ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

Journal of the Royal Statistical Society
Statistics in Society
Series A

J. R. Statist. Soc. A (2020)

On the estimation of population size from a dependent triple-record system

Kiranmoy Chatterjee

Indian Statistical Institute, and Bidhannagar College, Kolkata, India

and Prajamitra Bhuyan

Imperial College London, UK

[Received April 2018. Final revision April 2019]

Summary. Population size estimation based on a capture–recapture experiment under a triple-record system is an interesting problem in various fields including epidemiology and population studies. In many real life scenarios, there is inherent dependence between capture and recapture attempts. We propose a novel model that successfully incorporates the possible dependence and the associated parameters have nice interpretations. We provide estimation methodology for the population size and the associated model parameters based on the maximum likelihood method. The model proposed is applied to analyse real data sets from public health and census coverage evaluation studies. The performance of the estimate proposed is evaluated through extensive simulation study and the results are compared with existing competitors. The results exhibit superiority of the model over the existing competitors both in real data analysis and in a simulation study.

Keywords: Behavioural dependence; Disease surveillance; Maximum likelihood; Time-ordered capture; Trivariate Bernoulli model

1. Introduction

Estimation of population size or the number of vital events that have occurred, during a given time span, is a relevant statistical problem in various scientific disciplines including epidemiology, population studies and life sciences. Federal agencies are generally interested in such estimates for planning and policy formulation. In general, a census or any registration system often fails to capture all the individuals and that leads to undercoverage of the population under consideration. However, in some instances, duplicate records or members outside the target population are included in the census or any other registers because of erroneous enumeration. This issue is known as overcoverage, and it is a common practice to identify and remove the erroneous inclusions through administrative follow-up actions (Chipperfield *et al.*, 2017) or to adjust the census data on the basis of an estimate of the overcoverage rate (Zhang, 2015). In this paper, we focus only on the issues that are related to the commonly encountered problem of undercoverage, assuming that the available data are free from any erroneous inclusion. To reduce the undercoverage error, information from more than one attempt needs to be considered. The data that are obtained from various sources are summarized by matching the lists of captured individuals and analysed to obtain an estimate of the unknown population size (Rastogi and

Address for correspondence: Prajamitra Bhuyan, Department of Mathematics, Imperial College London, South Kensington Campus, London, SW7 2AZ, UK.
E-mail: bhuyan.prajamitra@gmail.com

 0964–1998/20/183000

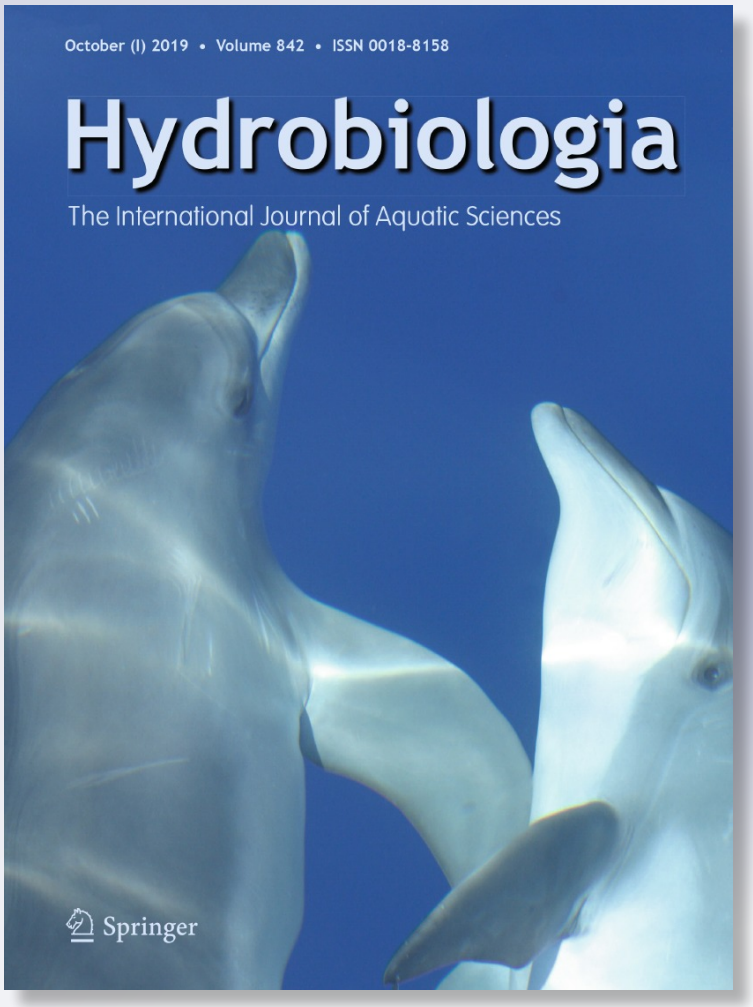
Effects of nutrient limitation, salinity increase, and associated stressors on mangrove forest cover, structure, and zonation across Indian Sundarbans

Rajojit Chowdhury, Tapan Sutradhar, Mst. Momtaj Begam, Chandan Mukherjee, Kiranmoy Chatterjee, Sandip Kumar Basak & Krishna Ray

Hydrobiologia
The International Journal of Aquatic Sciences

ISSN 0018-8158
Volume 842
Number 1

Hydrobiologia (2019) 842:191-217
DOI 10.1007/s10750-019-04036-9



Springer