Supplementary Material S5: Simulation validation and diagnostics

To help create and validate the realism of our simulated DGN fishery we compared between observed and simulated data: 1) the seasonality of swordfish catches (Fig. S5.1), b) the frequency distribution of swordfish catch rates (number per set; Fig. S5.2), c) the offshore distribution of fishing effort (Fig. S5.3), d) step distances between consecutive fishing sets (Fig. S5.4), and the distribution of fishing effort (Fig. S5.5). Observed data was taken from the 1990-2000 period, which was pre-PLCA, so we compared this to our 'no turtle closure' scenario which was designed to represent the 1990-2000 period. We focus here on swordfish catches, because swordfish was the driver of profit-maximizing fisher location decisions. The agreement between observed and simulated seasonality of catch (Fig. S5.1) validates our allocation of fishing effort and simulated catch rates. The agreement between catch frequencies (Fig. S5.2) helps validate our approach for sampling integer catches given predicted mean catch rates; this comparison was also used to tune the 'accuracy' term in our ABM (m, Table S4.7). Comparisons of offshore distances, step distances, and general effort distribution were used to tune the ABM parameters D_{set}, h, N_{max} (Table S4.7), helped validate the initial step times (Table S4.6), and were crucial in applying additional constraints on offshore movement (Fig. S4.1).
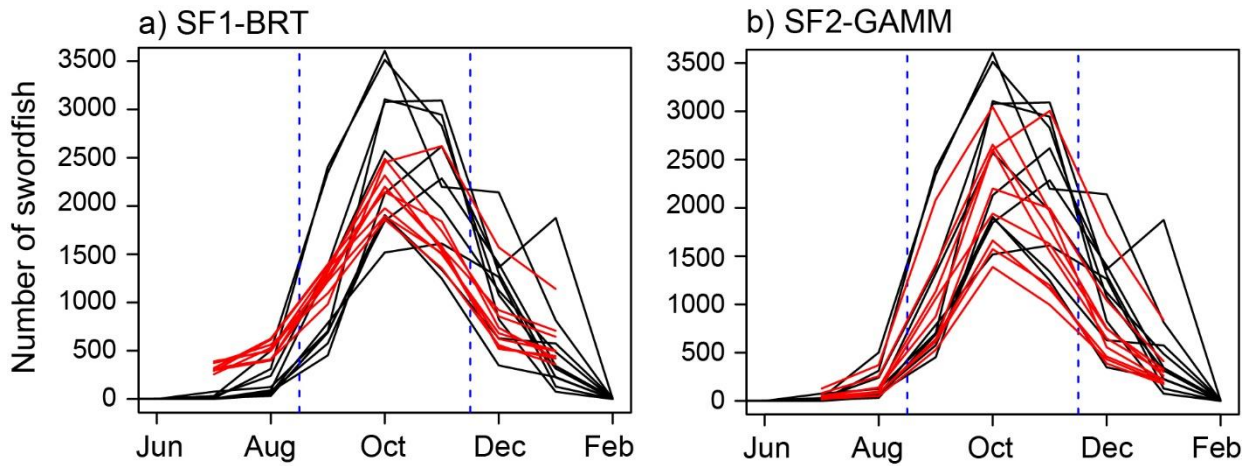
Fig. S5.1. Comparison of seasonal swordfish landings for simulated (red lines) and observed data (black lines), for the SF1 (a) and SF2 (b) swordfish catch models. Each line is one of the 1991-92 to 1999-00 fishing seasons. Simulated catches were taken from the 'no turtle closure' scenario, and observed catches were taken from the DGN logbooks. Simulated landings were expected to be more similar among seasons because fishing effort was constant (3000 sets per season), whereas observed fishing effort varied considerably (~2000-4500 per season). For reference, the start and end dates of the actual PLCA are indicated (dashed vertical lines).
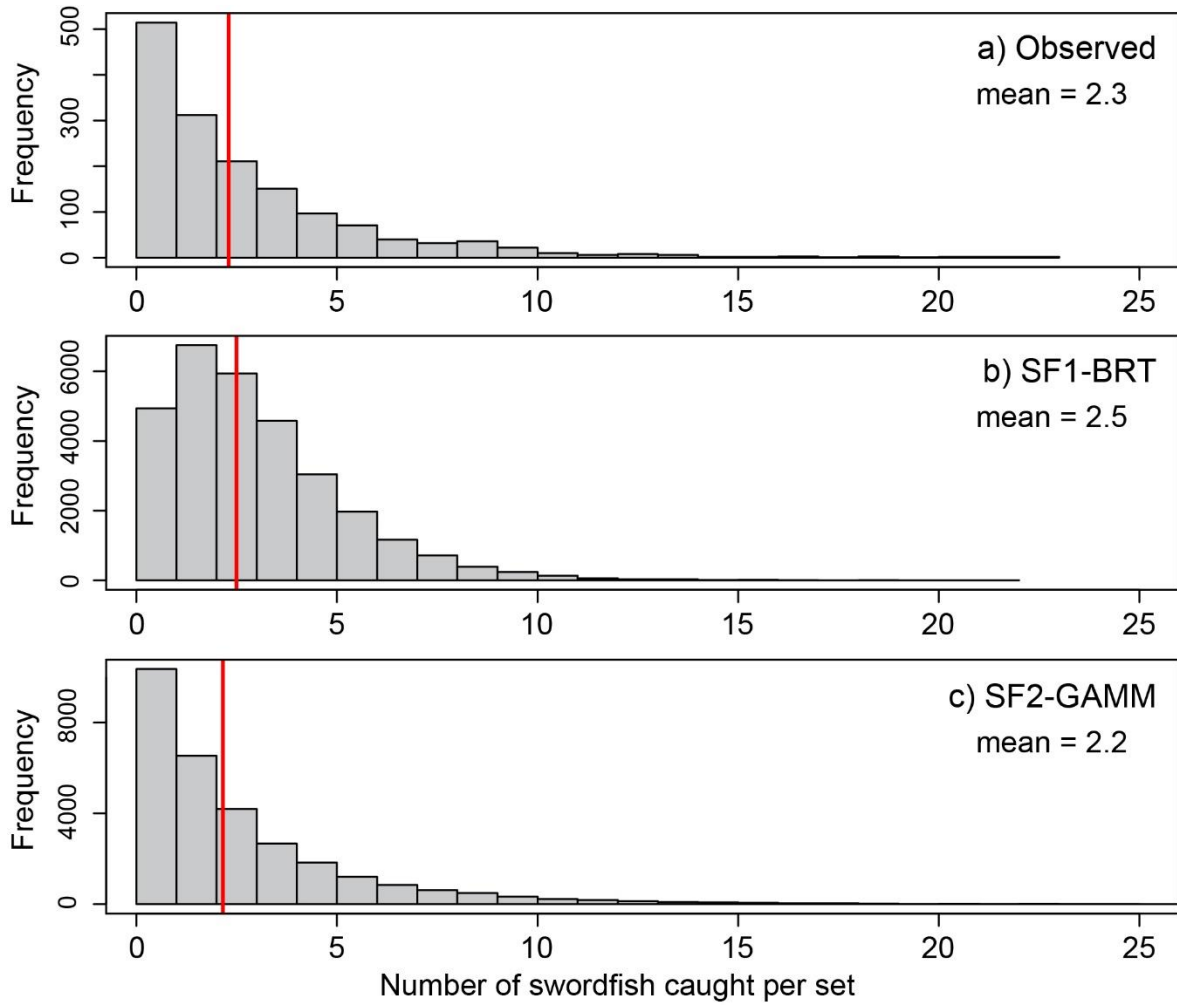
Fig. S5.2. Histograms of observed (a) and simulated (b-c) swordfish catch rates per set. Observed data was from 1990-2000, and the simulated was from the 'no turtle closure' scenario for both the swordfish catch models. The red line indicates the mean. Generally, the mean and frequency distribution of simulated swordfish catches per set were similar to observed, although the BRT (fitted using a Poisson family) underestimated the number of zeros.
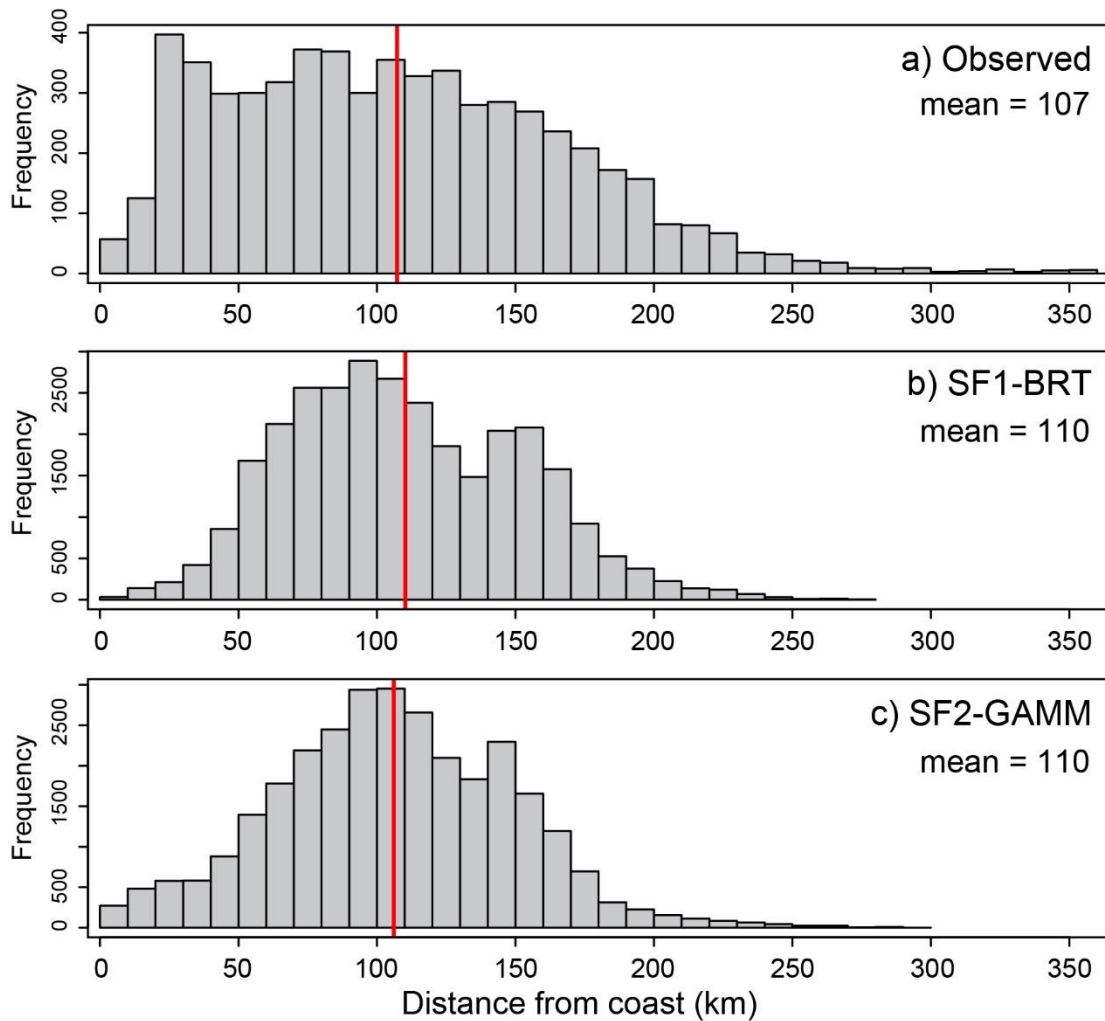
Fig. S5.3. Histograms of the distance offshore of DGN sets for the observed (a) and simulated (b-c) data. Observed data was from 1990-2000, and the simulated was from the 'no turtle closure' scenario for the two swordfish catch models. The red line indicates the mean. The mean distance offshore was similar between observed and simulated, but simulated vessels fished less often closer to shore than observed vessels. This was largely driven by observed vessels fishing close to shore in Southern California, which was an area predicted by our model to have relatively poor catch rates, and was thus not often fished by the dynamic agents.
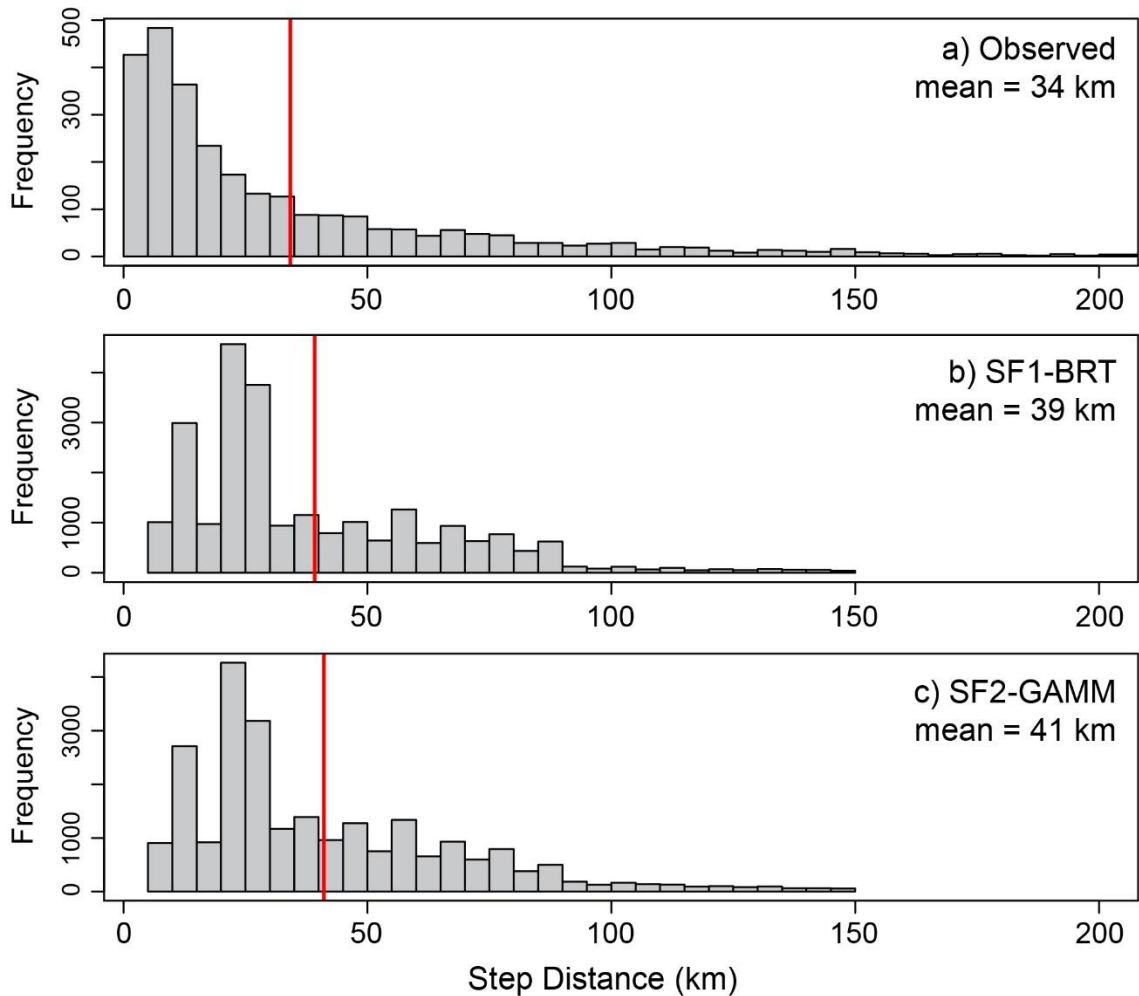
Fig. S5.4. Histograms of the distance vessels moved between consecutive DGN sets (i.e. steps) for the observed (a) and simulated (b-c) data. Observed data was from 1990-2000, and the simulated was from the 'no turtle closure' scenario for both swordfish catch models. The red line indicates the mean. Simulated vessels on average made slightly longer steps than observed vessels, and made no very large steps, which were impossible given our restrictions on set and travel durations. The 'stepped' shape of the simulated histograms is due to the discrete space in our model (i.e. step distances must be the fixed distances between cells) compared to the continuous space in the observer data. We also specified that vessels could not fish the same cell consecutively, which may made steps in the first histogram bin impossible.
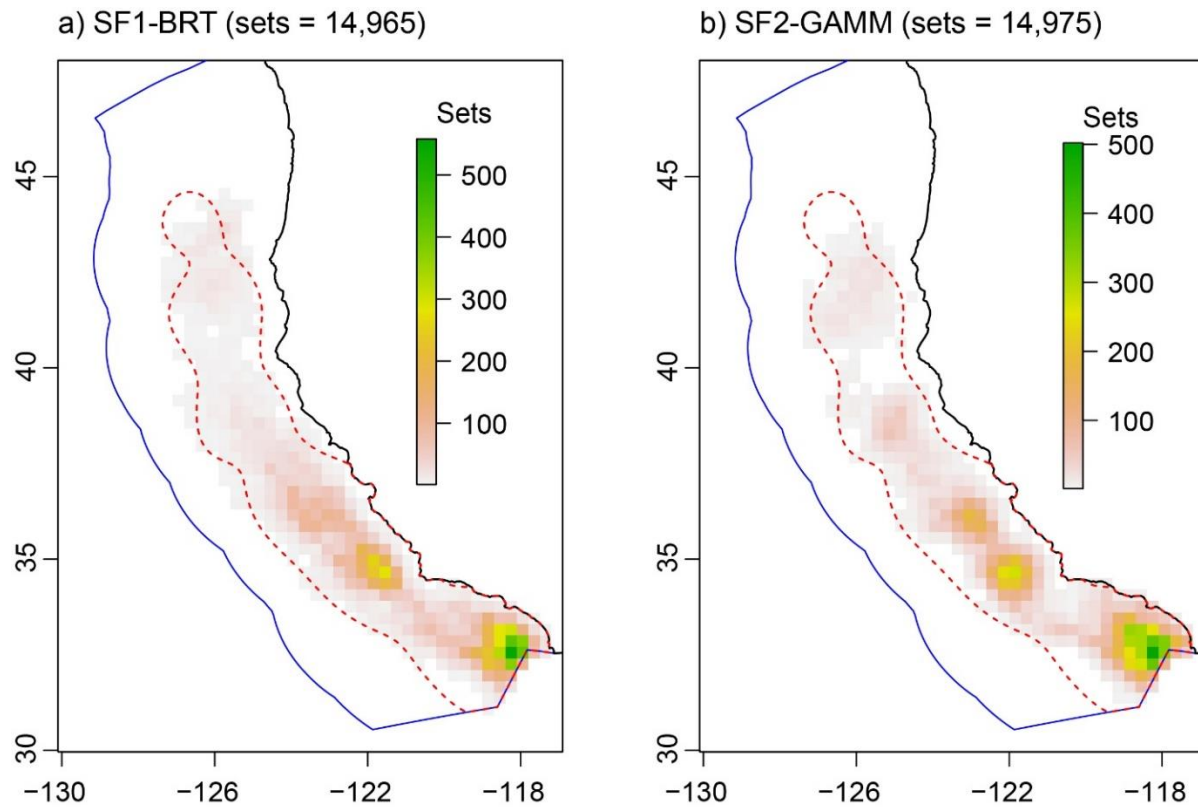
Fig. S5.5. The distribution of simulated fishing effort for the two swordfish catch models. This represents the total number of sets per 0.3°-degree cells during Phase 3 of our simulation (i.e. 5 fishing seasons and ~15,000 sets). Simulated data were from the 'no turtle closure' scenario, and for one iteration of the LB1 OM. There was little difference among iterations. The same resolution data for the observer data may contain confidential information, so we show only the 95% effort contour (from a KDE) for the 1990-2000 observer data (red line). The hot-spot of fishing effort in Southern California near the southern EEZ agrees with observed data (see Fig. 1 in Scales *et al.* 2017).

Allocating sources of variation

We were able to interrogate our simulation runs to approximate broad sources of variation in simulation results. The key to attributing variation is controlling for some sources while varying others. More simulation runs were not feasible in our study, but our simulation structure was such that some broad sources of variation were controlled. We chose two key model outputs to examine: total swordfish catch per season, and total leatherback turtle bycatch per season. For each closure strategy, we calculated pairwise differences in simulated fishing seasons accordingly:

$$Diff_{all} = |C_{all1} - C_{all2}|$$
$$Diff_{its} = |C_{its1} - C_{its2}|$$
$$Diff_{seas} = |C_{seas1} - C_{seas2}|$$

Where C is total swordfish catch (or turtle bycatch), and $Diff$ if the absolute difference in catch. C_{all1} and C_{all2} represent recorded catch in two randomly selected fishing seasons, that vary in both fishing season and simulation iteration (e.g. season 1996 iteration 1, and season 1998 iteration 3). Because these '*all*' draws vary across both season and iteration they include all sources of variation, including ocean conditions and stochasticity associated with closure creation, and thus represent, on average, the maximum pair-wise difference in catch that we simulated. C_{its1} and C_{its2} represent two recorded catches from the same randomly selected fishing season, but from different iterations (e.g. season 1995, iterations 2 and 4). Because '*its*' draws keep the same fishing season, differences in ocean conditions are excluded from these pairwise differences, and thus the difference between $Diff_{all}$ and $Diff_{its}$ approximates the variation due to this 'environmental' variation (we simulate constant fishing effort among years, so this is excluded as a source of variation). C_{seas1} and C_{seas2} represent recorded catch in two randomly selected seasons, but from the same iteration (e.g. 1993 iteration 2, and 1997 iteration 2). Because '*seas*' draws keep the same iteration, they share the same simulated closure and observer program, thus the difference between $Diff_{all}$ and $Diff_{seas}$ approximates the variation due to which catch and bycatch events were observed and how this affected closure creation.

We bootstrapped 1000 pairwise comparisons and calculated the mean for each *Diff*, and then calculated the differences of these means to apportion sources of variation:

$$Var_{enviro} = \frac{Diff_{all} - Diff_{its}}{Diff_{all}}$$

$$Var_{closure} = \frac{Diff_{all} - Diff_{seas}}{Diff_{seas}}$$

$$Var_{other} = 1 - (Var_{enviro} + Var_{closure})$$

Where Var_{enviro} is the proportion of observed variation in catch that is due to variation in the environment, and likewise for the *closure* creation process, and the remaining *other* (mostly stochastic) variation. These values were calculated for each combination of turtle and swordfish model, and for each closure scenario. For simplicity, and because the results were similar, for the comparison of static and dynamic closures we took the mean of the Static-obs$_s$ and Static-pred closures (static), and the mean of the Dyn-multi$_s$ and Dyn-turt$_s$ closures (dynamic). We note that these proportions are approximations only, and a more thorough sensitivity analysis could be constructed to better identify sources of variation.

References

Scales, K.L., Hazen, E.L., Maxwell, S.M., Dewar, H., Kohin, S., Jacox, M.G., Edwards, C.A., Briscoe, D.K., Crowder, L.B., Lewison, R.L., and Bograd, S.J. (2017) Fit to predict? Ecoinformatics for predicting the catchability of a pelagic fish in near real-time. *Ecological Applications* **27**(8), 2313-2329. doi:10.1002/eap.1610