*Supplementary Material*

**Remote underwater video for monitoring reef fish spawning aggregations**

Erwan Saulnier1\*, Annette Breckwoldt2, Marianne Robert1, Dominique Pelletier1,3

1DECOD (Ecosystem Dynamics and Sustainability), IFREMER, INRAE, L’Institut Agro, Lorient, France

2Leibniz Centre for Tropical Marine Research (ZMT) GmbH, Bremen, Germany

3LEAD, IFREMER, Nouméa, New Caledonia

\*Corresponding author: [esaulnier.ecology@gmail.com](mailto:esaulnier.ecology@gmail.com)

*ICES Journal of Marine Science*

Supplementary Text

Text S1. Description of the sampling stations

Station M1 was located at the bottom of the reef channel at a depth of 20 m, approximately 5 m from the southeastern slope. This habitat was rather flat and characterized by low live coral cover (20%). Stony coral colonies were sparse and small, although a few large *Acropora* spp. tables were present. Gorgonians and soft corals were the most abundant sessile organisms. Both substrate composition and sessile fauna resulted in relatively high habitat complexity at this station. Station M2, located on the north side of the channel at a depth of 14 m, was characterized by a hard substrate composed mainly of slabs and dead coral, typical of seascapes eroded by currents and sand transport. Sessile fauna were scarce, especially hard corals (Scleractinians) and to a lesser extent soft corals (Alcyonacea). In contrast, station M3 (located at a depth of 17m) was typical of outer slopes with exclusively hard coral substrate and a large diversity of hard corals including *Acropora* spp., *Montipora* spp., *Turbinaria* spp. and *Pocillopora* spp. The live coral cover was high (40%-60%), but the colonies were small to medium in size, and both gorgonians and soft corals were scarce.

Text S2. Video analysis and sampling effort

The videos included in this study correspond to the first batch of videos that were analysed; this batch contained approximately twice as many videos from station M2 as from stations M1 and M3 (Fig. S1). The second batch was not processed due to the limited resources that were ultimately available. A sensitivity analysis was performed to assess the robustness of our results to this unbalanced sampling effort (see the Material and Methods section and Fig. S7). In total, 64% of the videos recorded at stations M1, M2 and M3 were processed, of which 84% were validated and subsequently analysed (n=521); the remaining videos were not validated due to technical problems (e.g., incorrect camera settings or low battery).

Supplementary figures



Fig. S1. Sampling design. Each coloured cell (yellow to dark blue) represents a 10-min video analysed at stations M1, M2, or M3, while a light grey cell indicates no data available for this study. The MICADO systems were generally deployed late in the morning or around midday on the first sampling day of each month. Colors denote the five levels of the covariate ‘time of day’ included in generalized linear mixed effect models (GLMMs) (early morning: 06:00-07:00, morning: 08:00-10:30, midday: 12:00-13:30, afternoon: 15:00-16:00, late afternoon: 16:30-17:30).



Fig. S2. Relative abundance (number of fish per rotation) throughout the year at station M3. Very few individuals were observed at this station, with at most, one fish per rotation. No data in August, December and February.



Fig. S3. Model validation plots created with the DHARMa R package for the best Poisson GLMM. (a) QQ-plot to detect overall deviations from the expected distribution, (b) residuals vs predicted values, (c) dispersion test, and (d) zero-inflation test.



Fig. S4. Residuals versus each covariate included (month, station, time of day, tidal height) and not included (Days After the Full Moon, DAFM) in the best Poisson GLMM. Model validation plots created with the DHARMa R package.



Fig. S5. Model validation plots created with the DHARMa R package for the best binomial GLMM. (a) QQ-plot to detect overall deviations from the expected distribution, (b) residuals vs predicted values and (c) dispersion test.



Fig. S6. Residuals versus each covariate included (month, station, time of day) and not included (tidal height and Days After the Full Moon, DAFM) in the best binomial GLMM. Model validation plots created with the DHARMa R package.



Fig. S7. Standardized coefficient estimates (dots) and 95% confidence intervals (solid lines) for the best Poisson GLMM fitted either to the entire dataset or to a subset of data including only those collected at the dates and times at which they were available at station M3.

Supplementary tables

Table S1. The top 10 Poisson GLMMs used to estimate spatiotemporal variations in blacksaddled coralgrouper relative abundance. The covariates, degrees of freedom (df), log-likelihood (logLik), and Akaike’s information criterion corrected for small sample size (AICc) of each model are listed. As the top three GLMMs had similar support from the data (ΔAICc ≤ 2), the most parsimonious one (denoted in bold) was selected as the best model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Covariates | df | logLik | AICc | ΔAICc |
| month + station + time of day + tidal height + month:station + (1 | video\_ID) | 25 | -894,7 | 1840,5 | 0,0 |
| **month + station + time of day + tidal height + (1 | video\_ID)** | **18** | **-902,4** | **1841,3** | **0,8** |
| month + station + time of day + tidal height + month:station + tidal height:station + (1 | video\_ID) | 26 | -894,7 | 1842,6 | 2,0 |
| month + station + time of day + tidal height + tidal height:station + (1 | video\_ID) | 19 | -902,4 | 1843,4 | 2,8 |
| month + station + time of day + month:station + (1 | video\_ID) | 24 | -897,7 | 1844,4 | 3,9 |
| month + station + time of day + (1 | video\_ID) | 17 | -905,3 | 1845,0 | 4,5 |
| month + station + time of day + tidal height + month:station + time of day:station + (1 | video\_ID) | 29 | -893,1 | 1845,7 | 5,2 |
| month + station + time of day + tidal height + time of day:station + (1 | video\_ID) | 22 | -900,7 | 1846,3 | 5,8 |
| month + station + tidal height + month:station + (1 | video\_ID) | 21 | -902,3 | 1847,5 | 6,9 |
| month + station + tidal height + (1 | video\_ID) | 14 | -909,6 | 1847,6 | 7,1 |

Table S2. The top 10 Binomial GLMMs used to estimate spatiotemporal variations in the probability of occurrence of males displaying a courtship coloration. The covariates, the degrees of freedom (df), the log-likelihood (logLik), and the Akaike’s information criterion corrected for small sample size (AICc) of each model are listed. The best model (lowest AICc) is denoted in bold.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Covariates | df | logLik | AICc | ΔAICc |
| **month + station + time of day + (1 | video\_ID)** | **17** | **-360,7** | **755,9** | **0,0** |
| month + station + time of day + tidal height + (1 | video\_ID) | 18 | -360,7 | 758,0 | 2,1 |
| month + station + time of day + month:station + (1 | video\_ID) | 24 | -354,6 | 758,1 | 2,2 |
| month + station + time of day + time of day:station + (1 | video\_ID) | 21 | -358,0 | 758,9 | 2,9 |
| month + station + time of day + tidal height + tidal height:station + (1 | video\_ID) | 19 | -360,3 | 759,2 | 3,3 |
| month + station + tidal height + (1 | video\_ID) | 14 | -365,9 | 760,2 | 4,2 |
| month + station + time of day + tidal height + month:station + (1 | video\_ID) | 25 | -354,6 | 760,2 | 4,3 |
| month + station + time of day + tidal height + time of day:station + (1 | video\_ID) | 22 | -358,0 | 760,9 | 5,0 |
| month + station + (1 | video\_ID) | 13 | -367,4 | 761,2 | 5,3 |
| month + station + time of day + month:station + time of day:station + (1 | video\_ID) | 28 | -352,2 | 761,7 | 5,8 |