# A new framework for estimating abundance of animals using a network of cameras

Camille Magneville, Capucine Brissaud, Valentine Fleuré, Nicolas Loiseau, Thomas Claverie, Sébastien Villéger

# A new framework for estimating abundance of animals using a network of cameras

**Authors:**

Camille Magneville [1,2*] (https://orcid.org/0000-0003-0489-3822)

Capucine Brissaud [1]

Valentine Fleuré [3,1] (https://orcid.org/0000-0002-6567-1986)

Nicolas Loiseau [1] (https://orcid.org/0000-0002-2469-1980)

Thomas Claverie [1, 4, 5] (https://orcid.org/0000-0002-6258-4991)

Sébastien Villéger [1] (https://orcid.org/0000-0002-2362-7178)

[1] MARBEC, Univ Montpellier, CNRS, Ifremer, IRD, Montpellier, France.

[2] Center for Ecological Dynamics in a Novel Biosphere (ECONOVO), Department of Biology, Aarhus University, Ny Munkegade 114, DK-8000 Aarhus, Denmark

[3]ZooParc de Beauval & Beauval Nature, 41110, Saint Aignan, France

[4]Centre Universitaire de Formation et de Recherche de Mayotte, France.

[5] UMR ENTROPIE, Univ La Réunion, IRD, IFREMER, Univ Nouvelle-Calédonie, CNRS, Saint-Denis, Réunion, France

[*] corresponding author (camille.magneville@gmail.com)

**Running head: Estimating abundance with a camera network**

**Keywords**: abundance estimation, fishes, maxN, remote underwater videos, R package

**Abstract (<250 words)**

While many ecology studies require estimations of species abundance, doing so for mobile animals in an accurate, non-invasive manner remains a challenge. One popular stopgap method involves the use of remote video-based surveys using several cameras, but abundance estimates derived from this method are computed with conservative metrics (e.g. *maxN* computed as the maximum number of individuals seen simultaneously on a single video). We propose a novel methodological framework based on a remote-camera network characterised by known positions and non-overlapping field-of-views. This approach involves a temporal synchronisation of videos and a maximal speed estimate for studied species. Such a design allows computing a new abundance metric called *Synchronised maxN* (*SmaxN*). We provide a proof-of-concept of this approach with a network of nine remote underwater cameras that recorded fishes for three periods of one hour on a fringing reef in Mayotte (Western Indian Ocean). We found that abundance estimation with *SmaxN* yielded up to four times higher values than *maxN* among the six fish species studied. *SmaxN* performed better with an increasing number of cameras or longer recordings. We also found that using a network of synchronised cameras for a short time period performed better than using few cameras for a long duration. The *SmaxN* algorithm can be applied to many video-based approaches. We built an open-sourced R package to encourage its use by ecologists and managers using video-based censuses, as well as to allow for replicability with *SmaxN* metric.

## Introduction

Measuring the abundance of species and their size classes is the cornerstone of many ecological studies and management of protected areas and fisheries (Langlois, Harvey, and Meeuwig 2012). In fact, species abundance distribution provides insights into ecosystem functioning, as it underlies key ecological phenomena such as resource availability (Liu et al. 2021), biotic interactions (Boulangeat et al. 2012) and colonisation capacities (Verberk 2011). Counting individuals of a mobile species over a given area often presents challenges due to the inherent difficulty in detecting individuals as they move (Birt et al. 2012). In reef ecosystems, the abundance of mobile organisms is mostly estimated using Underwater Visual Census (UVC) by scuba divers (Brock 1954, Harmelin-Vivien et al. 1985). UVC data have been the core of worldwide conservation programs (Murphy and Jenkins 2010) and fish stocks management (Labrosse et al. 2002). Yet this method is not without problems, as the divers presence could affect species communities through sound (Radford et al. 2005) and visual stimuli (Dickens et al. 2011). Studies have indeed found lower fish richness and species abundance while using UVC compared to least invasive methods such as Remote Underwater Videos (RUVs) (Dearden et al. 2010; Zarco-Perello and Enríquez 2019). Moreover, RUVs can be set up in environments where divers are not able to perform long and numerous observations, such as on mesophotic reefs or in open water habitats (Mallet and Pelletier 2014).

However, estimating fish abundance is challenging because it is difficult to physically distinguish one individual fish from another. Hence, it is impossible to differentiate between a single individual passing multiple times in front of the camera and multiple individuals passing in front of the camera once (Cappo et al. 2003). To overcome this issue, the *maxN* metric, computed as the maximal number of individuals spotted simultaneously (Ellis and De Martini 1995), has been used by most RUV studies (e.g. 81% of baited RUV studies between

1950 and 2016 (Whitmarsh et al. 2017)). The *maxN* metric is thus conservative and provides a non-linear underestimation with true abundance (Schobernd et al. 2014; Campbell et al. 2015). Other metrics have been proposed to estimate abundances such as the mean count over a time step (MeanCount, (Schobernd et al. 2014)) but they have also been shown to be non-linearly related with true abundance (Kilfoil et al. 2017).

To improve the accuracy of *maxN*, it has been suggested to expand the filmed area (Campbell et al. 2015). Indeed, Campbell et al. (2018) demonstrated that expanding the camera field of view from 90° to 360° allows the relationship between the *maxN* estimate to be linearly related to actual species abundance. It thus leads towards the use of several cameras instead of just one. In fact, using several cameras increases the chance of observing more individuals on a single frame (Schobernd et al. 2014). Using Baited RUV (BRUV) and four cameras facing different directions, Whitmarsh et al. (2017) highlighted the ability of several cameras to increase *maxN* estimates especially for highly abundant species. Indeed, the use of a network of cameras has recently increased (Harvey et al. 2007, Widmer et al. 2019, Lopez-Marcano et al. 2021, Letessier et al. 2021). However, the abundance of a species is still estimated as the maximum number of individuals recorded within a specific time frame by a single camera.

In addition, extending the duration of video recordings enhances the probability of capturing all individuals within the spatial area in front of a camera at least once, thereby leading to more accurate estimates of abundance (Campbell et al. 2015). And yet, because analysing long-duration video data is a time-consuming process, a trade-off must be found between video-duration and the number of cameras to reduce individual counting time. Garcia et al. (2021), studied this trade-off effect on species number count (richness) with 46 videos of 10 minutes. They found that a minimum of five videos was adequate to sample the majority of species richness, with most species recorded within the initial five minutes. .

Nevertheless, to our knowledge, no study has explored the impact of the number of videos and extended recording duration on estimation of abundance of each species.

In the present study, we propose a novel methodological framework based on a network of synchronised remote cameras and we introduce a new metric for estimating species abundance, the *Synchronised maxN* (*SmaxN*). The *SmaxN* index considers the maximal speed of the studied species and the distances between the cameras. We proposed a proof-of-concept of this framework for six reef fish species using a network of nine underwater cameras, each recording for one hour during three different time periods on a fringing reef in Mayotte (Western Indian Ocean). We specifically address the following questions: (i) How does using synchronised cameras and including species maximal speed influence the estimation of abundance? and (ii) How does the number of underwater cameras and recording duration influence the estimation of abundance?

## Materials and Procedures

### Computation of the *SmaxN* metric

To implement the *SmaxN* framework, surveys should involve a set of at least two fixed cameras positioned at specified geographical coordinates within a habitat, with non-overlapping fields of view (see Box 1- Step 1). All videos recorded by these cameras should be temporally synchronised at a given precision (e.g. same second).

Once videos have been recorded in the field, abundances of the studied species through time should be assessed by experts. The frequency of time steps for which individuals are counted must align with the study's objectives and be greater than the precision of camera synchronisation. Therefore, for each species, there exists a matrix known as the camera time abundance matrix, which records the abundance for each time step (rows) across each camera (columns) (*see* Box 1 - Step 3).

Using camera positions, the distances *d* between each pair of cameras are computed. The maximal speed $v_{max}$ of the studied species should then be determined by retrieving the critical swimming speed $U_{crit}$ established through laboratory experiments with velocity tests or based on field speed performance (Fulton 2007) which could also be estimated from present video measurements if stereo cameras are used (Satterfield et al. 2022). Then, the minimal time $t_{min}$ needed for an individual to move from one camera to another is computed (*see* Box 1 - Step 2) as $t_{min} = d / v_{max}$ . Minimal times for each pair of cameras are gathered in the time-lag matrix.

Since the cameras are synchronised, and their fields-of-view do not overlap, an individual can only be recorded by one camera at a given time. Abundance can thus be estimated using the *instantaneous Synchronised maxN (iSmaxN)* metric calculated as the maximum sum of abundances across all cameras for a specific time step. By definition, *iSmaxN* is at least equal to the maximum of the *maxN* estimates computed independently for each camera. However, the *iSmaxN* metric is still conservative as it only hypotheses that an individual cannot move instantaneously from one camera field-of-view to another.

Therefore, we propose to expand the concept of non-duplicity of individuals across both space and time, that is: individuals from a given species recorded by different cameras are different individuals if they are seen during a time span shorter than the time required to move between those cameras, considering species' maximal speed and the distances between cameras. Hence, abundance can be more accurately estimated with the *SmaxN* metric computed as the maximum number of individuals recorded during a time span defined according to species' maximal speed and the distances between cameras. The challenge of the *SmaxN* approach is to find the maximum abundance possible within the camera×time abundance matrix, given the distance between cameras and species' maximal speed. The number of combinations possible increases with the number of cameras and the distance

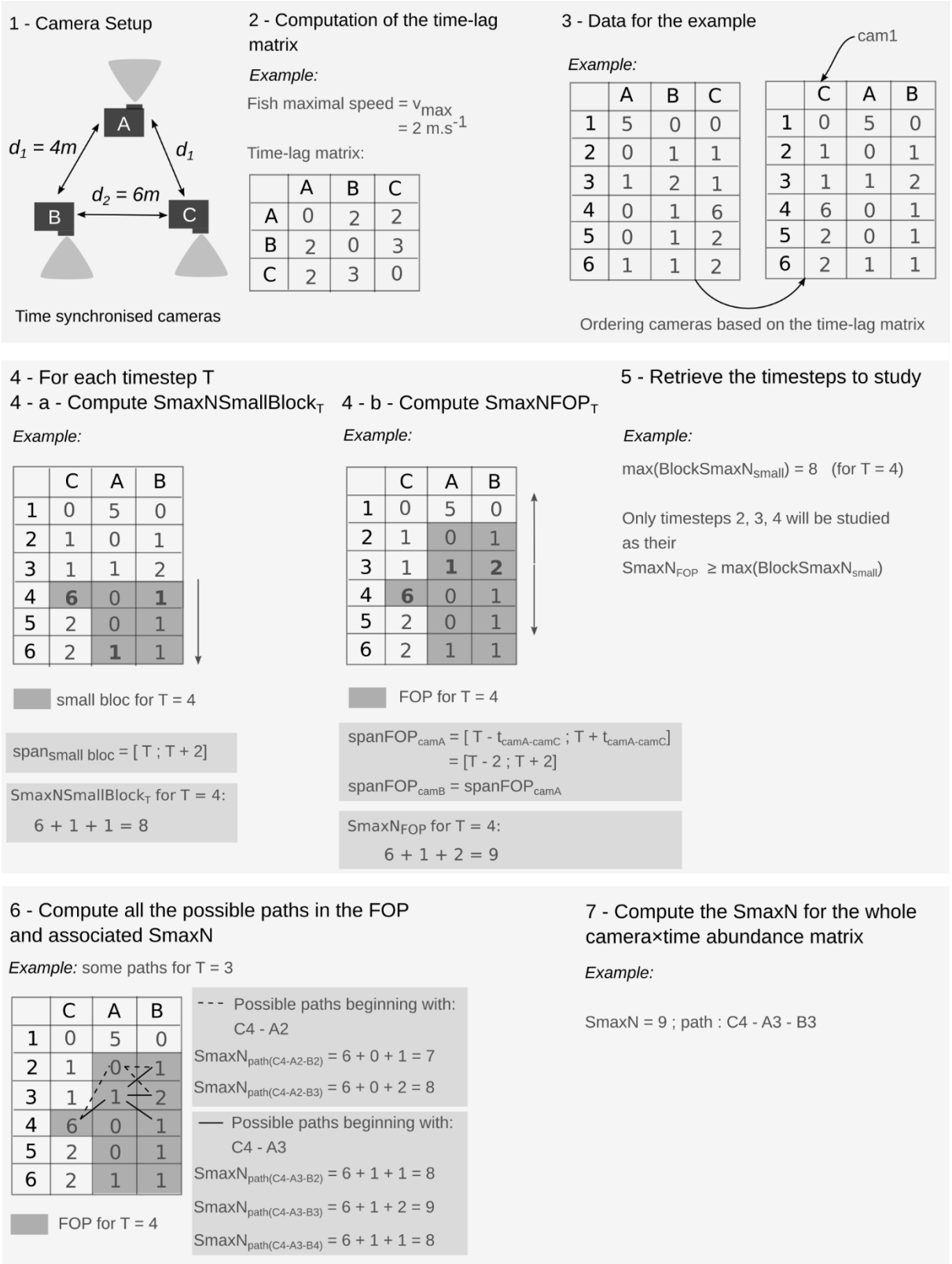
between them for a given speed, and it increases with decreasing speed for a given survey design.

To reduce computation time, we designed an iterative pipeline which prevents exploring time steps that have no chance to provide an estimate higher than the *iSmaxN* metric, or no chance to increase estimates computed for previous time steps (see details in Box 1 - Step 4 to 7).

For the same sampling effort, the *SmaxN* metric equals the *maxN* metric when the maximal number of individuals across all cameras and time steps is obtained on a single camera on a timestep surrounded by the absence of the studied species. In any other cases, *SmaxN* is higher than *maxN,* for instance in Box 1 *SmaxN* equals 9 and *maxN* equals 6. *SmaxN* increases with the number of cameras as the chance of observing more individuals increases with the recorded area. *SmaxN* also varies with species maximal speed as the possible paths vary with different maximal speeds.

*SmaxN*, *maxN* and *iSmaxN* were computed using the *SmaxN* R package currently available on Github (https://github.com/CmlMagneville/SmaxN).

## BOX 1: STEPS TO COMPUTE THE SMAXN METRIC




We illustrate in the figure below the computation of *SmaxN* for a simple case with only three cameras and six time steps.

In the camera×time abundance matrix, cameras are ordered with the first column being the most central camera (lowest mean distance to others according to their position; step 1) and then with increasing distances to the central camera (step 3).

To look for the maximum abundance possible given the camera×time abundance matrix, we propose an algorithm exploring "paths" for each time step. A "path" is defined for a given time step as a combination of time steps for all cameras except the central one, that checks the condition that these time steps are less distant than the time required by the studied species to move between cameras (given species' maximal speed and the distances between cameras). Hence, the number of paths increases with an increasing number of cameras, time steps and distances between cameras for a given species' maximal speed.

To reduce computation time, all paths are not computed. We first define for each time step T, a downward-moving window called the *small block* , which span is defined given species' maximal speed and the minimum distance between cameras (step 2) represented as [T; T + min(time-lag matrix)] (step 4a). For each time step T, the $SmaxSmallBlock_T$ is computed as the sum of the maximal abundance value of each camera in the small block of the time step T (step 4a). The maximum value of the $SmaxSmallBlock_T$, computed for all time steps, represents the minimum abundance estimate considering species speed and the minimum distance between the synchronised cameras (that is a conservative estimate as actually some cameras are more distant to each other than the minimum distance). Then, we define for each timestep a *Frame Of Possible* (FOP) (*s*tep 4b) gathering the abundance values which can be chosen given the distance between the central camera and the other cameras and species' maximal speed. FOP span is defined for each timestep T and each non-central camera j as $spanFOP_{t, camj} = [T - t_{cam1\text{-}camj} ; T+ t_{cam1\text{-}camj}]$ where $t_{cam1\text{-}camj}$ is the minimal time it takes for an individual of the studied species to go from the central camera (the first camera in the camera×time abundance matrix) and the camera j. We then compute for each timestep

T, $SmaxN_{FOP,\,T}$ defined as the sum of the maximal abundance value of each camera in the FOP (step 4b). $SmaxN_{FOP,\,T}$ represents, for each time step, the potential highest sum of abundance possible, given species speed and distance to the central camera. Only the time steps for which $SmaxN_{FOP,\,T} \geq \max(SmaxSmallBlock_T)$ are retained kept for subsequent steps (step 5) as they have the potential to yield an abundance estimate equal to or higher than the smallest abundance estimate achievable in the camera×time abundance matrix (step 5).

For each of the selected time steps, paths within the FOP are computed iteratively and conditionally to further reduce computation time (step 6). A path is defined by starting from the central camera and selecting a possible cell for the next camera based on the time-lag matrix (step 6).

For each path, $SmaxN_{path}$ is computed as the sum of the abundance values along the path (step 6). After each addition of a cell within a path, the maximum possible value of $SmaxN_{path}$, considering selected cells and those remaining for other cameras, is computed and compared to $SmaxN_{path}$ computed for previous time steps. If this potential maximum value is lower than an observed one, the path is abandoned.

Lastly, the *SmaxN* metric of the abundance matrix is computed as the maximal $SmaxN_{path}$ value found among all complete paths (*see* Box 1 - Step 7).

**Application of the *SmaxN* metric to a reef fishes case study**

The *SmaxN* framework was applied to estimate the abundance of six species of fishes over a coral reef in Mayotte (Western Indian Ocean) (*see* Supplementary Figure 1 and 2). The studied fringing reef was located in the Marine Protected Area of N'Gouja (-12.96° lat ; 45.08° long) and consisted of a mix of branching and star living corals, along with turf and detritic substrates, with an average depth of 3m. A network of nine GoPro Hero 5 (GoPro Inc, United States) in waterproof housings was set up on the 08th of November 2020. Six cameras

were paired, and each pair of cameras recorded in opposite directions. Cameras recorded high-definition videos (1920 by 1080 pixels at 25 frames per second) and were synchronised with a one second precision. Synchronisation was achieved using a watch passed in front of each camera, establishing a link between camera time and real time. Each camera has a 90° field-of-view and was mounted on a 20 cm high tripod. Immediately after the start of the recording, a 2m² quadrat was placed in front of each camera for 30 seconds and subsequently removed to avoid the disturbance of fish behaviour. This quadrat deployment allowed us to measure fish abundance over this standardised area (Longo et al. 2014) by marking the quadrat shape on the computer interface.

Distances between quadrats spanned between one meter (between two cameras mounted on the same tripod) and 110 meters for the most distant cameras (*see* Supp Info Table 1 for distances between cameras). The cameras recorded for about two hours during three time slots, and we retained only videos starting 45 minutes after the divers left the surveyed area and finishing 15 minutes before divers returned near the camera: this was done to reduce the impact of divers on fish detection. Overall, one hour of recording was thus used for the three time slots: 7:30 - 8:30 ; 11:30 - 12:30 ; 15:30 - 16:30. We studied six species representing five combinations of gregariousness and mobility: *Chaetodon trifasciatus, Gomphosus caeruleus, Parapercis hexophtalma, Parupeneus macronemus, Thalassoma hardwicke* and *Ctenochaetus striatus* (*see* Supplementary Table 2 for their traits).

For each species, the number of individuals present above the 2m² surveyed area was counted on each frame (1s precision). The maximal swimming speeds of the five species was estimated to be 0.5 $m.s^{-1}$, that is a conservative estimate since most of the critical speeds reported in Fulton (2007) for seven reef fish families were below this value.

*SmaxN, iSmaxN* and *maxN* metrics were computed for each camera's tuple going from one camera to nine cameras to test the effect of an increase in camera number on

abundance estimates, using the *SmaxN* R package. The *maxN* metric was computed as the maximal number of individuals on a given time step from a given camera over all cameras and timesteps. The *iSmaxN* and *SmaxN* metrics were computed as detailed above. The three metrics were also computed using the set of nine cameras for an increasing amount of time, ranging from ten minutes to one hour, to assess the effect of recording duration. These time sequences started at the beginning of each recording period and thus overlapped. Lastly, the three metrics were computed for a 1 $m.s^{-1}$ swimming speed to test for the effect of species speed on abundance estimates.

To test the effect of the number of cameras and recording duration on *SmaxN* and the difference between the *SmaxN* and the *maxN* metrics, we used Generalised Linear Models (GLMM) with Negative Binomial and Quasi Poisson distributions for the number of cameras and recording duration, respectively. Species identity and recordings were used as random effects. The three camera recording periods were used as replicates. GLMM were computed using the *glmm* R package and checked using the *performance* R package.

All data were analysed using R 4.1.2 and analysis are available on Github (https://github.com/CmlMagneville/SmaxNanalysis).

## Assessment

### Influence of using a network of synchronised cameras and species maximal speed to estimate the abundance of a given species

The *SmaxN* and the *maxN* metrics were significantly positively correlated (Spearman's correlation coefficient = 0.74, *p-value* < 0.05). The *SmaxN* metric was 1.3 to 4 times higher than the *maxN* metric (Figure 1A) with a mean *SmaxN*/*maxN* value of 2.58 ± 0.94 (mean ± sd) across all species and recordings. *SmaxN* was equal or up to three times higher than the

*iSmaxN* metric (Figure 1B) with a mean *SmaxN*/*iSmaxN* value of 1.77 ± 0.42 (mean ± sd) across all species and recordings. *SmaxN* and the *iSmaxN* were significantly correlated (Spearman's correlation coefficient = 0.82, *p-value* < 0.05).

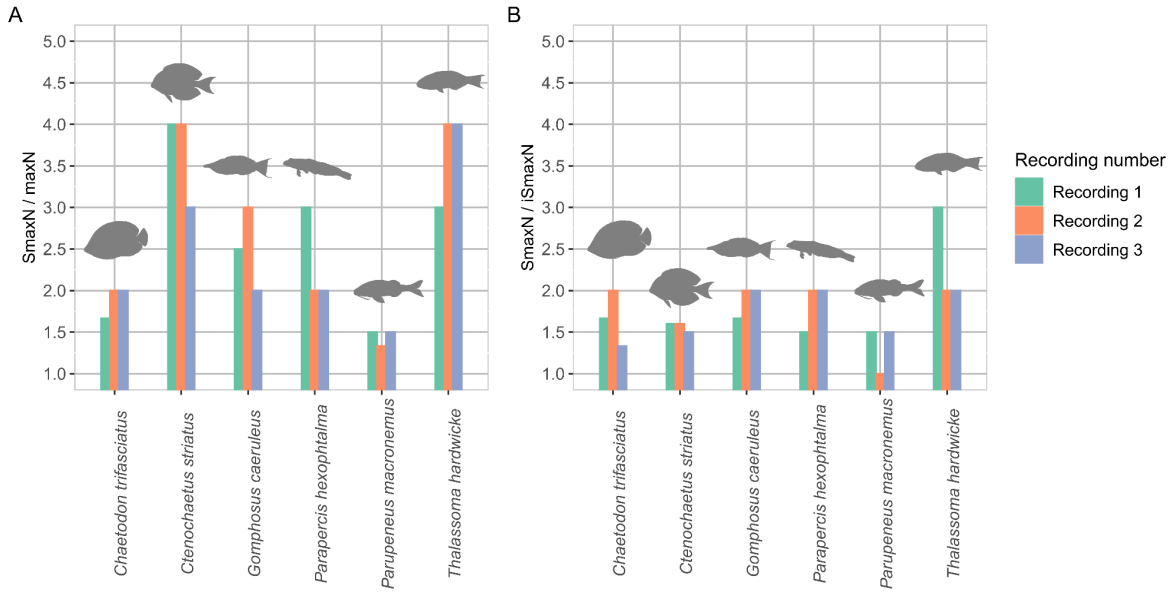
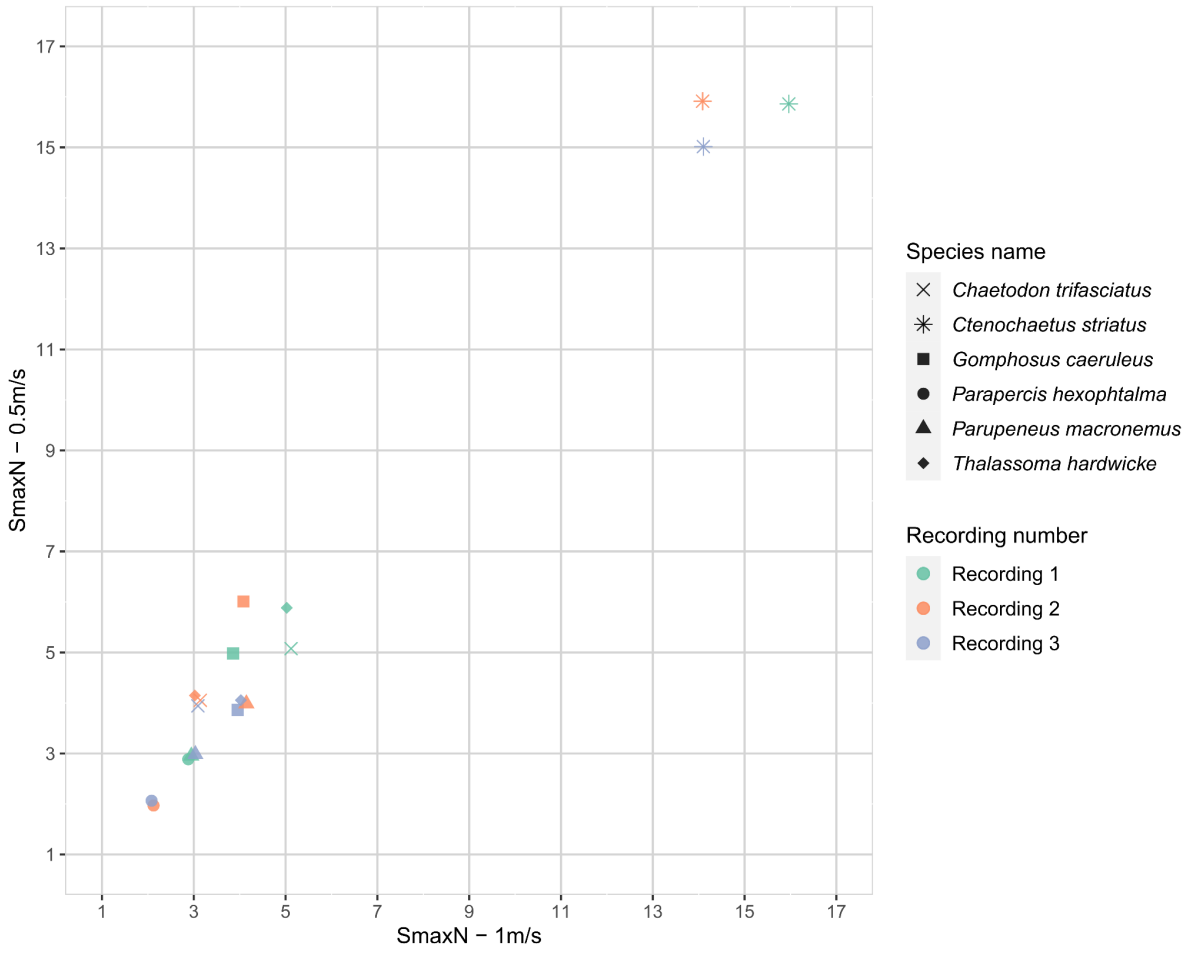



Figure 1: Ratios of *SmaxN*/*maxN* (A) and *SmaxN*/*iSmaxN* (B) for the set of six species across the three recordings (colours): *SmaxN* takes into account species maximal speed and distances between the synchronised cameras, *iSmaxN* takes only into account synchronised cameras and *maxN* is the maximal abundance retrieved on a single camera.

A significant positive correlation was found between the *SmaxN* metrics computed at different fish maximal speed (0.5m.s$^{-1}$ and 1m.s$^{-1}$) (Spearman's correlation coefficient = 0.92, *p-value* < 0.05). The *SmaxN* metric computed with the lowest fish speed was up to 1.5 higher than the *SmaxN* metric computed with the highest fish speed (Figure 2) across the five species (mean value of 1.12 and standard deviation of 0.16). However, *SmaxN* was not affected by fish speed in 10 out of the 18 combinations of Species×Recordings.




291

292 Figure 2: Variation of *SmaxN* for two different fish speeds (0.5m.s$^{-1}$ on the y-axis and

293 1 m.s$^{-1}$ on the x-axis) for the set of six species (shapes) and the three recordings (colours).

294

295

296 **Effect of the number of recording cameras on *SmaxN* and *maxN* metrics**

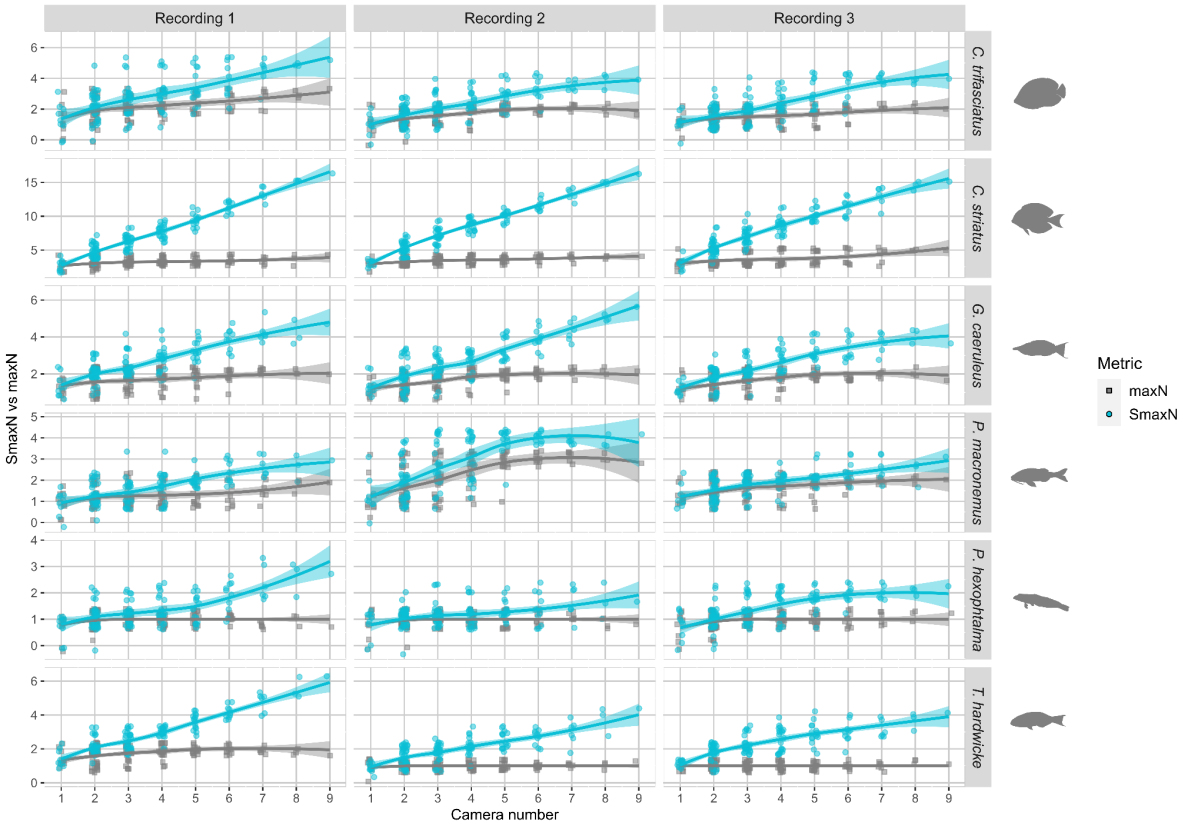
297 The *SmaxN* metric was significantly affected by the number of cameras with differences
298 between species (GLMM results - Supp. Table 3). *SmaxN* increased with the number of
299 cameras (Figure 3) with a mean increase of 72.83% from one camera to nine cameras over all
300 species and all recordings. For 78% of all species × recordings combinations, *SmaxN* stopped
301 increasing before nine cameras.

302 The difference between the *SmaxN* and *maxN* metrics was significant among camera numbers
303 and species (GLMM results - Supp. Table 4). For most species, it increased with the number
304 of cameras (Figure 3). The two metrics showed no difference for one camera, a mean
305 advantage of *SmaxN* over *maxN* of 39.85% for five cameras and a mean advantage of *SmaxN*
306 over *maxN* of 56.02 % for nine cameras.

307

308

309




310

311 Figure 3: *maxN* (grey) and *SmaxN* (blue) evolution across an increasing number of

312 cameras for six species. The lines are local polynomial regression fitting (2 degrees)

313 estimations surrounded by their confidence interval.

314

315 The minimal number of cameras needed to obtain the highest *SmaxN* value with a

316 one-hour recording was highly variable between species and between recordings for four out

317 of six species (Figure 4).

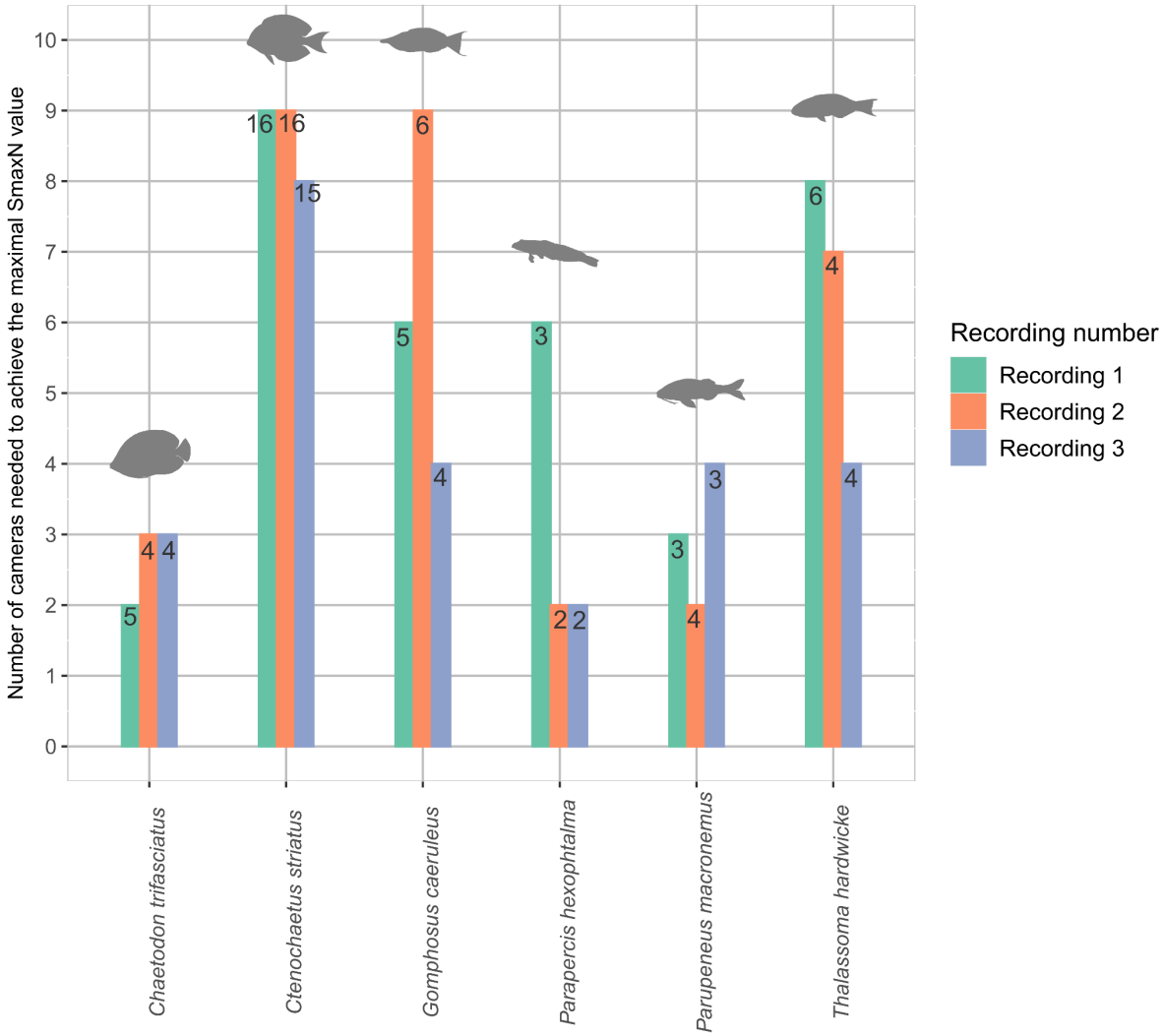



Figure 4: Minimal number of cameras needed to achieve the maximal *SmaxN* value (numbers in barplots) for each species and each recording (colours) with a recording time of one hour and a maximal number of nine cameras.

**Effect of the recording duration on the *SmaxN* and the *maxN* metrics**

The *SmaxN* metric was significantly different among recording durations and species (GLMM results - Supp. Table 5). The *SmaxN* metric increased with the recording duration (Figure 5) with a mean increase of 36.78% between 10 minutes and one hour over all species and all recordings.

The difference between the *SmaxN* and *maxN* metrics was significant among recording durations and species (GLMM results - Supp. Table 6). The deviation between *SmaxN* and *maxN* increased with the recording duration (Figure 5), showing a mean increase across species and recordings of 48.16% at 10 minutes, a mean increase of 55.58% at 30 minutes and a mean increase of 56.01% at one hour.

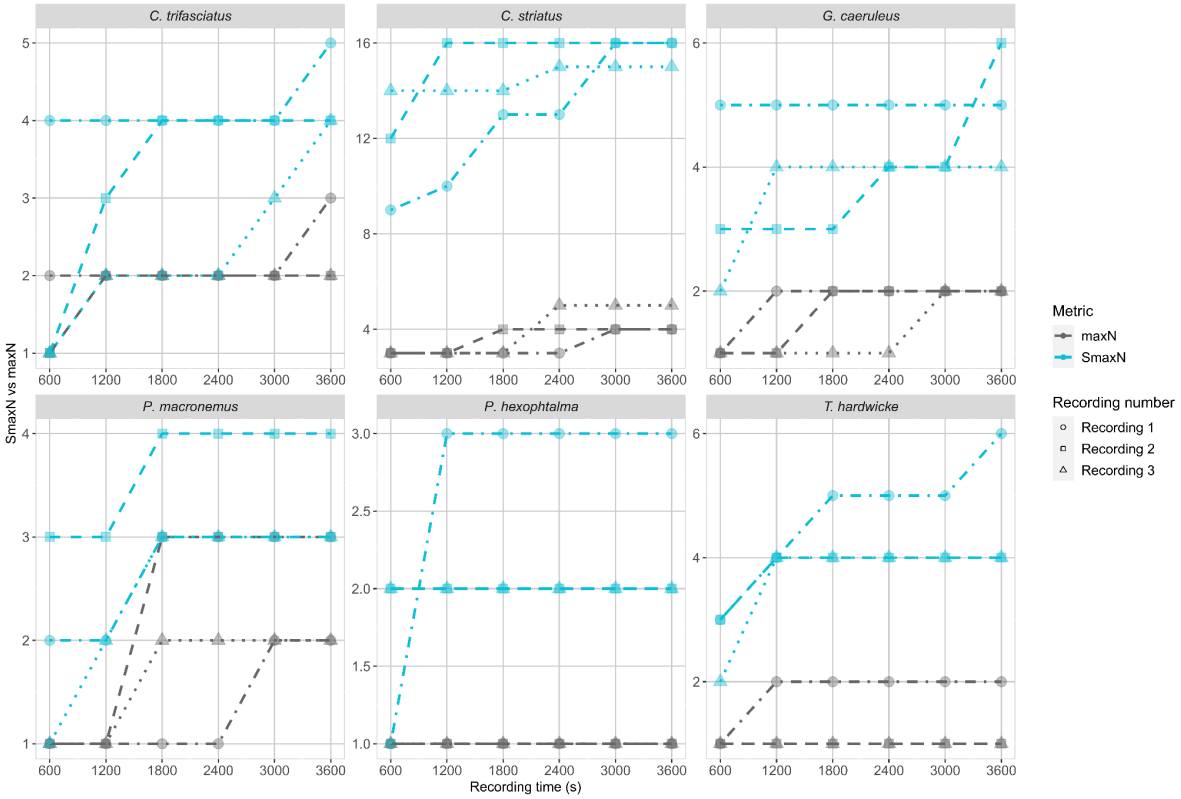
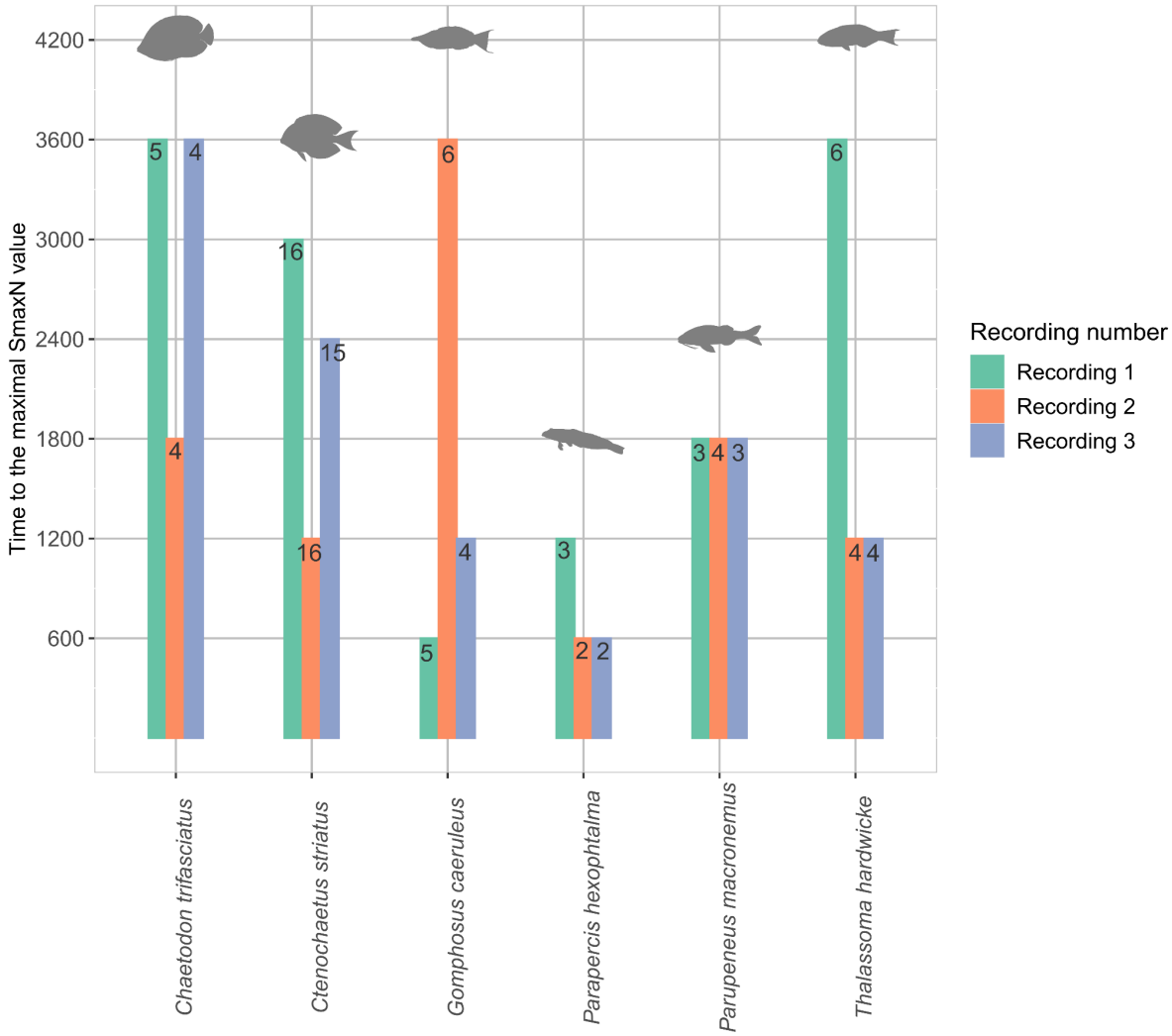



Figure 5: Abundance estimates of six fish species according to *maxN* (grey) and *SmaxN* (blue) indices across an increasing recording duration with nine cameras for three recordings (shapes and line types).

The minimal recording duration needed to obtain the highest *SmaxN* value with nine cameras was variable between species and among recordings for four out of six species (Figure 6).




340

341 Figure 6: Minimal recording duration to the maximal *SmaxN* value in seconds

342 (numbers in barplots) for each species and each recording (colours) with a network of nine

343 cameras and a maximal recording duration of one hour.

344

345

## Discussion

We developed a reproducible framework to quantify species abundance based on a network of synchronised cameras and an associated open-source algorithm. We then provided a proof-of-concept of this framework based on a network of nine remote underwater cameras deployed along Mayotte's fringing reef (Western Indian Ocean) for estimating the abundance of six fish species.

The *SmaxN* framework can be applied to all remote video-based surveys provided that cameras are synchronised, the distance between them is known, and their field-of-view do not overlap. Camera timestamping can be achieved through embedded softwares or by physically showing the same watch in front of each camera during recording. As for computing the distances between cameras, they could be measured either as geographical distances using GPS coordinates when cameras are positioned at significant distances and at the same depth, or directly measured on the field when cameras are positioned at small distances and/or at different depths. Cameras field-of-view should be selected as a trade-off between coverage and ability to identify species on video given its resolution: a large field-of-view allows to detect elusive species (e.g., large predators), yet it reduces apparent object size which could prevent the identification of small species. In addition, the filmed area can be standardised by placing a quadrat in front of each camera and subsequently removing them to minimise disturbance to the animals (Longo et al. 2014). Such area-based surveys allow computing abundance-based indices of biodiversity with a standardised protocol.

Our framework can also be applied with baited or unbaited stereo-cameras that allow for measuring individual sizes per unit area as well as their distance from the cameras, thus providing abundance estimates per class of size for each species. When designing a camera

network, it is important to consider the ecology of the studied species, their mobility and the micro habitat distribution. For instance, if the studied species is solitary and highly mobile, placing cameras in close proximity may result in a low probability of detecting multiple individuals. If the species is known to undertake diel migration -- changing habitats between day and night (Hitt et al. 2011; Courbin et al. 2019; Juby et al. 2021) --, placing cameras at the boundary between the two habitats could increase the detection of individuals in the studied area.

Another aspect to take into consideration is that computing *SmaxN* requires knowledge of the maximal speed of the species under study, which can be challenging to measure (Gilbert et al. 2021). In the work of Fulton (2007), captive fish individuals were exposed to an increasing water flow, and their maximum swimming speed was estimated as the current velocity when the fish became exhausted and stopped swimming. Data on maximal swimming speed are only available for a limited number of species (e.g., 117 coral reef fish species belonging to 10 families (Fulton 2007), 474 terrestrial and aquatic species (Hirt et al. 2017)). If data on maximal speeds are missing, we recommend using conservative estimates, such as the maximal speed of the fastest species within the same clade (e.g., family or order). Moreover, habitat characteristics that could affect travel time between cameras, and hence distort *Smax* computation, should be recorded during fieldwork. For instance, strong current could increase actual fish swimming speed. In such cases, we advocate for a conservative approach to avoid double counting. Therefore we suggest adding current speed to the maximal swimming speed of the fish when computing *SmaxN*.

In our study case, the network of synchronised cameras and the associated *SmaxN* metric yielded higher estimates of species abundance than *maxN* for the same level of sampling effort. In fact, the *SmaxN* metric counted up to four times more individuals than the

*maxN* metric, which does not account for speed and camera network. In fact, *maxN* does not differentiate between individuals recorded with slight temporal spacing on distant cameras, whereas *SmaxN* do confirm that these are different individuals. The gain of using *SmaxN* over *maxN* increased non-linearly with the number of cameras and recording durations. Indeed, for the same recording effort, the *maxN* metric only takes into account the highest abundance value across all cameras and time steps, whereas the *SmaxN* metric considers the sum of the highest abundance values across all cameras within a given time span. This difference in abundance estimation between *maxN* and *SmaxN* affects the estimated distribution of abundance among species assemblages. For instance, with a network of nine cameras recording for one hour, the *maxN* metric estimated that there were twice as many individuals of the butterflyfish *Chaetodon trifasciatus* compared to the surgeonfish *Ctenochaetus striatus,* while the *SmaxN* metric estimated that there were four times as many individuals of *C. trifasciatus* compared to individuals of *C. striatus*. Our metric will thus improve relative abundance estimates that are key to understand the drivers of assemblage diversity (e.g., relative strength of abiotic constraints and biotic interactions) as well as impact of species on ecosystem functioning (e.g., control of trophic network and nutrient fluxes).

The *SmaxN* estimates increased with the number of cameras in the network. Using a network of nine cameras captured over three times as many individuals as using a single camera for the same total recording time of one hour. Such a marked increase in abundance estimates with increasing cameras' total field of view has been documented in both terrestrial and marine ecosystems (O'Connor et al. 2017; Campbell et al. 2018), and is expected because the addition of view points increases the detection probability and thus the abundance estimation. The impact of increasing the number of cameras varied among species, with some species exhibiting a more pronounced effect (*Ctenochaetus striatus*, *Gomphosus caeruleus*, *Thalassoma hardwicke*) compared to others (*Parapercis hexophtalma*, *Parupeneus macronemus*). In fact, *C. striatus*, *G. caeruleus* and *T. hardwicke* were recorded mostly swimming across camera field of views while *P. hexophtalma* and *P. macronemus* were frequently observed foraging or remaining stationary in front of the cameras. Some

426 *SmaxN* values computed with one or two cameras were equal to zero, underscoring the
427 stochastic nature of species detection when observation effort is low. Therefore, using a
428 network of cameras optimises the detection probability by covering different habitats
429 (Verberk 2011). Moreover, since the *SmaxN* estimate varied across the number of cameras,
430 when comparing multiple sites with the *SmaxN*, it is advisable to use the same number of
431 cameras in each environment, with comparable distances between them, and in front of the
432 same type of microhabitats.
433
434 As for recording duration, the performance of *SmaxN* improved with longer recording
435 durations, although the magnitude of this improvement varied among species. A recording
436 duration of 10 minutes captured on average about four individuals, whereas a recording
437 duration of one hour captured on average about six individuals. In fact, recording for a longer
438 period helps to detect more individuals of the studied species (Campbell et al. 2015). Yet, the
439 six species studied in this environment were found to be common, with an average presence
440 ranging from 12% to 98% of the recording time. It would thus be informative to test whether
441 the abundance of rare species also increases with recording duration. As recording durations
442 increase, the *SmaxN* may also increase. Therefore, it is essential to maintain consistent
443 recording durations when assessing the abundance of a species across different environments.
444 If subsampling is employed, it must be ensured that subsampled recordings share the same
445 overall duration across various environments.
446
447 Overall, only one third of the species×recordings combinations reached the maximal
448 *SmaxN* value after more than 30 minutes (1800s) of recording and about half of the
449 species×recordings combinations reached the maximal *SmaxN* value with more than four
450 cameras. Because frame or video analysis is time-consuming, we recommend to set up a
451 network of many (i.e. more than five) cameras filming for a short amount of time (i.e. about
452 30 minutes) on a given habitat, rather than using a single camera or a small number of
453 cameras filming for an extended duration to estimate species abundance using the *SmaxN*
454 framework.

455

456 Establishing a network of cameras following *SmaxN* requirements is not demanding,
457 especially in shallow environments. In our underwater case study, it took two divers less than
458 15 minutes to set up the network of nine cameras. A similar camera network could be
459 employed with baited cameras dropped from a boat, either to record pelagic or benthic
460 habitats, and which are usually set hundreds of meters from each other (Whitmarsch et al.
461 2017). The *SmaxN* framework can be applied to a continuous recording of biodiversity with
462 camera videos or to a punctual recording of biodiversity with camera traps. We here provide
463 an open-source R package to ensure a reproducible use of this framework
464 (https://github.com/CmlMagneville/SmaxN).
465

466 The use of a punctual recording is common in terrestrial environments, where camera
467 traps are often used in numbers exceeding 50 (78 cameras on each site on average as reported
468 in a compilation of about 100 papers by Steenweg et al. (2017)), while in the marine
469 environment the use of continuous recording has become increasingly popular (Tebbett et al.
470 2020; Marques et al. 2021; Magneville et al. 2022). In both cases, counting individuals on
471 video frames is a time-consuming process. This process could be sped up by using annotation
472 software such as the Behavioral Observation Research Interactive Software (BORIS) (Friard
473 and Gamba 2016). In this study, the occurrence of individuals of six species were annotated
474 on 27 hours of videos, which took about 150 hours on a similar annotation software.
475 Reducing the annotation frequency could help to decrease the annotation time. However, to
476 prevent double-counting individuals on different cameras, it is crucial to keep a frequency
477 higher than the ratio between minimal distance between cameras and fish maximal speed. If
478 cameras are positioned at a distance (e.g., with BRUVs for large marine predators), a lower
479 frequency (e.g., every 5 seconds) could be considered. However, this may result in missed

480 occurrences of fast-moving individuals, particularly if the field of view is too narrow. In our
481 study, we could have reduced the annotation frequency to two images per second because
482 it takes two seconds for an individual swimming at $0.5 m.s^{-1}$ to pass between our closest
483 cameras (1 meter). The rise of deep-learning algorithms to automatically detect individuals of
484 some species would thus be a milestone to speed up processing video to get abundance
485 through time data (Ditria et al. 2020).
486
487
488
489
490
491
492
493
494
495
496
497
498
499

500 **References**

501


502 Alexander, R. McN., V. A. Langman, and A. S. Jayes. 1977. Fast locomotion of some
503 African ungulates. J. Zool. **183**: 291–300. doi:10.1111/j.1469-7998.1977.tb04188.x

504 Birt, M. J., E. S. Harvey, and T. J. Langlois. 2012. Within and between day variability in
505 temperate reef fish assemblages: Learned response to baited video. J. Exp. Mar. Biol.
506 Ecol. **416–417**: 92–100. doi:10.1016/j.jembe.2012.02.011

507 Boulangeat, I., D. Gravel, and W. Thuiller. 2012. Accounting for dispersal and biotic
508 interactions to disentangle the drivers of species distributions and their abundances.
509 Ecol. Lett. **15**: 584–593. doi:10.1111/j.1461-0248.2012.01772.x

510 Brock, V. E. 1954. A Preliminary Report on a Method of Estimating Reef Fish Populations.
511 J. Wildl. Manag. **18**: 297–308. doi:10.2307/3797016

512 Campbell, M. D., A. G. Pollack, C. T. Gledhill, T. S. Switzer, and D. A. DeVries. 2015.
513 Comparison of relative abundance indices calculated from two methods of generating
514 video count data. Fish. Res. **170**: 125–133. doi:10.1016/j.fishres.2015.05.011

515 Campbell, M. D., J. Salisbury, R. Caillouet, W. B. Driggers, and J. Kilfoil. 2018. Camera
516 field-of-view and fish abundance estimation: A comparison of individual-based model
517 output and empirical data. J. Exp. Mar. Biol. Ecol. **501**: 46–53.
518 doi:10.1016/j.jembe.2018.01.004

519 Cappo, M., E. Harvey, H. Malcolm, and P. Speare. 2003. Potential of video techniques to
520 monitor diversity, abundance and size of fish in studies of Marine Protected Areas. p
521 455–464. In: Beumer JP, Grant A and Smith DC (eds) Aquatic Protected Areas - what
522 works best and how do we know? World Congress on Aquatic Protected Areas
523 proceedings, Cairns, Australia, August 2002. Australian Society of Fish Biology.

524 Cinner, J. E., C. Huchery, M. A. MacNeil, and others. 2016. Bright spots among the world's

coral reefs. Nature **535**: 416–419. doi:10.1038/nature18607

Courbin, N., A. J. Loveridge, H. Fritz, D. W. Macdonald, R. Patin, M. Valeix, and S. Chamaillé-Jammes. 2019. Zebra diel migrations reduce encounter risk with lions at night. J. Anim. Ecol. **88**: 92–101. doi:10.1111/1365-2656.12910

Dearden, P., M. Theberge, and M. Yasué. 2010. Using underwater cameras to assess the effects of snorkeler and SCUBA diver presence on coral reef fish abundance, family richness, and species composition. Environ. Monit. Assess. **163**: 531–538. doi:10.1007/s10661-009-0855-3

Dickens, L. C., C. H. R. Goatley, J. K. Tanner, and D. R. Bellwood. 2011. Quantifying Relative Diver Effects in Underwater Visual Censuses. PLOS ONE **6**: e18965. doi:10.1371/journal.pone.0018965

Ditria, E. M., S. Lopez-Marcano, M. Sievers, E. L. Jinks, C. J. Brown, and R. M. Connolly. 2020. Automating the Analysis of Fish Abundance Using Object Detection: Optimizing Animal Ecology With Deep Learning. Front. Mar. Sci. **7**.

Ellis, D. M., and E. E. De Martini. 1995. Evaluation of a video camera technique for indexing the abundances of juvenile pink snapper, Pristipomoides filamentosus, and other Hawaiian insular shelf fishes. Fish. Bull. **93**: 67–77.

Friard, O., and M. Gamba. 2016. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. Methods in Ecology and Evolution 7: 1325–1330. doi:10.1111/2041-210X.12584

Fulton, C. 2007. Swimming speed performance in coral reef fishes: Field validations reveal distinct functional groups. Coral Reefs **26**: 217–228. doi:10.1007/s00338-007-0195-0

Garcia, G. S., M. S. Dias, and G. O. Longo. 2021. Trade-off between number and length of remote videos for rapid assessments of reef fish assemblages. J. Fish Biol. **99**: 896–

904. doi:10.1111/jfb.14776

Gilbert, N. A., J. D. J. Clare, J. L. Stenglein, and B. Zuckerberg. 2021. Abundance estimation of unmarked animals based on camera-trap data. Conserv. Biol. **35**: 88–100. doi:10.1111/cobi.13517

Harmelin-Vivien, M. L., J. G. Harmelin, C. Chauvet, and others. 1985. Evaluation visuelle des peuplements et populations de poissons méthodes et problèmes. Rev. Ecol. Terre Vie **40**: 467–539.

Harvey, E. S., M. Cappo, J. J. Butler, N. Hall, and G. A. Kendrick. 2007. Bait attraction affects the performance of remote underwater video stations in assessment of demersal fish community structure. Mar. Ecol. Prog. Ser. **350**: 245–254. doi:10.3354/meps07192

Hirt, M. R., W. Jetz, B. C. Rall, and U. Brose. 2017. A general scaling law reveals why the largest animals are not the fastest. Nat. Ecol. Evol. **1**: 1116–1122. doi:10.1038/s41559-017-0241-4

Hitt, S., S. Pittman, and R. Nemeth. 2011. Diel movements of fishes linked to benthic seascape structure in a Caribbean coral reef ecosystem. Mar. Ecol. Prog. Ser. **427**: 275–291. doi:10.3354/meps09093

Husak, J. F., S. F. Fox, M. B. Lovern, and R. A. V. D. Bussche. 2006. FASTER LIZARDS SIRE MORE OFFSPRING: SEXUAL SELECTION ON WHOLE-ANIMAL PERFORMANCE. Evolution **60**: 2122–2130. doi:10.1111/j.0014-3820.2006.tb01849.x

Juby, R., A. Bernard, and A. Götz. 2021. Day/night patterns of habitat use by dogfish sharks (Squalidae) at photic and subphotic warm-temperate reefs: evidence for diel movements and size- and sex-segregation. Afr. J. Mar. Sci. **43**: 325–336. doi:10.2989/1814232X.2021.1951839

575 Kilfoil, J., A. Wirsing, M. Campbell, J. Kiszka, K. Gastrich, M. Heithaus, Y. Zhang, and M.
576 Bond. 2017. Baited Remote Underwater Video surveys undercount sharks at high
577 densities: insights from full-spherical camera technologies. Mar. Ecol. Prog. Ser. **585**:
578 113–121. doi:10.3354/meps12395
579 Labrosse, P., M. Kulbicki, and J. Ferraris. 2002. Underwater visual fish census survey -
580 proper use and implementation. Noumea, New Caledonia: Secretariat of the Pacific
581 Community. REAT: Reef Resources Assessment Tools, vi, 54 p.
582 https://purl.org/spc/digilib/doc/fs6ca
583 Langlois, T. J., E. S. Harvey, and J. J. Meeuwig. 2012. Strong direct and inconsistent indirect
584 effects of fishing found using stereo-video: Testing indicators from fisheries closures.
585 Ecological Indicators 23: 524–534. doi:10.1016/j.ecolind.2012.04.030
586
587 Layne, J. N., and A. H. Benton. 1954. Some Speeds of Small Mammals. J. Mammal. **35**:
588 103–104. doi:10.2307/1376079
589 Letessier, T., R. Proud, J. Meeuwig, M. Cox, P. Hosegood, and A. Brierley. 2021. Estimating
590 pelagic fish biomass in a tropical seascape using echosounding and baited stereo-
591 videography.doi:10.1007/s10021-021-00723-8
592 Liu, Y., W. Qi, D. He, Y. Xiang, J. Liu, H. Huang, M. Chen, and J. Tao. 2021. Soil resource
593 availability is much more important than soil resource heterogeneity in determining
594 the species diversity and abundance of karst plant communities. Ecol. Evol. **11**:
595 16680–16692. doi:10.1002/ece3.8285
596 Lopez-Marcano, S., E. L. Jinks, C. A. Buelow, C. J. Brown, D. Wang, B. Kusy, E. M. Ditria,
597 and R. M. Connolly. 2021. Automatic detection of fish and tracking of movement for
598 ecology. Ecol. Evol. **11**: 8254–8263. doi:10.1002/ece3.7656
599 Magneville, C., M.-L. Leréec-le-Bricquir, T. Dailianis, G. Skouradakis, T. Claverie, and S.

Villéger. 2022. Long-duration remote underwater videos reveal that grazing by fishes is highly variable through time and dominated by non-indigenous species. Remote Sens. Ecol. Evol. **In Press**.

Mallet, D., and D. Pelletier. 2014. Underwater video techniques for observing coastal marine biodiversity: A review of sixty years of publications (1952–2012). Fish. Res. **154**: 44–62. doi:10.1016/j.fishres.2014.01.019

Marques, V., P. Castagné, A. Polanco Fernández, and others. 2021. Use of environmental DNA in assessment of fish functional and phylogenetic diversity. Conserv. Biol. **35**: 1944–1956. doi:10.1111/cobi.13802

Murphy, H. M., and G. P. Jenkins. 2010. Observational methods used in marine spatial monitoring of fishes and associated habitats: a review. Mar. Freshw. Res. 61(9), 1023–1028. https://doi.org/10.1071/MF09240O'Connor, K. M., L. R. Nathan, M. R. Liberati, M. W. Tingley, J. C. Vokoun, and T. A. G. Rittenhouse. 2017. Camera trap arrays improve detection probability of wildlife: Investigating study design considerations using an empirical dataset. PLOS ONE **12**: e0175684. doi:10.1371/journal.pone.0175684

Radford, C., A. Jeffs, C. Tindle, R. Cole, and J. Montgomery. 2005. Bubbled waters: The noise generated by underwater breathing apparatus. Mar. Freshw. Behav. Physiol. **38**: 259–267. doi:10.1080/10236240500333908

Satterfield, D. R., T. Claverie, and P. C. Wainwright. 2023. Body shape and mode of propulsion do not constrain routine swimming in coral reef fishes. Functional Ecology 37: 343–357. doi:10.1111/1365-2435.14227

Schobernd, Z., N. Bacheler, and P. Conn. 2014. Examining the utility of alternative video monitoring metrics for indexing reef fish abundance. Can. J. Fish. Aquat. Sci. **71**. doi:10.1139/cjfas-2013-0086

625 Steenweg, R., M. Hebblewhite, R. Kays, and others. 2017. Scaling-up camera traps:
626 monitoring the planet's biodiversity with networks of remote sensors. Front. Ecol.
627 Environ. **15**: 26–34. doi:10.1002/fee.1448
628 Tebbett, S., A. Hoey, M. Depczynski, S. Wismer, and D. Bellwood. 2020. Macroalgae
629 removal on coral reefs: realised ecosystem functions transcend biogeographic
630 locations. Coral Reefs **39**. doi:10.1007/s00338-019-01874-w
631 Verberk, W. 2011. Explaining General Patterns in Species Abundance and Distributions. Nat.
632 Educ. Knowledge **3**: 38.
633 Whitmarsh, S. K., P. G. Fairweather, and C. Huveneers. 2017. What is Big BRUVver up to?
634 Methods and uses of baited underwater video. Rev. Fish Biol. Fish. **27**: 53–73.
635 doi:10.1007/s11160-016-9450-1
636 Widmer, L., E. Heule, M. Colombo, A. Rueegg, A. Indermaur, F. Ronco, and W. Salzburger.
637 2019. Point-Combination Transect (PCT): Incorporation of small underwater cameras
638 to study fish communities. Methods Ecol. Evol. **10**: 891–901. doi:10.1111/2041-
639 210X.13163
640 Zarco-Perello, S., and S. Enríquez. 2019. Remote underwater video reveals higher fish
641 diversity and abundance in seagrass meadows, and habitat differences in trophic
642 interactions. Sci. Rep. **9**: 6596. doi:10.1038/s41598-019-43037-5
643
644
645
646
647
648
649

**Acknowledgments:**

This project was funded by the University of Montpellier (I-site MUSE program, BUBOT project) and European Union (H2020 program, MACOBIOS project).

**Conflict of Interest:**

The authors have no conflict of interest to declare.

**Authors Contribution Statement:**

C. Magneville, S. Villéger, V. Fleuré, N. Loiseau and T. Claverie conceived the ideas and designed the metric; CB, CM, SV, VF, NL and TC collected the data; CM, VF and SV conceived the algorithm on which the metric is computed; CB and CM tested the metric using collected data; CM and SV led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

**Data availability:**

Data and scripts are available on Github (https://github.com/CmlMagneville/SmaxNanalysis).

The SmaxN package is available on Github (https://github.com/CmlMagneville/SmaxN.