



**HAL**  
open science

## **U-CARE: Utilities for performing goodness of fit tests and manipulating CAPture-REcapture data**

Rémi Choquet, Jean-Dominique Lebreton, Anne-Marie Reboulet, Roger Pradel, Olivier Gimenez

► **To cite this version:**

Rémi Choquet, Jean-Dominique Lebreton, Anne-Marie Reboulet, Roger Pradel, Olivier Gimenez.  
U-CARE: Utilities for performing goodness of fit tests and manipulating CAPture-REcapture data.  
Ecography, 2009, 32 (6), pp.1071-1074. 10.1111/j.1600-0587.2009.05968.x . hal-03515210

**HAL Id: hal-03515210**

**<https://hal.science/hal-03515210v1>**

Submitted on 6 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# U-CARE: Utilities for performing goodness of fit tests and manipulating Capture–REcapture data

Rémi CHOQUET, Jean-Dominique LEBRETON, Olivier GIMENEZ, Anne-Marie REBOULET and Roger PRADEL.

Corresponding author : R. Choquet ([remi.choquet@cefe.cnrs.fr](mailto:remi.choquet@cefe.cnrs.fr)), Centre d'Ecologie Fonctionnelle et Evolutive - CNRS - UMR 5175 - 1919 Route de Mende - F34293 Montpellier cedex 5

Abstract: In any statistical analysis, assessing the goodness of fit of a model to the data is crucial to avoid drawing incorrect conclusions. U-CARE is a computer application that deals with the mandatory first steps of the analyses of capture-recapture data: the preparation of the data set and the assessment of the fit of a general model (Cormack-Jolly-Seber and variants for single-state data; Jolly-Move and variants for multi-state data). U-CARE implements the current state of the art in goodness-of-fit testing by incorporating components aimed at detecting the most likely departures from assumptions (Pradel et al. 2003), (Pradel et al. 2005). It is a free and stand-alone application for Windows.

Keywords: Arnason-Schwarz model, Contingency tables, Directional tests, Memory model, Overdispersion coefficient, Test of multinomial mixture, Test WBWA, Test 3G.SR, Test M.ITEC, Transient, Trap-dependence.

42  
43  
44  
45  
46  
47  
48  
49

50 The estimation and comparison of demographic rates in animal and plant populations (Lebreton et  
51 al. 1992), (Gregg and Kery 2006) is currently based on following marked individuals over time and  
52 possibly space and analyzing the resulting data with adequate statistical models. Such models, in  
53 which individuals may move among states (such as geographical sites or breeding statuses) over  
54 discrete time periods, must account for the non-exhaustive detection of individuals. These so-called  
55 capture-recapture models are still evolving (Pradel et al. 2005). In their current form, they are based  
56 on two key assumptions.

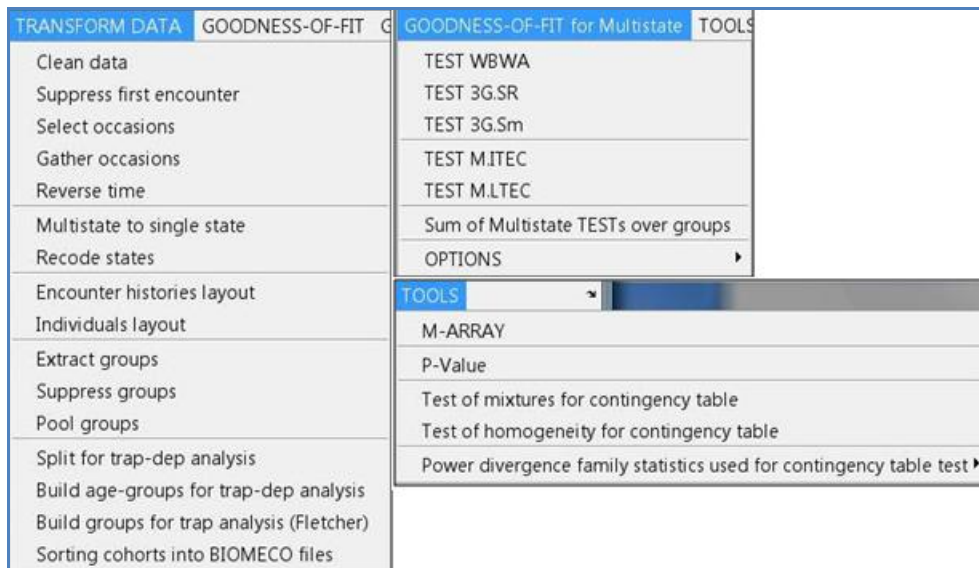
- 57 1) The individuals are assumed to be independent,
- 58 2) Within the history of each individual, the successive pairs of each release associated with its  
59 subsequent reencounter (if any) are assumed to be independent (Burnham 1991).

60 While fitting multi-state capture-recapture models can now be carried out without much difficulty  
61 with a variety of computer software applications (M-SURGE (Choquet et al. 2004 #340), E-SURGE  
62 (Rouan et al., Choquet et al. 2008) or MARK (White and Burnham 1999) for a frequentist approach,  
63 MARK for Bayesian inference), the examination of the above mentioned assumptions can be done in  
64 an optimal way (Pradel et al. 2003) only with program U-CARE. And yet this step is critical. For  
65 instance, if a major structural effect such as the presence of transient individuals on the study site is  
66 overlooked, survival will be severely underestimated. More generally, spurious effects will often be  
67 detected if no model fits the data. This is best seen when considering how model selection proceeds.  
68 In the frequentist paradigm, model selection is generally based on the Akaike information criterion  
69 (AIC, see (Burnham and Anderson 2002)). The AIC is equal to the deviance plus two times the number  
70 of estimable parameters. In presence of lack of fit, the deviance tends to be inflated, thus leading to  
71 the selection of over-parameterized models and potentially to erroneous biological conclusions.  
72 Moreover, the estimate of precision of the Maximum Likelihood Estimators (MLE) will be overly  
73 optimistic. The consequences of lack-of-fit are thus too deleterious to be ignored.

74 A preliminary assessment of Goodness-Of-Fit (GOF) is thus a crucial prerequisite, just as in any  
75 statistical analysis (D'Agostino and Stephens 1986). Unlike in regression, residuals are not easily  
76 available in CR to check the validity of a model. Furthermore, the omnibus approach to goodness-of-  
77 fit testing that consists of comparing expected vs. observed sufficient statistics is impractical due to  
78 the sparseness of the data. A specific approach was developed (reviewed in Pradel et al. 2003). By  
79 making this approach available in an easy-to-use software application, we aim at encouraging  
80 practitioners to assess the validity of their assumptions. Sound model selection, i.e. preceded by  
81 appropriate GOF assessment, is indeed becoming more and more common in the literature (e.g.  
82 (Henaux et al. 2007), (Jenouvrier et al. 2008)) but it is still far from being systematic. We describe  
83 here the main features of program U-CARE. More details on tests and tools can be found in the user's  
84 manual (Choquet et al. 2005).

85 U-CARE (<<http://purl.oclc.org/NET/U-CARE>>) is a free, easy-to-use, stand-alone, menu-driven  
 86 computer application for Windows, with a set of options for managing data and GOF tests  
 87 capabilities for both single-state and multi-state capture-recapture models (see Figure 1). The two  
 88 kinds of test are presented as the sum of components examining different aspects of the data  
 89 through a range of contingency tables. This structure guides the choice of an appropriate model. For  
 90 instance, if the only significant subcomponent of the multistate goodness-of-fit test is TEST 3G.SR, a  
 91 two age-class structure for survival should be used (Pradel et al. 1995), while, if this is TEST  
 92 WhereBeforeWhereAfter (WBWA), a memory model is recommended (Pradel et al. 2003). Due to  
 93 historical reasons, biological relevance is particularly furthered for single-state data through the  
 94 computation of directional tests. We briefly examine hereafter the following four aspects of data  
 95 analysis for which U-CARE can be very useful:

- 96 1) How to prepare the data?
- 97 2) How to detect overdispersion?
- 98 3) How to correct for overdispersion?
- 99 4) How to conduct specific tests?



100  
 101 **Figure 1 : U-CARE has 5 menus showing utilities for input-output of capture-recapture files (FILE menu), manipulation of**  
 102 **data (TRANSFORM DATA menu), goodness-of-fit testing for single-state (GOODNESS OF FIT menu) and multi-state**  
 103 **(GOODNESS OF FIT FOR MULTI-STATE menu) data and various additional tools (TOOLS menu). Three of them are shown.**

- 104 1) How to prepare the data?

105 First, the menu TRANSFORM DATA offers some tools for selecting a data subset, such as a particular  
 106 range of years or some groups (e.g. males or females, 'ringed as young' or 'ringed as adults'), and for  
 107 recoding data, such as pooling states, groups or years. Among other features, the menu FILE makes it  
 108 possible to convert data between two main formats (Biomeco (Lebreton and Roux 1989) and Mark  
 109 (White and Burnham 1999)).

- 110 2) How to detect overdispersion?

111 When the data are ready, the next step should be the assessment of the fit of a general model to the  
 112 data at hand. For single-state data, the classical model is the model with time dependent only  
 113 parameters (Cormack-Jolly-Seber) (CJS); for multistate data, this is a time and state dependent model  
 114 (Arnason-Schwarz) (AS). For both of them, the probability of the encounter history of each individual

115 is calculated conditional on its first capture. Optimal goodness of fit tests of the assumptions  
 116 inherent in the CJS or the JMV model, this last being a slight generalization of the AS model, are then  
 117 derived based on the classical partitioning, according to sufficient statistics T, of the likelihood  
 118  $P(\text{data}/\text{parameters})$ :

119 
$$P(\text{data}/\text{parameters})=P(\text{data}/T) \times P(T/\text{parameters}).$$

120 The compatibility of the data with the hypergeometric distributions in  $P(\text{data}/T)$  are tested  
 121 asymptotically by contingency table chi-squared tests. These tests are organized into several  
 122 interpretable components by further partitioning  $P(\text{data}/T)$  (Pradel et al. 2005). If systematic  
 123 departures from the general model can be ruled out, lack of fit may result from the lack of  
 124 independence among individuals (the two members of a pair, social groups, a contagious disease...  
 125 Then, the value of the Pearson statistic is much larger than the residual degrees of freedom. One  
 126 simple measure for overdispersion is  $\hat{c}$  defined as the ratio of the Pearson statistic  $X^2$  by its number  
 127 of degrees of freedom  $df$

128 
$$\hat{c} = \frac{\chi X^2}{df}$$

129 This ratio can be calculated for each component individually or overall. Although there is no clear cut  
 130 decision rule to decide that an observed lack of fit results solely from overdispersion, a reasonable  
 131 rule of thumb is: the  $\hat{c}$  ratio is greater than 1 for all components, and there is no component for  
 132 which the  $\hat{c}$  ratio greatly exceeds the others. This procedure, although not perfect, works generally  
 133 well as long as the overall ratio does not exceed 3, sometimes 5 (Burnham and Anderson 2002 -  
 134 numéro de page?); if it does, an important factor has likely been left out of the model and should be  
 135 identified. On the other hand, small structural effects that result in an overall  $\hat{c}$  between 1 and 3 can  
 136 without much damage be assimilated to overdispersion and treated as noise (see below for how to  
 137 do this in practice).

138 To illustrate these general ideas, we will now treat the example of the study of movements of  
 139 Canada geese *Branta canadensis* between three wintering regions, Mid-Atlantic, Chesapeake and  
 140 Carolinas, between 1984 and 1989 (Hestbeck et al. 1991).

141  
 142 The option 'GOODNESS-OF-FIT for Multi-state' of the main menu of U-CARE opens onto the  
 143 goodness-of-fit test of the JollyMove (JMV) model (Brownie et al. 1993). Multi-state models  
 144 (Hestbeck et al. 1991) allow for transitions between states (the wintering sites for the geese),  
 145 survival probabilities and encounter probabilities. In the JMV model, transitions vary by state of  
 146 departure, state of arrival and time interval. Survival probabilities vary by state of departure and time  
 147 interval. Encounter probabilities vary by previous state, current state and date. This model, in  
 148 contrast with the better-known Arnason-Schwarz (AS) model, allows encounter probabilities to vary  
 149 by previous state. However, there is currently no optimal GOF test available for the AS model.  
 150 Program U-CARE contains specific tests for transience (Test 3G.SR, null hypothesis H0: 'There is no  
 151 difference in the probability of being later reencountered between "new" and "old" individuals  
 152 encountered simultaneously. '), trap-dependence (Test M.ITEC, H0(i): 'There is no difference in the  
 153 probabilities of being reencountered in the different states at i+1 between the animals in the same  
 154 state at occasion i whether encountered or not encountered at this date, conditional on presence at  
 155 both occasions. '), and memory (Test WBWA H0: 'There is no difference in the expected state of next  
 156 reencounter among individuals previously encountered in the different states'). The  $\hat{c}$  ratio for the  
 157 JMV model is computed from these three main components plus two complementary tests 3G.SM  
 158 and M.LTEC (Pradel et al. 2005).

159 
$$\hat{c} = \frac{\chi X_{WBWA}^2 + \chi X_{aG.SR}^2 + \chi X_{aG.SM}^2 + \chi X_{M.ITEC}^2 + \chi X_{M.LTEC}^2}{df_{WBWA} + df_{aG.SR} + df_{aG.SM} + df_{M.ITEC} + df_{M.LTEC}}$$

160 If we were to consider that the lack of fit is due solely to overdispersion, the overall  $\hat{c}$  ratio for the  
 161 geese would be:

162 
$$\hat{c} = \frac{472.8 + 117.7 + 302.7 + 68.2 + 21.9}{20 + 12 + 119 + 27 + 19} = 5$$

163 However, the  $\hat{c}$  ratio relative to the WBWA component is huge: 23.6 (see Figure 2)!

164

TEST WBWA (test of memory: Where Before vs Where After) , group =1						
===== Summary of TEST WBWA results =====						
sta	pval	df	occ	state	1 if fisher	0 otherwise
19.591	0.000	2.000	2.000	1.000	0.000	
37.868	0.000	2.000	2.000	2.000	0.000	
4.487	0.034	1.000	2.000	3.000	1.000	
80.590	0.000	1.000	3.000	1.000	0.000	
98.761	0.000	4.000	3.000	2.000	0.000	
0.807	0.369	1.000	3.000	3.000	1.000	
27.705	0.000	1.000	4.000	1.000	0.000	
53.694	0.000	2.000	4.000	2.000	0.000	
25.293	0.000	1.000	4.000	3.000	0.000	
43.655	0.000	1.000	5.000	1.000	0.000	
50.926	0.000	2.000	5.000	2.000	0.000	
29.476	0.000	2.000	5.000	3.000	0.000	
----- TEST WBWA (sta, pval, df) -----						
472.855	0.000	20.000				

165

166 **Figure 2 Details of Test WBWA for the Canada goose (*Branta Canadensis*) data from U-CARE. The Pearson statistic (*sta*) is**  
 167 **calculated per occasion and state. The  $\hat{c}$  ratio is 23.6. This suggests strong memory effects in accordance with the overall**  
 168 **p-value (*pval*).**

169

170 3) How to correct for overdispersion?

171 Given the large value of the overall  $\hat{c}$  ratio obtained for the Canada geese, we must try and  $\hat{c}$  find a  
 172 more general model that takes into account the effects unveiled by the GOF test:  $\hat{c}$  the strong  
 173 memory effect and a transient effect. The  $\hat{c}$  ratio of Test 3G.SR is indeed 9.8. A model studied in  
 174 (Rouan et al.) combines these two effects. After discarding the components corresponding to the  
 175 effects incorporated in this model (memory: component WBWA; transience: component 3G.SR), the  
 176 new  $\hat{c}$  ratio is

177 
$$\hat{c} = \frac{302.7 + 68.2 + 21.9}{119 + 27 + 19} = 2.4,$$

178 This value is typical of large data sets with several thousands of individuals where individual  
 179 differences inherent in any animal population are inevitably detected. This ratio can be used as a  
 180 variance inflation factor for the model of Rouan and derived models. There is an option in M-SURGE,  
 181 E-SURGE and MARK to introduce this factor. The impact of the correction factor  $\hat{c}$  is null on  
 182 parameter estimates, but estimated (co)variances are multiplied by  $\hat{c}$ , widths of the Wald-CI-  
 183 intervals are  $\sqrt{\hat{c}}$  times larger and the deviance is divided by  $\hat{c}$ . This in turn affects model selection  
 184 where the AIC is replaced with the QAIC (Burnham and Anderson 2002). This procedure of identifying  
 185 a starting model is valid more generally and should be pursued until an acceptable value for the  $\hat{c}$   
 186 ratio is reached.

187 4) How to conduct more specific tests?

188 U-CARE provides details that may give some additional information, notably the tables of observed  
 189 and expected numbers. For example, Figure 3 shows that newly marked individuals are less  
 190 recaptured than previously marked individuals at occasion 5 on sites 1 and 3 but not 2. This is a  
 191 general pattern over all occasions which suggests that transients may be absent from the central site  
 192 2. A more specific model could be considered. For an application to seasonal data, see ((Gauthier et

193 al. 2001)). U-CARE also gives directional tests (REF) focusing more closely on the detection of  
 194 transience and trap-dependence.

TABLE 3G.SR(5,1): animals caught at occasion5 in state1 seen earlier and later.	
Seen again - Never seen again	
New	93 317
Old	279 602
Table of expected values	
Seen again - Never seen again	
New	118.141 291.859
Old	253.859 627.141
TABLE 3G.SR(5,2): animals caught at occasion5 in state2 seen earlier and later.	
Seen again - Never seen again	
New	297 748
Old	512 1330
Table of expected values	
Seen again - Never seen again	
New	292.8317 752.1683
Old	516.16834 1325.8317
TABLE 3G.SR(5,3): animals caught at occasion5 in state3 seen earlier and later.	
Seen again - Never seen again	
New	157 504
Old	134 197
Table of expected values	
Seen again - Never seen again	
New	193.9022 467.0978
Old	97.09778 233.9022

195

196 **Figure 3 Tables of expected and observed values for Test 3G.SR at occasion 5 for sites 1, 2, 3**

197 Eventually, the option 'TOOLS' contains some specific tools, not easily found, like the ability to test  
 198 for mixtures of multinomial in a contingency table (Pradel et al. 2003), with an improved algorithm.

199 To cite U-CARE or acknowledge its use, cite this Software Note as follows, substituting the version of  
 200 the application that you used for "Version 2.3":

201 CHOQUET R., LEBRETON J.-D., GIMENEZ O., REBOULET A.-M. and PRADEL R. (200?). U-CARE: Utilities  
 202 for performing goodness of fit tests and manipulating Capture-REcapture data. *Ecography* ? : ?-?.  
 203 (Version 2.3)

204 References

205

206

207 Brownie, C., J. E. Hines, J. D. Nichols, K. H. Pollock, and J. B. Hestbeck. 1993. Capture-Recapture  
 208 studies for multiple strata including non-Markovian transitions. *Biometrics*:1173-1187.

209 Burnham, K. P. 1991. On a unified theory for release-resampling of animal populations. Pages 11-35  
 210 *in* C. Chao, editor. *Symposium in Statistics*. Institute of Statistical Science, Academia sinica,  
 211 Taipei.

212 Burnham, K. P., and D. R. Anderson. 2002. *Model selection and Inference: A practical Information-*  
 213 *theoretic approach.*, 2nd edition, New York.

214 Choquet, R., A. M. Reboulet, J. D. Lebreton, O. Gimenez, and R. Pradel. 2005. U-CARE 2.2 User's  
 215 Manual. CEFE, UMR 5175, Montpellier.

216 Choquet, R., A. M. Reboulet, R. Pradel, O. Gimenez, and J. D. Lebreton. 2004. {M-SURGE}: New  
 217 software specifically designed for multistate recapture models. *Animal Biodiversity and*  
 218 *Conservation* **27**:207-215.

219 Choquet, R., L. Rouan, and R. Pradel. 2008. Program {E-SURGE}: A software application for fitting  
 220 multievent models. Pages 847-868 *in* D. L. Thomson, E. G. Cooch, and M. J. Conroy, editors.  
 221 *Modeling Demographic Processes in Marked Populations*. Springer.

222 D'Agostino, R., B., and M. Stephens, A. 1986. Goodness-of-fit techniques. Marcel Dekker, Inc.  
223 Gauthier, G., R. Pradel, S. Menu, and J. D. Lebreton. 2001. Seasonal survival of Greater Snow Geese  
224 and effect of hunting under dependence in sighting probability. *Ecology* **82**:3105-3119.  
225 Gregg, K. B., and M. Kery. 2006. Comparison of size vs. life-state classification demographic models  
226 for the terrestrial orchid *Cleisthes bifaria*. *Biological Conservation* **129**:50-58.  
227 Hénaux, V., T. Bregnballe, and J. D. Lebreton. 2007. Dispersal and recruitment during population  
228 growth in a colonial bird, the great cormorant *Phalacrocorax carbo sinensis*. *Journal of Avian*  
229 *Biology* **38**:44-57.  
230 Hestbeck, J. B., J. D. Nichols, and R. A. Malecki. 1991. Estimates of movement and site fidelity using  
231 mark-resight data of wintering Canada Geese. *Ecology* **72**:523-533.  
232 Jenouvrier, S., G. Tavecchia, J.-C. Thibault, R. Choquet, and V. Bretagnolle. 2008. Recruitment  
233 processes in long-lived species with delayed maturity: estimating key demographic  
234 parameters. *Oikos* **117**:620-628.  
235 Lebreton, J.-D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling survival and testing  
236 biological hypotheses using marked animals: A unified approach with case studies. *Ecological*  
237 *Monographs* **62**:67-118.  
238 Lebreton, J. D., and M. Roux. 1989. BIOMECO (Biometrie-Ecologie). *in*. Avenix, Montpellier.  
239 Pradel, R., E. Cooch, and F. Cooke. 1995. Transient animals in a resident population of snow geese:  
240 local emigration or heterogeneity? *Journal of Applied Statistics* **22**.  
241 Pradel, R., O. Gimenez, and J. D. Lebreton. 2005. Principles and interest of GOF tests for multistate  
242 capture–recapture models. *Animal Biodiversity and Conservation* **28**:189-204.  
243 Pradel, R., C. M. A. Wintrebert, and O. Gimenez. 2003. A Proposal for a Goodness-of-Fit Test to the  
244 Arnason-Schwarz Multisite Capture-Recapture Model. *Biometrics* **59**:43-53.  
245 Rouan, L., R. Choquet, and R. Pradel. A general framework for modeling memory in capture-  
246 recapture data. *Jabes* **In press**.  
247 White, G. C., and K. P. Burnham. 1999. Program MARK: Survival estimation from populations of  
248 marked animals. *Bird Study* **46 (suppl.)**:120-139.  
249  
250