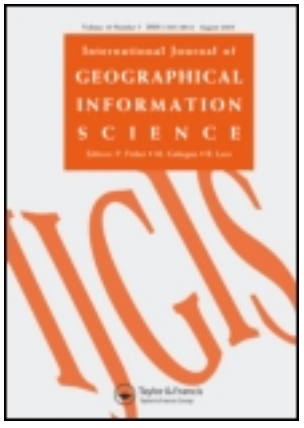

International Journal of Geographical Information Science

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tgis20

The total operating characteristic to measure diagnostic ability for multiple thresholds

Robert Gilmore Pontius Jr[a] & Kangping Si[a]

[a] Graduate School of Geography, Clark University, Worcester, MA, USA

Published online: 06 Jan 2014.

To cite this article: Robert Gilmore Pontius Jr & Kangping Si (2014) The total operating characteristic to measure diagnostic ability for multiple thresholds, International Journal of Geographical Information Science, 28:3, 570-583, DOI: 10.1080/13658816.2013.862623

To link to this article: http://dx.doi.org/10.1080/13658816.2013.862623

International Journal of Geographical Information Science, 2014
Vol. 28, No. 3, 570–583, http://dx.doi.org/10.1080/13658816.2013.862623

Taylor & Francis
Taylor & Francis Group

The total operating characteristic to measure diagnostic ability for multiple thresholds

Robert Gilmore Pontius Jr and Kangping Si*

Graduate School of Geography, Clark University, Worcester, MA, USA

(*Received 3 September 2013; final version received 23 October 2013*)

The relative operating characteristic (ROC) is a popular statistical method to measure the association between observed and diagnosed presence of a characteristic. The diagnosis of presence or absence depends on whether the value of an index variable is above a threshold. ROC considers multiple possible thresholds. Each threshold generates a two-by-two contingency table, which contains four central entries: hits, misses, false alarms, and correct rejections. ROC reveals for each threshold only two ratios, hits/(hits + misses) and false alarms/(false alarms + correct rejections). This article introduces the total operating characteristic (TOC), which shows the total information in the contingency table for each threshold. TOC maintains desirable properties of ROC, while TOC reveals strictly more information than ROC in a manner that makes TOC more useful than ROC. We illustrate the concepts with an application to land change science.

Keywords: area under the curve (AUC); diagnosis; land use and land cover change (LUCC); relative operating characteristic (ROC); total operating characteristic (TOC)

1. Introduction

The relative operating characteristic (ROC), also known as receiver operating characteristic, is a statistical tool that measures the ability of an index variable to diagnose either presence or absence of a characteristic (Green and Swets 1966, Egan 1975, Fawcett 2006). A threshold that is applied to the index variable determines the diagnosis, which is then compared to reference information concerning presence or absence for each observation. The ROC procedure considers multiple possible thresholds, and each threshold generates a two-by-two contingency table, which contains four central entries and additional marginal totals. ROC reveals two ratios of some of the numbers in each contingency table.

ROC has been a fundamental tool for signal detection and diagnostic systems. ROC is popular due in part to numerous articles in *Science*, especially articles by Swets (1961, 1963, 1973, 1979, 1988). ROC has been used in a wide variety of applications such as sensory perception (Carterette and Jones 1967, Cain 1977, Smith *et al.* 1982, Nuechterlein *et al*. 1983), genetics (Johnson *et al.* 2003, Johnson *et al.* 2007), radiology (Lusted 1971, Hanley and McNeil 1982), medicine (Knaus *et al*. 1991), memory (Reed 1973), brain structure (Fleming *et al.* 2010), animal science (Nevin 1965, Blough 1967), machine learning (Bradley 1997, Hand and Till 2001), meteorology (Jolliffe and Stephenson 2003), remote sensing (Saatchi *et al.* 2008), land change science (Pontius and Batchu 2003, Pontius and Pacheco 2004, Pérez-Vega *et al*. 2012), and species modeling (Fielding

*Corresponding author. Email: sikangping@gmail.com

and Bell 1997, Miller and Franklin 2002, Wiley *et al.* 2003, Hirzel *et al.* 2006). Many software packages compute ROC (Stephan *et al.* 2003, Mas *et al.* 2013). ROC has appeared at least 19 times during 2005–2013 in the *International Journal of Geographical Information Science* (Overmars and Verburg 2005, Paegelow and Olmedo 2005, Li *et al.* 2010, Lin *et al.* 2010, Wang and Mountrakis 2010, Conway and Wellen 2011, Gerçek *et al.* 2011, Jansen and Veldkamp 2011, Leitão *et al.* 2011, Peters *et al.* 2011, Gao and Zhang 2012, Golicher *et al.* 2012, Manzo *et al.* 2012, Camacho Olmedo *et al.* 2013, Chen *et al.* 2013, Jeong *et al.* 2013, Kolb *et al.* 2013, Mateo Sánchez *et al.* 2013, Wang *et al.* 2013).

Criticisms abound concerning ROC and its area under the curve (AUC) metric, which is used frequently to summarize the strength of the overall diagnostic ability. Scientists have cautioned against the sole use of AUC as a measure of performance and have urged more research into ROC methodology (Eastman *et al.* 2005, Cook 2007, Lobo *et al.* 2008, Peterson *et al.* 2008, Golicher *et al.* 2012). Partial-area ROC curves have been proposed as one way to focus on particularly important segments of the ROC curve, especially for cases where some types of errors are more important than other types of errors (Jiang *et al.* 1996, Dodd and Pepe 2003). We argue that one of the most important limitations to ROC's interpretation is that ROC fails to reveal the size of each entry in the contingency table for each threshold. The sizes of the entries in the contingency tables express directly the total information that generates the ROC. If scientists know the sizes of the table's entries for each threshold, then scientists can compute numerous potentially useful measurements for each threshold.

Therefore, our article introduces the total operating characteristic (TOC), which rectifies this important limitation of ROC. Both TOC and ROC consider multiple thresholds of the index variable, whereas TOC shows all the entries in each contingency table for each threshold. Our article shows that TOC maintains desirable properties of ROC, while TOC is more informative than ROC.

2. Methods

2.1. Example in land change science

We illustrate both TOC and ROC using a case study. Figure 1 shows the case study's data, which are raster maps of Chiquitania, Bolivia.

Figure 1a shows the reference variable of presence versus absence of disturbance, which means anthropogenic modification of the natural landscape. The observations of the study area are the pixels that were non-disturbance at 1986, which constitute 20.5 thousand square kilometers. TOC defines reference presence so that the number of observations of presence is less than or equal to the number of observations of absence. Therefore, reference presence is gain of disturbance during 1986–2000 and absence is persistence of non-disturbance during 1986–2000. Disturbance gain during 1986–2000 occupies 1.3 thousand square kilometers, which is 6% of the area that was non-disturbance at 1986.

Figure 1b shows the index variable, which is the proximity from disturbance at 1986. We hypothesize that the gain of disturbance tends to occur relatively near previous disturbance, thus smaller distances are ranked higher in terms of proximity. Both TOC and ROC are designed to address the question 'Does the gain of disturbance during 1986–2000 tend to be concentrated at the places that are relatively near to disturbance at 1986?' This type of question is typical in land change science where there is a need to measure

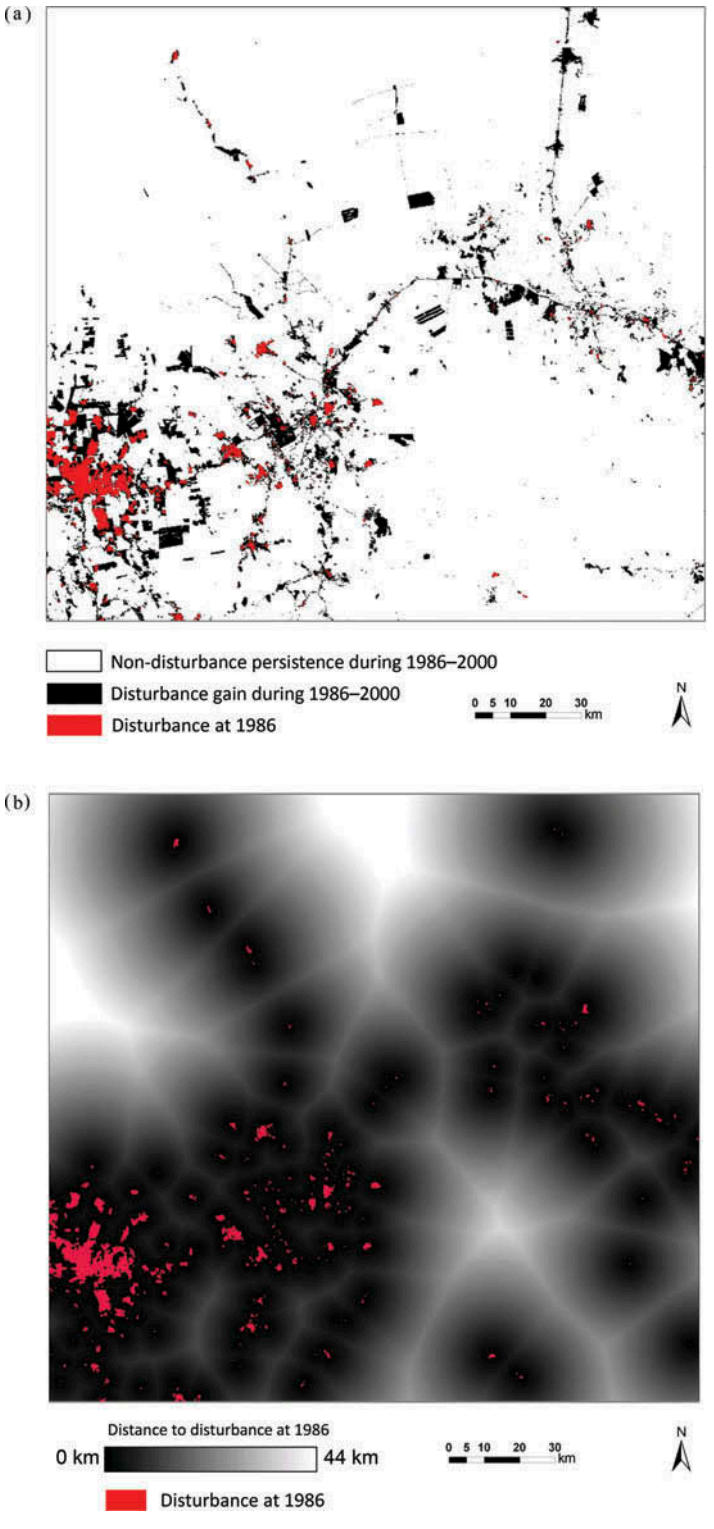

Figure 1. (a) Reference variable where presence is disturbance gain during 1986–2000 and (b) proximity index variable based on distance to disturbance at 1986. Smaller distances are ranked higher in terms of proximity. Disturbance at 1986 is eliminated from the measurement of association.

the diagnostic ability of simulation models that predict the spatial allocation of a characteristic.

2.2. *Thresholds and contingency tables*

The procedures for both TOC and ROC compare the reference variable to the index variable by diagnosing each observation as either presence or absence. A threshold determines the diagnosis. If an observation's index value is greater than or equal to the threshold, then the observation is diagnosed as presence, otherwise the observation is diagnosed as absence. Table 1 shows the contingency table that derives from the comparison between the reference variable and the diagnosis for threshold t. Table 2 defines the mathematical notation. The four central entries in Table 1 are hits H_t, misses M_t, false alarms F_t, and correct rejections C_t. The total number of observations is $P + Q$. The

Table 1. Contingency table showing the number of observations for a threshold.

		Reference		
		Presence	Absence	Diagnosed Total
Diagnosis	Presence	H_t	F_t	$H_t + F_t$
	Absence	M_t	C_t	$M_t + C_t$
	Reference Total	$H_t + M_t = P$	$F_t + C_t = Q$	$P + Q$

Table 2. Mathematical notation.

Symbol	Meaning
t	Denotation for a threshold
H_t	Hits, which is the number of observations that are reference presence and have an index value greater than or equal to threshold t
M_t	Misses, which is the number of observations that are reference presence and have an index value less than threshold t
F_t	False Alarms, which is the number of observations that are reference absence and have an index value greater than or equal to threshold t
C_t	Correct Rejections, which is the number of observations that are reference absence and have an index value less than threshold t
P	Number of observations that are reference presence
Q	Number of observations that are reference absence
X_t	False positive ratio F_t/Q, which is the coordinate on the horizontal axis of the point on the ROC curve for threshold t
Y_t	True positive ratio H_t/P, which is the coordinate on the vertical axis of the point on the ROC curve for threshold t
$\hat{P}$	Estimated number of observations that are reference presence for applications to convert a sample ROC curve to an estimated TOC curve
$\hat{Q}$	Estimated number of observations that are reference absence for applications to convert a sample ROC curve to an estimated TOC curve
$\widehat{H_t + F_t}$	Estimated number of observations in the population that are diagnosed as presence for threshold t
$\hat{H}_t$	Estimated number of observations in the population that are diagnosed correctly as presence for threshold t

selection of the threshold influences the diagnosis but not the reference variable, thus the selection of the threshold influences neither P nor Q.

Four independent bits of information determine all the entries in the contingency table, including its marginal totals. For example, if we know H_t, M_t, F_t, and C_t, then we can compute all the marginal totals for threshold t. Alternatively, if we know H_t/P, F_t/Q, P and Q, then we can compute all the entries in the table. Two bits of information are not sufficient to complete the contingency table. For example, if we know only H_t/P and F_t/Q, which is what ROC shows, then it is impossible to know any entries in the table.

2.3. *Construction of TOC and ROC plots*

The procedure for both TOC and ROC considers multiple thresholds, therefore generates multiple contingency tables of the format of Table 1. It is helpful to have a method to reveal the information of the entire contingency table for all thresholds efficiently in one graphical plot, which is what TOC accomplishes and ROC does not.

Figure 2a shows the TOC plot. The unit on both axes indicates the number of observations, which we express as area for our case study. The horizontal axis ranges from 0 to 20.5 thousand square kilometers, which is the size of the study area $P + Q$. The coordinate on the horizontal axis indicates the size of presence in the diagnosis for a given threshold $H_t + F_t$. The vertical axis ranges from 0 to 1.3 thousand square kilometers, which is the size of presence in the reference variable P. The vertical axis indicates the hits for a given threshold H_t. A horizontal line called 'Hits + Misses' through the vertical coordinate P defines an upper boundary of the TOC space, because the vertical coordinate for all points in the TOC space must be less than or equal to P. To enhance visual clarity, we stretch the vertical axis so that its length equals the length of the horizontal axis, in spite of the fact that $P \neq P + Q$.

The maximum and minimum boundaries in Figure 2a form a parallelogram that contains the mathematically possible space where the TOC curve appears. This mathematically possible space is dictated by P and Q. The maximum boundary represents a hypothetical index variable for which the P highest ranking observations of the index variable correspond with presence for the reference variable. The maximum boundary has a corner at the point (P, P). The left segment of the maximum boundary is the hypotenuse of a right triangle formed by the points (0,0), $(P,0)$, and (P,P). The maximum boundary forms an acute angle with the horizontal axis. This angle is Arctan$[(P + Q)/P]$, because the length of the vertical axis has been stretched to equal $P + Q$. Arctan$[(P + Q)/P]$ is between 63° and 90° depending on the size of P relative to Q. The angle is 63° when half of the observations are reference presence; the angle approaches 90° as $(P + Q)/P$ approaches infinity. The minimum boundary represents a hypothetical index variable for which the Q highest ranking observations of the index variable are associated with absence for the reference variable. The minimum boundary has a corner at the point $(Q, 0)$. Q is equal to the horizontal distance from any point on the left side of the maximum boundary to the corresponding point on the right side of the minimum boundary.

Figure 2a highlights the threshold for 10 kilometers from the disturbance at 1986, in order to illustrate how TOC shows for each threshold the hits, misses, false alarms, and correct rejections. The vertical distance from the horizontal axis to the TOC curve equals H_t, while the vertical distance from the TOC curve and the Hits + Misses line equals M_t. The horizontal distance from the maximum boundary to the TOC curve equals F_t, while the horizontal distance from the TOC curve to the minimum boundary equals C_t.

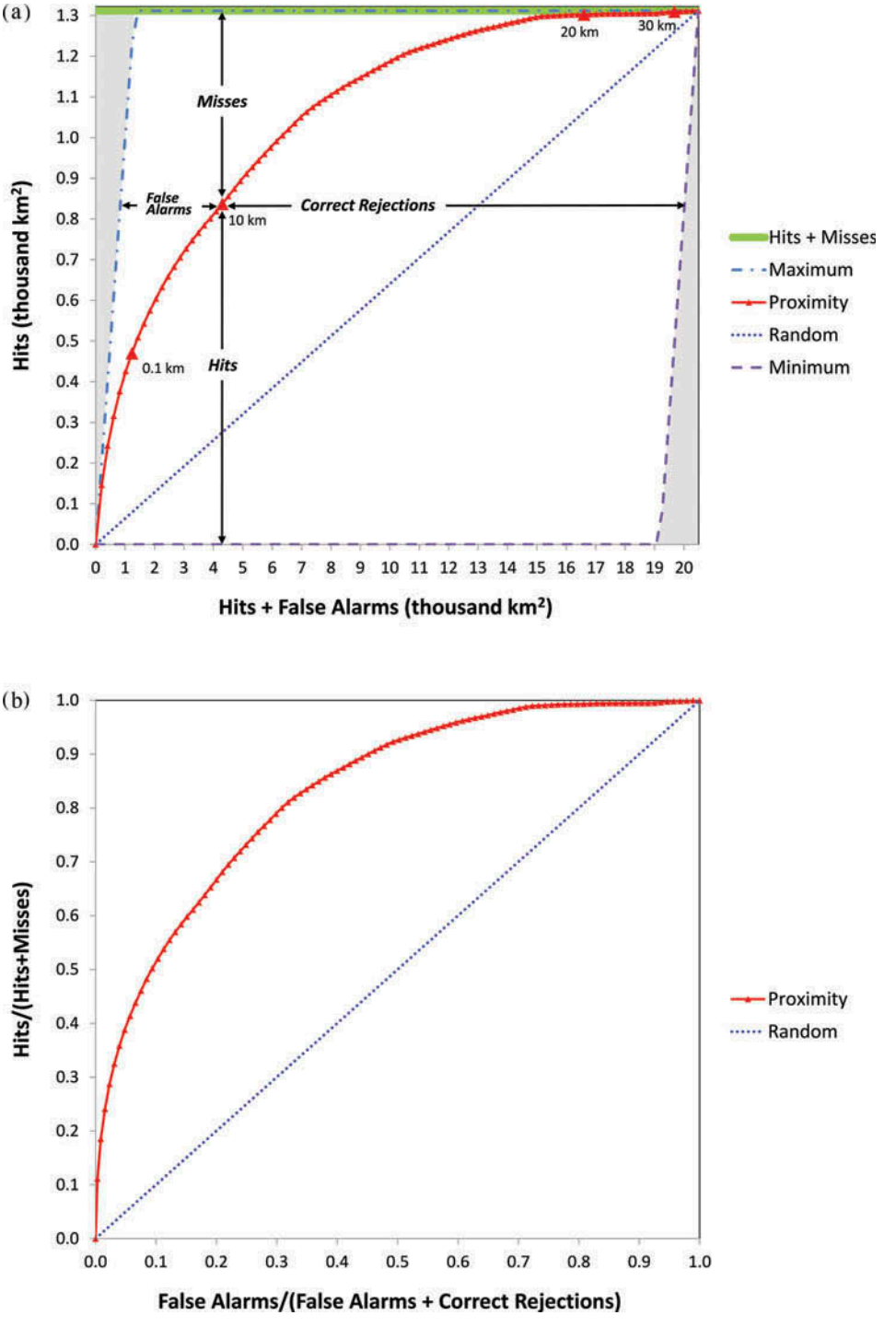

Figure 2. (a) Total operating characteristic (TOC) and (b) relative operating characteristic (ROC) for proximity to disturbance at 1986. The gray triangles are mathematically impossible areas for the TOC curve.

A diagonal dotted line joins the lower left corner to the upper right corner within TOC's parallelogram. This dotted line shows the statistically expected results for a hypothetical index variable that has values assigned at random. For all points along the random line, the ratio of H_t to P equals the ratio of F_t to Q. The random line offers a baseline that might be relevant for some applications. If higher ranking values for a given index have positive association with reference presence, then the given index's TOC curve is above the random line.

Figure 2b shows the ROC curves for the proximity index and the random index. The ROC curve graphs Y_t on the vertical axis versus X_t on the horizontal axis, where Equations (1) and (2) define each point (X_t, Y_t). X_t is also known as 1 – specificity. Y_t is also known as sensitivity.

$$X_t = \frac{F_t}{F_t + C_t} = \frac{F_t}{Q} = 1 - \text{specificity} \tag{1}$$

$$Y_t = \frac{H_t}{H_t + M_t} = \frac{H_t}{P} = \text{sensitivity} \tag{2}$$

Line segments join successive points on both the TOC and ROC curves. The AUC gives a measurement that synthesizes the results across multiple thresholds, such that larger AUCs mean stronger associations.

3. Results

TOC shows all the information that ROC shows, while TOC shows additional information. Both TOC and ROC show that the gain of disturbance during 1986–2000 tends to be concentrated at the places that are relatively near to disturbance at 1986, because both TOC and ROC curves are above their respective random lines. The AUC of the ROC is 0.83, which is greater than the random baseline AUC of 0.5. The supplementary materials give a mathematical proof that the AUC of the ROC is identical to the ratio of the area under the TOC curve that is within the TOC's parallelogram to the area of the TOC's parallelogram. Thus the TOC reveals the AUC of the ROC.

Figure 2 shows how the TOC plot reveals the total information on which the ROC plot is based. First, the TOC plot shows the size of the number of observations as indicated by the right bound on the horizontal axis $P + Q$. Second, the TOC plot shows the size of reference presence as indicated by the upper bound on the vertical axis P. Third, each point on the TOC curve shows the four central entries in the contingency table for each threshold. We have labeled thresholds for 0.1, 10, 20, and 30 kilometers from the disturbance at 1986. When the threshold is 0.1 kilometers, then the number of observations diagnosed as presence matches the number of observations of reference presence. The ROC plot shows neither the size of reference presence nor the size of diagnosed presence for any of the thresholds.

4. Discussion

The point (X_t, Y_t) on the ROC curve corresponds to the point ($PY_t + QX_t$, PY_t) on the TOC curve. This correspondence shows how TOC contains two more bits of information than ROC. P and Q are these two bits of information, which are the keys that unlock the information for each point on the ROC curve, so that the ROC curve can transform into

the TOC curve. This transformation is the reason why TOC has many of the properties that ROC has. For example, scientists can interpret the shape of the TOC curve in a manner identical to how scientists can interpret the shape of the ROC curve. Furthermore, the TOC plot allows the reader to visualize ROC's X_t ratio for each threshold, because the TOC plot shows X_t's numerator, which is F_t, and X_t's denominator, which is Q. Similarly, the TOC plot allows the reader to visualize ROC's Y_t ratio for each threshold, because the TOC plot shows Y_t's numerator, which is H_t, and Y_t's denominator, which is P.

If a scientist knows the entries in the contingency table for each threshold, then the scientist can compute a wide variety of potentially useful measurements for each threshold (Fielding and Bell 1997). For example, knowledge of the entries in the contingency table allows a scientist to compute the conditional probability of a false alarm given a diagnosis of presence $F_t/(H_t + F_t)$, and the conditional probability of a miss given a diagnosis of absence $M_t/(M_t + C_t)$. TOC allows scientists to visualize these two ratios, while ROC cannot reveal these ratios.

Knowledge of the entries in the contingency table can help to identify segments of the curve that are particularly important for a specific application. For example, the point (P, H_t) on the TOC plot is where the diagnosed presence equals the amount of presence in the reference variable. This point is directly beneath the corner of the maximum boundary. Figure 2b highlights this point on the TOC curve by showing the threshold where the distance is 0.1 kilometers from disturbance at 1986. This point separates a left side from a right side of the curve. The left side of the TOC plot shows points for which the number of observations of diagnosed presence is less than the number of observations of reference presence. The right side of the TOC plot shows points for which the number of observations of diagnosed presence is greater than the number of observations of reference presence. ROC does not reveal this potentially important point where the number of observations of reference presence equals the number of observations of diagnosed presence.

TOC can be helpful to communicate the implications when extrapolating from a sample to an entire population. For example, suppose a sample of people participate in a laboratory study to determine the ability of an index to diagnose a disease, where the diagnosis depends on a threshold applied to the index. Sample data can generate the ROC curve to measure the diagnostic ability of the index. The ROC curve reveals how the selection of the threshold influences the conditional probability of a hit given that the person has the disease, and the conditional probability of a false alarm given that the person does not have the disease. If there were plans to apply the test to a larger population, then a scientist could use the information from the ROC curve to produce a TOC curve that estimates how the selection of the threshold influences the number of people in the population who would experience hits, misses, false alarms, and correct rejections. It is crucial for interpretation to know whether a study concerns a few observations or several thousand observations. TOC shows the number of observations, whereas ROC does not.

Equations (3) and (4) show how to convert from an ROC curve that derives from a sample to the corresponding estimated TOC curve that describes the population. The transformation requires two bits of information, $\hat{P}$ and $\hat{Q}$. The hats above the variables denote that the variables could be estimates.

$$\widehat{H_t + F_t} = \hat{P}Y_t + \hat{Q}X_t \tag{3}$$

$$\widehat{H}_t = \hat{P}Y_t \tag{4}$$

5. Conclusion

Our article proposes TOC as new approach to measure the diagnostic ability of an index variable for multiple thresholds. TOC reveals the total information that forms the basis of ROC. TOC presents its information in terms of the total number of observations, which facilitates interpretation beyond the two ratios that ROC offers for each threshold. We recommend scientists use TOC instead of ROC because TOC gives more information using the same amount of graphical space.

Acknowledgments

The United States' National Science Foundation (NSF) supported this research via award # OEC-1238212 entitled The PIE-LTER. Researchers at the Noel Kempff Mercado Natural History Museum developed the case study's data from Landsat TM imagery. The data are available at www.clarku.edu/~rpontius and in the tutorial for the software Idrisi (www.clarklabs.org).

References

Blough, D.S., 1967. Stimulus generalization as signal detection in pigeons. *Science*, 158 (3803), 940–941.

Bradley, A.P., 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30 (7), 1145–1159.

Cain, W.S., 1977. Differential sensitivity for smell: 'noise' at the nose. *Science*, 195 (4280), 796–798.

Camacho Olmedo, M.T., Paegelow, M., and Mas, J.F., 2013. Interest in intermediate soft-classified maps in land change model validation: suitability versus transition potential. *International Journal of Geographical Information Science*, 27 (12), 2343–2361.

Carterette, E.D. and Jones, M.H., 1967. Visual and auditory information processing in children and adults. *Science*, 156 (3777), 986–988.

Chen, Y., *et al*., 2013. Modeling urban land-use dynamics in a fast developing city using the modified logistic cellular automaton with a patch-based simulation strategy. *International Journal of Geographical Information Science*, 28 (2), 234–255.

Conway, T.M. and Wellen, C.C., 2011. Not developed yet? Alternative ways to include locations without changes in land use change models. *International Journal of Geographical Information Science*, 25 (10), 1613–1631.

Cook, N.R., 2007. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*, 115 (7), 928–935.

Dodd, L.E. and Pepe, M.S., 2003. Partial AUC estimation and regression. *Biometrics*, 59 (3), 614–623.

Eastman, J.R., Van Fossen, M., and Solorzano, L.A., 2005. Transition potential modeling for land cover change. *In*: D. Maguire, M. Batty, and M.F. Goodchild, eds. *GIS, Spatial analysis and modeling*. Redlands, CA: ESRI Press, 339–368.

Egan, J.P., 1975. *Signal detection theory and ROC analysis. Series in cognition and perception*. New York: Academic Press.

Fawcett, T., 2006. An introduction to ROC analysis. *Pattern Recognition Letters*, 27 (8), 861–874.

Fielding, A.H. and Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24 (1), 38–49.

Fleming, S.M., *et al*., 2010. Relating introspective accuracy to individual differences in brain structure. *Science*, 329 (5998), 1541–1543.

Gao, J. and Zhang, Y., 2012. Incorporating spectral data into logistic regression model to classify land cover: a case study in Mt. Qomolangma (Everest) National Nature Preserve. *International Journal of Geographical Information Science*, 26 (10), 1–18.

Gerçek, D., Toprak, V., and Strobl, J., 2011. Object-based classification of landforms based on their local geometry and geomorphometric context. *International Journal of Geographical Information Science*, 25 (6), 1011–1023.

Golicher, D., *et al*., 2012. Pseudo-absences, pseudo-models and pseudo-niches: pitfalls of model selection based on the area under the curve. *International Journal of Geographical Information Science*, 26 (11), 2049–2063.

Green, D.M. and Swets, J.A., 1966. *Signal detection theory and psychophysics*. New York: Wiley Press.

Hand, D.J. and Till, R.J., 2001. A simple generalization of the area under the ROC curve to multiple class classification problems. *Machine Learning*, 45 (2), 171–186.

Hanley, J.A. and McNeil, B.J., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143 (1), 29–36.

Hirzel, A.H., *et al*., 2006. Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modeling*, 199 (2), 142–152.

Jansen, L.J.M. and Veldkamp, T., 2011. Evaluation of the variation in semantic contents of class sets on modelling dynamics of land-use changes. *International Journal of Geographical Information Science*, 26 (4), 717–746.

Jeong, M.-H., *et al*., 2013. Decentralized and coordinate-free computation of critical points and surface networks in a discretized scalar field. *International Journal of Geographical Information Science*, 28 (1), 1–21.

Jiang, Y.L., Metz, C.E., and Nishikawa, R.M., 1996. A receiver operating characteristic partial area index for highly sensitive diagnostic tests. *Radiology*, 201 (3), 745–750.

Johnson, D.S., *et al*., 2007. Genome-wide mapping of in vivo protein-DNA interactions. *Science*, 316 (5830), 1497–1502.

Johnson, J.M., *et al*., 2003. Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*, 302 (5653), 2141–2144.

Jolliffe, I.T. and Stephenson, D.B., 2003. *Forecast verification a practitioner's guide in atmospheric science*. Chichester: Wiley.

Knaus, W.A., Wagner, D.P., and Lynn, J., 1991. Short-term mortality predictions for critically ill hospitalized adults: science and ethics. *Science*, 254 (5030), 389–394.

Kolb, M., Mas, J.-F., and Galicia, L., 2013. Evaluating drivers of land-use change and transition potential models in a complex landscape in Southern Mexico. *International Journal of Geographical Information Science*, 27 (9), 1804–1827.

Leitão, P.J., Moreira, F., and Osborne, P.E., 2011. Effects of geographical data sampling bias on habitat models of species distributions: a case study with steppe birds in southern Portugal. *International Journal of Geographical Information Science*, 25 (3), 439–454.

Li, L., Wang, J., and Leung, H., 2010. Using spatial analysis and Bayesian network to model the vulnerability and make insurance pricing of catastrophic risk. *International Journal of Geographical Information Science*, 24 (12), 1759–1784.

Lin, Y., *et al*., 2010. Predictive ability of logistic regression, auto-logistic regression and neural network models in empirical land-use change modeling – a case study. *International Journal of Geographical Information Science*, 25 (1), 65–87.

Lobo, J.M., Jimenez-Valverde, A., and Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17 (2), 145–151.

Lusted, L.B., 1971. Signal detectability and medical decision-making. *Science*, 171 (3977), 1217–1219.

Manzo, G., *et al*., 2012. GIS techniques for regional-scale landslide susceptibility assessment: the Sicily (Italy) case study. *International Journal of Geographical Information Science*, 27 (7), 1433–1452.

Mas, J.-F., *et al*., 2013. A suite of tools for ROC analysis of spatial models. *ISPRS International Journal of Geo-Information*, 2 (3), 869–887.

Mateo Sánchez, M.C., Cushman, S.A., and Saura, S., 2013. Scale dependence in habitat selection: the case of the endangered brown bear (*Ursus arctos*) in the Cantabrian Range (NW Spain). *International Journal of Geographical Information Science*. doi:10.1080/13658816.2013.776684

Miller, J. and Franklin, J., 2002. Modeling the distribution of four vegetation alliances using generalized linear models and classification trees with spatial dependence. *Ecological Modeling*, 157 (2), 227–247.

Nevin, J.A., 1965. Decision theory in studies of discrimination in animals. *Science*, 150 (3699), 1057.

Nuechterlein, K.H., Parasuraman, R., and Jiang, Q., 1983. Visual sustained attention: image degradation produces rapid sensitivity decrement over time. *Science*, 220 (4594), 327–329.
Overmars, K.P. and Verburg, P.H., 2005. Analysis of land use drivers at the watershed and household level: linking two paradigms at the Philippine forest fringe. *International Journal of Geographical Information Science*, 19 (2), 125–152.
Paegelow, M. and Olmedo, M.T.C., 2005. Possibilities and limits of prospective GIS land cover modelling – a compared case study: Garrotxes (France) and Alta Alpujarra Granadina (Spain). *International Journal of Geographical Information Science*, 19 (6), 697–722.
Pérez-Vega, A., Mas, J.-F., and Ligmann-Zielinska, A., 2012. Comparing two approaches to land use/cover change modeling and their implications for the assessment of biodiversity loss in a deciduous tropical forest. *Environmental Modelling & Software*, 29 (1), 11–23.
Peters, J., *et al*., 2011. Synergy of very high resolution optical and radar data for object-based olive grove mapping. *International Journal of Geographical Information Science*, 25 (6), 971–989.
Peterson, A.T., Papes, M., and Soberón, J., 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modeling*, 213 (1), 63–72.
Pontius, R.G. and Batchu, K., 2003. Using the relative operating characteristic to quantify certainty in prediction of location of land cover change in India. *Transactions in GIS*, 7 (4), 467–484.
Pontius, R.G. and Pacheco, P., 2004. Calibration and validation of a model of forest disturbance in the Western Ghats, India 1920–1990. *GeoJournal*, 61 (4), 325–334.
Reed, A.V., 1973. Speed-accuracy trade-off in recognition memory. *Science*, 181 (4099), 574–576.
Saatchi, S., *et al*., 2008. Modeling distribution of Amazonian tree species and diversity using remote sensing measurements. *Remote Sensing of Environment*, 112 (5), 2000–2017.
Smith, E.L.D., *et al*., 1982. Color vision is altered during the suppression phase of binocular rivalry. *Science*, 218 (4574), 802–804.
Stephan, C., *et al*., 2003. Comparison of eight computer programs for receiver-operating characteristic analysis. *Clinical Chemistry*, 49 (3), 433–439.
Swets, J.A., 1961. Is there a sensory threshold? *Science*, 134 (3473), 168–177.
Swets, J.A., 1963. Information retrieval systems. *Science*, 141 (3577), 245–250.
Swets, J.A., 1973. The relative operating characteristic in psychology. *Science*, 182 (4116), 990–1000.
Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. *Science*, 240 (4857), 1285–1293.
Swets, J.A., *et al*., 1979. Assessment of diagnostic technologies. *Science*, 205 (4408), 753–759.
Wang, J. and Mountrakis, G., 2010. Developing a multi-network urbanization model: a case study of urban growth in Denver, Colorado. *International Journal of Geographical Information Science*, 25 (2), 229–253.
Wang, N., *et al*., 2013. Comparative performance of logistic regression and survival analysis for detecting spatial predictors of land-use change. *International Journal of Geographical Information Science*, 27 (10), 1960–1982.
Wiley, E.O., *et al*., 2003. Niche modeling and geographic range predictions in the marine environment using machine-learning algorithm. *Oceanography*, 16 (3), 120–127.

Appendix

This is a mathematical proof that the AUC of the ROC equals the ratio of the area under the TOC curve that is within the TOC's parallelogram to the area of the TOC's parallelogram. Figure 3 illustrates the concepts that the proof uses.

We use the notation in Table 2 of the manuscript and the following additional notation:

T = number of thresholds.
G = area under TOC curve.
B = ratio of area under the TOC curve that is within TOC's bounding parallelogram to the area of TOC's bounding parallelogram.
A = area under ROC curve.
The goal is to show $B = A$.

The area under the TOC curve that is also within the parallelogram is $G - P^2/2$, because the area in the lower right corner of the TOC plot that is below the minimum line is $P^2/2$.

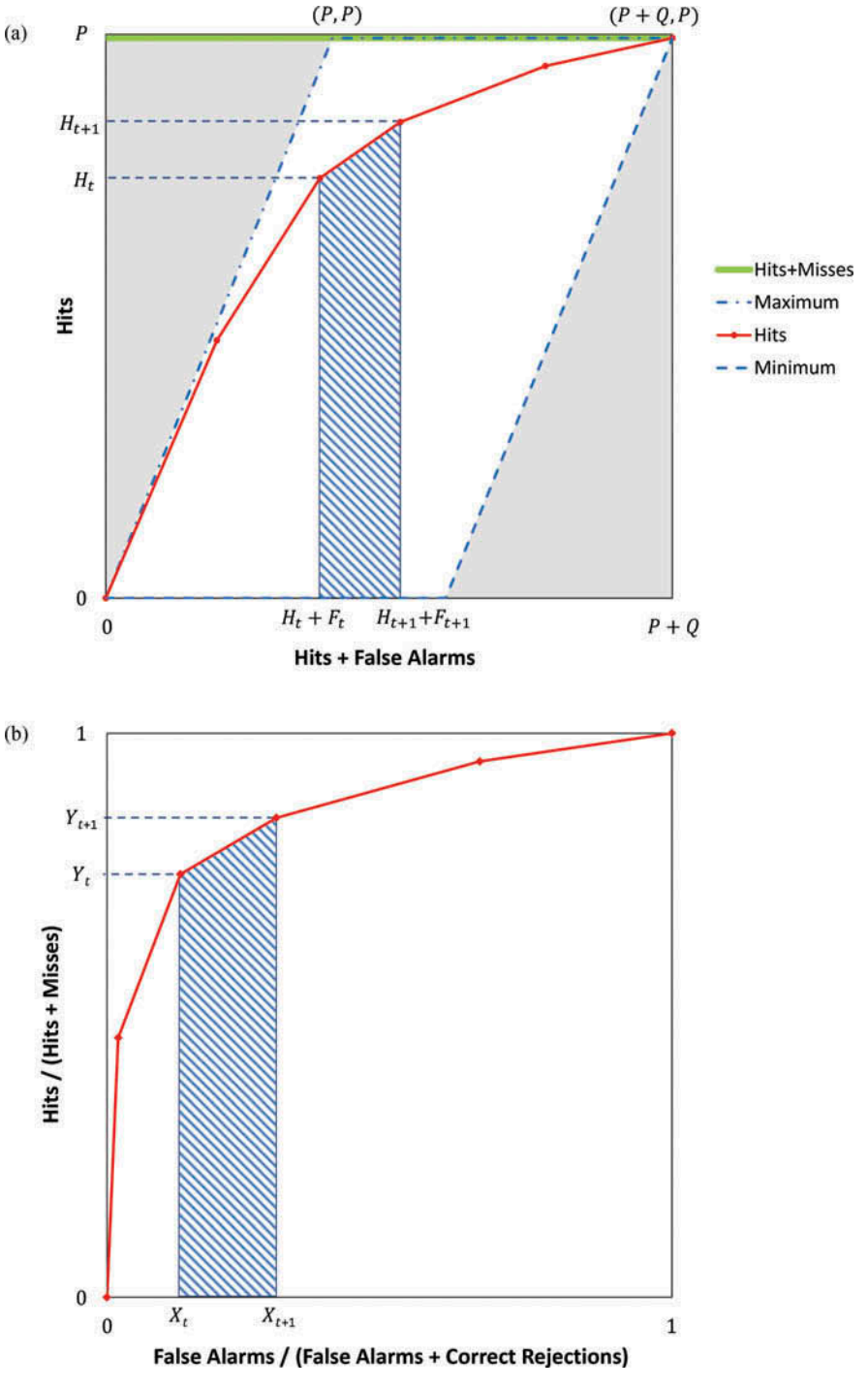

Figure 3. (a) TOC and (b) ROC for the mathematical proof.

The area within the bounding parallelogram is $P(P + Q) - P^2$, because the rectangular space of the TOC plot is $P(P + Q)$ and the space outside the parallelogram is P^2. Thus,

$$B = \frac{G - P^2/2}{P(P+Q) - P^2}$$

G denotes the area under the TOC curve, which equals the sum of areas in each trapezoid defined by each interval $[H_t + F_t, H_{t+1} + F_{t+1}]$ on the horizontal axis in Figure 3a. The area under a single trapezoid of the TOC is the area of the lower rectangular portion of the trapezoid plus the upper triangle of the trapezoid, thus is

$$\{H_t[(H_{t+1} + F_{t+1}) - (H_t + F_t)]\} + \{[H_{t+1} - H_t][(H_{t+1} + F_{t+1}) - (H_t + F_t)]\}/2$$

which is equal to $[H_{t+1} + H_t][(H_{t+1} + F_{t+1}) - (H_t + F_t)]/2$.

When we sum over all T-1 trapezoids of the TOC, we get

$$G = \sum_{t=1}^{T-1}\{[H_{t+1} + H_t][(H_{t+1} + F_{t+1}) - (H_t + F_t)]/2\}$$

By substituting G into the equation above for B, we get

$$\begin{aligned}
B &= \frac{\sum_{t=1}^{T-1}\{[H_{t+1} + H_t][(H_{t+1} + F_{t+1}) - (H_t + F_t)]/2\} - P^2/2}{P(P+Q) - P^2} \\
&= \frac{\sum_{t=1}^{T-1}\{[H_{t+1} + H_t][(H_{t+1} + F_{t+1}) - (H_t + F_t)]\} - P^2}{2[P^2 + PQ - P^2]} \\
&= \frac{\sum_{t=1}^{T-1}\{[H_{t+1} + H_t][(H_{t+1} - H_t) + (F_{t+1} - F_t)]\} - P^2}{2PQ} \\
&= \frac{\sum_{t=1}^{T-1}\{[H_{t+1} + H_t][H_{t+1} - H_t] + [H_{t+1} + H_t][F_{t+1} - F_t]\} - P^2}{2PQ} \\
&= \frac{\sum_{t=1}^{T-1}\left\{\left[(H_{t+1})^2 - (H_t)^2\right] + [H_{t+1} + H_t][F_{t+1} - F_t]\right\} - P^2}{2PQ} \\
&= \frac{(H_T)^2 - (H_1)^2 + \sum_{t=1}^{T-1}\{[H_{t+1} + H_t][F_{t+1} - F_t]\} - P^2}{2PQ} \\
&= \frac{P^2 - 0^2 + \sum_{t=1}^{T-1}\{[H_{t+1} + H_t][F_{t+1} - F_t]\} - P^2}{2PQ} \\
&= \frac{\sum_{t=1}^{T-1}\{[H_{t+1} + H_t][F_{t+1} - F_t]\}}{2PQ} \\
&= \sum_{t=1}^{T-1}\left[\frac{H_{t+1}}{P} + \frac{H_t}{P}\right]\left[\frac{F_{t+1}}{Q} - \frac{F_t}{Q}\right]/2
\end{aligned}$$

Since $X_t = \frac{F_t}{Q}$ and $Y_t = \frac{H_t}{P}$, we get

$$B = \sum_{t=1}^{T-1} \left[\frac{H_{t+1}}{P} + \frac{H_t}{P}\right] \left[\frac{F_{t+1}}{Q} - \frac{F_t}{Q}\right] /2 = \sum_{t=1}^{T-1} [Y_{t+1} + Y_t][X_{t+1} - X_t]/2 = A$$

Consequently, $B = A$, which is what we wanted to show.