Biological Conservation 191 (2015) 692–699

Contents lists available at ScienceDirect

Biological Conservation

journal homepage: www.elsevier.com/locate/bioc

Changes in protected area management effectiveness over time: A global analysis

Jonas Geldmann [a,*], Lauren Coad [b,c], Megan Barnes [d], Ian D. Craigie [e], Marc Hockings [c,f], Kathryn Knights [g], Fiona Leverington [f,g], Ivon C. Cuadros [g], Camilo Zamora [g], Stephen Woodley [h], Neil D. Burgess [a,c]

[a] Center for Macroecology, Evolution and Climate, Natural History Museum of Denmark, University of Copenhagen, Universitetsparken 15, 2100 Copenhagen E, Denmark
[b] Environmental Change Institute, School of Geography, University of Oxford, OX1 3QY, UK
[c] United Nations Environment Programme World Conservation Monitoring Centre, 219 Huntingdon Road, Cambridge CB3 0DL, UK
[d] Centre of Excellence for Environmental Decisions, The University of Queensland, Brisbane, QLD 4072, Australia
[e] ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, QLD 4811, Australia
[f] School of Geography, Planning and Environmental Management, University of Queensland, St Lucia, Brisbane, Australia
[g] Protected Area Solutions, Brisbane, Australia
[h] World Commission on Protected Areas, IUCN, 64 Juniper Road, Chelsea, QC J9B1T3, Canada

ARTICLE INFO

Article history:
Received 18 May 2015
Received in revised form 16 July 2015
Accepted 9 August 2015
Available online 1 September 2015

Keywords:
Management effectiveness
Protected area
Aichi target 11
Governance
Management effectiveness tracking tool

ABSTRACT

Protected area coverage has reached over 15% of the global land area. However, the quality of management of the vast majority of reserves remains unknown, and many are suspected to be "paper parks". Moreover, the degree to which management can be enhanced through targeted conservation projects remains broadly speculative. Proven links between improved reserve management and the delivery of conservation outcomes are even more elusive. In this paper we present results on how management effectiveness scores change in protected areas receiving conservation investment, using a globally expanded database of protected area management effectiveness, focusing on the "management effectiveness tracking tool" (METT). Of 1934 protected areas with METT data, 722 sites have at least two assessments. Mean METT scores increased in 69.5% of sites while 25.1% experienced decreases and 5.4% experienced no change over project periods (median 4 years). Low initial METT scores and longer implementation time were both found to positively correlate with larger increases in management effectiveness. Performance metrics related to planning and context as well as monitoring and enforcement systems increased the most while protected area outcomes showed least improvement. Using a general linear mixed model we tested the correlation between change in METT scores and matrices of 1) landscape and protected area properties (i.e. topography and size), 2) human threats (i.e. road and human population density), and 3) socio-economics (i.e. infant mortality rate). Protected areas under greater threat and larger protected areas showed greatest improvements in METT. Our results suggest that when funding and resources are targeted at protected areas under greater threat they have a greater impact, potentially including slowing the loss of biodiversity.

1. Introduction

Protected areas are one of the most important conservation tools for protecting biodiversity and ecosystem services (Naidoo et al., 2006; Rodrigues, 2006; Klein et al., 2007; Coad et al., 2008; Scharlemann et al., 2010). This has led to the development of a global network of protected areas covering more than 15.4% of the terrestrial land surface (Juffe-Bignoli et al., 2014). Despite this extensive coverage, biodiversity continues to decline (Butchart et al., 2010; Tittensor et al., 2014) and protected areas are not immune to biodiversity and habitat loss (Craigie et al., 2010; Laurance et al., 2012; Geldmann et al., 2013), or increases in human-caused pressure (Geldmann et al., 2014).

Expanding the coverage of protected areas has been suggested as a strategy to mitigate the present negative biodiversity trajectories (Target 11, Aichi Biodiversity Targets, Convention on Biological Diversity, 2010) and as much as a third of the total global terrestrial area is estimated to be necessary to fully meet all elements of Target 11 (Butchart et al., 2015). However, coverage is only one aspect of protected area performance and effectiveness. Protected areas need to be managed effectively within appropriate legal frameworks and governance structures to meaningfully contribute to halting the loss of biodiversity (Leverington et al., 2010; Watson et al., 2014). Given declines in biodiversity continue even within protected area boundaries (Butchart et al., 2010; Tittensor et al., 2014) it is probable that current levels of

* Corresponding author.
E-mail address: Jgeldmann@snm.ku.dk (J. Geldmann).

http://dx.doi.org/10.1016/j.biocon.2015.08.029

management within protected areas at a global scale are insufficient to "*halt the loss of biodiversity*" (Watson et al., 2014). Allocating conservation funds cost-effectively to achieve maximum conservation benefit is therefore a key question in conservation science (Wilson et al., 2006).

Measuring whether protected area management improves over time, as well as understanding what external factors affect the observed changes in management, is a crucial benchmark for Aichi Target 11 and the overall delivery of the CBD Strategic Plan (Convention on Biological Diversity, 2010). While tools such as the World Database on Protected Areas (WDPA) provide information on the locations, number, and size of more than 210,000 protected areas, information on the quality of management, or biological outcomes within the same sites is much scarcer. Protected Area Management Effectiveness (PAME) assessments have been used in many countries to evaluate the strengths and weaknesses of protected area management, and help guide improvement to the conservation delivery of these areas (Leverington et al., 2010). The IUCN World Commission on Protected Areas (WCPA) has developed an evaluation framework for management effectiveness assessment allowing specific evaluation methodologies to be designed within a consistent overall approach (Hockings, 2003; Hockings et al., 2006). In general, PAME assessments are conducted by one or more of: protected area managers, government agency employees and donor institutions including NGOs. Most PAME tools are questionnaires measuring the management inputs, activities, and outputs associated with a conservation intervention, in order to assess management strengths, weaknesses, and needs (Mascia et al., 2014). Evaluation tools generally rely on qualitative indicators to assess management success and are therefore heavily dependent on knowledge amongst protected area stakeholders (Cook and Hockings, 2011; Cook et al., 2014).

To date more than 18,000 PAME evaluations have been conducted using 95 methodologies in over 9000 protected areas across 180 countries (Coad et al., accepted). These provide baseline data to evaluate management performance and are also used as one of the indicators for tracking international commitments to halt the loss of biodiversity (that is, the 2020 Aichi targets inviting "…Parties to…expand and institutionalize management effectiveness assessments to work towards assessing 60% of the total area of protected areas by 2015 using various national and regional tools and report the results into the global database on management effectiveness…" (CBD Aichi Targets, COP 10 Decision X/31, 19a)). This target has only been reached by 17.5% of all countries (Coad et al., accepted).

Protected areas undergoing multiple and systematic evaluations often represent protected areas with outside investments from donor organizations (e.g. the World Bank, the Global Environment Facility (GEF), WWF) or reserves where there is increased national focus on improving the management and governance foundation. Additionally, some countries, for example Australia, Colombia and South Africa, have implemented systematic repeated PAME assessments to track changes in management. However, the implementation of PAME evaluations in itself is no panacea for improving or fully understanding protected area delivery of ecological and social outcomes (Coad et al., accepted). It does however provide valuable information on the potential of protected areas to secure biodiversity and, in the absence of appropriate data on the status of and trends in biological attributes, can serve as a proxy of protected area performance (Kleiman et al., 2000; Hockings et al., 2006). In addition, anecdotal evidence suggest that the process of evaluation often leads to management improvements, through protected area managers sharing information and redirecting resources to the most serious issues.

Previous analyses have looked at the global coverage of PAME evaluations (Coad et al., 2013) as well as mean management effectiveness scores and strengths and weaknesses (Leverington et al., 2010). These analyses address whether protected areas are being evaluated for management effectiveness and calculate average total evaluation scores, as well as average scores for individual elements, at a global scale. While we are aware of repeat evaluations being analyzed at agency or protected area level in a number of cases, most of this information is unpublished, and the scarcity of repeat evaluations has meant that only limited analysis of trends in scores has so far been possible at a global scale.

Here we use one of the most widely used PAME tools; the management effectiveness tracking tool (METT) (Stolton et al., 2007) to complete a global analysis of relevance to international policy and practices. We restrict our analysis to protected areas where METT assessments have been conducted multiple times so that we can investigate how management and governance change over time. We map the global distribution of sites where METTs have been repeated and use these sites to derive general statistics on the general direction of changes in management, and the characteristics of countries where these assessments occur. Using theories of management and governance we further analyze which dimensions of management and governance have changed most substantially. Finally we use a suite of globally collected and validated contextual variables covering protected area attributes, landscape, human pressure, and socioeconomic context to understand what determines changes in management effectiveness.

2. Methods

2.1. The management effectiveness tracking tool

METT assessments collect information on 1) objectives, 2) threats, 3) budgets, 4) staffing, 5) size, and 6) designations of protected areas. METT also documents the status of 30 specific management-elements ranging from legal status, equipment, and quality of management plans, to outreach programs and tourist facilities (Table A1). Each METT assessment is conducted by local assessors who assign scores on a four point scale from 0 to 3 depending on the status of the specific management element (for example law enforcement: 0 = The staff have no effective capacity/resources to enforce protected area legislation and regulations, 1 = There are major deficiencies in staff capacity/resources to enforce protected area legislation and regulations (e.g. lack of skills, no patrol budget), 2 = The staff have acceptable capacity/resources to enforce protected area legislation and regulations but some deficiencies remain, and 3 = The staff have excellent capacity/resources to enforce protected area legislation and regulations) (Stolton et al., 2007). Several local adaptations of the METT evaluation exist, based on experiences and needs from protected area managers, organizations and country officials (Coad et al., 2013).

We extracted all METT assessments from the global database on PAME assessments (Coad et al., accepted). This database was started in 2006 as a research project with the University of Queensland (Leverington et al., 2008, 2010) and has been used by UNEP-World Conservation Monitoring Centre to provide data on management effectiveness through the Biodiversity Indicator Partnership (Walpole et al., 2009) and as a key tool for measuring CBD Aichi Target 11 (e.g. Tittensor et al., 2014). METT evaluations originate from a range of sources, including NGOs, national governments and international agencies (e.g. WWF, the World Bank and the GEF). From the METT assessments included in this analysis we selected a random subset of 88 from the database for which we calculated the error rate between the original data sheet and the database entry. We found an error rate of 2.5% (Table A2). New METT assessments are still being collected and entered into the database.

From all available METT assessments (n = 4748) we identified all PAs that had multiple entries (n = 933). From these we kept only sites with at least one year between first and last assessment. Where more than two METT assessments existed from different years we used the earliest and most recent to provide the greatest number of years between assessments. Subsequently we removed all protected areas where year of assessment was missing, or where less than 10 of the 30 questions were answered (n = 722).

The total number of METT questions is 30, but as two questions have been changed these 30 are not comparable over time. We removed these two questions — leaving 28 questions for analysis. We then calculated a total standardized management score (TSMS) for each protected area as follows:

$$\mathrm{TSMS} = \frac{\left(\sum_{i=0}^{n=30} Score\right) \cdot \mathrm{Questions_{max}}}{(number\ of\ questions \neq 0) \cdot \mathrm{Score_{max}}} \cdot 100 \qquad (1)$$

Where $\mathrm{Questions_{max}}$ is the maximum number of possible questions included (n = 28) and $\mathrm{Score_{max}}$ is the maximum score possible from the 28 questions (n = 84). The TSMS is adjusted for the completeness of individual assessments to avoid deflating scores where questions not applicable to the individual protected area are missing or have been omitted for unknown reasons. TSMS is reported on a scale of 0–100.

2.2. Grouping METT variables

We assessed the 28 questions based on six categories drawn from established theory and dealing with different aspects of management and governance (Table A1). We first evaluated which questions could be attributed to either of the three of the four dimensions within the common pool resource framework included in the METT: 1) decision making arrangements, 2) resource use rights, and 3) monitoring and enforcement systems (Ostrom, 1990, 2009). Subsequently we used the IUCN World Commission on Protected Areas (WCPA) frameworks for management effectiveness covering: 1) design and planning, 2) appropriateness of management systems and processes, and 3) delivery of protected area objectives (Hockings, 2003; Hockings et al., 2006). These categories represent different aspects and successive steps in the management cycle and are all crucial in achieving effectively managed protected area. Visitor facilities (question 24) and fees (question 26) were excluded as these did not fit the framework. This division of questions in to overall theoretically recognized units allows us to separate elements of management to investigate how these differ in speed and success of implementation over time.

2.3. Contextual variables

To understand under what circumstances METT scores changed we compiled independent spatially explicit variables across four domains: 1) protected area attributes, 2) landscape, 3) human pressure, and 4) socioeconomic context. In total eight variables were used in the modeling: 1) the size of the protected areas (IUCN and UNEP-WCMC, 2014), 2) mean elevation and 3) median slope from SRTM v.4 (Jarvis et al., 2008), 4) mean human influence index (HII) from the human footprint (Sanderson et al., 2002), 5) density of roads in a 10 km buffer around the protected areas from gRoads (Center for International Earth Science Information Network, 2013), 6) human population density from the Global Rural–Urban Mapping Project, Version 1 (GRUMPv1) in a 10 km buffer around the protected area (Center for International Earth Science Information Network – CIESIN – Columbia University et al., 2011), 7) infant mortality rate in a 10 km buffer around the protected area, and 8) the national Human Development Index (HDI).

2.4. Statistical analysis

All analyses were conducted in R 3.20.1 (R Development Core Team, 2015). Paired Students t-tests were used to compare the difference in accumulated METT scores between the first and last evaluation for each site, as well as the difference between individual questions. A general linear model (GLM) was used to investigate the correlation between the difference in accumulated METT scores between the first and last assessment for each site, and the length of time in years between the first and last assessment. To account for the fact that initially well managed protected areas can increase less than initially poorly managed protected area; changes in METT scores for individual questions and the six management dimensions were standardized based on the initial METT scores:

$$Standardized\ change = \frac{\mathrm{TSMS_{t=1}} - \mathrm{TSMS_{t=0}}}{100 - \mathrm{TSMS_{t=0}}} \cdot 100 \qquad (2)$$

Where $\mathrm{TSMS_{t=1}}$ is the final score and $\mathrm{TSMS_{t=0}}$ the initial score. A linear mixed effects model (GLMM) assuming a Gaussian error distribution was used to investigate the correlation between changes in METT scores and the chosen contextual variables, with country ID as a random factor. Independent variables were transformed based on a Box Cox power test (Fig. A3). All independent variables were standardized by subtracting the mean and dividing by the standard deviation to make effect sizes comparable. We assessed the collinearity between all independent variables and only used variables in combination without significant inter-correlation. Where variables were inter-correlated we used the variable with the best fit based on univariate models. The initial TSMS of each protected area as well as the number of years between first and last assessment were included in the model to account for the duration as well as the potential larger increase in protected areas with initially lower scores. Substituting one collinear variable for another did not change the overall direction or magnitude of the model. However, a stepwise reduction, removing all non-significant terms was conducted, showing no change in direction or magnitude of significant variables between the two models (Table A6).

3. Results

Multiple METT assessments were available for 722 protected areas from 74 countries (Fig. 1, Table A3). Of these, 502 (69.5%) experienced improvements in the total METT score while 181 (25.1%) experienced decreases and 39 (5.4%) experienced no change. The mean duration between first and last assessment was 4.46 years (S.E. = 0.08 years, median = 4 years). There was a significant increase ($t = 8.8$, $p < 0.0001$) in overall scores (TSMS) between first (mean = 46.10, S.E. = 0.66) and last (mean = 54.66, S.E. = 0.66) assessment for the entire METT portfolio.

When individual METT questions were grouped based on management (Hockings, 2003; Hockings et al., 2006) and governance (Ostrom, 1990, 2009); dimensions related to initial steps in protected area establishment (i.e. designs and planning and resource user rights) increased the most, followed by aspects often requiring more time to implement (i.e. improving monitoring, enforcements, and management processes). Delivery of protected area objectives increased the least, and improvements were primarily driven by improved educational and outreach activities rather than improved conservation outcomes (Fig. 2).

For the individual questions the mean score improved significantly in 27 of 28 questions. The exception was the indicator on condition assessment (e.g. the state of biodiversity, question 27) (Table A4). The greatest improvements were observed for legal gazettement (question 1) followed by the adequacy of management plans (question 7), protected area boundaries (question 6), protected area objectives (question 4), and protected area regulation (question 2) (Fig. 3). We found a significant negative correlation between starting score and change in scores, so that protected areas with a lower initial score were significantly more likely to increase (Fig. A2).

To test whether increased time between assessments resulted in greater improvements in management and governance scores we examined the correlation between change in total score and years between the first and last METT assessment. There was a highly significant positive effect of time, measured in years, ($t = 5.469$, $p < 0.001$), which varied across governance and management dimensions. Decision making arrangements as well as design and planning showed the greatest improvements, while delivery of protected area objectives showed the smallest effect of implementation time (Fig. 4). For all six dimensions considerable variation was observed.

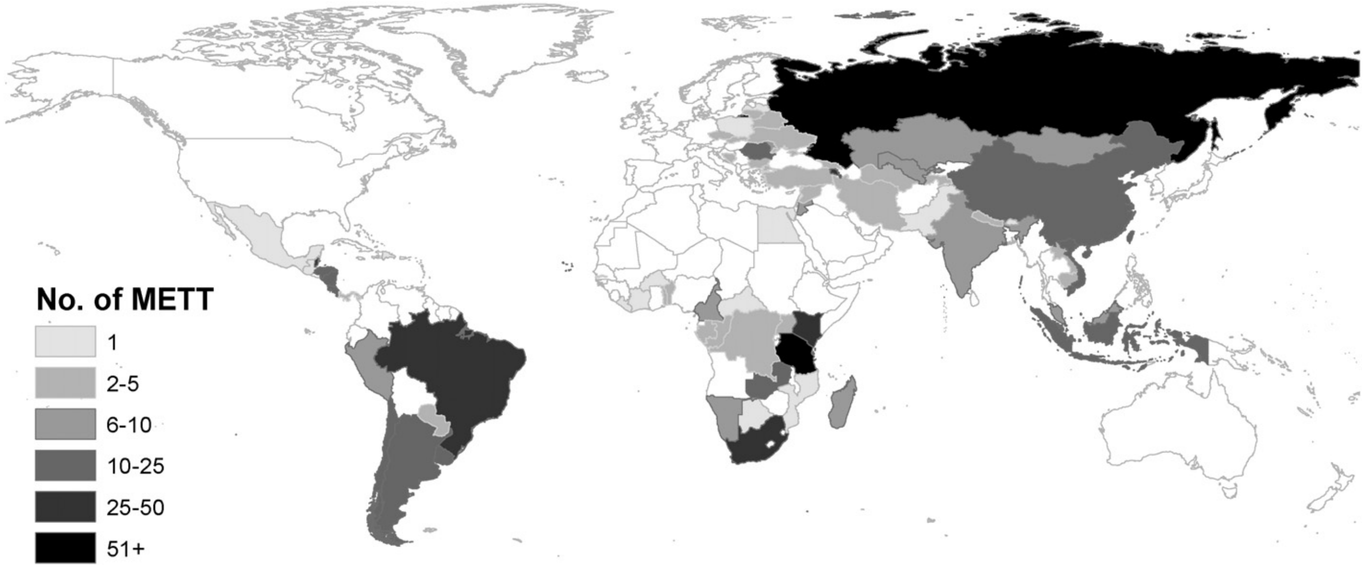

Fig. 1. The 575 protected areas in the WDPA for which multiple METT assessments existed. Ascending darker colors show the number of repeated METT assessments in individual countries. White are countries for which no repeated METT assessments existed. For full details on individual countries see Table A3.

As changes in management effectiveness are not independent of external factors we tested correlation between changes in scores and nine factors covering protected area characteristics, human pressures, and socio-economic factors. The full model, after removing collinear variables, contained 1) mean elevation, 2) roads density in the buffer, 3) human population density in the buffer, 4) the size of the protected area, 5) local infant mortality rate as well as 6) the initial METT score of the protected area and 7) the time in years between first and last assessment (Table 1).

The initial score had a small negative effect, while increased time between first and last assessment, the size of the protected areas and surrounding human population density positively affected the change in score. Infant mortality rate, road density and the elevation of the protected area had no effect in our model (Fig. 5).

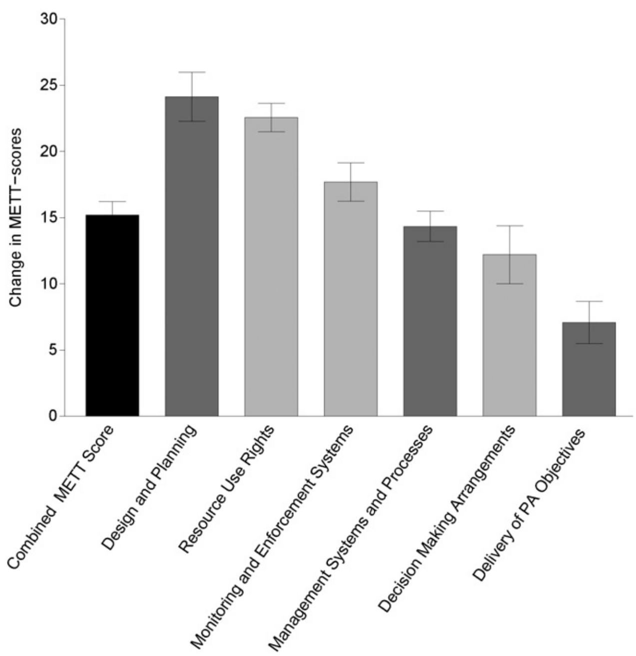

Fig. 2. Mean score change for the combined assessment and each of the six elements from the first and last assessment at a reserve. The error bars are standard error. All scales have been normalized to reflect a possible score between 0 and 100 for each element. Scores are normalized based on the mean starting value of the category. Black is the mean overall score for all 28 questions, light gray colors indicate governance dimensions and dark gray indicates management dimensions.

4. Discussion

4.1. A progression in interventions

The observed patterns of improvements follow general management theory, where the project management cycle begins with an initial conceptualization phase, followed by planning, implementation, analysis and adaptation, and learning (Conservation Measures Partnership, 2007) leading to hypothesized causal chain of events where project or program inputs increase, this results in activities, outputs, outcomes, and impacts (Kellogg Foundation W.K., 2004). Changes across the three dimensions of management (Hockings et al., 2003, 2006) and three dimensions of governance (Ostrom, 1990, 2009) suggest a similar progression. The largest increases in scores were observed in elements of context and planning, as well as the establishment of formal resource user rights around the protected area. These were followed by smaller improvements in elements related to enforcement and monitoring relaying on increased staffing capacity, equipment, and education. Elements related to stakeholder involvements (decision making arrangements) as well as improvements in actual conservation outcomes changed the least.

We found no significant change in the score for biological outcomes (question 27), as measured by the METT between the first and last assessment. While this is potentially of concern, it is important to highlight that the METT assessment is primarily a tool for capturing processes and inputs, rather than outcomes (Mascia et al., 2014) Biological outcomes has only one question in the METT and this is unlikely to capture the complexity of assessing ecological condition in a protected area. Further, change in biological indices generally occurs over longer time than the 2–8 years sample period we have for most sampled protected areas (Mace et al., 2010). To fully evaluate progress towards outcomes and objectives, it requires independent measures such as remotely sensed land cover change or changes in species abundance should be collected and used as part of the evaluation process (Nolte and Agrawal, 2013; Carranza et al., 2014; Henschel et al., 2014). In addition, increased information on outcomes collected through improved monitoring and an increased focus on evaluation may lead to decreased scores over time, as the poor state of biodiversity becomes more apparent to assessors.

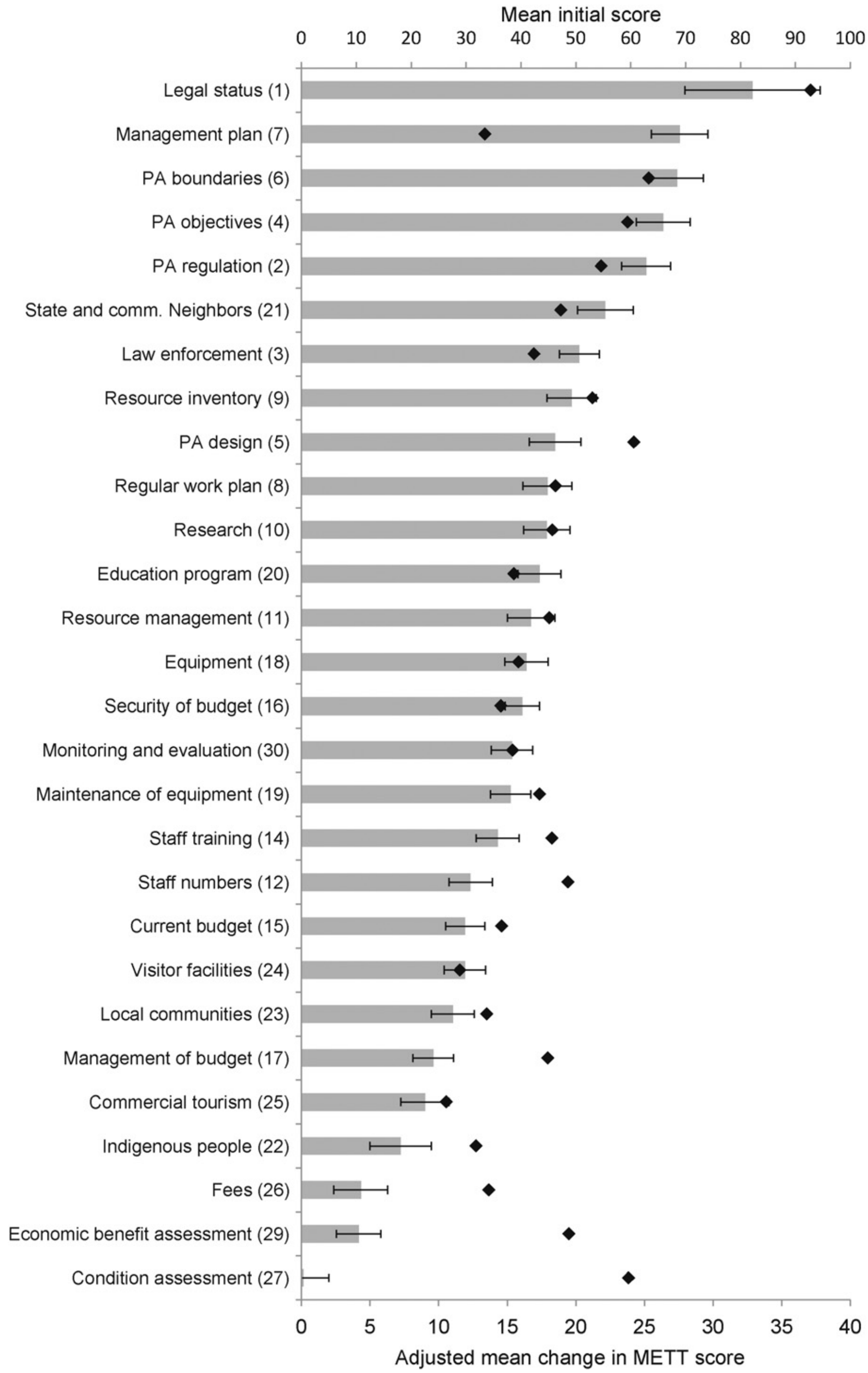

Fig. 3. Standardized mean change in scores between first and last assessment from the individual questions ranked by the amount of change (bottom axis). Diamonds show the mean initial TSMS value for the individual question (top axis). Numbers in parentheses represent the question number as they appear in the METT score card v. 1. Questions 13 and 28 were not kept in the updated version of the METT. As a consequence these could not be analyzed and are not shown on the figure. Error bars are based on the standard error.

4.2. What determines change?

Protected area size had a significant correlation with increases in management effectiveness scores. Similar patterns have been found both for terrestrial (Struhsaker et al., 2005; Joppa and Pfaff, 2011) and marine protected areas (Edgar et al., 2014). However whether available resources are better invested in larger protected areas, often located in remote areas (Joppa and Pfaff, 2009) is a heated discussion. Larger protected areas have a proportionately smaller edge to pressures and through their remote location often result in relatively intact biodiversity values (Mittermeier et al., 2003) making them attractive for further conservation investment. However, large and remote protected areas often experience little pressure, protected or not (Joppa and Pfaff, 2011) suggesting resources may be better invested in smaller protected areas which would otherwise experience greater biodiversity loss (Craigie et al., 2014).

We found a significant positive correlation between local human population density and changes in management effectiveness scores. This effect remained when substituting population density with the human footprint HII (Sanderson et al., 2002). This pattern is perhaps not surprising. Greater human density works as an increased pressure on biodiversity, suggesting a higher need for protected area

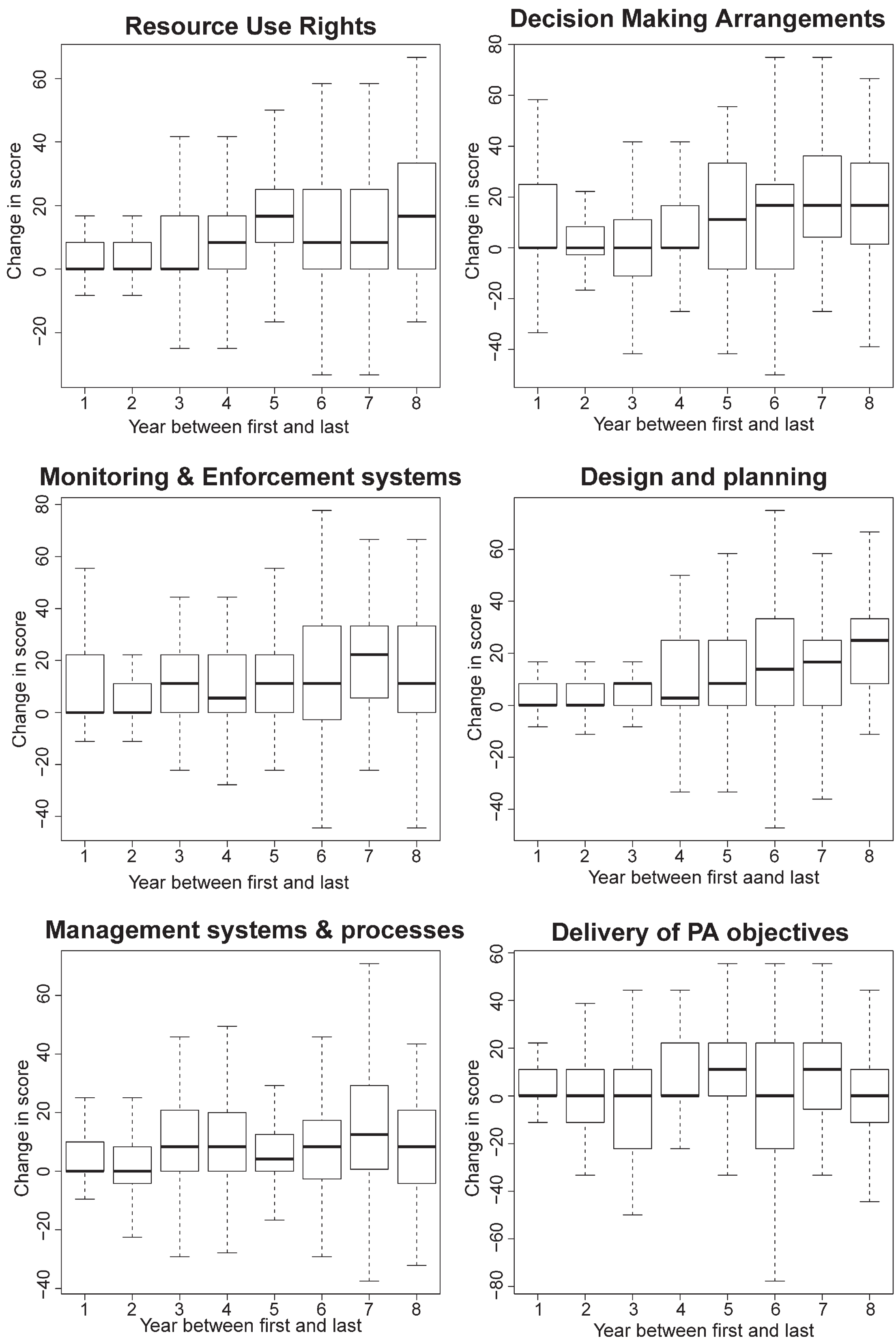

Fig. 4. Change over time for the total score and the six categories. All categories increased statistically significant with time. The horizontal line shows the median increase. The bottom and top of the box show the 25th and 75th percentiles, respectively. The top and bottom of the bars mark the maximum and minimum values.

management inputs. At the same time, resources to improve management effectiveness may very likely be directed towards highly populated areas, to improve management effectiveness in protected areas which are used and valued by people. Finally, protected areas with high surrounding populations often contain exceptionally rich biodiversity (Balmford et al., 2001), thus making them obvious targets for

Table 1
Model output from the best fit general linear model testing for the impact of contextual variables on the scores changes in METT assessments over time.

Variable	Estimate	S.E.	*t*-value
Intercept	33.102	4.123	8.029
Log mean elevation	0.637	0.990	0.644
Log density of roads (10 km buffer)	0.665	0.549	1.211
Log mean population density (10 km buffer)	2.438	0.692	3.521
Log of protected area size	1.051	0.363	2.890
Log infant mortality rate (10 km buffer)	0.080	2.685	0.030
Initial score	−0.467	0.037	−12.661
Years between first and last assessment	1.359	0.276	4.931

improved management effectiveness by conservation donors and national governments. However with funding for conservation being much lower than required (Waldron et al., 2013) prioritization is vital to optimize the return on investment (Wilson et al., 2006).

This study alone does not demonstrate whether improved scores in management effectiveness evaluations truly reflect an improvement in protected area effectiveness in conserving biodiversity. However improvement in management effectiveness is vital in ensuring that protected areas can adapt to new situations and is capable of addressing pressures and threats more effectively in the long term.

4.3. Validity of METT scores

In our sample of protected areas, METT scores generally improved over time, both looking at the overall performance of protected areas as well as the different dimensions of management (Hockings, 2003; Hockings et al., 2006) and governance (Ostrom, 1990; Ostrom et al., 1999). This is what would be expected in protected areas receiving donor support given that 1) the time between first and last evaluation represents a monetary and resource investment in management activities and possibly 2) the evaluations are conducted on site often by people who are dependent on showing improvements to secure further resources. There have long been concerns that the identity of the METT assessor significantly influences evaluation scores, and hence the suitability of METT evaluations for impact assessment (Coad et al., accepted). However, some of the findings of this study suggest that despite these potential biases, METT score do provide a useful reflection of management realities. First, accumulated scores were significantly correlated with the time between the first and the last assessment suggesting implementation time positively affects changes in METT. This result would be expected if improvements in METT scores reflected true improvements. Second, 30.5% of the protected areas experienced no change, or even declines, in overall scores, which is a considerable

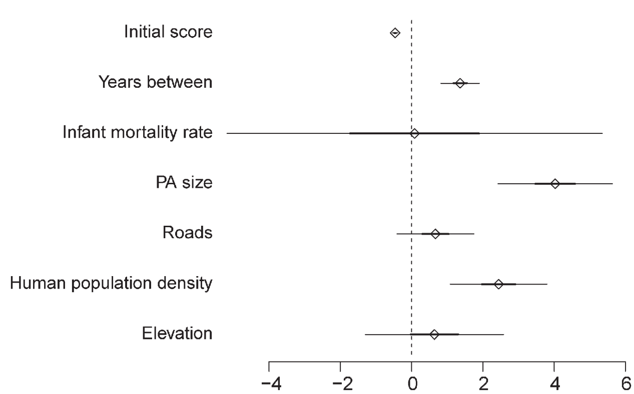

Fig. 5. Standardized regression estimates from best fit model. Diamonds indicate the mean effect size. The thick and thin lines represent the inner and outer 95% confidence interval respectively. Variables where the thick lines do not intersect the y-axis are significant. The mean effect size of HDI, roads and the intercept are not displayed as these are much larger and not significant.

proportion had there been systematic manipulation of scores. Although the above does not represent definitive causal evidence that scores are not manipulated, it does suggest that at least some of the observed changes can be attributable to actual changes in management effectiveness on the ground. This finding is also supported by previous studies looking at national level changes in governance across 41 countries (Dearden et al., 2005).

Our results are encouraging, suggesting that adaptive processes in the protected areas may be leading to improved management and that funding and resources are targeted at protected areas under greater threat where they will likely have a greater impact. However, this is at best a proxy for the true objective of protected areas: to assist in halting the loss of biodiversity. To understand whether protected areas are truly effective we need to understand whether they are maintaining species and habitats and how improved management effectiveness contributes towards this end.

Acknowledgments

We thank C Rahbek, RJ Smith, A Ahrends, K Redford and B Child who delivered valuable input to the early version of the MS. We thank M Mascia for inspiration. We thank Luke Harrison for the help on the extraction and generation of contextual variables. This study was dependent on the collective effort of hundreds if not thousands of people using the same tool despite working within multiple organizations, donors and governments, and this analysis would not have been possible if organizations had not continued to use the same methodology across time and across organizations. We thank the Danish National Research Foundation (DNRF096) for financial support. We also thank UNDP, the IUCN WCPA-SSC Joint Task-Force on Biodiversity and Protected Areas, and UNEP-WCMC for financial and institutional support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.biocon.2015.08.029.

References

Balmford, A., Moore, J.L., Brooks, T., Burgess, N., Hansen, L.A., Williams, P., Rahbek, C., 2001. Conservation conflicts across Africa. Science 291, 2616–2619.

Butchart, S.H.M., et al., 2010. Global biodiversity: indicators of recent declines. Science 328, 1164–1168.

Butchart, S.H.M., et al., 2015. Shortfalls and solutions for meeting national and global conservation area targets. Conserv. Lett. (http://onlinelibrary.wiley.com/doi/10.1111/conl.12158/full).

Carranza, T., Manica, A., Kapos, V., Balmford, A., 2014. Mismatches between conservation outcomes and management evaluation in protected areas: a case study in the Brazilian Cerrado. Biol. Conserv. 173, 10–16.

Center for International Earth Science Information Network, 2013. Global Roads Open Access Data Set, Version 1 (gROADSv1). NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY, USA (Available from http://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1).

Center for International Earth Science Information Network - CIESIN - Columbia University, International Food Policy Research Institute - IFPRI, The World Bank, and Centro Internacional de Agricultura Tropical - CIAT, 2011. Global Rural–Urban Mapping Project, Version 1 (GRUMPv1): Urban Extents Grid. NASA Socioeconomic Data and Applications Center (SEDAC).

Coad, L., Burgess, N., Fish, L., Ravilious, C., Corrigan, C., Pavese, H., Granziera, A., Besancon, C., 2008. Progress Towards the Convention on Biological Diversity Terrestrial 2010 and Marine 2012 Targets for Protected Area Coverage. NatureBureau, UK, Gland, Switzerland.

Coad, L., et al., 2013. Progress towards the CBD protected area management effectiveness targets. Parks 19, 13–24.

Coad, L., et al., 2015d. Towards the assessment of protected area management effectiveness. Philos. Trans. R. Soc. B Biol. Sci. (accepted).

Conservation Measures Partnership, 2007. Open Standards for the Practice of Conservation, Version 2.0. Conservation Measures Partnership, Washington DC, USA.

Convention on Biological Diversity, 2010. Strategic Plan for Biodiversity 2011–2020 — COP 10, decision X/2. Convention on Biological Diversity (Available from http://www.cbd.int/decision/cop/?id=12268).

Cook, C.N., Hockings, M., 2011. Opportunities for improving the rigor of management effectiveness evaluations in protected areas. Conserv. Lett. 4, 372–382.

Cook, C.N., Carter, R.W., Hockings, M., 2014. Measuring the accuracy of management effectiveness evaluations of protected areas. J. Environ. Manag. 139, 164–171.

Craigie, I.D., Baillie, J.E.M., Balmford, A., Carbone, C., Collen, B., Green, R.E., Hutton, J.M., 2010. Large mammal population declines in Africa's protected areas. Biol. Conserv. 143, 2221–2228.

Craigie, I.D., Pressey, R.L., Barnes, M., 2014. Remote regions — the last places where conservation efforts should be intensified. A reply to McCauley et al. (2013). Biol. Conserv. 172, 221–222.

Dearden, P., Bennett, M., Johnston, J., 2005. Trends in global protected area governance, 1992–2002. Environ. Manag. 36, 89–100.

Edgar, G.J., et al., 2014. Global conservation outcomes depend on marine protected areas with five key features. Nature 506, 216–220.

Geldmann, J., Barnes, M., Coad, L., Craigie, I.D., Hockings, M., Burgess, N.D., 2013. Effectiveness of terrestrial protected areas in reducing habitat loss and population declines. Biol. Conserv. 161, 230–238.

Geldmann, J., Joppa, L.N., Burgess, N.D., 2014. Mapping change in human pressure globally on land and within protected areas. Conserv. Biol. 28, 1604–1616.

Henschel, P., Coad, L., Burton, C., Chataigner, B., Dunn, A., MacDonald, D., Saidu, Y., Hunter, L.T.B., 2014. The lion in West Africa is critically endangered. PLoS One 9, e83500.

Hockings, M., 2003. Systems for assessing the effectiveness of management in protected areas. Bioscience 53, 823–832.

Hockings, M., Leverington, F., James, R., 2003. Evaluating management effectiveness: maintaining protected areas for now and for the future. Briefing Paper Prepared for the Vth World Parks Congress, Durban, September 2003. IUCN WCPA.

Hockings, M., Stolton, S., Leverington, F., Dudley, N., Courrau, J., Valentine, P., Editor, S., 2006. Evaluating Effectiveness: A Framework for Assessing Management Effectiveness of Protected Areas, Gland, Switzerland.

IUCN and UNEP-WCMC, 2014. The World Database on Protected Areas (WDPA) February 2014. UNEP-WCMC, Cambridge, UK.

Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E., 2008. Hole-filled SRTM for the Globe Version 4 Available from http://srtm.csi.cgiar.org.

Joppa, L.N., Pfaff, A., 2009. High and far: biases in the location of protected areas. PLoS One 4, e8273.

Joppa, L.N., Pfaff, A., 2011. Global protected area impacts. Proc. R. Soc. B Biol. Sci. 278, 1633–1638.

Juffe-Bignoli, D., et al., 2014. Protected Planet Report 2014. UNEP-World Conservation Monitoring Centre, Cambridge, UK.

Kellogg Foundation W.K., 2004. Using Logic Models to Bring Together Planning, Evaluation, and Action: Logic Model Development Guide. W.K. Kellogg Foundation, Battle Creek, Michigan, USA.

Kleiman, D.G., Reading, R.P., Miller, B.J., Clark, T.W., Scott, M., Robinson, J., Wallace, R.L., Cabin, R.J., Felleman, F., 2000. Improving the evaluation of conservation programs. Conserv. Biol. 14, 356–365.

Klein, A.M., Vaissiere, B.E., Cane, J.H., Steffan-Dewenter, I., Cunningham, S.A., Kremen, C., Tscharntke, T., 2007. Importance of pollinators in changing landscapes for world crops. Proc. R. Soc. B Biol. Sci. 274, 303–313.

Laurance, W.F., et al., 2012. Averting biodiversity collapse in tropical forest protected areas. Nature 489, 290–294.

Leverington, F., Hockings, M., Lemos-Costa, K., 2008. Management Effectiveness Evaluation In Protected Areas: Report for the Project "Global Study Into Management Effectiveness Evaluation of Protected Areas". The University of Queensland, Gatton, IUCN WCPA, TNC, WWF, Australia.

Leverington, F., Costa, K.L., Pavese, H., Lisle, A., Hockings, M., 2010. A global analysis of protected area management effectiveness. Environ. Manag. 46, 685–698.

Mace, G.M., Collen, B., Fuller, R.A., Boakes, E.H., 2010. Population and geographic range dynamics: implications for conservation planning. Philos. Trans. R. Soc. B Biol. Sci. 365, 3743–3751.

Mascia, M.B., et al., 2014. Commonalities and complementarities among approaches to conservation monitoring and evaluation. Biol. Conserv. 169, 258–267.

Mittermeier, R.A., Mittermeier, C.G., Brooks, T.M., Pilgrim, J.D., Konstant, W.R., da Fonseca, G.A.B., Kormos, C., 2003. Wilderness and biodiversity conservation. Proc. Natl. Acad. Sci. 100, 10309–10313.

Naidoo, R., Balmford, A., Ferraro, P.J., Polasky, S., Ricketts, T.H., Rouget, M., 2006. Integrating economic costs into conservation planning. Trends Ecol. Evol. 21, 681–687.

Nolte, C., Agrawal, A., 2013. linking management effectiveness indicators to observed effects of protected areas on fire occurrence in the Amazon rainforest. Conserv. Biol. 27, 155–165.

Ostrom, E., 1990. Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press, Cambridge, UK.

Ostrom, E., 2009. A general framework for analyzing sustainability of social–ecological systems. Science 325, 419–422.

Ostrom, E., Burger, J., Field, C.B., Norgaard, R.B., Policansky, D., 1999. Revisiting the commons: local lessons, global challenges. Science 284, 278–282.

R Development Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rodrigues, A.S.L., 2006. Are global conservation efforts successful? Science 313, 1051–1052.

Sanderson, E.W., Jaiteh, M., Levy, M.A., Redford, K.H., Wannebo, A.V., Woolmer, G., 2002. The human footprint and the last of the wild. Bioscience 52, 891–904.

Scharlemann, J.P.W., Kapos, V., Campbell, A., Lysenko, I., Burgess, N.D., Hansen, M.C., Gibbs, H.K., Dickson, B., Miles, L., 2010. Securing tropical forest carbon: the contribution of protected areas to REDD. Oryx 44, 352–357.

Stolton, S., Hockings, M., Dudley, N., MacKinnon, K., Whitten, T., L. F., 2007. Reporting Progress in Protected Areas. A Site-Level Management Effectiveness Tracking Tool. second edition. World Bank/WWF Forest Alliance, Gland, Switzerland.

Struhsaker, T.T., Struhsaker, P.J., Siex, K.S., 2005. Conserving Africa's rain forests: problems in protected areas and possible solutions. Biol. Conserv. 123, 45–54.

Tittensor, D.P., et al., 2014. A mid-term analysis of progress toward international biodiversity targets. Science 346, 241–244.

Waldron, A., Mooers, A.O., Miller, D.C., Nibbelink, N., Redding, D., Kuhn, T.S., Roberts, J.T., Gittleman, J.L., 2013. Targeting global conservation funding to limit immediate biodiversity declines. Proc. Natl. Acad. Sci. 110 (29), 12144–12148.

Walpole, M., et al., 2009. Tracking progress toward the 2010 biodiversity target and beyond. Science 325, 1503–1504.

Watson, J.E.M., Dudley, N., Segan, D.B., Hockings, M., 2014. The performance and potential of protected areas. Nature 515, 67–73.

Wilson, K.A., McBride, M.F., Bode, M., Possingham, H.P., 2006. Prioritizing global conservation efforts. Nature 440, 337–340.