**A comparison of three statistical methods for analysing extinction threat status**

HEATHER R. TAFT, DEREK A. ROFF, ATTE KOMONEN AND JANNE S. KOTIAHO

**APPENDIX 1**

R code for assessing the chance that a given number of correctly assigned species or more could occur randomly to determine whether the analyses presented here are able to correctly classify threatened species more accurately than random assignment.

set.seed(20)

Max.rep <- 10000

N <- matrix(0,Max.rep,3)

for ( i in 1:Max.rep)

{

 X <- matrix(0,80) # First put zeros (= non-threatened)

 X[1:17] <- 1 # Set Threatened =1 Assuming that 17 rows are threatened

 Y <- sample(X) # Pick random sample of length X without replacement

 Z.total <- Y

 Z.threatened <- Z.total[1:17] # 1 to 17 are those threatened 1s are correct zeros are not

 Z.not <- Z.total[18:80] # 18 to 80 are those not threated 0s are correct 1s are not

 N[i,2:3] <- c(length(Z.threatened[Z.threatened==1]), length(Z.not[Z.not==0]))

# Number correct

 N[i,1] <- sum(N[i,2:3]) # The sum of correct assignments

}

print(N)

# Now find the number in each category

N.total <- unique(N[,1]) # Find total number correctly assigned

n <- length(N.total)

P.total <- matrix(0,n,2)

P.total[,1] <- N.total

for ( i in 1:n){ P.total[i,2] <-length( N[N[,1]==N.total[i],1])/Max.rep}

P.total

sum(P.total[,2])

# Threatened correctly assigned

N.threatened <- unique(N[,2]) # Distribution of correctly assigned threatened

n <- length(N.threatened)

P.threatened <- matrix(0,n,2)

P.threatened[,1] <- N.threatened

for ( i in 1:n){ P.threatened[i,2] <-length( N[N[,2]==N.threatened[i],1])/Max.rep}

P.threatened

sum(P.threatened[,2])

# Not threatened correctly assigned

N.not <- unique(N[,3])

n <- length(N.not)

P.not <- matrix(0,n,2)

P.not[,1] <- N.not

for ( i in 1:n){ P.not[i,2] <-length( N[N[,3]==N.not[i],1])/Max.rep}

P.not

sum(P.not[,2])

# What is the probabilty of correctly assigning at least N.threatened correctly?

# Suppose observed N.threatened = 11 Have to pick all cases greater than 11

# But suppose that the largest in output is 10 (as it is here). This means that probability is less than 1 in 10,000

# Suppose the observed values is 5. Then have to pick all values greater than 5

 N.threatened <- 5

 P.more.than.obs <- P.threatened[P.threatened[,1]>=N.threatened,]

 P <- sum(P.more.than.obs[,2])