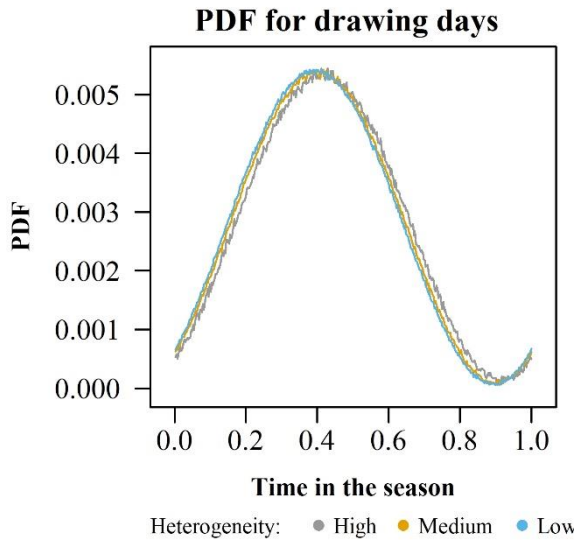


Figure D.3. Probability function for drawing the fishing days.

Since vessels are identical, we make N_{draws} draws (with replacement) of vessels from a uniform distribution.

4. Fishing locations

Whereas a continuous approach could have been taken to generate draws of fishing locations (e.g., approximating the spatial probability distribution function using the Metropolis-Hastings algorithm), we chose to take a discrete approach consisting in discretizing space at a refine scale and generating a field of random errors.

Thus, for each cell of the grid, a random error $\varepsilon(lon, lat, t)$ is drawn from a Gumbel distribution centered in 0 and with a scale of 1. We assume that errors are independent and

identically distributed which implies that the difference of two different error terms follows a logistic distribution with mean $\mu_{\Delta\varepsilon} = 0$ and a standard deviation $\sigma_{\Delta\varepsilon} = \frac{\pi}{\sqrt{3}}$. Since only differences in utilities matter, we scale the magnitude of the error terms relative to the distribution of the differences in the deterministic utilities across a whole season $\{\Delta V_{j,k}(t)\}_{j,k}$. Given the high skewness of the distributions of the utility differences (Figures D.4.a and D.4.b), we take the bottom 0.1% of the differences $P_{0.1}(\Delta V)$[21] – which is 116 alternatives in average per day - as the reference point.

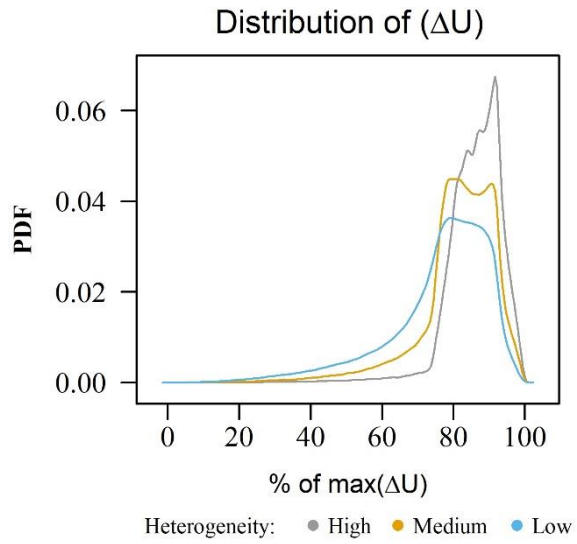


Figure D.4.a.: PDF of the differences in the deterministic utilities, for each level of spatial heterogeneity of the VPUE and over a whole season.

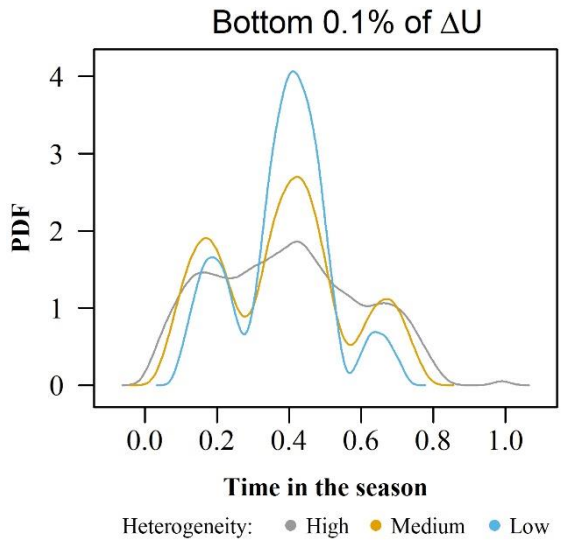
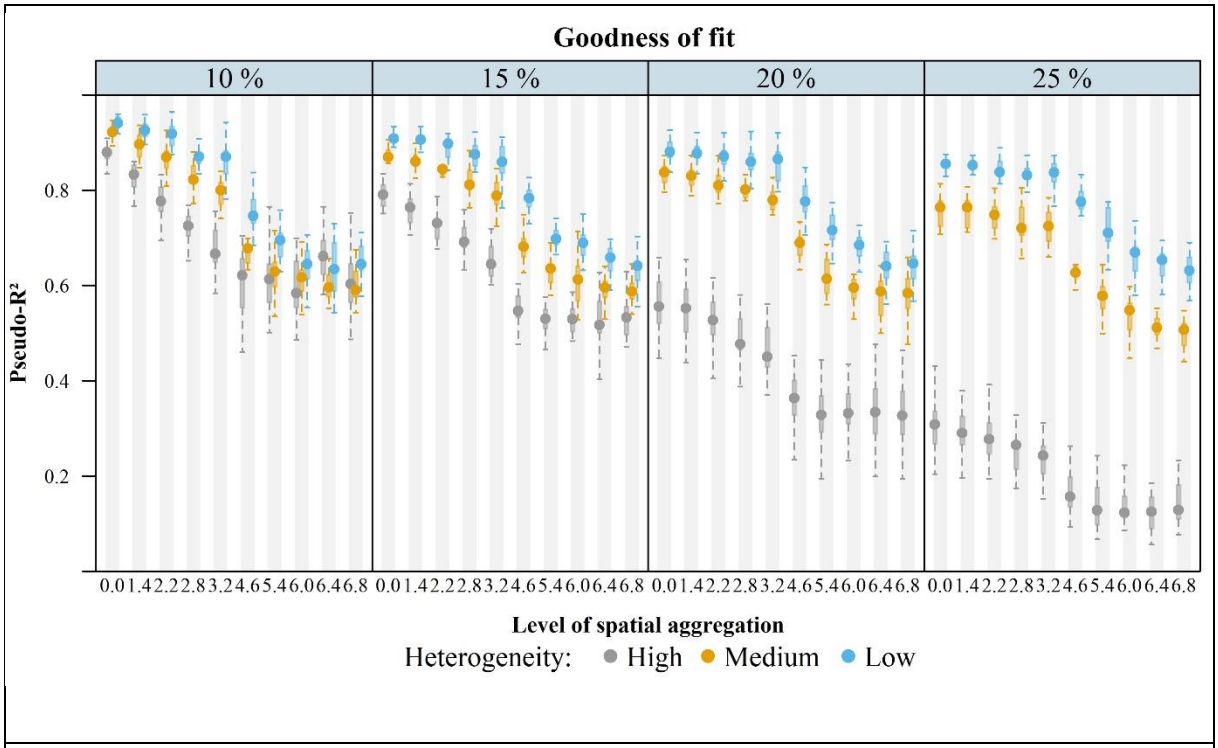


Figure D.4.b.: PDF of the temporal distribution of the bottom 0.1% of the differences in the deterministic utilities, for each level of spatial heterogeneity of the VPUE and over a whole season.

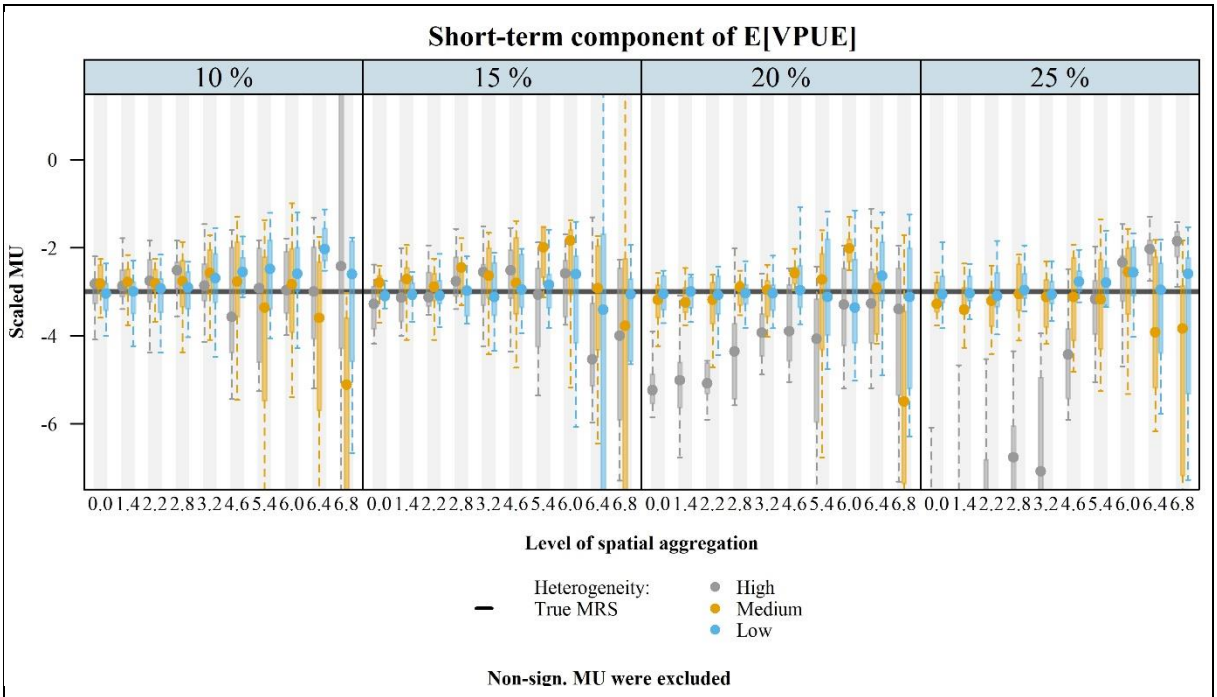
[21] i.e., the threshold for the 0.1% alternatives that are the closest to the alternative with the highest utility

After having tested for differences of error terms (Δε) having a standard deviation of 10%, 15%, 20%, and 25% of the bottom 0.1% of the differences in the deterministic utilities, we chose to scale the standard deviation of Δε as 15%, having found that starting at 15%[22] and increasing the magnitude led to more and more random simulated fishing positions inducing significant drops in the capacity of the estimated DCM to fit the data and to recover the true parameters up to a scale factor (Figure D.4.c).



A

[22] In practice, we multiply the errors terms ε by $0.15 * \frac{P_{0.1}(\Delta V)}{\frac{\pi}{\sqrt{3}}}$.



B

Figure D.4.c. Goodness-of-fit (Panel A) and estimated marginal utility of the short-term component of the expected VPUE (Panel B) of 15 Monte Carlo draws estimated for magnitudes of the stochastic part of the utilities varying from 10% to 25% (panels from left to right) of the bottom 0.1% of the distribution of differences in the deterministic part of the utility. Magnitudes of the error term higher than 15% lead to a significant drop in the capacity of the RUMs to fit the data, included at the proper spatial scale (level of spatial aggregation = 0).

5. Model estimation

We estimate the simple conditional logit model corresponding to the data-generating process, assuming that the weights for forming fishers' expectations are unknown to the researcher, and distinguish configurations of information availability using dummy variables:

$$U_{ijt}=\beta_{dist}*Dist_{ijt}+\beta_{VPUE}*E\left[VPUE_{ijt}\right]+\varepsilon_{ijt}$$

$$\beta_{VPUE}*E\left[VPUE_{ijt}\right] = \begin{cases} \beta_{VPUE}^{Full\ info-short-term} * \overline{VPUE}_{m-1}^{ft} + \beta_{VPUE}^{Full\ info-long-term} * \overline{VPUE}_{ym-1}^{ft}\ if\ case\ 1 \\ \beta_{VPUE}^{Short-term\ only} * \overline{VPUE}_{m-1}^{ft} if\ case\ 2 \\ \beta_{VPUE}^{Long-term\ only} * \overline{VPUE}_{ym-1}^{ft} if\ case\ 3 \\ \beta_{VPUE}^{No\ info}\ if\ case\ 4 \end{cases}$$

With:

- case 1: both short-term **and** long-term historical VPUE are available
- case 2: **only** short-term historical VPUE are available
- case 3: **only** long-term historical VPUE are available
- case 4: **neither** short-term **or** long-term historical VPUE are available

The hypothesis is that we should be able to recover the weights as well as β_{VPUE}:

$$H: \begin{cases} \hat{\beta}_{VPUE}^{Full\ info-short-term} = \beta_{VPUE} * \alpha_{ft,m-1} \\ \hat{\beta}_{VPUE}^{Full\ info-long-term} = \beta_{VPUE} * \alpha_{ft,ym-1} \\ \hat{\beta}_{VPUE}^{Short-term\ only} = \hat{\beta}_{VPUE}^{Long-term\ only} = \beta_{VPUE} * \left(\alpha_{ft,m-1} + \alpha_{ft,y-1}\right) = \beta_{VPUE} \\ \hat{\beta}_{VPUE}^{No\ info} = 0 \end{cases}$$

E. Applied case study

1. Choice of the empirical specification

The selection process for the choice of the empirical specification for the applied case consisted of the following steps.

First, we undertook an extensive series of pre-analyses on the bottom longline section of the GoMRF to identify possible emerging patterns of the spatial and temporal dynamics of the fishery. Besides performing statistical descriptions of the fishery (e.g., vessel characteristics, targeted species, gear, revenue structure, homeports), we combined vessels' trajectories with reported catches, revenues, and fishing effort per trip to produce monthly and yearly maps of catch rates, revenue rates and fishing activity. In addition to combining these maps with the history of management measures and our knowledge of the ecological dynamics of the fish resources (groupers and tilefishes), we ran correlation tests between reported statistical areas of fishing locations and vessel and trip characteristics as well as with landing prices. We confirmed the information in the literature that describes longline fishers as primarily single-gear, single-species fishers, and we did not find evidence of seasonal patterns regarding the location of fishing effort or the price for groupers and tilefish species. However, there did appear to be a break in the distribution of fishing effort after the implementation of new spatial restrictions in 2009 and 2010 and this motivated our choice of splitting the dataset into two periods. Overall, we did not find obvious environment-related patterns emerging, having the impression that habits and distance seemed to be the most important drivers of fishing locations, as has been reported in the literature.

Second, based on this pre-model analysis and an extensive literature review about discrete-choice models of fishing behavior, we selected a set of variables that we deemed the most relevant. Specifically, we retain: Dist_{ijt}, the distance from one location to another ;

$\text{Eff.own}_{\text{ijt-1}}$ the level of fishing effort the day before in each site for a given vessel and $\text{Eff.oth}_{\text{ijt-1}}$, for the entire fleet ; and $\overline{\text{VPUE}}_{\text{w}}^{\text{s}}$ the historical records of value per unit of effort in each site for a given vessel ($s = \text{ind}$) and for the entire fleet ($s = \text{ft}$), and for a time window including either the 30 ($w = \text{m}$) or the 365 ($w = \text{y}$) days prior the fishing trip or the 30 days the year prior ($w = y\text{m} - 1$), surrounding the date of the fishing trip.

Working with this set of variables, we explore 12 possible specifications of the GoM model (cf. Table E.1) that we estimated with both the 2007-2008 and 2011-2012 datasets for all the 9 levels of spatial aggregations described in the paper. We also assumed either constant or random coefficients (i.e., either considering a conditional logit model or a mixed logit model) in order to account for vessels' heterogeneity.

Table E.1 Combinations of information signals considered for the specification of the expected VPUE

Model #	Info. Source	Individual level			Fleet level		
	Time span	[t;t-30] (m-1)	[t;t-365] (y-1)	[t-350;t-370] (ym-1)	[t;t-30] (m-1)	[t;t-365] (y-1)	[t-350;t-370] (ym-1)
1		N	N	N	Y	N	N
2		N	N	N	N	Y	N
3		N	N	N	N	N	Y
4		Y	N	N	Y	N	N
5		N	Y	N	N	Y	N
6		N	N	Y	N	N	Y
7		N	N	N	Y	Y	N
8		N	N	N	Y	N	Y
9		N	N	N	N	Y	Y
10		N	N	N	Y	Y	Y
11		Y	N	Y	Y	N	Y
12		N	N	N	N	N	N

Shaded line: Specification presented in the main body of the paper. t-30 for example means we used the information from the month prior.

Figure E.1.a shows the goodness of fit for the 12 model's specifications estimated with 2007-2008 data, by level of spatial aggregation and for either fixed (left panel) or random parameters (right panel). Results are similar using 2011-2012 data and comparing models' prediction errors instead of models' goodness-of-fit. Figure E.1.b. shows the AIC differences of the 12 models relatively to the lowest AIC each level of spatial aggregation. Model #11 has (almost) systematically the lowest AIC, despite having the largest number of variables (50). Overall, we found very little difference between the conditional and mixed logit models, and between models distinguishing individual and fleet-level information. In order not to lose the reader with unnecessary complexity and to focus the paper on the issue of spatial aggregation, we chose to present only the parsimonious model (model #8) that we describe in the paper.

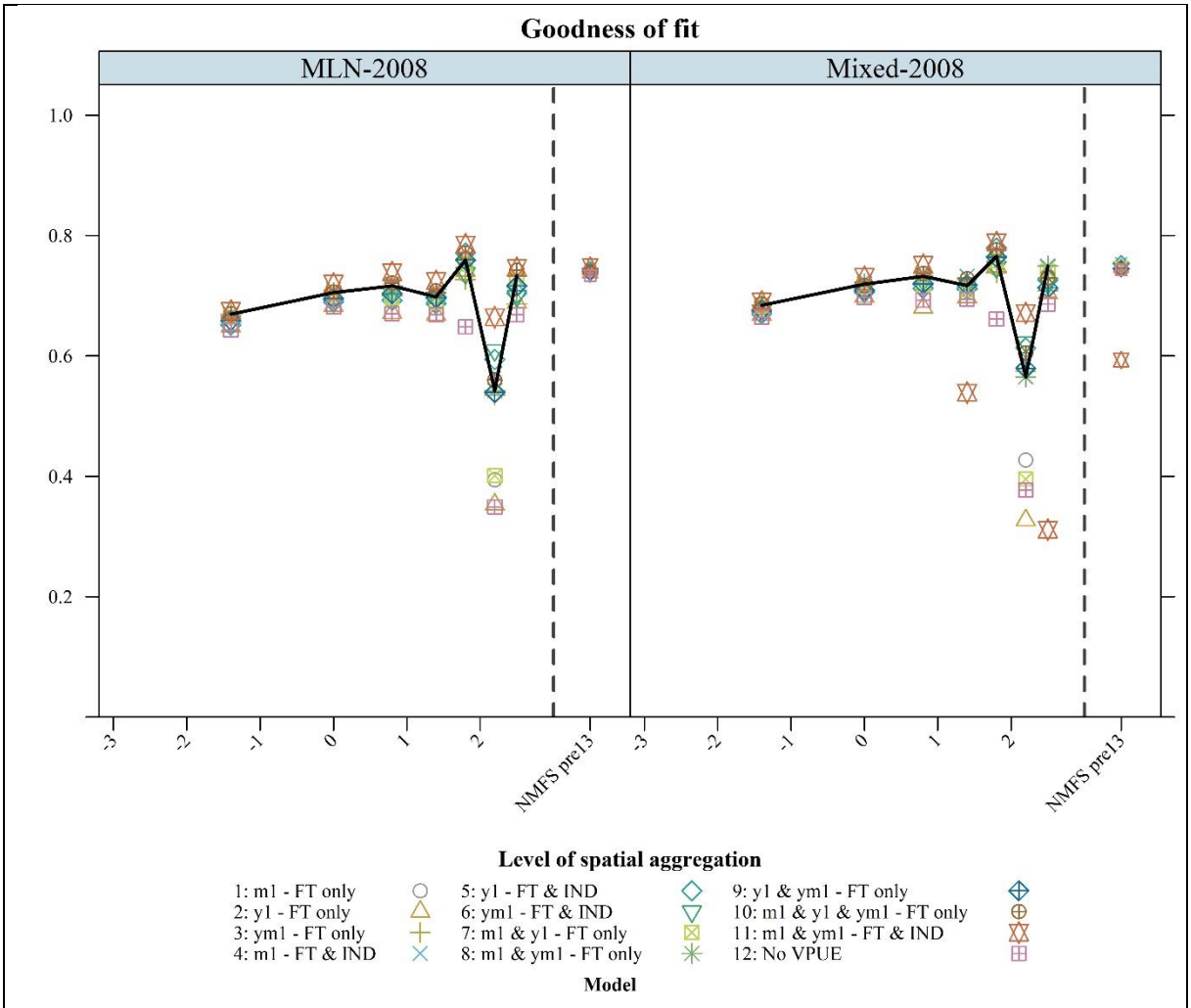


Figure E.1.a. Goodness-of-fit for the 12 model's specifications estimated with 2007-2008 data, by level of spatial aggregation and for either fixed (left panel) or random parameters (right panel). Model #8 with fixed parameters is the model presented in the main body of the paper and is highlighted with the black line.

Computational limitations did not allow to present the estimates for the most refined level of spatial aggregation for all model specifications. Convergence issues in the estimation algorithm for the most complex specification (model 11) suggest that estimates for this specification may not correspond to a global optimum.

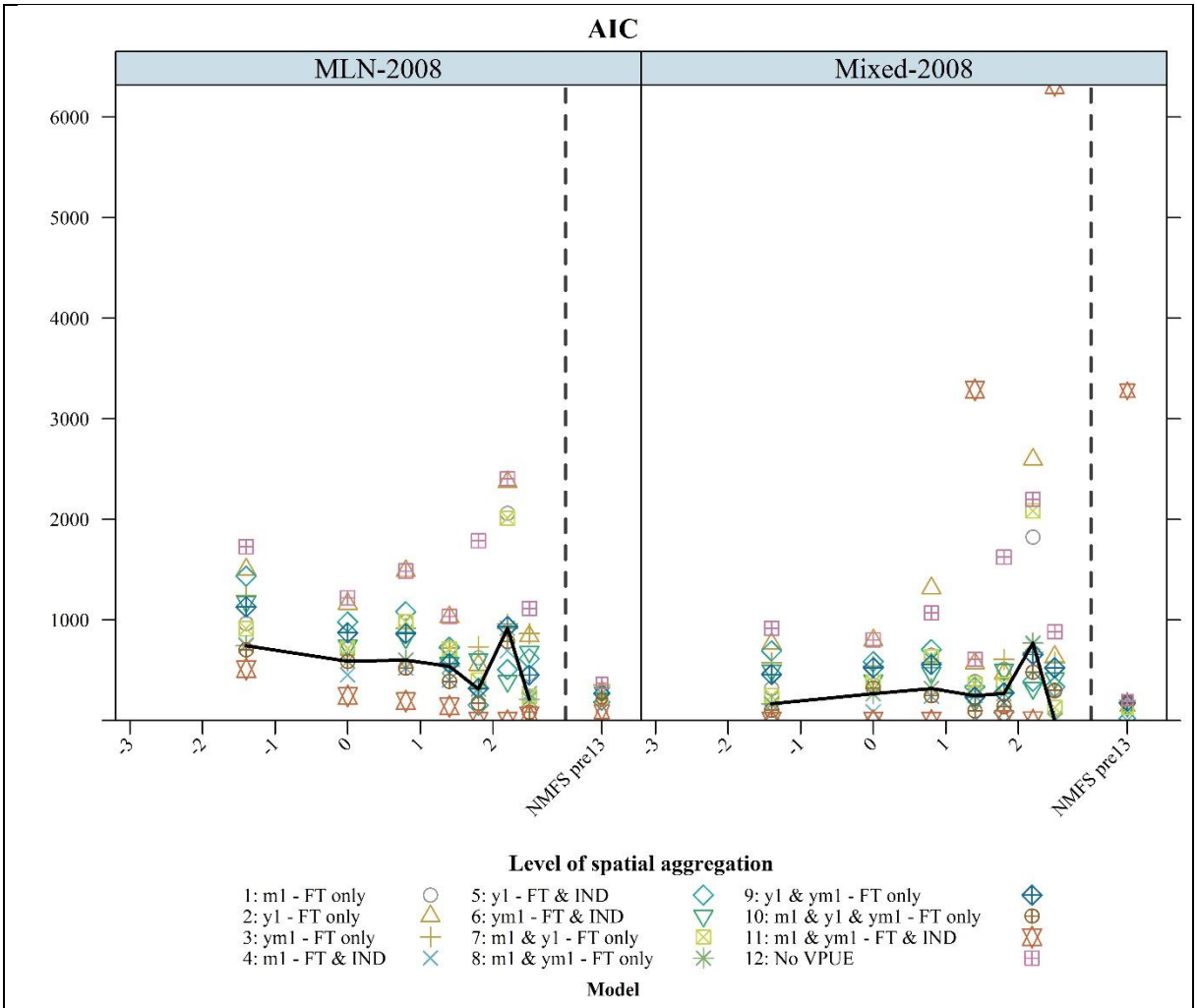


Figure E.1.b. AIC differences for the 12 model's specifications estimated with 2007-2008 data, by level of spatial aggregation and for either fixed (left panel) or random parameters (right panel). Differences are computed relatively to the lowest AIC obtained for each level of spatial aggregation and each model type. Model #8 with fixed parameters is the model presented in the main body of the paper.

Computational limitations did not allow to present the estimates for the most refined level of spatial aggregation for all model specifications. Convergence issues in the estimation algorithm for the most complex specification (model 11) suggest that estimates for this specification may not correspond to a global optimum.

2. Complementary estimation assumptions

A couple of assumptions were necessary to estimate the RUM. To begin with, we chose a daily time scale for choice occasions. This time scale is a compromise between the most refined time scale that would be based on fishing sets (but that would be much more data intensive and require a more refined analysis of the VMS pings) - and the coarser time scale of trips. Although this latter time scale may be very well suited for single-day trip fisheries (such as the urchin fishery studied extensively by Smith, e.g., Smith 2002, Smith, 2005), it is not appropriate here given the average duration of a fishing trip is approximately of one week. However, as the resolution of models becomes more spatially refined, the assumption of the uniqueness of choice becomes sometime violated (thereby emphasizing once more the trade-off between the spatial resolution of models and estimation issues). We followed a standard assumption in the literature (Girardin et al., 2015) of designating the "chosen" alternative as the one where most of the fishing effort was allocated[23]. Effort, catches and revenues were re-assigned accordingly. Depending on the tessellations, between 4% and 7% of effort was re-assigned according to that process. In total, we obtained 6,406 and 2,944 unique choice occasions (for 2008 and 2012 respectively), defined as a combination of a logbook trip with a day of the year.

We assume that the fishing effort remains constant over each fishing trip. Whereas the number of hooks per line is clearly fixed for a trip, the number of fishing sets[24] per day may vary from. However, we assumed it did not affect the decision of where to fish and, when allocating effort on a daily basis, that the total number of fishing sets was homogenously distributed across the different days of the same trip.

[23] When different sites had the same levels of effort, we randomly selected one of them.

[24] i.e., the number of times the longline is soaked and hauled back.

Finally, we assume that the decision to go fishing and the decision on effort level fishing were independent from the decision of the fishing location. We tested those assumptions using 2008 logbooks and analyzed the correlation between fishing sites (reported for a given trip as the statistical area that yielded the highest revenue) and landing prices (a major driver for the decision to start a fishing trip or not) as well as the correlation between fishing sites and effort levels. In both cases, we found only a weak correlation.

Complementary results

1. Monte Carlo experiments

Absence of correlations between spatial indexes and model's performance

Figure F.1.a, F.1.b and F.1.c show the results of the correlation analyses that we carried out between the spatial indexes of data heterogeneity that we considered – the Shannon entropy index and the Shannon equity index (see Section B of the Appendix) – and each model's performance in terms of goodness-of-fit (Figure F.1.a), prediction capability (Figure F.1.b) and capacity to recover the model's true parameters (Figure F.1.c).

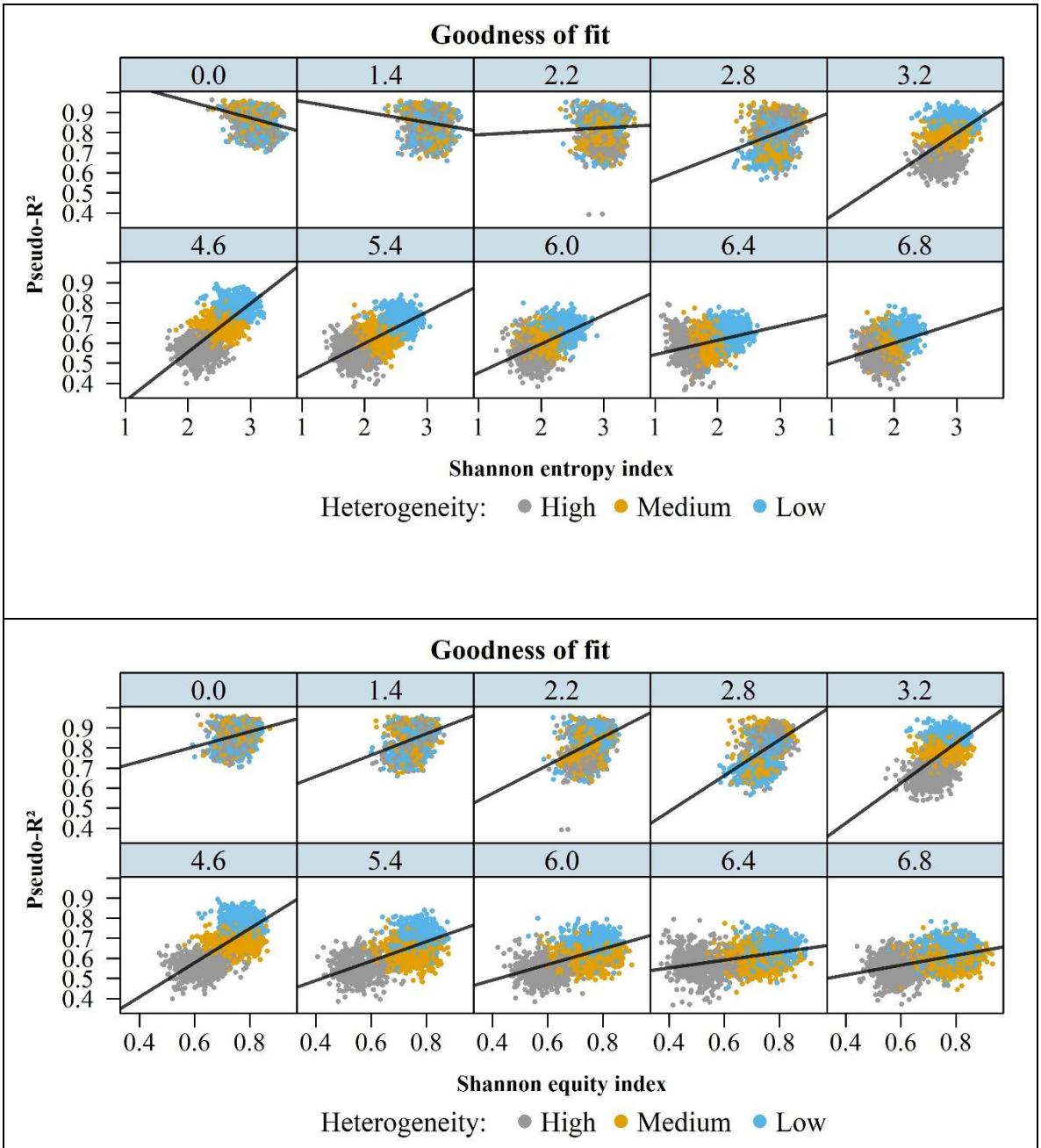


Figure F.1.a. Goodness-of-fit of 600 Monte Carlo experiments in function of Shannon entropy (upper panels) and equity (lower panels) index. Panels reflect for the ten levels of spatial aggregation tested. The solid black lines show the results of a simple linear regression.

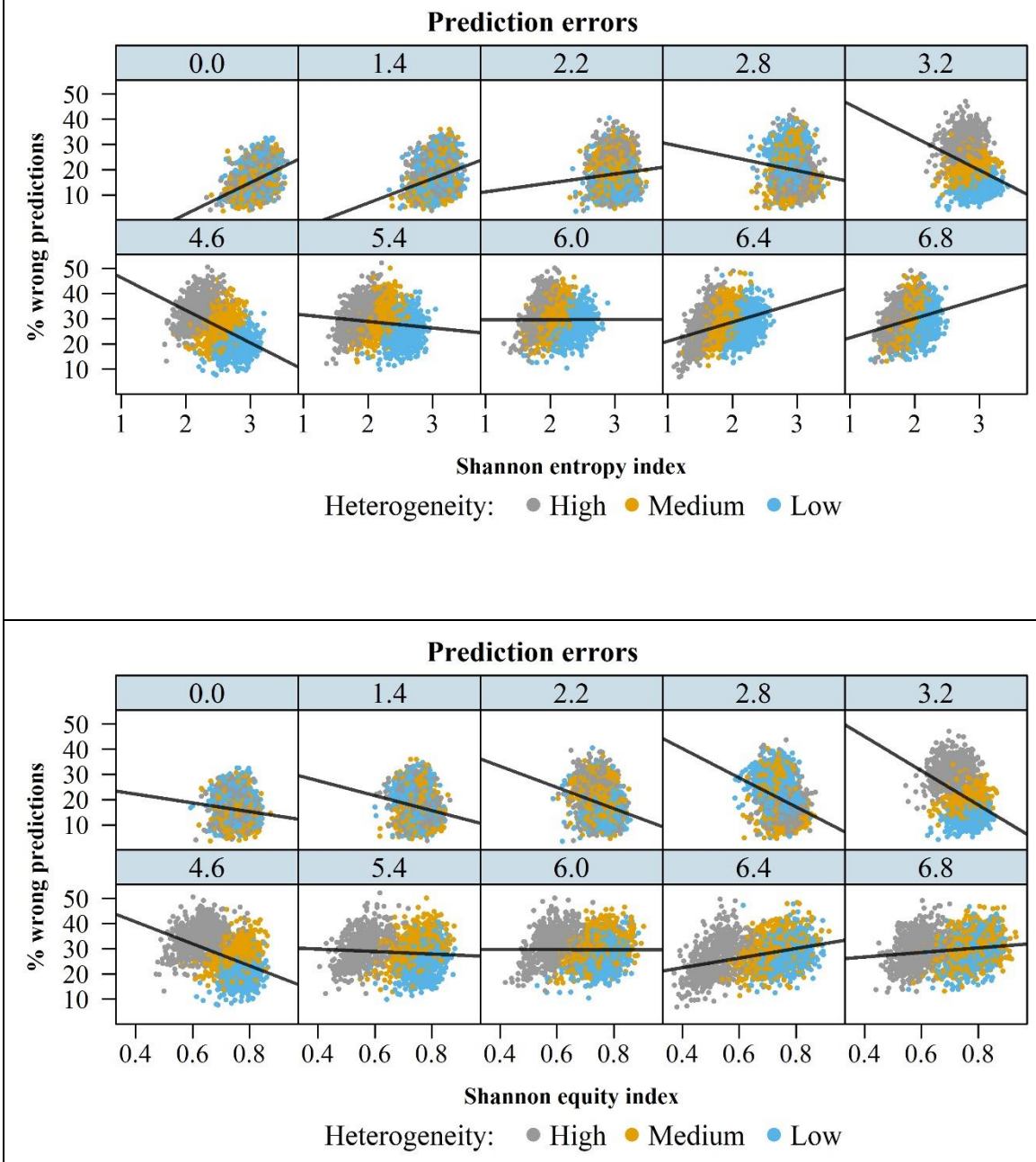


Figure F.1.b. Prediction errors of 600 Monte Carlo experiments in function of Shannon entropy (upper panels) and equity (lower panels) index. Panels reflect the ten levels of spatial aggregation tested. The solid black lines show the results of a simple linear regression.

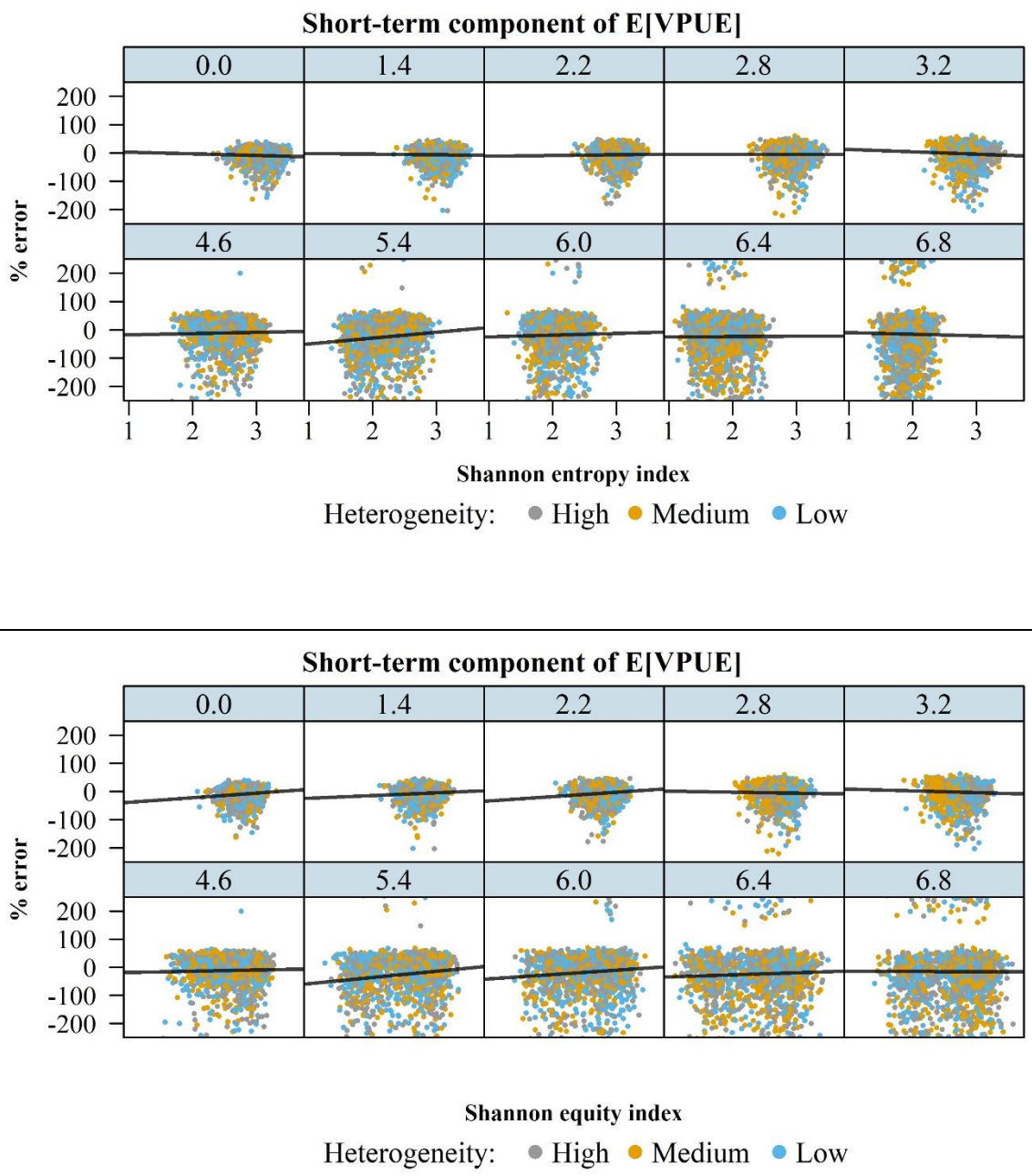


Figure F.1.c. % error in the estimation of the short-term component of the expected VPUE for each of 600 Monte Carlo experiments with respect to Shannon entropy (upper panels) and equity (lower panels) index. Panels reflect the ten levels of spatial aggregation tested. The solid black lines show the results of a simple linear regression.

Average Marginal Effects of explanatory variables

Here we show the average marginal effects of explanatory variables in estimating Eq (1).

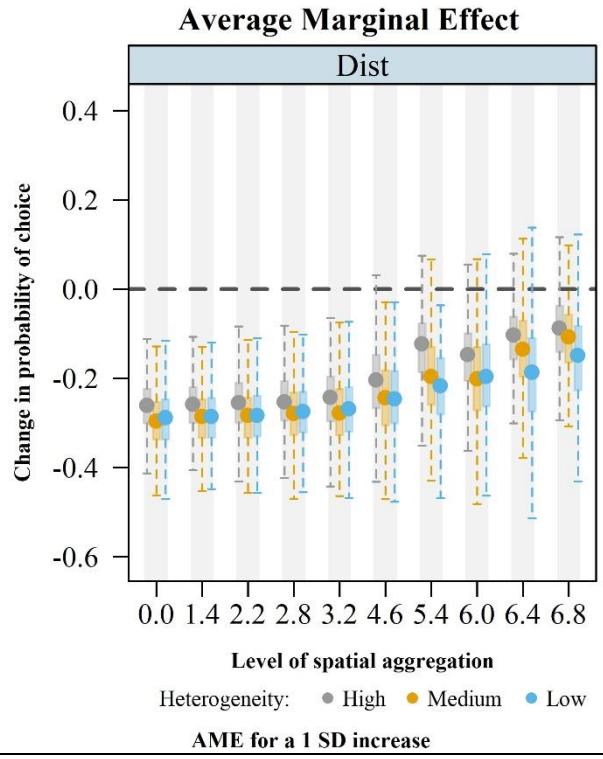


Figure F.1.d. Average marginal effects (AME) on choice probabilities of the distance variable estimated in the Monte Carlo experiments. Effects are computed for an increase of one standard deviation. The box edges are the 25th and 75th percentiles of the distribution of the AME and whiskers are located at +/- 1.5 times the interquartile ranges.

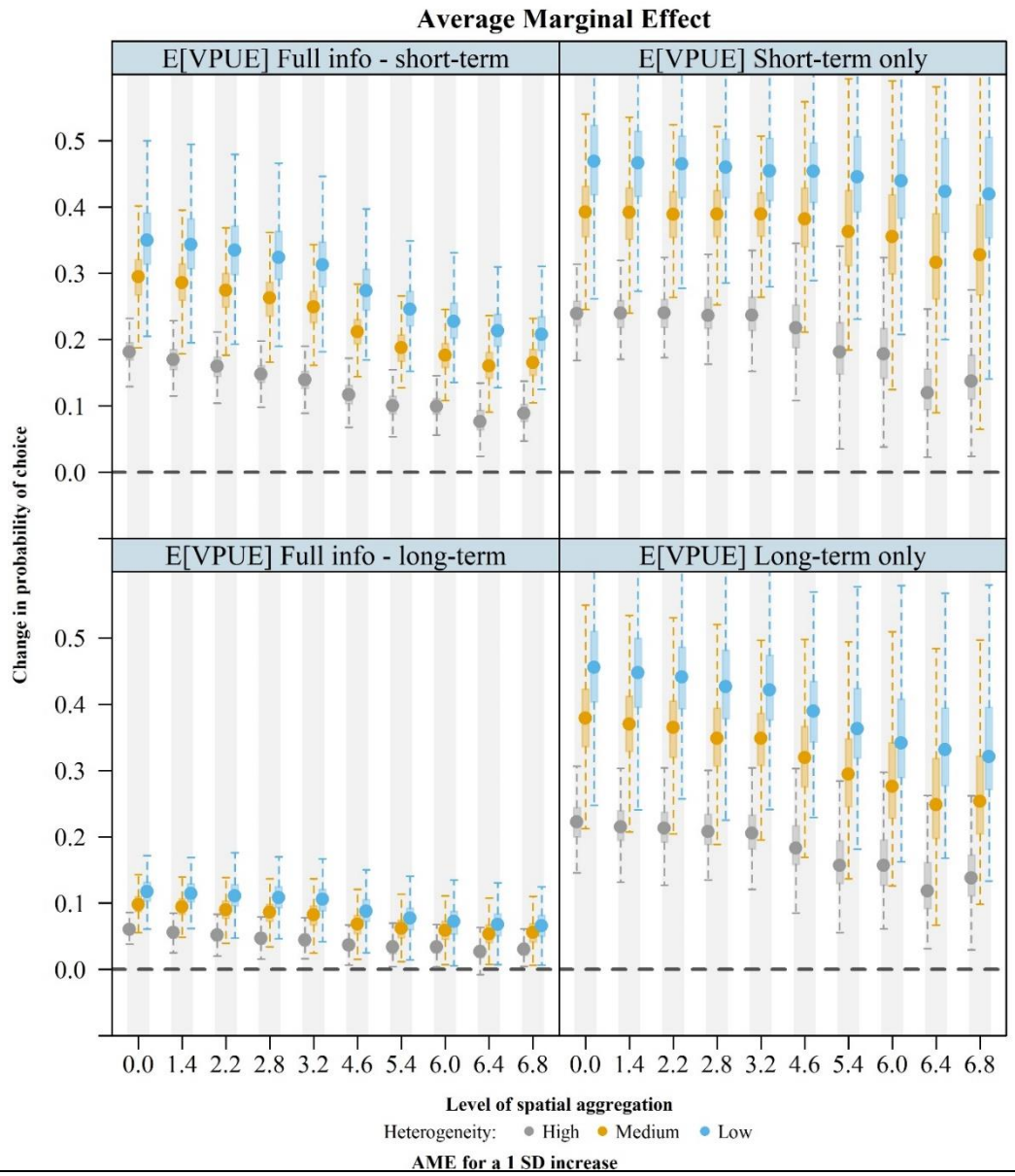


Figure F.1.e. Average marginal effects (AME) on choice probabilities of the four components of the expected VPUE estimated in the Monte Carlo experiments. The top and bottom left panels show, respectively, the short-term and long-term components of the expected VPUE when both information types are available, the top right panel shows the case when only short-term information is available, and the bottom right panel shows the case when only long-term information is available. Effects are computed for an increase of one standard deviation of the corresponding variable. The box edges are the 25th and 75th percentiles of the distribution of the AME and whiskers are located at +/- 1.5 times the interquartile ranges.

Root-mean squared-errors of parameter estimates

Table F.1 Root-mean squared errors (RMSE) of parameter estimates for each level of spatial heterogeneity and level of spatial aggregation. RMSE are computed taking the mean over all the 600 simulations of the squared difference between a parameter estimate and its true value.

Het.	Parameters	Spatial aggregation									
		0	**1.4**	**2.2**	**2.8**	**3.2**	**4.6**	**5.4**	**6**	**6.4**	**6.8**
High	$\beta_{VPUE}^{Full\ info\ -\ short-term}$	1821	957	12	3	1	165	10610	2676	17509	112284
	$\beta_{VPUE}^{Full\ info\ -\ long-term}$	110	323	9	9	9	26	1743	150	844	18267
	$\beta_{VPUE}^{Short-term\ only}$	570	413	24	23	21	965	2904306	8561	81660	150951
	$\beta_{VPUE}^{Long-term\ only}$	3803	2068	28	7	3	810	35144	6175	25791	432064
	$\beta_{VPUE}^{No\ info}$	6199	4316	33	630	5	27	1137	1005	88075	142373
Med.	$\beta_{VPUE}^{Full\ info\ -\ short-term}$	0	0	1	1	2	23	307	171	14399	20259
	$\beta_{VPUE}^{Full\ info\ -\ long-term}$	9	9	9	9	9	11	46	28	2555	2475
	$\beta_{VPUE}^{Short-term\ only}$	11	12	13	15	20	100	1443	773	66996	83385
	$\beta_{VPUE}^{Long-term\ only}$	1	1	1	2	5	62	806	366	45013	23347
	$\beta_{VPUE}^{No\ info}$	8	7	6	7	10	109	4492	12636	155390	88426
Low	$\beta_{VPUE}^{Full\ info\ -\ short-term}$	0	0	1	2	3	23	586	23325	1360	2956
	$\beta_{VPUE}^{Full\ info\ -\ long-term}$	9	9	9	9	9	11	52	4574	127	401
	$\beta_{VPUE}^{Short-term\ only}$	11	11	14	16	22	106	1758	192534	5271	13058
	$\beta_{VPUE}^{Long-term\ only}$	1	1	2	3	4	58	937	141904	2697	9316
	$\beta_{VPUE}^{No\ info}$	6	6	5	5	8	612	1468	60123	9892285	174352

Het.	Parameters	Spatial aggregation									
		-4.6	**-3.2**	**-2.4**	**-1.8**	**-1.4**	**0**	**0.8**	**1.4**	**1.8**	**2.2**
High	$\beta_{VPUE}^{Full\ info\ -\ short-term}$	81400	18	1970	2672	7350	1	1063	140	286	3279
	$\beta_{VPUE}^{Full\ info\ -\ long-term}$	1756	28	1017	2580	3044	9	317	23	67	419
	$\beta_{VPUE}^{Short-term\ only}$	3	13	2695	4650	14628	12	607	321	1121	7040
	$\beta_{VPUE}^{Long-term\ only}$	23	15	8689	2820	5045	2	913	523	403	5553
	$\beta_{VPUE}^{No\ info}$	14252	115	64243	18654	62282	6	845	27	384	2709
Med.	$\beta_{VPUE}^{Full\ info\ -\ short-term}$	1114	31	5612	226	328	0	1275	20	1063	2916
	$\beta_{VPUE}^{Full\ info\ -\ long-term}$	633	43	9706	110	63	9	141	10	218	451
	$\beta_{VPUE}^{Short-term\ only}$	1	0	11798	98	273	11	1785	85	1130	13661
	$\beta_{VPUE}^{Long-term\ only}$	15	13	8776	231	439	2	3537	56	1603	89888
	$\beta_{VPUE}^{No\ info}$	3418	1151	35286	2932	3861	4	10904	51	17988	2710
Low	$\beta_{VPUE}^{Full\ info\ -\ short-term}$	2086	37	54	1948	3054	1	22876	1526	3098	371
	$\beta_{VPUE}^{Full\ info\ -\ long-term}$	265	26	77	186	347	9	1911	298	416	61
	$\beta_{VPUE}^{Short-term\ only}$	1	7	217	969	2719	12	15931	4032	8780	1840
	$\beta_{VPUE}^{Long-term\ only}$	15	31	28	666	4428	2	26396	2975	4335	1426
	$\beta_{VPUE}^{No\ info}$	6	6	5	5	8	612	1468	60123	9892285	174352

2. Applied case study

Here we show the average marginal effects of explanatory variables in estimating Eq (2).

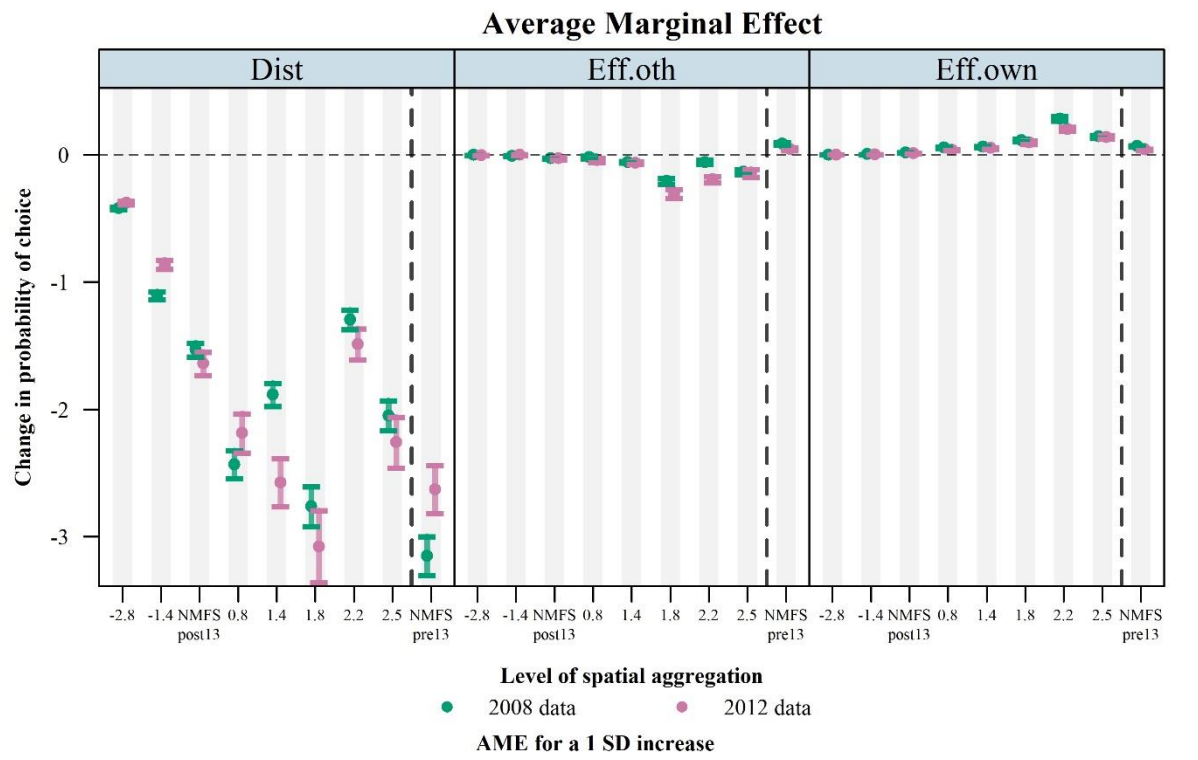


Figure F.2.a. Average marginal effects on choice probabilities of the distance to fishing grounds, of the fishing effort of other vessels, and of vessel's own fishing effort the day before estimated with GoMRF data. Segments indicate the 95% confidence interval of the corresponding point estimate. Effects are computed for an increase of one standard deviation of the corresponding variable.

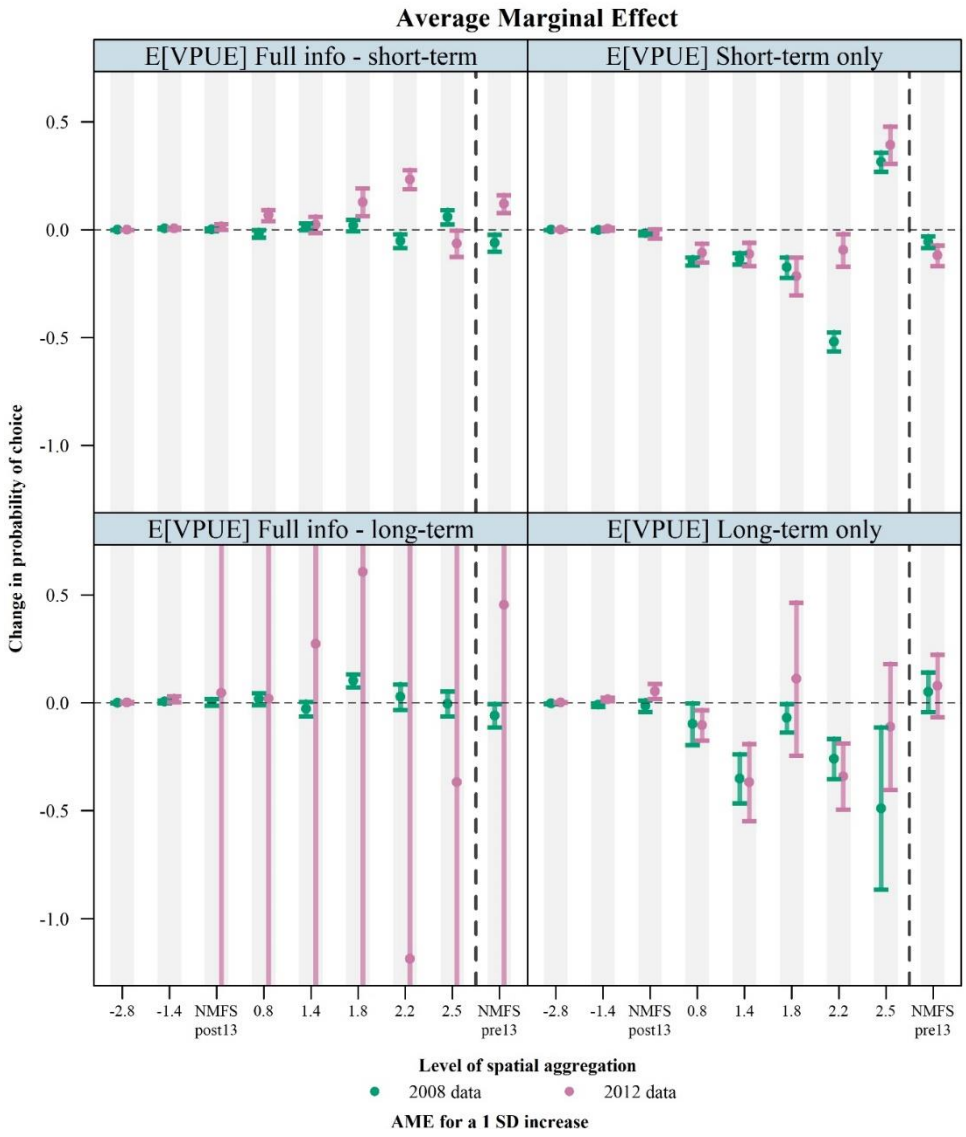


Figure F.2.b. Average marginal effects on choice probabilities of the four components of the expected VPUE estimated with GoMRF data. The top and bottom left panels show, respectively, the short-term and long-term components of the expected VPUE when both information are available, the top right panel shows the case when only short-term information is available, and the bottom right panel shows the case when only long-term information is available. Segments indicate the 95% confidence interval of the corresponding point estimate. Effects are computed for an increase of one standard deviation of the corresponding variable.