## Conditional spatially explicit simulations of strongly 0-inflated data to estimate biodiversity indices

N. Bez<sup>a</sup>, V. Suntov<sup>b</sup>, B. Mérigot<sup>c</sup>, J.M. Fromentin<sup>b</sup>, and M. Woillez<sup>c</sup>

<sup>a</sup> IRD UMR 212 EME (IRD, Ifremer, UM2) Sète, France nicolas.bez@ird.fr

<sup>b</sup> Ifremer UMR 212 EME (IRD, Ifremer, UM2) Sète, France victoria.suntov@partenaire-exterieur.ifremer.fr jean.marc.fromentin@ifremer.fr

> <sup>c</sup> Université Montpellier 2 UMR 212 EME (IRD, Ifremer, UM2) Sète, France bastien.merigot@univ-montp2.fr

> > <sup>d</sup> Ifremer Brest, France Mathieu.Woillez@ifremer.fr

**Keywords:** Indicators of biodiversity, spatial ecology, monitoring of biodiversity, geostatistical simulations.

Abstract: Upscaling diversity indices is recognized as a problem as the diversity of a large region is not the average of the diversities of its sub-divisions. This led to the definition of the concept of  $\beta$ -diversity which corresponds to the difference between the mean of local  $\alpha$ diversities and the global  $\gamma$ -diversity. However, its expression is not known in practice which lets the problem unsolved unless some relationships between  $\alpha$ - and  $\gamma$ -diversities are postulated. The non additivity of diversity indices makes it necessary to estimate them via indirect approaches where each species' distribution or presence is estimated before computing the index. Moreover, being non linear transformations of the species densities, these indices must be estimated on the basis of conditional simulations of species densities (and not on the basis of kriging maps). However, in marine ecosystems, the vast majority of species are rare or extremely rare (50% of the 139 species distributions get more than 95% of 0 in the present case study). This paper presents a general framework based on conditional simulation spatially consistent that allows producing accurate maps of diversity. A particular focus is devoted to the simulations of strongly 0-inflated distributions (Gibbs sampler with the variogram of the hidden Gaussian variable). Simulated data are used for demonstration purposes, but a case study is also presented where the Rao entropy is mapped over the Gulf of Lion (France) on the basis of scientific trawl survey data.