Marine Biodiversity
August 2020, Volume 50 Issue 4, Pages 52 (15p.)
https://doi.org/10.1007/s12526-020-01068-6
https://archimer.ifremer.fr/doc/00637/74890/

Spatial and interannual variability of presettlement tropical fish assemblages explained by remote sensing oceanic conditions

Jaonalison Henitsoa [1, *] , Durand Jean-Dominique [2], Mahafina Jamal [1], Demarcq Herve [3], Lagarde Raphaël [4], Ponton Dominique [5]

[1] Institut Halieutique et des Sciences Marines, Université de Toliara, Rue Dr. Rabesandratana, 601, Toliara, Madagascar
[2] MARBEC, IRD, Univ. Montpellier, CNRS, Ifremer, Montpellier, Franc
[3] MARBEC, (IRD, Univ Montpellier, CNRS, Ifremer), Centre de Sète, France
[4] Université de Perpignan Via Domitia-CNRS, Centre de Formation et de Recherche sur les Environnements Méditerranéens, UMR 5110, F 66860, Perpignan, France
[5] ENTROPIE, IRD-Université de La Réunion-CNRS, Laboratoire d'Excellence CORAIL, c/o Institut Halieutique et des Sciences Marines, Université de Toliara, Rue Dr. Rabesandratana,, 601, Toliara, Madagascar

* Corresponding author : Henitsoa Jaonalison, email address : jaonasat@gmail.com

Abstract :

Understanding the interannual effect of various environmental factors on biodiversity distribution is fundamental for developing biological monitoring tools. The interannual variability of environmental factors on presettlement fish assemblages (PFAs) has been so far under investigated, especially in Madagascar. Numerous explanatory variables including local hydro-dynamic conditions recorded during the sampling night, characteristics of the benthic substrate and remotely sensed oceanic conditions (RSOC) were used to explain the spatio-temporal variability of PFAs in southwestern Madagascar. Gradient forest analyses were used to hierarchically classify the effect of these explanatory variables on the PFAs for two sites and during two different recruitment seasons. RSOC variables appeared to better explain the PFAs than the local variable and the characteristics of the benthic substrate. The PFAs caught in water masses with coastal characteristics were better explained than those with open water characteristics. This spatial variability is hypothesised to be linked to differences in feeding conditions among water masses. The gradient forest analyses also highlighted the complexity of predicting PFAs as the species for which abundances were better explained by RSOC variables varied between years. This interannual variability was mainly explained by the interannual variation of chlorophyll a (Chl a) concentration, wind and surface current, with better prediction obtained during the year with high Chl a values associated with high averaged sea surface temperature. These findings suggest the importance of forecasting Chl a concentrations, taking into account the impact of tropical storms and climate variability in order to predict PFAs in the future.

Keywords : Fish post-larvae, DNA barcoding, Structure, Environmental conditions, Hierarchical classification, Gradient forest

Introduction

Several studies pointed the importance of studying presettlement fish assemblages (aka post-larvae) to better understand and predict how adult biomass varies (Lockwood et al. 1981; Nichols et al. 1987; Hsieh et al. 2005; Koslow and Davison 2016; Chen et al. 2018) or more generally to monitor marine fish communities (Koslow and Wright 2016). Indeed, post-larvae represent survivors of the larval phase, and despite the important mortality that can occur after, throughout the juvenile stage, they correspond to the individuals that can replenish adult fish populations (Takahashi and Watanabe 2004). Presettlement fish assemblages (PFAs) also represent a promising tool for monitoring coastal marine ecosystems as their structures relate to environmental factors and to different anthropogenic pressures (Smith et al. 2018). Developing PFAs surveys is thus considered as an urgent need in areas experiencing climate change and anthropogenic pressure (Koslow and Wright 2016).

Furthermore, in a context of climate change it seems vital to better understand how future environmental conditions will impact PFAs and thus fish communities. This highlights the necessity to develop predictive models able to anticipate the PFAs dynamic and to identify the hierarchical effect of explanatory variables on PFAs structure. In tropical waters, this field of research has remained largely under investigated. Carassou et al. (2008) conducted one of the rare studies aiming at hierarchically classifying the explanatory variables recorded at different temporal and spatial scales structuring the PFAs. These authors demonstrated that abundances of presettlement fish responded to large scale meteorological measured over a few days before sampling, or measured the day of sampling, and small-scale water column factors measured locally, according to threshold values that varied among families. However, they found no significant difference between past-day and daily measures. Their study neither addressed the effect of interannual variation of explanatory variables on PFAs nor incorporated the remotely sensed oceanic conditions. To our knowledge, the eventual importance of remotely sensed oceanic conditions (RSOC) extracted over few days preceding fish sampling in structuring PFAs has never been investigated in tropical waters.

In Madagascar, small-scale fishermen and coastal populations highly depend on marine resources for their subsistence (Cooke et al. 2000). Their increasing numbers (INSTAT and ICF 2016, Le Manach et al. 2012) and the declining quality of coastal habitats (IOTC 2009) may lead to modifications in the taxonomic composition of fish assemblages (Folke et al. 2004). In order to better record or even anticipate these modifications, PFAs need to be efficiently investigated. In SW Madagascar, PFAs surveys started about a decade ago (Mahafina 2011; Jaonalison et al. 2016) but employed different sampling gears and were based on different taxonomic resolutions. Recently, standardized protocols for sampling and identifying presettlement fish at the species level have been developed (Jaonalison et al. in review). Using these protocols, these authors demonstrated that regression tree models based on remotely sensed oceanic conditions (RSOC) were able to predict the richness and abundance of tropical PFAs. As a follow-up, the present study aimed to further expand the approach and test if the structure of PFAs described at the species level could also be explained by RSOC. More precisely, the specific objective of this study was to assess the interannual variability of the PFAs by characterizing: (i) the variability of the RSOC and PFAs among years, and (ii) how

the interannual changes in RSOC impact PFAs structure. Our hypothesis was that the PFAs interannual variability was induced by the RSOC interannual variability.

Materials and Methods

Study Sites, Sample Collection and Processing

Presettlement fish were sampled in two sites in the Great Barrier Reef of Toliara (GRT) and in the reef off Anakao (ANA) on the southwestern coast of Madagascar (**Fig.1**). A detailed description of each site is provided in Jaonalison et al. (in review) but rapidly these two sites mainly differ by the characteristics of their water masses. Indeed, the GRT site is mostly under the influence of oceanic water masses due to the diurnal tidal inflow that renews up to 80% of water in Toliara Lagoon (Chevalier et al. 2015). Oppositely, the ANA site, closer to the coast, is influenced by the Onilahy River flows when the local northerly wind blows (Jaonalison et al. in review). Larval collection was performed monthly from November 2016 to April 2017 (Campaign #1), and from November 2017 to April 2018 (Campaign #2). The monthly sampling was conducted during three consecutive nights of the new moon period, three samplings were collected during one night in each site (9 samples per month and per site). One light trap known as SLEEP (Collet et al. 2018) was set at each of the three approx. 200 to 300m apart stations of each site. As the catch of each light trap in each station correspond to a sample, a total 108 samples per campaign for the two sites were intended to be obtained. In the laboratory, all fishes were sorted alive by morphospecies. One individual per morphospecies was photographed with a "Nikon D90" camera and a piece of its caudal fin was preserved in ethanol 90° for DNA barcoding.

Identification Process

The whole process of larval identification from DNA extraction till sequencing is detailed in Collet et al. (2018). The sequences were uploaded to Barcode Of Life Datasystems (BOLD) for assigning a Barcode Index Number (BIN) for each specimen sequence. Due to the ambiguity in the assignment of specie name to each BIN, different labels were used for identifying the species. First, if the BIN corresponded to only one species in BOLD, and this species was observed in this BIN only, the specimen was identified as "Genus+species" (e.g. *Ostorhinchus angustatus*). Second, if the BIN corresponded to different species from the same genus, or if the species corresponding to this BIN was also assigned to different BINs in BOLD, the specimen was identified as "Genus+BIN" (e.g. *Ostorhinchus* [BOLD:ACV9601]). Third, if the BIN corresponded to species from different genera, but belonging to the same family, the specimen was identified as "Family+BIN". Note that identification such as "Genus+BIN", or "Family+BIN", do correspond to identifications at the species level as each BIN corresponds to an operational taxonomic unit, and thus to a putative species (Ratnasingham and Hebert 2013).

Environmental Data

Seven local variables were recorded: sea surface temperature (using a thermometer), water transparency (using a Secchi disk), wind speed and direction (using an anemometer and a compass). Due to technical problems, sea surface salinity was recorded for the second campaign

only and was thus not retained for the analyses. The difference between the time of sunset and light-trap-setting, and between the time of sunrise and light trap retrieving were also considered as explanatory variables for explaining the eventual effect of high tide which always occurred around sunset during the sampling periods. The abundance of small pelagic fish species in light-trap was also considered among local variables as their numbers may affect the way light traps attract and retain assemblages of presettlement fish.

A specific campaign was conducted once a year in order to analyse the characteristics of the benthic substrate at each sampling station. The approach was based on high precision photo-quadrats following Dumas et al. (2009). Photo-quadrats were taken using a waterproof Olympus TG 860 camera along three 20 m transects forming of star placed randomly at each sampling station. In the laboratory, percentage cover of sediment, coral, seagrass, and macroalgae were measured using Coral Point Count with Excel® extension (CPCe, Kohler and Gill 2006).

The remotely sensed oceanic conditions (RSOC) were extracted for an eight-day period preceding each sampling over a 50 by 120 km area (Lon.max=43.611403, Lon.min=43.022609, Lat.max=-22.900648, Lat.min=-24.115845, **Fig.1**). This spatial and temporal scale was chosen as it was identified as the best predictor of presettlement fish richness and abundance at the two studied sites (Jaonalison et al. in review). The RSOC included the Sea Surface Temperature (SST), the concentration of chlorophyll *a* (Chla), the cross-shelf wind velocity U (Wind_U) and the alongshore wind velocity V (Wind_V), and the cross-shelf current velocity U (Current_U) and the alongshore current velocity V (Current_V), see **Table 1** for details.

Data Analysis and Modeling

Multivariate analyses were performed for depicting the spatial and interannual variability of each of the three types of explanatory variables. The differences between sites in benthic substrate characteristics and presettlement fish assemblages (PFAs) were visualized with Mutlidimensional scaling (MDS) and tested with Permutational Multivariate Analysis of Variance (Anderson 2017). The most discriminative characteristic of the benthic substrate, or species within assemblages, was then identified with a Similarity of Percentage analysis (Clarke et al. 1994). The spatial and interannual variability of local variables and the remotely sensed oceanic conditions were analysed with Principal Component Analysis (PCA) for visualising the parameter discriminating both sites and campaigns. The analysis for RSOC was run independently from local variables as they concern merely the interannual variability. The differences in mean variation of remotely sensed oceanic conditions between campaigns were tested with analyses of variance (ANOVA).

A total of 19 explanatory variables were used for explaining the PFAs. Among the different methods developed for identifying response of plant or animal assemblages faced to environmental gradients, Gradient Forest (GF) appears to be the most robust and well suited among the exploratory machine learning techniques. Indeed, GF can deal with non-linear relationships (Ellis et al. 2012), and can analyse large databases with numerous species and explanatory variables (Roubeix et al. 2017). GF, which is an extension of Random Forest method, deal with the prediction of multi-species simultaneously (Ellis et al. 2012) by applying the approach called “assemble-predict-together” defined by Ferrier and Guisan, (2006). This

approach consists in analysing simultaneously the response of several species to the explanatory variables within a common modeling framework.

Given the strong expected differences in water masses characteristics between the two sites, the PFAs were analysed separately. For each site, three datasets were thus created: one for the two campaigns considered together, one for campaign#1 and one for campaign#2. The datasets for campaign#1 and campaign#2 allowed to compare among years the explained presettlement fish assemblages (PFAs), and the importance of explanatory variables. As GF analyses are sensitive to the presence of rare species, only species exceeding 3% of occurrence were retained. For each dataset, the GF analyses were performed following two steps. During a first step, three models were trained using either local variable, characteristics of the benthic substrate, or remotely sensed oceanic conditions separately. During a second step, GF analyses were performed using all the explanatory variables (Table. 1). This two steps approach allowed to compare the goodness of fit of each model for the explained PFAs, based on the average of R^2 value ($R^2_{aver.}$) of the species for which the abundances were better described as applied by Roubeix et al. (2017). $R^2_{aver.}$ is the key diagnostic for Gradient Forest: the higher value of $R^2_{aver.}$ is the more predictable variability of the PFAs.

GF analyses generate four types of graphs that are of central importance for understanding the results: the overall variable importance plots, the species performance plot, the split density plots and the specific cumulative importance curves. The overall variable importance plot corresponds to barplots presenting the explanatory variable in decreasing R^2 weighted importance. Higher R^2 weighted importance values indicate important explanatory variables for explaining the PFAs variability. The species performance plot which is also a barplot presents the goodness of fit performance measures (R^2) for the species for which the variations in abundance are better described by the GF model. The split density plots allow identifying the thresholds value for each explanatory variable structuring the explained PFAs. Finally, the specific cumulative importance curves are used for identifying the thresholds of the explanatory variables for each species. Concerning the split density plot, the thresholds of the important explanatory variables structuring the explained PFAs were obtained through the analysis of the ratio between density of splits and density of data. The effect of an explanatory variable on the explained PFAs is important when this ratio is greater than one. This indicates that the range of values associated to ratios greater than 1 are associated to obvious change in the structure of PFAs. Regarding the specific cumulative importance curve, the significant thresholds influencing each of the species for which the abundances were better described were identified based on the three criteria for threshold selection applied by Roubeix et al. (2017). According to these authors, the threshold of a given explanatory variables is ecologically significant if (1) the explanatory variable is important, based on R^2 weighted importance, for explaining the abundances; (2) the threshold of this explanatory variable (indicated by the potential peak) is clearly defined based on the specific cumulative importance curves; and (3) the threshold is clearly highlighted for several species, based on the specific cumulative importance curves.

Data analyses were performed with R programming software R.5.1 (R Core Team 2018), using "vegan" package (Version 2.5.4, Oksanen et al., 2019) for statistical analysis, the

"gradientForest" (Version 0.1-17, Ellis et al., 2012) and "extendedForest" (Version 1.6.1, Ellis et al., 2012) packages for modeling the PFAs.

Results

Spatio-Temporal Change of Local and RSOC Variables

PCA analysis indicated that the local variables varied between sites and campaigns. Campaign#1 was characterized by a higher water transparency at GRT, while cross-shelf wind and SST were more important at ANA (**Fig.2b** and **c**). For campaign#2, no clear spatial variation was observed.

Concerning the remotely sensed oceanic conditions (RSOC), their pattern differed significantly among campaigns (**Fig.2f** and **Fig.3**). The difference among campaigns were mostly associated to Chla and the SST (**Fig.2e** and **Fig.2f**) which were significantly higher (ANOVA, $p < 0.01$) during campaign#1 (average±SD of 0.24 ± 0.06 mg m^{-3} and 28.06 ± 1.03 °C, respectively) than during campaign#2 (average of 0.18 ± 0.04 mg m^{-3} and 27.20 ± 1.01 °C, respectively). In terms of Chla concentrations, campaign#2 was characterized by a high number of observations exceeding 0.5 mg m^{-3} in February, while only one observation exceeded this value during campaign#1 for the same month (**Fig.3**). Differences between campaigns were also observed for the onshore winds speed (Wind_U, **Fig.2e** and **Fig.2f**) that was significantly lower (ANOVA, $p < 0.01$) in campaign#1 (average: 1.04 ± 0.27 m s^{-1}) than in campaign#2 (1.29 ± 0.27 m s^{-1}). The alongshore wind (Wind_V) consisted mostly of northerly wind that blew at an average of 1.06 ± 0.50 m s^{-1}) during campaign#1 and 0.79 ± 0.50 m s^{-1} only during campaign#2. During campaign#1, up to 60% of currents flew offshore (with an average speed of 0.04 ± 0.02 m s^{-1}) and 40% were onshore (average: 0.02 ± 0.004 m s^{-1}). During campaign#2, currents flew entirely onshore (average: 0.03 ± 0.03 m s^{-1}). During campaign#1, over 80% of alongshore currents were southward (with an average speed of 0.12 ± 0.06 m s^{-1}), while this percentage dropped to around 67% for campaign#2 (average: 0.09 ± 0.06 m s^{-1}). Only 20% of currents were northward for campaign#1 (with a steady speed of 0.30 m s^{-1}) and about 33% only for campaign 2 (average: 0.07 ± 0.02 m s^{-1}).

Spatial Characteristics of the Benthic Substrate

The characteristics of the benthic substrate at GRT and ANA were different (PERMANOVA test, $R^2 = 0.30$, $p < 0.01$, **Fig.4a**). SIMPER analysis revealed that this dissimilarity was mainly due to macroalgae (Table. 2). Indeed, macroalgae represented 52% of the benthic substrate in ANA but ~5% only at GRT (**Fig.4b)**. The higher coverage by live coral (~18%) and sand (~42%) at GRT against ~11% and ~19% at ANA, respectively, also contributed to the dissimilarity in the characteristics of the benthic substrate between the two sites (**Fig.4b** and **Table. 2**).

Spatio-Temporal Variability in Presettlement Fish Assemblages

A total of 165 species belonging to 42 families were collected during the two sampling campaigns, with 99 species caught at GRT and 96 at ANA (Online resource 1). About 60 species from 20 families had an occurrence greater than 3%, with 46 species for GRT and 51 species for ANA. At GRT, the most frequent species were *Ostorhinchus* [BOLD:AAC2084], *Dascyllus trimaculatus*, and *Apogonichthys perdix* during campaign#1 (**Fig.5a**), *A. perdix,*

Pomacentrus trilineatus, and *Petroscirtes* [BOLD:AAE6131] during campaign#2 (**Fig.5c**). At ANA, the most frequent species were *Chaetodon auriga, D. trimaculatus, and Lutjanus fulviflamma* during campaign#1 (**Fig.5b**), *C. auriga, A. perdix*, and *Chromis viridis* during campaign#2 (**Fig.5d**). At GRT, the most abundant species were *Ostorhinchus* [BOLD:AAC2084], *A. perdix* and *Nectamia* [BOLD:AAL9262] during campaign#1 (**Fig.5e**), *Lethrinus mahsena, Ostorhinchus* [BOLD:AAC2084], and *A. perdix* during campaign#2 (**Fig.5g**). At ANA, the most abundant species were *Siganus sutor, Lethrinus* [BOLD:AAB0511], and *L. fulviflamma* during campaign#1 (**Fig.5f**), *Lethrinus mahsena, C. viridis*, and *C. auriga* during campaign#2 (**Fig.5h**). The difference in terms of the most frequent and abundant species reflects the dissimilarity of presettlement fish assemblages between sites and campaigns, a difference that was confirmed by the MDS analyses (**Fig.6**). The results of PERMANOVA tests show that these differences were significant at $p < 0.01$.

Relative Contribution of the Three Types of Explanatory Variables on PFAs

When the two campaigns were considered together, the result of Gradient Forest analyses indicates that the variations of presettlement fish assemblages (PFAs) were not explained by the characteristic of the benthic substrates. Indeed, the R^2 value corresponding to habitat variables was only 0.06 for ANA (**Table.3**) and explained the variations of *S. sutor* only. For GRT, the R^2 value was more important, reaching 0.35, but it explained the variation of *Fistularia commersonii* abundances only. Therefore, the characteristics of the benthic substrate were excluded from subsequent analyses.

The variables measured locally during sampling (i.e. local variables) were also not relevant in explaining the PFAs at GRT ($R^2_{max.}$=0.11, $R^2_{aver.}$ =0.05), and at ANA ($R^2_{max.}$=0.21, $R^2_{aver.}$ = 0.09, **Table.3**).

When the sites were considered separately, the PFAs appeared more predictable based on $R^2_{aver.}$ values (**Table.4**). Compared to the characteristics of the benthic substrate and the local variables, the remotely sensed oceanic conditions (RSOC) explained better the structuration of the PFAs with $R^2_{aver.}$ = 0.22 for GRT and 0.34 for ANA during campaign#1 and $R^2_{aver.}$ = 0.18 for GRT and 0.32 for ANA during campaign#2 (**Table.4**). Interestingly, based on these "$R^2_{aver.}$" RSOC appeared to explain only a low part of the variations of the explained PFAs for GRT, while about a moderate part for ANA.

Spatio-Temporal Variability of RSOC Contribution

The species for which the abundances were better explained by RSOC differed between campaigns and between sites for a given campaign (**Fig.7)**. The number of these species was higher during campaign#2 (**Fig.7f and h**) than during campaign#1 (**Fig.7b and d**). Moreover, the number of these species were also more important in ANA with nine species during campaign#1 (**Fig.7d**) and 13 species during campaign#2 (**Fig.7h**) than in GRT with four and seven species respectively (**Fig.7b** and **Fig.7f**). At GRT, *Chromis viridis* and *C. auriga* were the species for which the abundances were better described in campaign#1 (**Fig.7b**) while it was *L. mahsena* and *C. petersii* in campaign#2 (**Fig.7f**). At ANA, the species for which the abundances were better described were different among campaigns, expect for *L. fulviflamma* which was always among these species (**Fig.7d** and **Fig.7h**).

The rank and importance of the explanatory variables (based on R^2 weighted importance) varied between years and sites. During campaign#1, Chla and the northerly Wind_V were the most important explanatory variables of presettlement fish assemblages (PFAs) at GRT and ANA sites (**Fig.7a** and **c**), Current_V also being among the most important explanatory variables at ANA (**Fig.7c**). For campaign#2, five explanatory variables appeared to be the most important, but the rank of these explanatory variables differed at each site (**Fig.7e** and **g**).

RSOC Threshold for Presettlement Fish Assemblages

During campaign#1, three important RSOC variables presented thresholds for the PFAs at ANA, and up to five variables at GRT. Important shifts in the PFAs structure occurred at 0.24 mg m^{-3} and 0.27 mg m^{-3} for Chla, and at 0.8 and 1.4 m s^{-1} for the northerly wind for the both sites, and at 0 and 0.1 m s^{-1} for southward Current_V, and at 0.05 m s^{-1} for northward Current_V at ANA only (**Fig.8**). During campaign#2, important shifts in the PFAs structure observed at GRT at 0.5 m s^{-1} and 1.2 m s^{-1} for northerly Wind_V, at 0 and 0.065 m s^{-1} for northward Current_V, at 0.03 m s^{-1} for onshore Current_U, at 1.1 and 1.5 m s^{-1} for westerly Wind_U, and at 26.5°C , 27.75 and 28.25°C for SST (**Fig.8**). At ANA, important turnover in the PFAs structure were detected at 26.5°C and 28.25°C for SST, at 1.5 m s^{-1} for westerly Wind_U, at 1.2 m s^{-1} for northerly Wind_V, at 0.035 m s^{-1} for onshore Current_U, at 0.1 m s^{-1} for southward Current_V, and at 0 and 0.05 m s^{-1} for northward Current_V (**Fig.8**).

RSOC Thresholds for Species

During campaign#1, the northerly wind speed at 1.2 m s^{-1} induced the important shifts in the abundances of all the species for which the abundances were better described (**Fig.9a** and **b**), the Chla concentrations at 0.29 mg m^{-3} played a role in the abundances of *C. viridis* at GRT (**Fig.9a**), and *S. sutor*, *Pomacentrus trilineatus*, and *L. fulviflamma* at ANA (**Fig.9b**). A common threshold of northward currents speed at 0.12 m s^{-1} was detected during campaign#1 for the five species at ANA (**Fig.9b**).

During campaign#2, no clear common threshold was observed for the two species at GRT (**Fig.9c**) or for the five species at ANA (**Fig.9d**). GF models detected thresholds for some species only. For ex., the onshore wind at 1.65 m.s^{-1} generated a shift in the abundances of *L. mahsena* at GRT, and in the abundances of *Sphyraena barracuda* and *L. fulviflamma* at ANA. At ANA, onshore currents at 0.06 m s^{-1} also induced shifts in the abundance of *L. notatus*.

Discussion

To our knowledge, this study is one of the first describing the structure of tropical presettlement fish assemblages (PFAs) based on accurate species identification obtained through DNA barcoding, and using a) Gradient Forest (GF), one the most robust exploratory machine learning technique, and b) remotely sensed oceanic conditions (RSOC) observed before sampling (i.e. past-day RSOC). The results of GF analyses highlighted the importance of past-day RSOC in explaining the PFAs compared to information recorded during sampling or the characteristics of the benthic substrate.

RSOC: More Useful for Explaining the PFAs

In the present study, the PFAs differed significantly among sites. Such differences in PFAs between GRT and ANA sites had already been observed by Jaonalison et al. (2016). These authors hypothesized that the spatial differences of PFAs might be explained by the characteristics of the benthic substrate as such link had been clearly defined for post-settlement stage by Levin et al. (1997). Although the characteristics of the benthic substrate were different between the two sites, GF analyses demonstrated that they were of low importance for explaining the PFAs structure. This weak relationship between the characteristics of the benthic substrate and PFAs structure may suggest the water masses characteristics play an important role for PFAs. Using Canonical Correspondence Analysis (CCA), Chen et al. (2018) found that the larval fish assemblages collected with plankton nets in northern South China Sea can be defined by water mass characteristics, mainly SST and sea surface salinity. In the current study, the water masses characteristics differed among sites. In ANA, the water masses corresponded to a more coastal environment under the influence of the nearby Onilahy River with salinities varying from 31.5 to 35.7 (Jaonalison et al. in review). Oppositely, at GRT the water masses were less variable and more similar to oceanic water conditions due to the important tidal inflow that regularly renews lagoonal water masses (Chevalier et al. 2015). Although water masses characteristics have been found to explain the spatial differences in abundances (N) of presettlement fish (Jaonalison et al. in review), they appeared to be less important for explaining PFAs structure in the current study. Indeed, GF models obtained using locally recorded variables had the goodness of fit varying between 5 and 9% only.

The remotely sensed oceanic conditions (RSOC) appeared to be the most important variables explaining the observed presettlement fish assemblages (PFAs) - reflecting the importance of coastal oceanic conditions in nearshore area for shaping PFAs. Out of the 60 species for which presettlement stage occurrence was more than 3%, the abundances of 23 species only were explained by gradient forest based on RSOC. Among these 23 species, 15 were among the most frequent species observed in the both sites but the eight others (*Abudefduf vaingiensis*, *Canthigaster petersii*, *Dascyllus abudafar*, *Fowleria* [BOLD:AAD8726], *Lutjanus notatus*, *Paramonacanthus pusillus*, *Pomacentrus agassizi*, and *Sphyraena barracuda*) were not among the most frequent species. Thus, the occurrence of a species does not appear to be the main parameter making its abundances better described by gradient forest. Moreover, the abundances of species belonging to Blenniidae, Caesionidae, Fistularidae, Holocentridae, Ostracidae, Plesiopidae, Pomacanthidae, Scorpaenidae, and Syngnathidae were not described by GF models although some of them such as *Caesio caerulaurea*, *Fistularia commersonii* were among the most frequent species in the samples. These species may be sensitive to environmental factors that have not been considered in this study such as sea surface salinity, and SST anomaly.

Spatial and Interannual Variabilities of GF Model Goodness of Fit Based on RSOC

The goodness of fit of the GF models based on RSOC, and the species for which the abundances were better described by the GF models differed between sites and years. The spatial variability could be linked to the difference in PFAs structure among sites as species may respond differently to environmental conditions. Surprisingly, the number of these species and the goodness of fit of models based on RSOC were always higher at ANA although this site

appeared to be a more variable environment due to the influence of the nearby Onilahy River (Jaonalison et al. in review). This result suggests that other characteristics of coastal water masses not recorded in this present study may play an important role. Some previous studies have suggested the hypothesis that coastal waters could present better feeding conditions for presettlement fish than open waters (Nagelkerken et al. 2001; Cocheret de la Morinière et al. 2002). In New Caledonia, the biomass of zooplankton has been observed to be strongly influenced by terrigenous inputs from rivers due to their high concentrations of particulate organic matter (Le Borgne et al. 1989). These findings led Carassou et al. (2010) to suggest that terrigenous inputs may provide best feeding conditions to presettlement fish in areas influenced by nearby rivers. The highest number of species at ANA could be thus linked to the best feeding conditions of presettlement fish in ANA (i.e. in coastal water masses) than at GRT (i.e. in open water mass). Concerning the interannual variability of the PFAs and of the species, both appeared to be associated to the importance rank of Chla. Indeed, Chla was the highest explanatory variables for PFAs during campaign#1 when it presented an average concentration of 0.24 ± 0.06 mg m^{-3}, and the lowest explanatory variables during campaign#2 with an average concentration of 0.18 ± 0.04 mg m^{-3}. In New Caledonia, the Chla concentration measured at the moment of sampling was found to be the main factor determining the spatial structure of PFAs (Carassou et al. 2008). High Chla concentrations are known to provide the main food source for zooplankton (Chassot et al. 2010). Thus, food availability for fish larvae, induced by high Chla, could be the main factor determining both the spatial and interannual variability of PFAs in tropical areas. This seems logical as a higher food availability increases the survival of pre-settlement fish (Olivar et al. 2010) and allow them to efficiently resist predation (Owen et al. 1989). These findings also give some insight to the tolerance of fish larvae to food availability, or to the dependency of their preys to primary production. In this present study, *S. sutor* was among the species during campaign#1 when Chla concentrations were high. The abundances of Siganidae larvae were also linked to high Chla concentration in New Caledonia coastal waters (Carassou et al. 2008). Oppositely, during campaign#2 Chla concentrations were generally low, except in February, but the abundances of Sphyraenidae (*Sphyraena barracuda*), Lutjanidae (*L. fulviflamma* and *L. notatus*) and Apogonidae (*A. perdix*) were highly explained. This suggests these species consume prey that are not directly linked to primary production.

Based on criteria applied by Roubeix et al. (2017) for significant threshold selection, three significant thresholds were clearly defined during campaign#1 only: 0.29 mg m^{-3} for Chla, 1.2 m s^{-1} for northerly alongshore wind speed, and 0.12 m s^{-1} for northward alongshore current in (**Fig.9**). In contrast, no clear thresholds were detected during campaign#2 characterized by an important number of observations with Chla concentration exceeding 0.5 mg m^{-3} in February (**Fig.3**). This highly skewed distribution of the values of Chla during campaign#2 may explain the absence of correlation between this parameter and the PFAs. The fact that high concentrations of Chla occurred in February during campaign#2 suggests they were induced by tropical storms that were more frequent than during campaign#1 according to Météo-France (2017). Indeed, tropical storms are known to increase the Chla concentrations (Lin et al. 2003). Moreover, a significant reduction of the average SST was also observed during campaign#2 compared to campaign#1. Strong winds associated to tropical storm have been demonstrated to decrease SST (Price et al. 2008) from a minimum of 1° (Cione et al. 2000) to a maximum of 9°

(Lin et al. 2003). The cold surface water related to such storms can extend to hundreds of kilometres (Emanuel 2001), and the return to initial conditions can take between 5 and 30 days (Dare and McBride 2011). While inducing lower SST, tropical storms may weaken (Morsink 2018), or delay (Reynalte-Tataje et al. 2012), the reproductive activity of some fish species. Moreover, tropical storms may also induce eggs and larvae drifting away as no fish larvae can swim against the maximum water current ~1 m s^{-1} (Fisher 2005), i.e. observed during a tropical storm.

Research Improvements and Perspectives

In conclusion, this work highlighted the complexity of predicting PFAs in the future as their structure appeared to depend mainly on Chla concentrations and hydrodynamic conditions that can dramatically change from year to year. Moreover, the eventual influence of tropical storms on Chla concentrations and hydrodynamic conditions reinforces this complexity and explains the non-consistency of gradient forest results. Two years of survey does not seem sufficient for obtaining a consistent output and not useful for getting insights about interannual variability. According to Brodeur et al. (2008), at least ten years of data are needed to detect a significant change in PFAs. This period of time could be sufficient for addressing interannual variability in PFAs structure (Koslow and Wright 2016) and to document the effects of El Nino and La Nina events as well as those of tropical storms.

In parallel with the acquisition of longer time series of presettlement fish, forecasting Chla concentrations appears to be among the important steps to avoid the difficulty in predicting PFAs. Forecasting the Chla seems, however, a challenging step in a complex marine system as its concentration linked to the hydrodynamic conditions. The Wavelet-transform and Artificial Neural Network (WANN) appeared to be a promising method for forecasting the Chla concentration 1 month ahead in South San Francisco Bay (Rajaee and Boroumand 2015) using time series data from 1994 to 2013. The performance of this approach to predict the target variable with high accuracy was confirmed by Alizadeh et al. (2017). However, the forecast should be initially conditioned by silica, dissolved iron, nitrate, and forcing data including zonal and meridional wind stress, SST and shortwave radiation (Rousseaux and Gregg 2017). Most of this information remain to be acquired in Western Indian Ocean region.

Acknowledgments

Our acknowledgments going to the research assistants (J. J. Marcellin, D. Fiandria, R. Tsipy, Tovondrainy, and Noelson) for conducting fish sampling, and to the team of the 'Institut des Sciences de l'Evolution de Montpellier' (ISEM) for sequencing our fish tissues, and the Institut Halieutique et des Sciences Marines for issuing research permit. We also thank the three reviewers for their constructive remarks leading to the improvement our manuscript quality.

Funding

This work was co-funded by the French National Research Institute for Sustainable Development (JEAI-ACOM/IH.SM-IRD) in France, and the Critical Ecosystem Partnership Fund (MG-66341) in United States.

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical approval

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed by the authors.

Sampling and field studies

All necessary permits for sampling and observational field studies have been obtained by the authors from the competent authorities.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

Alizadeh MJ, Kavianpour MR, Kisi O, Nourani V (2017) A new approach for simulating and forecasting the rainfall-runoff process within the next two months. J Hydrol 548:588–597

Anderson MJ (2017) Permutational Multivariate Analysis of Variance (PERMANOVA). In: Balakrishnan N, Colton T, Everitt B, et al. (eds) Wiley StatsRef: Statistics Reference Online. John Wiley & Sons, Ltd, Chichester, UK, pp 1–15

Brodeur R, Peterson W, Auth T, Soulen H, Parnel M, Emerson A (2008) Abundance and diversity of coastal fish larvae as indicators of recent changes in ocean and climate conditions in the Oregon upwelling zone. Mar Ecol Prog Ser 366:187–202

Carassou L, Le Borgne R, Rolland E, Ponton D (2010) Spatial and temporal distribution of zooplankton related to the environmental conditions in the coral reef lagoon of New Caledonia, Southwest Pacific. Mar Pollut Bull 61:367–374

Carassou L, Ponton D, Mellin C, Galzin R (2008) Predicting the structure of larval fish assemblages by a hierarchical classification of meteorological and water column forcing factors. Coral Reefs 27:867–880

Chassot E, Bonhommeau S, Dulvy NK, Mélin F, Watson R, Gascuel D, Pape OL (2010) Global marine primary production constrains fisheries catches. Ecol Lett 13:495–505

Chen L-C, Lan K-W, Chang Y, Chen W-Y (2018) Summer Assemblages and Biodiversity of Larval Fish Associated with Hydrography in the Northern South China Sea. Mar Coast Fish 10:467–480

Chevalier C, Devenon J-L, Rougier G, Blanchot J (2015) Hydrodynamics of the Toliara Reef Lagoon (Madagascar): Example of a Lagoon Influenced by Waves and Tides. J Coast Res 31:1403–1416

Cione JJ, Molina P, Kaplan J, Black PG (2000) SST time series directly under tropical cyclones: Observations and implications. In: American Meteorolical Society.

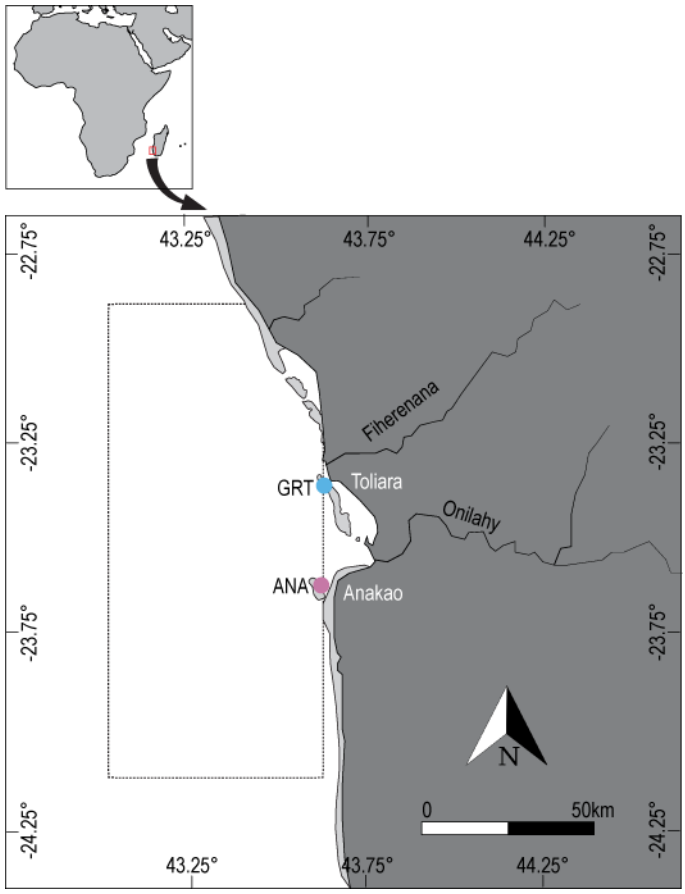
496 https://ams.confex.com/ams/last2000/webprogram/Paper12807.html. Accessed 28
497 May 2019
498 Clarke KR, Warwick RM, Laboratory PM (1994) Change in Marine Communities: An
499 Approach to Statistical Analysis and Interpretation, 2nd edn. Plymouth marine
500 laboratory, Natural environment research council
501 Cocheret de la Morinière E, Pollux BJA, Nagelkerken I, van der Velde G (2002) Post-
502 settlement Life Cycle Migration Patterns and Habitat Preference of Coral Reef Fish
503 that use Seagrass and Mangrove Habitats as Nurseries. Estuar Coast Shelf Sci 55:309–
504 321
505 Collet A, Durand J-D, Desmarais E, Cerqueira F, Cantinelli T, Valade P, Ponton D (2018)
506 DNA barcoding post-larvae can improve the knowledge about fish biodiversity: an
507 example from La Reunion, SW Indian Ocean. Mitochondrial DNA Part A 29:905–918
508 Cooke A, Ratomahenina O, Ranaivosoin E, Razafindraibe H (2000) Madagascar. In:
509 Sheppard CRC (eds) Seas at the millennium: an environmental evaluation.
510 Amsterdam, pp 113–131
511 Dare RA, McBride JL (2011) Sea Surface Temperature Response to Tropical Cyclones. Mon
512 Weather Rev 139:3798–3808
513 Dumas P, Bertaud A, Peignon C, Léopold M, Pelletier D (2009) A "quick and clean"
514 photographic method for the description of coral reef habitats. J Exp Mar Biol Ecol
515 368:161–168
516 Ellis N, Smith SJ, Pitcher CR (2012) Gradient forests: calculating importance gradients on
517 physical predictors. Ecology 93:156–168
518 Emanuel K (2001) Contribution of tropical cyclones to meridional heat transport by the
519 oceans. J Geophys Res Atmospheres 106:14771–14781
520 Ferrier S, Guisan A (2006) Spatial modelling of biodiversity at the community level. J Appl
521 Ecol 43:393–404
522 Fisher R (2005) Swimming speeds of larval coral reef fishes: impacts on self-recruitment and
523 dispersal. Mar Ecol Prog Ser 285:223–232
524 Folke C, Carpenter S, Walker B, Scheffer M, Elmqvist T, Gunderson L, Holling CS (2004)
525 Regime Shifts, Resilience, and Biodiversity in Ecosystem Management. Annu Rev
526 Ecol Evol Syst 35:557–581
527 Hsieh C, Reiss C, Watson W, Allen MJ, Hunter JR, Lea RN, Rosenblatt RH, Smith PE,
528 Sugihara G (2005) A comparison of long-term trends and variability in populations of
529 larvae of exploited and unexploited fishes in the Southern California region: A
530 community approach. Prog Oceanogr 67:160–185
531 INSTAT M, ICF M (2016) Enquête Démographique et de Santé de Madagascar 2008-2009.
532 In: World Bank Microdata Library.
533 https://microdata.worldbank.org/index.php/catalog/1435. Accessed on 08 July 2019
534 IOTC (2009) Madagascar National Report. In: Food and Agricultural Organizations of
535 United Nations. https://iotc.org/documents/madagascar-national-report. Accessed 19
536 Mar 2019
537 Jaonalison H, Mahafina J, Ponton D (2016) Fish post-larvae assemblages at two contrasted
538 coral reef habitats in southwest Madagascar. Reg Stud Mar Sci 6:62–74

539 Kohler KE, Gill SM (2006) Coral Point Count with Excel extensions (CPCe): A Visual Basic
540 program for the determination of coral and substrate coverage using random point
541 count methodology. Comput Geosci 32:1259–1269
542 Koslow JA, Davison PC (2016) Productivity and biomass of fishes in the California Current
543 Large Marine Ecosystem: Comparison of fishery-dependent and -independent time
544 series. Environ Dev 17:23–32
545 Koslow JA, Wright M (2016) Ichthyoplankton sampling design to monitor marine fish
546 populations and communities. Mar Policy 68:55–64
547 Le Borgne R, Blanchot J, Charpy L (1989) Zooplankton of tikehau atoll (Tuamotu
548 archipelago) and its relationship to particulate matter. Mar Biol 102:341–353
549 Levin PS, Chiasson W, Green JM (1997) Geographic differences in recruitment and
550 population structure of a temperate reef fish. Mar Ecol Prog Ser 161:23–35
551 Lin I, Liu WT, Wu C-C, Wong GTF, Hu C, Chen Z, Liang W-D, Yang Y, Liu K-K (2003)
552 New evidence for enhanced ocean primary production triggered by tropical cyclone.
553 Geophys Res Lett 30:1718
554 Lockwood SJ, Nichols JH, Dawson WA (1981) The estimation of a mackerel (*Scomber*
555 *scombrus L.*) spawning stock size by plankton survey. J Plankton Res 3:217–233
556 Mahafina J (2011) Perception et comportement des pêcheurs pour une gestion durable de la
557 biodiversité et de la pêcherie récifale du Sud-Ouest de Madagascar. Dissertation,
558 University of La Réunion
559 Météo-France OI (2017) Bilan des activités cycloniques dans le bassin Sud-ouest de l'Océan
560 Indien. In Cyclone Océan Indien. http://www.cycloneoi.com/archives-
561 blog/cyclone/2016-2017-une-nouvelle-saison-cyclonique-peu-active.html. Accessed
562 on 30 April 2019
563 Morsink K (2018) Hurricanes, Typhoons, and Cyclones. In: Smithsonian Institution Ocean.
564 http://ocean.si.edu/planet-ocean/waves-storms-tsunamis/hurricanes-typhoons-and-
565 cyclones. Accessed 28 May 2019
566 Nagelkerken I, Kleijnen S, Klop T, van den Brand R, de la Morinière E, van der Velde G
567 (2001) Dependence of Caribbean reef fishes on mangroves and seagrass beds as
568 nursery habitats: a comparison of fish faunas between bays with and without
569 mangroves/seagrass beds. Mar Ecol Prog Ser 214:225–235
570 Nichols JH, Bennett DB, Symonds DJ, Grainger R (1987) Estimation of the stock size of
571 adult *Nephrops norvegicus* (L.) from larvae surveys in the western Irish Sea in 1982. J
572 Nat Hist 21:1433–1450
573 Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara
574 RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2019) Community
575 Ecology Package. In: Cran r. https://cran.r-project.org/web/packages/vegan/vegan.pdf.
576 Accessed on 29 April 2019
577 Olivar MP, Emelianov M, Villate F, Uriarte I, Maynou F, Álvarez I, Morote E (2010) The
578 role of oceanographic conditions and plankton availability in larval fish assemblages
579 off the Catalan coast (NW Mediterranean). Fish Oceanogr 19:209–229
580 Owen RW, Lo NCH, Butler JL, Theilacker GH, Alvariño A, Hunter JR, Watanabe Y (1989)
581 Spawning and Survival Patterns of Larval Northern Anchovy, *Engraulis mordax*, in

582 Contrasting Environments-A Site-Intensive Study. Fishery bulletin (Washington, D.C.
583 : 1971) 87:673–688
584 Price JF, Morzel J, Niiler PP (2008) Warming of SST in the cool wake of a moving hurricane.
585 J Geophys Res 113:C07010
586 R Core Team (2018) R: A Language and Environment for Statistical Computing. In: R
587 Foundation for Statistical Computing. https://www.R-project.org/. Accessed on 11
588 June 2018
589 Rajaee T, Boroumand A (2015) Forecasting of chlorophyll-a concentrations in South San
590 Francisco Bay using five different models. Appl Ocean Res 53:208–217
591 Ratnasingham S, Hebert PDN (2013) A DNA-Based Registry for All Animal Species: The
592 Barcode Index Number (BIN) System. PLOS ONE 8:e66213
593 Reynalte-Tataje DA, Zaniboni-Filho E, Bialetzki A, Agostinho AA (2012) Temporal
594 variability of fish larvae assemblages: influence of natural and anthropogenic
595 disturbances. Neotropical Ichthyol 10:837–846
596 Roubeix V, Daufresne M, Argillier C, Dublon J, Maire A, Nicolas D, Raymond J-C, Danis P-
597 A (2017) Physico-chemical thresholds in the distribution of fish species among French
598 lakes. Knowl Manag Aquat Ecosyst 418:1–14
599 Rousseaux CS, Gregg WW (2017) Forecasting Ocean Chlorophyll in the Equatorial Pacific.
600 Front Mar Sci 4:236
601 Smith JA, Miskiewicz AG, Beckley LE, Everett JD, Garcia V, Gray CA, Holliday D, Jordan
602 AR, Keane J, Lara-Lopez A, Leis JM, Matis PA, Muhling BA, Neira FJ, Richardson
603 AJ, Smith KA, Swadling KM, Syahailatua A, Taylor MD, van Ruth PD, Ward TM,
604 Suthers IM (2018) A database of marine larval fish assemblages in Australian
605 temperate and subtropical waters. Sci Data 5:180207
606 Takahashi M, Watanabe Y (2004) Growth rate-dependent recruitment of Japanese anchovy
607 *Engraulis japonicus* in the Kuroshio-Oyashio transitional waters. Mar Ecol Prog Ser
608 266:227–238
609

610

611 **Figures**



612

613 **Fig.1** Sampling sites with blue dot: Great Barrier Reef of Toliara (GRT) and pink dot: Anakao
614 Reef (ANA). The dotted rectangle corresponds to the extraction area for the remotely sensed
615 oceanic conditions

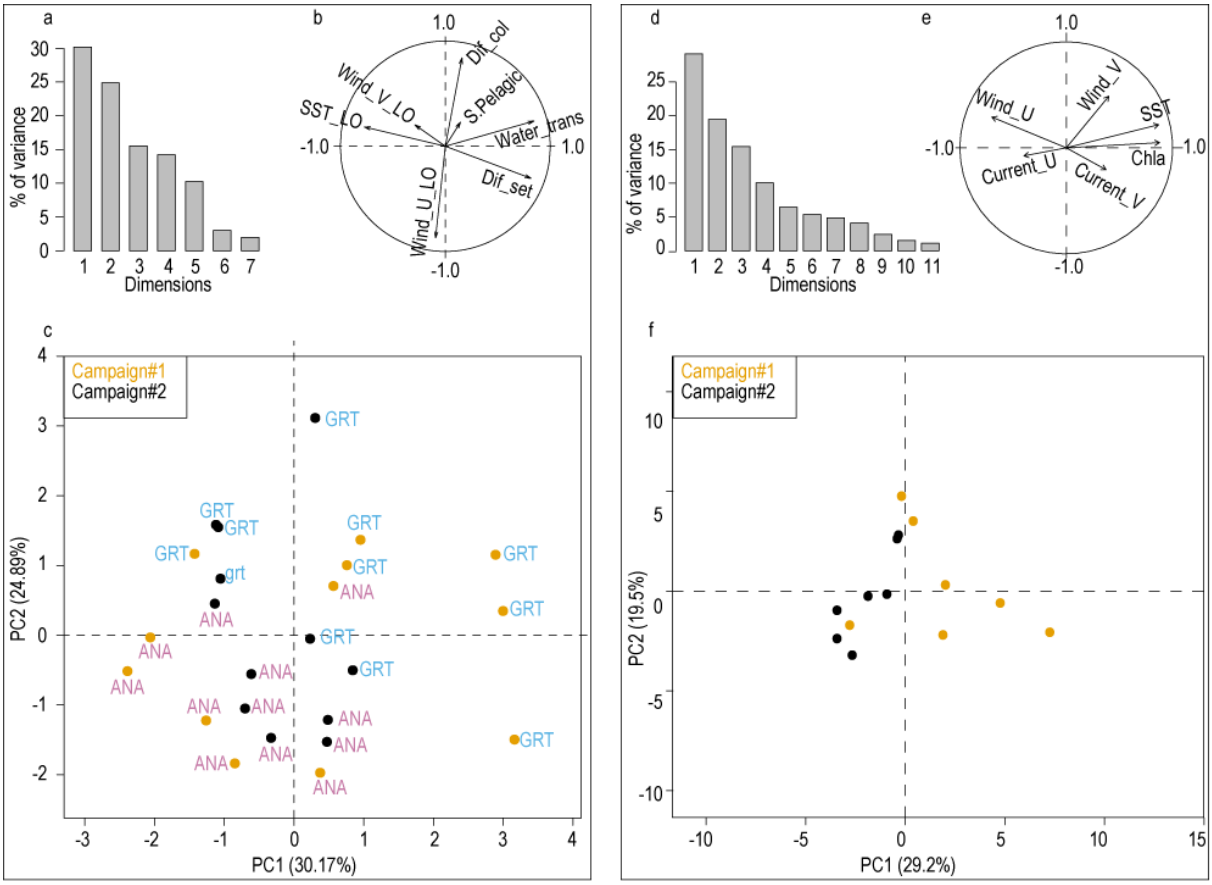
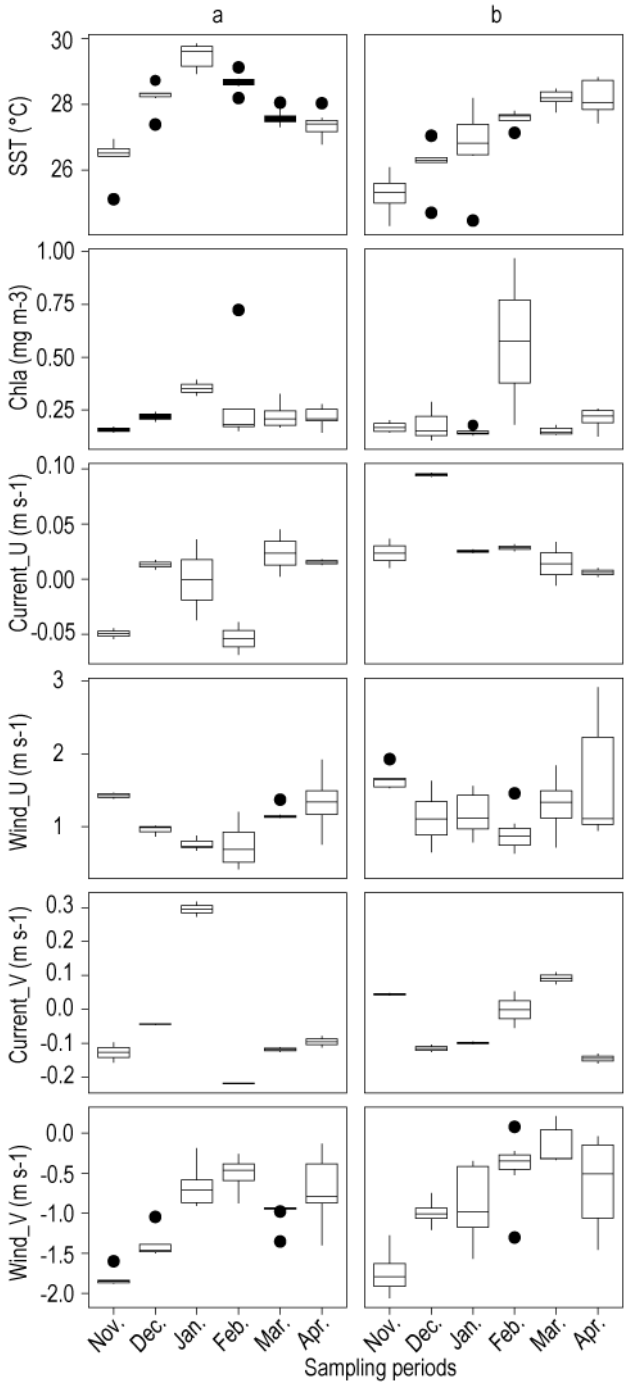


Fig.2 PCA of local variables (in left) with a: percentage of variance for each dimension, b: factor variables, c: PC1-PC2 plane with black dots for campaign#1 and yellow for campaign#2. Each point corresponds to the environmental record per site (in blue for GRT and pink for ANA) and per month. PCA of remotely sensed oceanic conditions (in right). With d: percentage of variance for each dimension, e: factor variables, f: PC1-PC2 plane with black dots for campaign#1 and yellow dots for campaign#2



623

624 **Fig.3** Interannual variability of the remotely sensed oceanic conditions for each month of
625 sampling, with a: campaign#1 (2016-2017) and b: campaign#2 (2017-2018). The lower and
626 upper boundaries of the boxes correspond to the 25th and 75th percentile, respectively. The
627 horizontal lines within the box correspond to the median values and the vertical lines show the
628 range of values that fall within 1.5 times the interquartile range individual points correspond to
629 values outside three times of the interquartile range

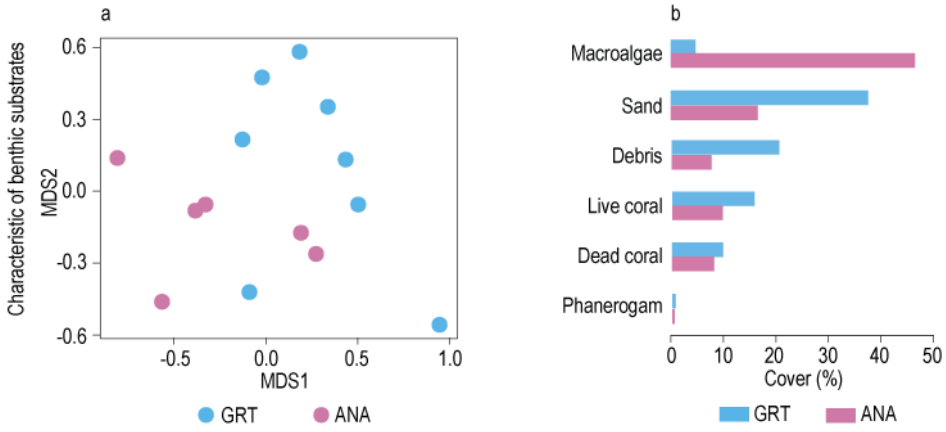


Fig.4 Distribution of the characteristics of benthic substrate of coral reef habitat in GRT (blue bars and dots) and ANA (pink bars and dots), with a: barplot of the percentage cover of the characteristics of benthic substrate and b: ordination of the characteristics of benthic substrate with multidimensional scaling (MDS) ordination

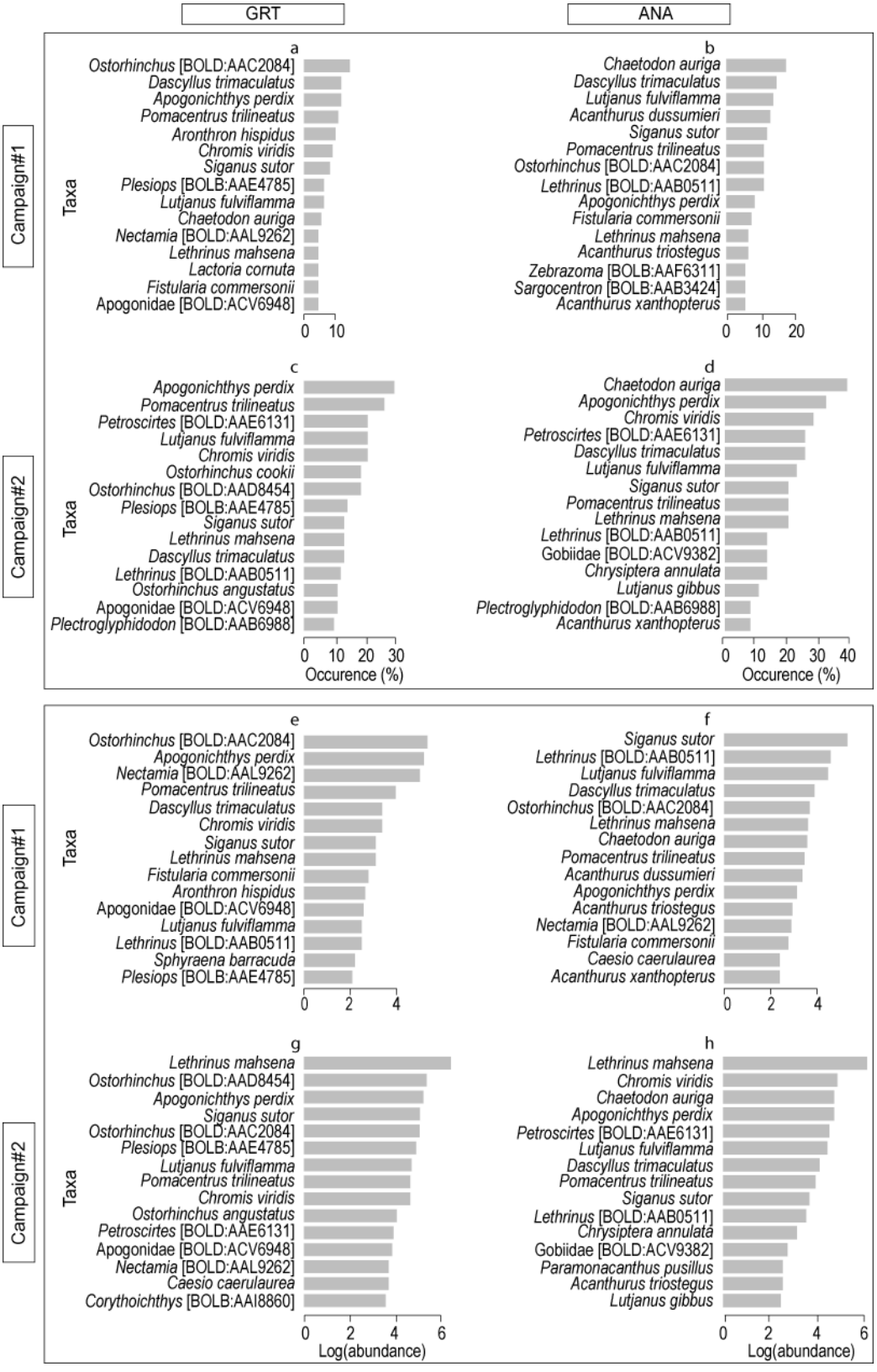
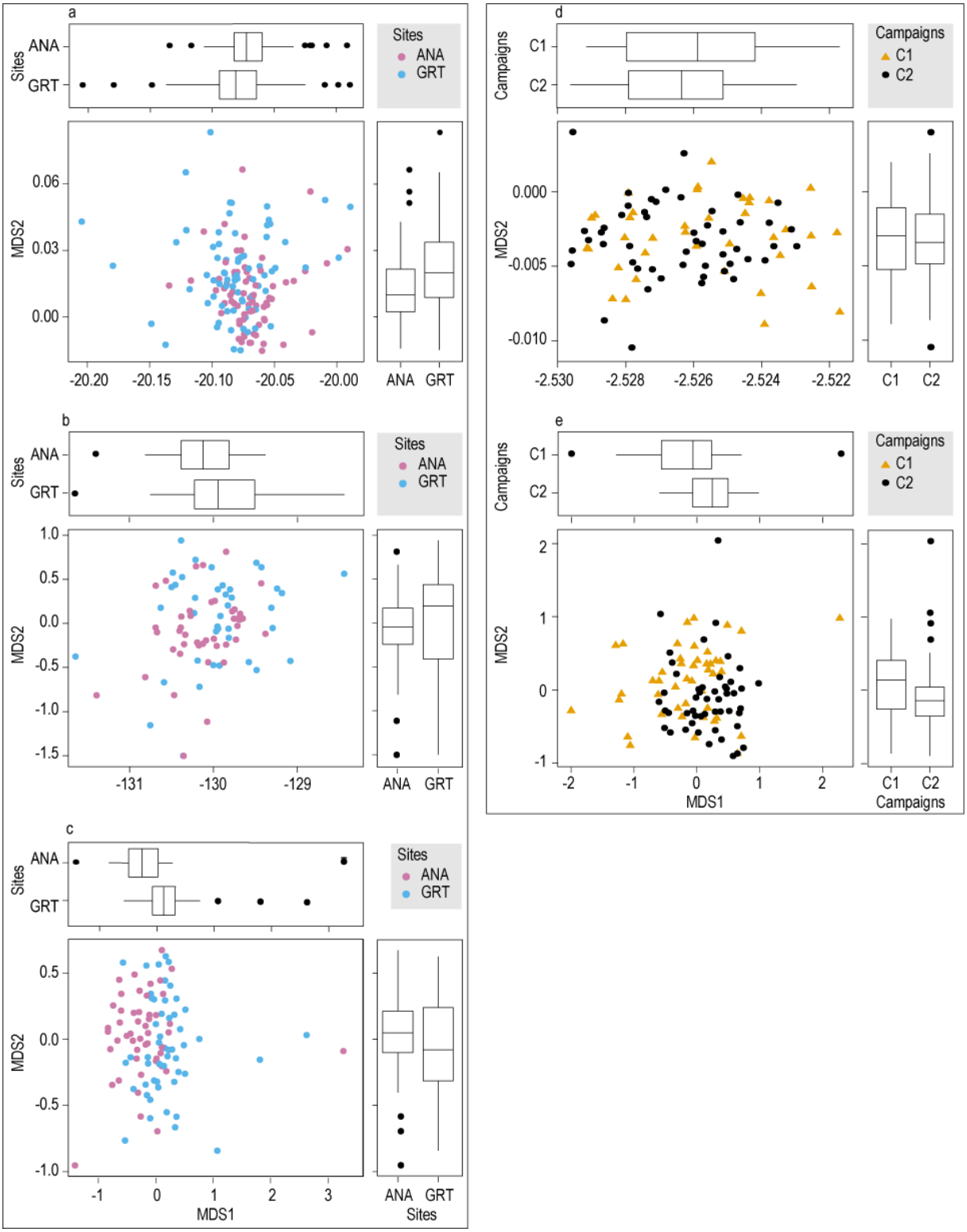


Fig.5 Occurrence of the first 20 most-occurring fish species observed at GRT during campaign#1 (a) and campaign#2 (c) and ANA during campaign#1 (b) and campaign#2 (d) in the top panel; and log-transformed abundances of the 20 most-abundant fish species observed at GRT for campaign#1 (e) and campaign#2 (g) and ANA for campaign#1 (f) and campaign#2 (h) in the down panel



641

642 **Fig.6** Multidimensional scaling (MDS) enlightening the spatial variability of the presettlement
643 fish assemblages (left panel), a: for both campaigns together, b: for campaign#1, and c: for
644 campaign#2 with blue points for GRT and pink points for ANA, and the interannual variability
645 of the presettlement fish assemblages (right panel), d: for GRT, e: for ANA with yellow
646 triangles corresponding to campaign#1 (C1) and black circles to campaign#2 (C2). Boxplots
647 indicate the spatial (left panel) and interannual (right panel) variability of presettlement fish

assemblages along each axis based on the distance between each observation and its corresponding centroid

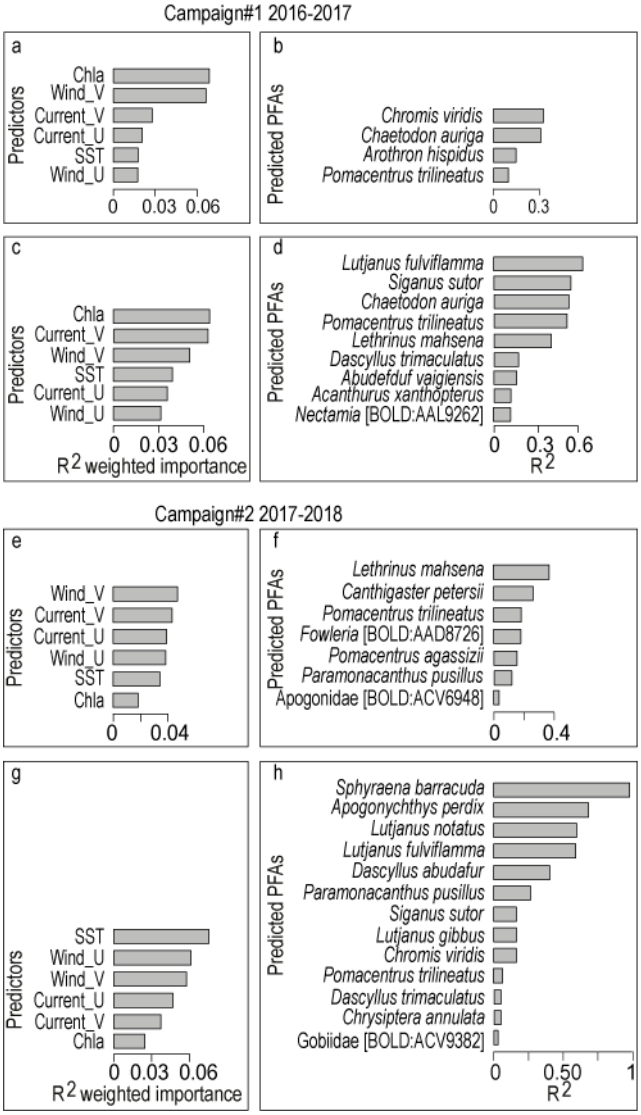
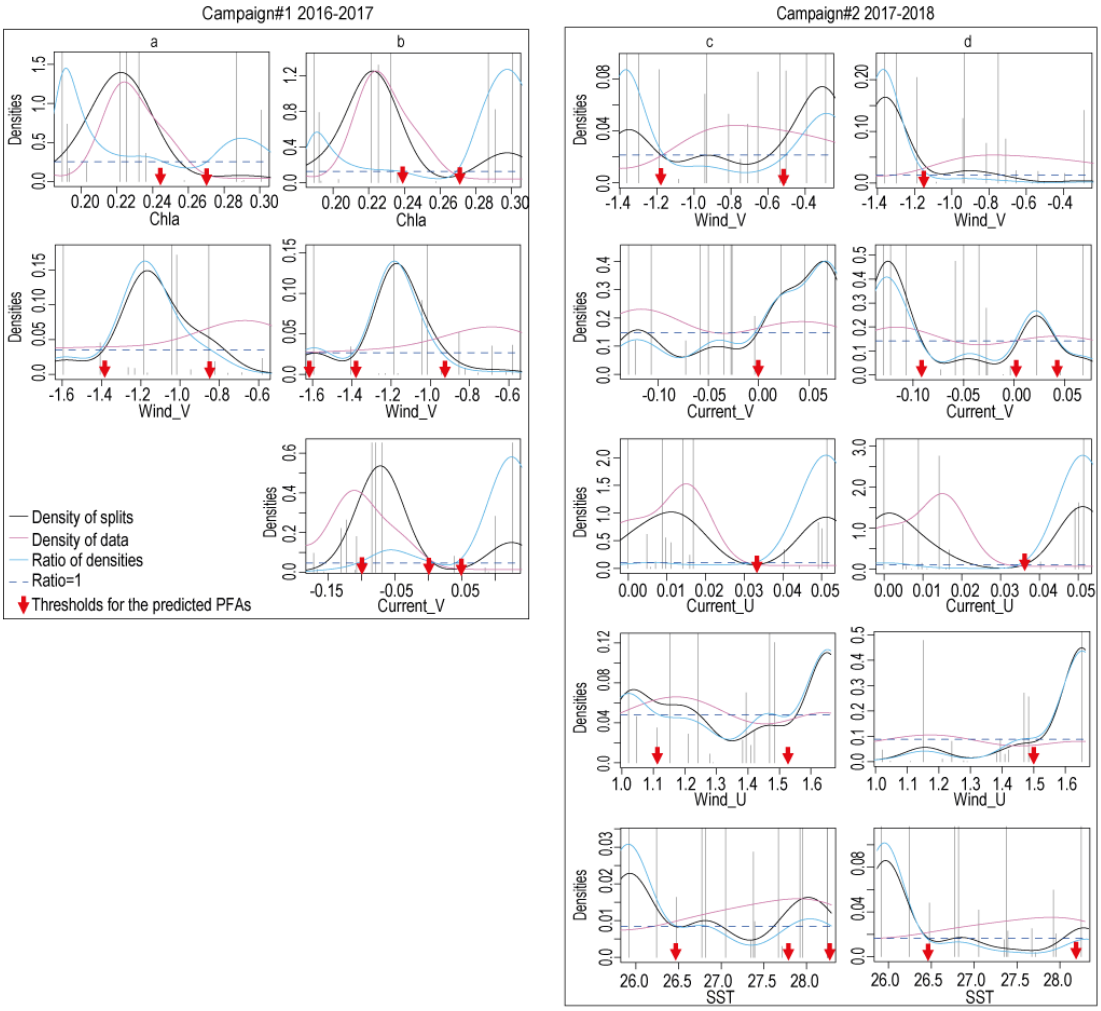


Fig.7 Gradient forest results for GRT (a, b, e, and f) and ANA (c, d, g, and h) with the overall RSOC variables importance plot (a and c for campaign#1 and e and g for campaign#2), and the species performance plot showing the goodness of fit R^2 for overall species (b and d for campaign#1 and f and h for campaign#2). Species with $R^2 \leq 0$ are not represented



657

658 **Fig.8** Split density plot for campaign#1 (a, b) and campaign#2 (c, d) showing the thresholds
659 (outlined by the red arrows) for the explained presettlement fish assemblages corresponding to
660 the range of value of the ratio of densities (blue full line) superior to the ratio=1 (blue dashed
661 line), the black and pink full line denote the density of splits and the density of data,
662 respectively. With a and c for GRT, and b and d for ANA

663

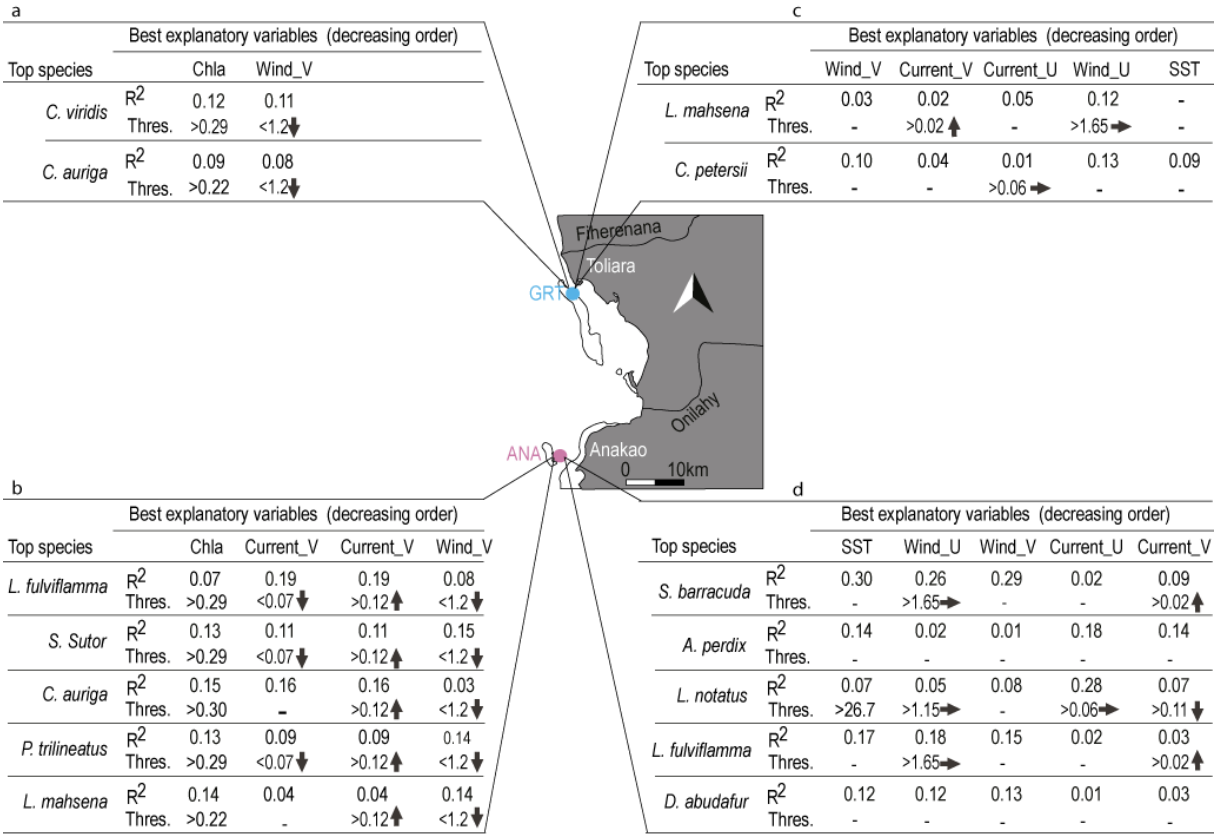


Fig.9 Interannual and spatial variability of accuracy importance (based on R^2 value) of the most important explanatory variables (in decreasing importance) for each of the top species for which the abundances were better described (in decreasing order) with a and b for GRT and ANA during campaign#1 and c and d for GRT and ANA during campaign#2. The thresholds (thres.) values were obtained and synthetized from the specific cumulative importance curves in **Online Resource3** for campaign#1, and **Online Resource4** for campaign#2. Dash (-) denotes the absence of thresholds and the arrows relate to the direction of wind and currents vectors

Supplementary materials

Online resource1: https://static-content.springer.com/esm/art%3A10.1007%2Fs12526-020-01068-6/MediaObjects/12526_2020_1068_MOESM1_ESM.pdf

Online resource2: https://static-content.springer.com/esm/art%3A10.1007%2Fs12526-020-01068-6/MediaObjects/12526_2020_1068_MOESM2_ESM.pdf

Online resource3: https://static-content.springer.com/esm/art%3A10.1007%2Fs12526-020-01068-6/MediaObjects/12526_2020_1068_MOESM3_ESM.pdf

Online resource4: https://static-content.springer.com/esm/art%3A10.1007%2Fs12526-020-01068-6/MediaObjects/12526_2020_1068_MOESM4_ESM.pdf