Ecography

E6354
Liu, C., White, M. and Newell, G. 2010. Measuring and comparing the accuracy of species distribution models with presence–absence data. – Ecography 33: xxx–xxx.

Supplementary material

Appendix S1 Methods for simple proportion measures

1. Methods for assessment of one model

1.1 Clopper-Pearson confidence interval (CI_{CP})

(θ_L, θ_U), with a coverage probability — the proportion of the time that the CI contains the true value of interest — of at least $1-\alpha$, can be obtained by solving (Seaman et al. 1996):

$$\sum_{i=k}^{m}\binom{m}{i}\theta_L^i(1-\theta_L)^{m-i}=\frac{\alpha}{2} \text{ and } \sum_{i=0}^{k}\binom{m}{i}\theta_U^i(1-\theta_U)^{m-i}=\frac{\alpha}{2}.$$

When $k=0$, $\theta_L=0$ and $\theta_U=1-(\alpha/2)^{1/n}$; when $k=m$, $\theta_L=(\alpha/2)^{1/n}$ and $\theta_U=1$; and when $0<k<m$, $\theta_L=Beta(k,m-k+1;\alpha/2)$ and $\theta_U=Beta(k+1,m-k;1-\alpha/2)$, where $Beta(\eta,\nu;\tau)$ is the τ percentile of *Beta* distribution with parameters η and ν (Pires and Amado 2008).

1.2 Wilson interval (CI_W)

$$CI_W=\left\{\left[\hat{\theta}+z_{1-\alpha/2}^2/(2m)\right]\mp z_{1-\alpha/2}\left[\hat{\theta}(1-\hat{\theta})m^{-1}+z_{1-\alpha/2}^2/(4m^2)\right]^{1/2}\right\}/\left(1+z_{1-\alpha/2}^2/m\right)$$

1.3 Agresti-Coull interval (CI_{AC})

$$CI_{AC}=\tilde{\theta}\mp z_{1-\alpha/2}\sqrt{\tilde{\theta}(1-\tilde{\theta})/\tilde{m}}\ ,$$

where $\tilde{m} = m + z_{1-\alpha/2}^2$ and $\tilde{\theta} = \left(\hat{\theta} + z_{1-\alpha/2}^2/(2m)\right)/\left(1 + z_{1-\alpha/2}^2/m\right)$.

1.4 Wilson interval with continuity correction $CI_{Wcc} = (\hat{\theta}_L, \hat{\theta}_U)$

$$\hat{\theta}_L = \left[2m\hat{\theta} + z_{1-\alpha/2}^2 - 1 - z_{1-\alpha/2}\sqrt{z_{1-\alpha/2}^2 - (2+1/m) + 4\hat{\theta}\left(m - m\hat{\theta} + 1\right)}\right] \Big/ \left[2\left(m + z_{1-\alpha/2}^2\right)\right]$$

and

$$\hat{\theta}_U = \left[2m\hat{\theta} + z_{1-\alpha/2}^2 + 1 - z_{1-\alpha/2}\sqrt{z_{1-\alpha/2}^2 + (2-1/m) + 4\hat{\theta}\left(m - m\hat{\theta} - 1\right)}\right] \Big/ \left[2\left(m + z_{1-\alpha/2}^2\right)\right].$$

2. Methods for comparison of two models

2.1 For independent data

2.1.1 Confidence interval of difference $d = \theta^{(1)} - \theta^{(2)}$ based on the Wilson score (CI_{Ws})

$$\left(L, U\right) = \left(\hat{d} - \delta, \hat{d} + \varepsilon\right),$$

where $\hat{d} = \hat{\theta}^{(1)} - \hat{\theta}^{(2)}$, $\delta = \sqrt{\left(k_1/m_1 - l_1\right)^2 + \left(u_2 - k_2/m_2\right)^2}$ and $\varepsilon = \sqrt{\left(u_1 - k_1/m_1\right)^2 + \left(k_2/m_2 - l_2\right)^2}$, l_1 and u_1 are the roots of $\left|x - k_1/m_1\right| = z_{1-\alpha/2}\sqrt{x(1-x)/m_1}$ (in terms of x), and l_2 and u_2 are the roots of $\left|x - k_2/m_2\right| = z_{1-\alpha/2}\sqrt{x(1-x)/m_2}$ (also in terms of x).

2.1.2 Confidence interval based on the Wilson score with continuity correction (CI_{Wscc})

The formula is similar to CI_{Ws}, but l_1 and u_1 delimit the interval $\left\{x: \left|x - k_1/m_1\right| - 1/(2m_1) \le z_{1-\alpha/2}\sqrt{x(1-x)/m_1}\right\}$, and $l_1 = 0$ if $k_1 = 0$ and $u_1 = 1$ if $k_1 = m_1$; l_2 and u_2 delimit the interval $\left\{x: \left|x - k_2/m_2\right| - 1/(2m_2) \le z_{1-\alpha/2}\sqrt{x(1-x)/m_2}\right\}$, and $l_2 = 0$ if $k_2 = 0$ and $u_2 = 1$ if $k_2 = m_2$.

2.2 For paired data

2.2.1 Confidence interval of difference based on Wilson score for paired data (CI_{WsP})

$(\hat{d} - \delta, \hat{d} + \varepsilon)$,

where $\hat{d} = \hat{\theta}^{(1)} - \hat{\theta}^{(2)}$, $\delta = \sqrt{l_1^2 - 2\hat{\phi} l_1 u_3 + u_3^2}$, $\varepsilon = \sqrt{u_1^2 - 2\hat{\phi} u_1 l_3 + l_3^2}$. Here $l_1 = (e+f)/m - l_2$, $u_1 = u_2 - (e+f)/m$ where l_2 and u_2 are roots of $\left|x - (e+f)/m\right| = z_{1-\alpha/2}\sqrt{x(1-x)/m}$; similarly, $l_3 = (e+g)/m - l_4$, $u_3 = u_4 - (e+g)/m$ where l_4 and u_4 are roots of $\left|x - (e+g)/m\right| = z_{1-\alpha/2}\sqrt{x(1-x)/m}$; and $\hat{\phi} = (eh - fg)\big/\sqrt{(e+f)(g+h)(e+g)(f+h)}$ with its numerator replaced by $\max(eg - fg - m/2, 0)$ if $eh > fg$.

Appendix S2 Methods for the Kappa statistic

1. The Asymptotic variance of Kappa proposed by Fleiss et al. (1969, see also Hanley 1987)

$$\hat{\sigma}^2_{FCE} = (T_1 + T_2 - T_3)/\left[n(1-a_e)^2\right],$$

where $T_1 = \sum_{i=0}^{1}(n_{ii}/n)[1-(n_{i+}+n_{+i})(1-\kappa)/n]^2$, $T_2 = (1-\kappa)^2 \sum_{i \neq j} n_{ij}(n_{i+}+n_{+j})^2 \Big/ n^3$,

$T_3 = [\kappa - a_e(1-\kappa)]^2$, and $a_e = (n_{1+}n_{+1} + n_{0+}n_{+0})/n^2$.

2. The Asymptotic variance of Kappa proposed by Blackman and Koval (2000)

$$\hat{\sigma}^2_{BK} = (1-\kappa)\{(1-\kappa)(1-2\kappa) + \kappa(2-\kappa)/[2P(1-P)]\}/n,$$

where P is the species prevalence, its maximum likelihood estimate from the common correlation model is $\hat{P} = (2n_{11} + n_{10} + n_{01})/(2n)$ (Bloch and Kraemer 1989).

3. The Asymptotic variance of Kappa proposed by Garner (1991)

$$\hat{\sigma}^2_G = 4\left[n^2(1-a_e)^2 \sum_{i,j=0}^{1}(n_{ij}+1)^{-1}\right]^{-1}$$

4. Cornfield's test-based method studied by Hale and Fleiss (1993)

The lower limit of the $100(1-\alpha)$ % confidence interval of Kappa is the root of the following equation $(n_{11} - N_{11})^2 V - z^2_{\alpha/2} = 0$ in terms of x which appears in N_{11}, where $V = N_{11}^{-1} + N_{10}^{-1} + N_{01}^{-1} + N_{00}^{-1}$, $N_{11} = [2n_{1+}n_{+1} + x(n_{1+}n_{+0} + n_{0+}n_{+1})]/(2n)$,

$N_{10} = n_{1+} - N_{11}$, $N_{01} = n_{+1} - N_{11}$, and $N_{00} = N_{11} + n_{0+} - n_{+1}$. The starting value of x can be set to the lower limit of the 100($1-\alpha$) % confidence interval with the FCE method. The condition $x < \hat{\kappa}$ must be satisfied in order to obtain the lower confidence limit, where $\hat{\kappa}$ is the estimated kappa value with the sample. The upper confidence limit can be obtained by replacing n_{11} with n_{00} and replacing N_{11} with N_{00} in the above equation and $x > \hat{\kappa}$ must be satisfied.

5. Profile-variance-based method proposed by Lee and Tu (1994)

The confidence interval of Kappa can be constructed by solving the following equation in terms of x:

$b_3x^3 + b_2x^2 + b_1x + b_0 = 0$,

where $b_3 = z_{\alpha/2}^2\left[-a_1 - a_2 + 2a_1^2 + 2a_2^2 + 6a_1a_2 - 8a_1a_2^2 - 8a_1^2a_2 + 8a_1^2a_2^2\right]/c$,

$b_2 = 1 + z_{\alpha/2}^2\left[3a_1 + 3a_2 - 6a_1^2 - 6a_2^2 - 14a_1a_2 + 20a_1a_2^2 + 20a_1^2a_2 - 20a_1^2a_2^2\right]/c$,

$b_1 = -2Kp + z_{\alpha/2}^2\left[-2a_1 - 2a_2 + 4a_1^2 + 4a_2^2 + 12a_1a_2 - 16a_1a_2^2 - 16a_1^2a_2 + 16a_1^2a_2^2\right]/c$,

$b_0 = z_{\alpha/2}^2\left[-4a_1a_2 + 4a_1a_2^2 + 4a_1^2a_2 - 4a_1^2a_2^2\right]/c$, $c = n(2a_1a_2 - a_1 - a_2)^2$, $a_1 = n_{+1}/n$, and $a_2 = n_{1+}/n$. The lower and upper limits of the confidence interval are the two roots of the above equation that are closest to $\hat{\kappa}$, and the third root is usually outside the range of (-1,1).

6. Exact bootstrap confidence intervals for κ (Klar et al. 2002)

The total number of points in the bootstrap sample space is $(n+3)(n+2)(n+1)/6$.

For each sample point ($m_{00}, m_{01}, m_{10}, m_{11}$), we can calculate the value of κ, which is denoted as $\hat{\kappa}(m_{00}, m_{01}, m_{10}, m_{11})$, and its associated probability $p(m_{00}, m_{01}, m_{10}, m_{11} \mid \pi) = p_{00}^{m_{00}} p_{01}^{m_{01}} p_{10}^{m_{10}} p_{11}^{m_{11}} n!/(m_{00}!m_{01}!m_{10}!m_{11}!)$, where $m_{00} \in \{0,1,2,\cdots,n\}$, $m_{01} \in \{0,1,2,\cdots,n-m_{00}\}$, $m_{10} \in \{0,1,2,\cdots,n-m_{00}-m_{01}\}$, $m_{11} = n - m_{00} - m_{01} - m_{10}$, and $p_{ij} = n_{ij}/n$ $(i, j = 0,1)$ are the observed proportions in the four cells of the observed confusion table. The $\hat{\kappa}(m_{00}, m_{01}, m_{10}, m_{11})$ and their associated probabilities comprise the bootstrap distribution, from which the bootstrap confidence interval can be constructed using the $(100\alpha/2)^{\text{th}}$ and $100(1-\alpha/2)^{\text{th}}$ percentiles. If necessary (because of the discreteness of the bootstrap distribution), linear interpolation can be used in the construction of the confidence intervals. From this bootstrap distribution, the variance of kappa can also be calculated.

7. Weighted least-squares (WLS) method introduced by Barnhart and Williamson (2002)

Suppose there are four categorical readings, Y_{11}, Y_{10}, Y_{01} and Y_{00}, assessed on each of n sites. The first two readings are corresponding to one method, among which Y_{11} is model prediction and Y_{10} is observation, and the latter two readings are corresponding to another method, among which Y_{01} is model prediction and Y_{00} is observation. Actually, in our current situation, Y_{10} and Y_{00} are the same. Cross-classifying Y_{11}, Y_{10}, Y_{01} and Y_{00}, we have a $2 \times 2 \times 2 \times 2$ contingency table (denoted by $Y_{11} \times Y_{10} \times Y_{01} \times Y_{00}$) with cell counts y_{ijkl} and cell probabilities $\pi_{ijkl} = prob(Y_{11} = i, Y_{10} = j, Y_{01} = k, Y_{00} = l)$

($i, j, k, l = 0,1$). Let $\pi = \left(\pi_{1111}, \pi_{1110}, \pi_{1101}, \pi_{1100}, \cdots, \pi_{0011}, \pi_{0010}, \pi_{0001}, \pi_{0000}\right)'$ denote the $2^4 \times 1$ vector of cell probabilities for the $Y_{11} \times Y_{10} \times Y_{01} \times Y_{00}$ contingency table and p its sample counterpart. Thus, the kappas for the two methods can be calculated from the collapsed cell probabilities of π_{ij++} and π_{++kl}, respectively. $\kappa = \left(\kappa^{(1)}, \kappa^{(2)}\right)'$ can be written as an explicit function of π: $\kappa = F(\pi) \equiv \exp A_4 \log A_3 \exp A_2 \log A_1 A_0\ \pi$, where A_i ($i = 0,1,2,3,4$) are matrices defined as:

A_0 is a 8×16 matrix $A_0 = \begin{pmatrix} e_4' & 0 & 0 & 0 \\ 0 & e_4' & 0 & 0 \\ 0 & 0 & e_4' & 0 \\ 0 & 0 & 0 & e_4' \\ I_4 & I_4 & I_4 & I_4 \end{pmatrix}$, where **e**4 is 4×1 matrix of all ones, **I**4 is 4×4 identity matrix, and **0** is a 1×4 matrix; the other matrices Ai ($i = 1,2,3,4$) has the form $A_i = \begin{pmatrix} A_{ii} & 0 \\ 0 & A_{ii} \end{pmatrix}$, where $A_{44} = (1 \quad -1)$, $A_{33} = \begin{pmatrix} -1 & -1 & 1 & 0 \\ -1 & -1 & 0 & 1 \end{pmatrix}$,

$$A_{22} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad A_{11} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

In the above equation, all the mathematical operations are taken from right to left. Its weighted least-squares estimator is $\hat{\kappa} = \exp A_4 \log A_3 \exp A_2 \log A_1 A_0\ p$.

The estimated variance-covariance matrix of $\hat{\kappa} = \left(\hat{\kappa}^{(1)}, \hat{\kappa}^{(2)}\right)'$ is

$\hat{C}ov(\hat{\kappa}^{(1)}, \hat{\kappa}^{(2)}) = \left(\frac{\partial F}{\partial p}\right) V \left(\frac{\partial F}{\partial p}\right)'$, where $V = \left(diag(p) - pp'\right)/n$ is the estimated variance-covariance matrix for **p**, diag(**p**) denotes the diagonal matrix with **p** in the diagonal entry, and $\partial F / \partial p$ is the partial derivative of **F** with respect to π evaluated at $\pi = p$, which is $\frac{\partial F}{\partial p} = \text{diag}(B_4)\ A_4\, \text{diag}(B_3)^{-1} A_3\, \text{diag}(B_2) A_2\, \text{diag}(B_1)^{-1} A_1 A_0$, where $B_1 = A_1 A_0\, \mathbf{p}$, $B_2 = \exp A_2 \log B_1$, $B_3 = A_3 B_2$ and $B_4 = \exp A_4 \log B_3$.

Appendix S3 Methods for other threshold-dependent measures

1. 100($1-\alpha$) % Confidence interval for PLR^{-1} and PLR with Fieller's approach

$\left(-B_1 \mp \sqrt{B_1^2 - 4A_1C_1}\right) / (2A_1)$ and

$\left(-B_1 \mp \sqrt{B_1^2 - 4A_1C_1}\right) / (2C_1)$,

where $A_1 = Se^2 - z_{1-\alpha/2}^2 \sigma_{Se}^2$, $B_1 = -2Se(1-Sp)$, $C_1 = (1-Sp)^2 - z_{1-\alpha/2}^2 \sigma_{Sp}^2$, and σ_{Se}^2 and σ_{Sp}^2 are the variance of Se and Sp respectively, which have been described in the text.

2. The 100($1-\alpha$) % confidence interval for NLR with Fieller's approach

$\left(-B_2 \mp \sqrt{B_2^2 - 4A_2C_2}\right) / (2A_2)$, where $A_2 = Sp^2 - z_{1-\alpha/2}^2 \sigma_{Sp}^2$, $B_2 = -2Sp(1-Se)$, $C_2 = (1-Se)^2 - z_{1-\alpha/2}^2 \sigma_{Se}^2$, and σ_{Se}^2 and σ_{Sp}^2 are the variances of Se and Sp respectively as described in the text.

3. The 100($1-\alpha$) % confidence intervals for PLR^{-1}, PLR and NLR using the delta method

$(1-Sp)/Se \mp z_{1-\alpha/2}\sigma_1/\sqrt{n}$,

$\left[(1-Sp)/Se \pm z_{1-\alpha/2}\sigma_1/\sqrt{n}\right]^{-1}$ and

$(1-Se)/Sp \mp z_{1-\alpha/2}\sigma_2/\sqrt{n}$,

where $\sigma_1 = \left[n(1-Sp)^2\sigma_{Se}^2/Se^4 + n\sigma_{Sp}^2/Se^2\right]^{1/2}$ and

$\sigma_2 = \left[n(1-Se)^2\sigma_{Sp}^2/Sp^4 + n\sigma_{Se}^2/Sp^2\right]^{1/2}$.

4. The 100($1-\alpha$) % confidence intervals for PPV and NPV

$\left[1+U_1(1-P)/P\right]^{-1} \le PPV \le \left[1+L_1(1-P)/P\right]^{-1}$ and

$\left[1+U_2(1-P)/P\right]^{-1} \le NPV \le \left[1+L_2(1-P)/P\right]^{-1}$,

where L_1 and U_1 are the lower and upper limits for PLR^{-1} with the above two methods, and L_2 and U_2 are the lower and upper limits for NLR; and where P is the species prevalence, which can be estimated as $\hat{P} = n_{1+}/n$.

5. The 100($1-\alpha$) % confidence interval for OR

(L, U) can be obtained by solving the following two equations:

$$\sum_{i=n_{11}}^{n_{+1}} \binom{n_{1+}}{i}\binom{n_{0+}}{n_{+1}-i} L^i \Big/ \sum_{i=0}^{n_{+1}} \binom{n_{1+}}{i}\binom{n_{0+}}{n_{+1}-i} L^i = \alpha/2 \text{ and}$$

$$\sum_{i=0}^{n_{11}} \binom{n_{1+}}{i}\binom{n_{0+}}{n_{+1}-i} U^i \Big/ \sum_{i=0}^{n_{+1}} \binom{n_{1+}}{i}\binom{n_{0+}}{n_{+1}-i} U^i = \alpha/2 .$$

When $n_{11} = 0$ or $n_{00} = 0$, the lower confidence limit is set to zero, and the upper limit is determined with level α. When $n_{10} = 0$ or $n_{01} = 0$, the upper confidence limit is set to infinity, and the lower limit is determined with level α.

6. Estimate of the variance of the true skill statistic (TSS)

$$\hat{\sigma}_1^2(TSS) = \frac{n_{11}}{n_{1+}^2}\left(1-\frac{n_{11}}{n_{1+}}\right) + \frac{n_{00}}{n_{0+}^2}\left(1-\frac{n_{00}}{n_{0+}}\right)$$

$$\hat{\sigma}_2^2(TSS) = \frac{n^2 - 4n_{1+}n_{0+}\left(\frac{n_{11}}{n_{1+}} + \frac{n_{00}}{n_{0+}} - 1\right)^2}{4nn_{1+}n_{0+}}$$

7. Estimate of the variance of normalized mutual information (NMI)

$$\hat{\sigma}^2(NMI) = \left\{ \sum_{i,j=0}^{1} (n_{ij}/n)\left[H_o \ln\left(n_{i+}n_{+j}/(n_{ij}n)\right) - \left(H_o + H_p - H_{op}\right)\ln(n_{+i}/n)\right]^2 \right\} \Big/ \left(nH_o^4\right),$$

where $H_o = \ln n - (n_{1+}\ln n_{1+} + n_{0+}\ln n_{0+})/n$, $H_p = \ln n - (n_{+1}\ln n_{+1} + n_{+0}\ln n_{+0})/n$,

$H_{op} = \ln n - (1/n)\sum_{i,j=0}^{1} n_{ij} \ln n_{ij}$.

8. Estimate of the standard deviation of extreme dependency score (EDS)

$$\hat{\sigma} = \sqrt{\frac{\hat{S}e\left(1-\hat{S}e\right)}{n\hat{P}} \frac{2\ln\hat{P}}{\hat{S}e\left[\ln\left(\hat{P}\hat{S}e\right)\right]^2}},$$

where $\hat{P} = n_{1+}/n$ is the estimated species prevalence.

Appendix S4 Methods for AUC

1. Delong et al.'s (1988) non-parametric method to estimate the variance-covariance matrix of a vector of AUCs

Let $\{X_i^{(k)}\}$ and $\{Y_j^{(k)}\}$ be the sets of predicted values from the k-th model that correspond to the n_{0+} absence sites and n_{1+} presence sites, respectively ($i = 1,2,\cdots,n_{0+}$; $j = 1,2,\cdots,n_{1+}$; $k = 1,2,\cdots,K$). The AUC for the k-th model can be estimated similarly as above, and is denoted as $\hat{\theta}^{(k)}$ ($k = 1,2,\cdots,K$).

The (s,t)-th element of the $K \times K$ variance-covariance matrix for the vector of AUCs $\hat{\theta} = \left(\hat{\theta}^{(1)},\hat{\theta}^{(2)},\cdots,\hat{\theta}^{(K)}\right)$ for the K models is estimated as

$$\text{cov}(\hat{\theta}^{(s)},\hat{\theta}^{(t)}) = \frac{1}{n_{0+}(n_{0+}-1)}\sum_{i=1}^{n_{0+}}\left[V_{10}^{(s)}(X_i^{(s)})-\hat{\theta}^{(s)}\right]\left[V_{10}^{(t)}(X_i^{(t)})-\hat{\theta}^{(t)}\right]$$

$$+\frac{1}{n_{1+}(n_{1+}-1)}\sum_{j=1}^{n_{1+}}\left[V_{01}^{(s)}(Y_j^{(s)})-\hat{\theta}^{(s)}\right]\left[V_{01}^{(t)}(Y_j^{(t)})-\hat{\theta}^{(t)}\right],$$ where

$V_{10}^{(k)}(X_i^{(k)}) = \frac{1}{n_{1+}}\sum_{j=1}^{n_{1+}}\phi(X_i^{(k)},Y_j^{(k)})$, $V_{01}^{(k)}(Y_j^{(k)}) = \frac{1}{n_{0+}}\sum_{i=1}^{n_{0+}}\phi(X_i^{(k)},Y_j^{(k)})$ ($k = s,t$), and the function $\phi(X,Y)$ has been defined in Table 3.

2. Variance (under the null hypothesis) of the difference between two dependent AUCs in Bandos et al. (2005) exact permutation test for the comparison of two dependent AUCs

$$V_{\Omega}\left(\hat{\theta}^{(1)} - \hat{\theta}^{(2)}\right) = \frac{1}{4n_{1+}^2 n_{0+}^2}\left[\sum_{i=1}^{n_{0+}}\left(w_{i+}^{(1,1)} + w_{i+}^{(1,2)}\right)^2 + \sum_{j=1}^{n_{1+}}\left(w_{+j}^{(1,1)} + w_{+j}^{(2,1)}\right)^2\right]$$, where

$w_{i+}^{(k,l)} = \sum_{j=1}^{n_{1+}} w_{ij}^{(k,l)}$, $w_{+j}^{(k,l)} = \sum_{i=1}^{n_{0+}} w_{ij}^{(k,l)}$, $w_{ij}^{(k,l)} = \phi(x_i^{(k)}, y_j^{(l)}) - \phi(x_i^{(3-k)}, y_j^{(3-l)})$, $x_i^{(k)}$ is the rank for absence site i in k-th model, $y_j^{(l)}$ is the rank for presence site j in l-th model ($k,l = 1,2$; $i = 1,2,\cdots,n_{0+}$; $j = 1,2,\cdots,n_{1+}$), and the function $\phi(X,Y)$ has been defined in Table 3.

Appendix S5 Methods for PAUC

Let $\{X_i \mid i = 1,2,\cdots,n_{0+}\}$ denote the predicted values for a sample of n_{0+} absence sites, and $\{Y_j \mid j = 1,2,\cdots,n_{1+}\}$ denote the predicted values for a sample of n_{1+} presence sites. In order to calculate the PAUC between two pre-specified values of false positive rates [FPR_l, FPR_h], the observed $\{X_i \mid i = 1,2,\cdots,n_{0+}\}$ are ordered with the highest predicted values assigned the highest ranking, and find the values of r_1 and r_0 corresponding to these two FPRs: $r_1 = F^{-1}(1 - FPR_l)$ and $r_0 = F^{-1}(1 - FPR_h)$, where $F(X)$ is the empirical distribution of X, then the PAUC can be estimated as $\hat{\theta} = \frac{1}{n_{0+}n_{1+}}\sum_{i=1}^{n_{0+}}\sum_{j=1}^{n_{1+}}\phi(X_i, Y_j)$, where $\phi(X,Y)$ equals 1 if $Y > X$ and $X \in [r_0, r_1]$, $1/2$ if $Y = X$ and $X \in [r_0, r_1]$, and 0 otherwise.

The variance-covariance matrix must be estimated in order to compare the PAUCs of K models. Let $\hat{\theta} = (\hat{\theta}^{(1)}, \hat{\theta}^{(2)}, \cdots, \hat{\theta}^{(K)})$ be a vector of PAUCs derived from the predicted values $\{X_i^{(k)} \mid i = 1,2,\cdots,n_{0+}\}$ and $\{Y_j^{(k)} \mid j = 1,2,\cdots n_{1+}\}$ ($k = 1,2,\cdots,K$) of K models. The estimated variance-covariance matrix has the (s,t)-th element

$$\mathrm{cov}(\hat{\theta}^{(s)}, \hat{\theta}^{(t)}) = \left(\frac{m_p}{n_{0+}}\right)^2 \left[\frac{1}{m_p(m_p-1)} \sum_{\substack{X_i^{(s)} \in [r_0^{(s)}, r_1^{(s)}] \\ X_i^{(t)} \in [r_0^{(t)}, r_1^{(t)}]}} \{V_{10}^{(s)}(X_i^{(s)}) - \hat{\tau}^{(s)}\}\{V_{10}^{(t)}(X_i^{(t)}) - \hat{\tau}^{(t)}\} \right.$$

$$\left. + \frac{1}{n_{1+}(n_{1+}-1)} \sum_{j=1}^{n_{1+}} \{V_{01}^{(s)}(Y_j^{(s)}) - \hat{\tau}^{(s)}\}\{V_{01}^{(t)}(Y_j^{(t)}) - \hat{\tau}^{(t)}\} \right],$$ where, m_p is the number of X 's

with the corresponding FPRs lying between FPR_l and FPR_h, inclusive;

$$V_{10}^{(k)}(X_i^{(k)}) = \frac{1}{n_{1+}} \sum_{j=1}^{n_{1+}} \phi(X_i^{(k)}, Y_j^{(k)}) \quad (X_i^{(k)} \in [r_0^{(k)}, r_1^{(k)}];\ k = s,t),$$

$$V_{01}^{(k)}(Y_j^{(k)}) = \frac{1}{m_p} \sum_{X_i^{(k)} \in [r_0^{(k)}, r_1^{(k)}]} \phi(X_i^{(k)}, Y_j^{(k)}) \quad (j = 1,2,\cdots,n_{1+}, k = s,t),$$ and

$$\hat{\tau}^{(k)} = \frac{1}{m_p n_{1+}} \sum_{X_i^{(k)} \in [r_0^{(k)}, r_1^{(k)}]} \sum_{j=1}^{n_{1+}} \phi(X_i^{(k)}, Y_j^{(k)}) \quad (k = s,t).$$

With these variance and covariance estimates, z statistics for both a single PAUC, and the difference between two PAUCs can be calculated, and therefore, the relevant confidence intervals can also be constructed.

Since the normalized partial area (nPAUC) is estimated as

$$\hat{\tau} = \frac{1}{m_p n_{1+}} \sum_{X_i \in [r_0, r_1]} \sum_{j=1}^{n_{1+}} \phi(X_i, Y_j) = \frac{n_{0+}}{m_p} \hat{\theta},$$ its variance-covariance matrix can be estimated

with the (s,t)-th element as $\mathrm{cov}(\hat{\tau}^{(s)}, \hat{\tau}^{(t)}) = \left(\frac{n_{0+}}{m_p}\right)^2 \mathrm{cov}(\hat{\theta}^{(s)}, \hat{\theta}^{(t)})$.

Appendix S6 Methods for the other threshold-independent measures

The sampling variance for the mean square error, i.e. Brier's score, and the R^2 based on sum-of-squares, also known as Brier's skill score:

$$\sigma^2\left(M\hat{S}E\right)=\frac{1}{n}\sum_{i=1}^{n}(p_i-o_i)^4-\left[\frac{1}{n}\sum_{i=1}^{n}(p_i-o_i)^2\right]^2 \text{ and}$$

$$\sigma^2(\hat{R}^2)=d_1\sigma^2(M\hat{S}E)+d_2\sigma^2(\hat{\sigma}_o^2)+d_3\,\mathrm{cov}(M\hat{S}E,\hat{\sigma}_o^2)\,,$$

where, $\sigma^2(\hat{\sigma}_o^2)=(n-1)\sigma_o^2[n-1+(6-4n)\sigma_o^2]/n^3$, $\sigma_o^2=n_{1+}n_{0+}/n^2$,

$$\mathrm{cov}(M\hat{S}E,\hat{\sigma}_o^2)=\frac{n-1}{n^2}\sigma_o^2\left(1-\frac{2n_{1+}}{n}\right)\left[\frac{1}{n}\left(\sum_{o_i=1}p_i^2-\sum_{o_i=0}p_i^2\right)+1-\frac{2}{n}\sum_{o_i=1}p_i\right],$$

$$d_1=\frac{1}{\sigma_o^4}\left(\frac{n}{n-1}\right)^2,\ d_2=\frac{\left(1-R^2\right)^2}{\sigma_o^4}\left(\frac{n}{n-1}\right)^4,\ d_3=-\frac{2\left(1-R^2\right)}{\sigma_o^4}\left(\frac{n}{n-1}\right)^3$$