



# Linking Management Effectiveness Indicators to Observed Effects of Protected Areas on Fire Occurrence in the Amazon Rainforest

CHRISTOPH NOLTE\* AND ARUN AGRAWAL

School of Natural Resources and Environment, University of Michigan, 440 Church Street, Ann Arbor MI 48109, U.S.A.

**Abstract:** *Management-effectiveness scores are used widely by donors and implementers of conservation projects to prioritize, track, and evaluate investments in protected areas. However, there is little evidence that these scores actually reflect the capacity of protected areas to deliver conservation outcomes. We examined the relation between indicators of management effectiveness in protected areas and the effectiveness of protected areas in reducing fire occurrence in the Amazon rainforest. We used data collected with the Management Effectiveness Tracking Tool (METT) scorecard, adopted by some of the world's largest conservation organizations to track management characteristics believed to be crucial for protected-area effectiveness. We used the occurrence of forest fires from 2000 through 2010 as a measure of the effect of protected areas on undesired land-cover change in the Amazon basin. We used matching to compare the estimated effect of protected areas with low versus high METT scores on fire occurrence. We also estimated effects of individual protected areas on fire occurrence and explored the relation between these effects and METT scores. The relations between METT scores and effects of protected areas on fire occurrence were weak. Protected areas with higher METT scores in 2005 did not seem to have performed better than protected areas with lower METT scores at reducing fire occurrence over the last 10 years. Further research into the relations between management-effectiveness indicators and conservation outcomes in protected areas seems necessary, and our results show that the careful application of matching methods can be a suitable method for that purpose.*

**Keywords:** avoided deforestation, conservation success, forest fires, impact assessment, protected areas

Vinculación de Indicadores de Efectividad de Manejo con los Efectos Observados de la Ocurrencia de Fuego en Áreas Protegidas en la Amazonia

**Resumen:** *Los valores de efectividad del manejo son utilizados ampliamente por donantes e implementadores de proyectos de conservación para priorizar, seguir, y evaluar inversiones en áreas protegidas. Sin embargo, hay poca evidencia de que estos valores reflejan la capacidad de las áreas protegidas para generar resultados de conservación. Examinamos la relación entre indicadores de efectividad del manejo en áreas protegidas y la efectividad de áreas protegidas para reducir la ocurrencia de fuego en bosques de la Amazonia. Utilizamos datos recolectados con el sistema de puntuación Herramienta para el Seguimiento de la Efectividad del Manejo (HSEM), adoptado por algunas de las organizaciones de conservación más grandes del mundo para monitorear atributos del manejo considerados cruciales para la efectividad de las áreas protegidas. Utilizamos la ocurrencia de los incendios forestales de 2000 a 2010 como una medida del efecto de las áreas protegidas sobre cambios no deseados en la cobertura de suelo en la Cuenca del Amazonas. Usamos combinaciones para comparar el efecto estimado de las áreas protegidas con valores HSEM de ocurrencia de fuego bajos versus altos. También estimamos los efectos de áreas protegidas individuales sobre la ocurrencia de fuego y exploramos la relación entre estos efectos y los valores HSEM. Las relaciones entre los valores HSEM*

\*email cbrnolte@umich.edu

Paper submitted September 23, 2011; revised manuscript accepted May 24, 2012.

*y los efectos de las áreas protegidas sobre la ocurrencia de fuego fueron débiles. Las áreas protegidas con valores HSEM más altos en 2005 no parecen haber reducido la ocurrencia de fuego en los últimos 10 años. Es necesario investigar más sobre las relaciones entre indicadores de efectividad del manejo y los resultados de conservación en áreas protegidas, y nuestros resultados muestran que la aplicación cuidadosa de métodos combinatorios es un método adecuado para ese propósito.*

**Palabras Clave:** Áreas protegidas, deforestación evitada, evaluación de impacto, éxito de conservación, incendios forestales

## Introduction

Protected areas are one of the most prominent tools for conserving biodiversity worldwide. By 2011 more than 160,000 sites were “recognized, dedicated and managed [...] to achieve the long term conservation of nature” (World Conservation Monitoring Centre 2011). Major conservation donors continue to dedicate substantial levels of financial resources to the establishment and consolidation of protected-area networks (Gordon and Betty Moore Foundation 2006; Global Environment Facility 2009; Kasperek et al. 2010). Put forward as a potentially effective means to reduce carbon emissions from deforestation and forest degradation (Trumper et al. 2009; Soares-Filho et al. 2010), protected areas can be expected to attract conservation funds for decades to come.

Given their substantial and growing financial commitments, donors and implementers of conservation projects have come under pressure to demonstrate that investments in protected areas deliver positive outcomes (Ferraro & Pattanayak 2006; Fuller et al. 2010; Mascia & Pailler 2011). Reacting to such concerns, parties to the Convention on Biological Diversity committed to evaluating the management effectiveness of 30% of their protected areas by 2010 (CBD 2004). Guiding this effort, the World Commission of Protected Areas defined protected-area management effectiveness (PAME) as being about more than conservation outcomes. Instead, PAME evaluations incorporate a wide range of management themes, including the adequacy of protected-area design, planning, resources, and management processes (Hockings et al. 2006). Major donors such as the Global Environment Facility (GEF), World Bank, Gordon and Betty Moore Foundation, and U.S. Agency for International Development, and implementing organizations of conservation projects such as The Nature Conservancy and World Wide Fund for Nature (WWF), have used PAME to prioritize, track, or evaluate their investments in protected areas. Results of recent global and regional surveys show over 9000 PAME assessments in 140 countries (Leverington et al. 2010; Nolte et al. 2010).

Although PAME evaluations have been driven by an ultimate interest in the effects of protected areas (i.e., changes in outcomes that can be attributed to the existence of a protected area and its management) (Ferraro 2009; Joppa & Pfaff 2011), in practice the focus

seems to have shifted. Among the 3 most widely used PAME evaluation approaches, only one contains an outcome indicator (Table 1). None provide a framework to interpret changes in outcomes as an effect of the examined protected area or its management strategies, for example, by estimating what outcomes would have been observed in the absence of protection or management (counterfactual) (Ferraro 2009). Instead of looking at effects, most PAME approaches collect data on a multitude of indicators of the management capacity of a protected area, such as the adequacy of its budget, staffing level, planning processes, stakeholder participation, and enforcement. The worldwide adoption of this approach seems to rest on the assumption that these indicators provide a reasonably good proxy for the extent to which a protected area is effective in delivering desired outcomes.

However, this assumption has not been thoroughly tested. There are studies in which PAME data were analyzed at a global extent to identify global patterns in protected-area management (e.g., chronic inadequacy of protected-area budgets, staffing levels, infrastructure, and community relations) (Dudley et al. 2007; Leverington et al. 2010). However, the absence of outcome indicators has limited the ability of these studies to illuminate the relation between PAME scores and the effectiveness of protected areas in delivering conservation outcomes. Zimsky et al. (2010) investigated this relation. They compared scores of the Management Effectiveness Tracking Tool (METT), the world’s most widely used PAME tool, to species data and expert opinion on 11 wildlife reserves in Zambia and found a positive correlation between an increase in METT scores and increases in species populations (Zimsky et al. 2010). However, the authors did not control for confounding factors that could be associated with the observed changes in outcomes.

We empirically examined the relations between METT scores and the effect of protected areas on the occurrence of forest fires in the Amazon rainforest. Recognized as a global priority area for the conservation of both biodiversity (Brooks et al. 2006) and carbon stocks (Saatchi et al. 2007), the Amazon basin has attracted considerable amounts of conservation funding over the last decade. Forest fires in the Amazon are predominantly of anthropogenic origin (Cochrane 2011; Pivello 2011) and exhibit close spatial relations with deforestation (Aragão

**Table 1.** Scoring systems used most widely (>10 countries) to assess protected-area management effectiveness.

<i>System name</i>	<i>Organization used by</i>	<i>No. countries used in, no. of assessments</i>	<i>No. of outcome indicators</i>	<i>Reference</i>
Management Effectiveness Tracking Tool	Global Environment Facility, World Bank	100, approximately 2000 <sup>a</sup>	1 <sup>c</sup>	Stolton et al. (2007)
Rapid Assessment and Prioritization of Protected Area Management Parks in Peril Site Consolidation Scorecard	World Wide Fund for Nature	49, >1600 <sup>b</sup>	0	Ervin (2003)
	U.S. Agency for International Development, The Nature Conservancy	15, 323 <sup>b</sup>	0	Martin and Rieger (2003)

<sup>a</sup>Data from Coad et al. (2011).

<sup>b</sup>Data from Leverington et al. (2010).

<sup>c</sup>Qualitative indicator. See Table 3 (indicator 27) for wording.

et al. 2008; Nelson & Chomitz 2011). Therefore forest fires in the Amazon provide a reasonable proxy for the type of land-cover change protected areas seek to reduce. For the purpose of this study, we define *effectiveness* of a protected area as the extent to which the occurrence of fires on forest parcels within its boundaries is lower than that on similar unprotected forest parcels. We used matching, a statistical technique widely used to estimate protected-area effectiveness (Andam et al. 2008; Joppa & Pfaff 2011; Nelson & Chomitz 2011), to compare the estimated effect of protected-area groups with low versus high METT scores on fire occurrence. We also estimated effects of individual protected areas on fire occurrence, and explored the associations of these estimates with METT scores.

## Methods

### Study Area

We focused on the tropical and subtropical moist broadleaf forests in the Amazon basin as defined by WWF's Terrestrial Ecoregions of the World (Olson et al. 2001). Within this region, we selected all approximately 1-km<sup>2</sup> parcels that contained at least 25% forest cover, as estimated by the vegetation continuous fields (VCF) algorithm (Hansen et al. 2003), and classified as forest or forest mosaic by the 1-km resolution global land cover (GLC) data set produced by the European System for Earth Observation (Bartholome & Belward 2005). The intersection of VCF and GLC produces a conservative estimate of tropical forest area that limits the risk of including observation of fire occurrence unrelated to deforestation (e.g., in tropical savannas or on land that was already cleared of forest or used predominantly for agriculture) (Nelson & Chomitz 2011). The GLC data set is only available for the year 2000, which we consequently chose as the starting year of our period of analysis (2000–2010).

### Fires in Forests

We extracted our fire data from the active fires product of the moderate resolution imaging spectroradiometer (MODIS). This product provides globally consistent daily estimates of the location and intensity of active fires at approximately 1-km resolution since October 2000 (Justice 2002). Following Morton et al. (2008), we extracted only fires occurring at night and daytime fires with >330 K brightness in the 4- $\mu$ m channel (high-confidence fires) observed between 2000 and 2010 on forest parcels in our study area. Our outcome variable was binary: MODIS detected or did not detect at least one high-confidence fire on a given forest parcel between 2000 and 2010.

### Effectiveness Data

Initially developed to assess PAME of forested protected areas (Stolton et al. 2007), the METT has been applied across a wide range of ecosystems as a reporting requirement for protected area projects supported by the World Bank and the GEF. Respondents are usually accountable to donor organizations and include protected-area managers, project staff, consultants, and management councils. A recent effort to compile all existing METT data (Coad et al. 2011) recorded >2000 METT assessments in >100 countries, which makes it one of the most widely used PAME assessment methods in the world.

The METT has 30 questions designed to provide information on management aspects believed to be crucial for effective protected-area management. Questions pertain to legal status and regulations, adequacy of budget, staff and resources, research and monitoring, and stakeholder relations. Respondents assign each indicator a score from 0 to 3, and qualitative statements provide indicator-specific guidance about the meaning of each number. Low values generally reflect lower relative performance (Supporting Information). Individual indicator scores are summed to create the composite METT score, which has been suggested as a possible proxy for overall management effectiveness (Dudley et al. 2007).

The METT allows for the exclusion of indicators that are not applicable and for subsequent rescaling of the composite score. To account for such missing indicators, we calculated METT composite scores as the average of all indicators for which a value was reported (0–3).

We extracted all assessments carried out in 2005 or earlier from the most recent version of the Global METT Database (Coad et al. 2011). We used only the most recent scores for each protected area in the case of repeat assessments. Because many GEF and World Bank projects in the region were launched before METT became a reporting requirement, our sample contained scores for 41 protected areas, 2 located in Bolivia, 6 in Peru, and 33 in Brazil. Although this sample is not random, it is geographically well distributed over the Amazon basin.

### Covariates and Treatment

The probability of protection and the probability of forest fire are affected by a number of variables that must be controlled when estimating the effectiveness of protected areas in reducing forest fires. Drawing on related assessments of the effectiveness of protected-area networks in reducing tropical deforestation and fires (Andam et al. 2008; Joppa & Pfaff 2011; Nelson & Chomitz 2011), we included 5 covariates in our analyses: elevation, slope, travel time, distance to forest edge, and rainfall.

Slope and elevation strongly affect whether a given location is suitable for different land uses; thus, they are expected to be associated with the probability that a given forest parcel will be converted to agriculture. We controlled for median elevation and average slope, which we extracted from spatial data layers of the Global Agro-Ecological Zones Assessment (Fischer et al. 2007).

The probability of timber extraction and agricultural use is strongly affected by access to markets, a function of a parcel's distance to roads, rivers, and major cities. To account for market access, we used estimates of travel time to major cities (>50,000 inhabitants) computed by the Joint Research Centre of the European Commission (Nelson 2008).

Deforestation is more likely to occur close to the forest edge and to locations that were deforested previously. We used the intersection of GLC and 25% VCF to define forest extent. We added a 1-km area to this extent to smooth out small nonforest patches surrounded by forest and therefore less likely to be part of the agricultural frontier. Distance to forest edge was the Euclidian distance from each forest parcel to the closest nonforest parcel.

Available moisture and precipitation affect the probability of fire occurrence and fire detection. Average annual rainfall affects the suitability of a given parcel for agricultural production and higher precipitation rates may be associated with greater cloud coverage (detection of fire by remote sensors may be inhibited by clouds). We used

average annual precipitation rates provided by WorldClim to control for this covariate (Hijmans et al. 2005).

We extracted geographical limits of protected areas from the World Database of Protected Areas (WDPA) (World Conservation Monitoring Centre 2011) and based boundaries of countries and states on the Vector Smart Map (VMAP) level 0 data set (NIMA 2000). We used ArcGIS 10.0 to resample spatial data layers for all 5 covariates, protected areas, and countries or states into the format used by MODIS active fires product (equal area sinusoidal projection, approximately 1-km resolution) and extracted all variables into one table with 5.26 million forest parcels for analyses.

### Effects of Protected-Area Groups

We divided our sample of protected areas with METT data into 2 groups—low and high composite METT scores—in order to compare their respective effects on fire occurrence. We conducted separate analyses for the full sample and for the Brazilian subsample. We included all protected areas designated in or prior to 2002 and chose thresholds of the composite METT scores that produced groups with a roughly similar number of protected areas (METT score threshold: 1.22 and 1.33 for the Brazilian sample and complete sample, respectively). Because the total area of forest cover varied among protected areas, the number of forest parcels in each group varied. To test the sensitivity of our analyses to the choice of these threshold parameters, we also explored alternative group definitions (e.g., use of different cutoff years, creation of groups with similar forest extent, and limiting analyses to the upper and lower quartiles of METT scores [Supporting Information]).

We used nearest-neighbor matching (NNM) to estimate effects of protected-area groups on fire occurrence. Widely used to estimate protected-area effectiveness (Andam et al. 2008; Joppa & Pfaff 2011; Nelson & Chomitz 2011), NNM is a statistical technique that mimics random assignment of treatment in observational data by matching each treated unit (forest parcel) to a unit from a pool of candidate control units that have covariates similar to the selected treatment unit. The difference in outcomes between the treatment and artificial control groups is assumed to reflect the average treatment effect on the treated (ATT). We measured our ATT of interest, the effect of protected-area groups on fire occurrence, as the difference in the percentage of forest parcels with observed fires between treatment and control group.

We estimated effects on fire occurrence of protected areas with high and low METT scores relative to the counterfactual of no protection by matching forest parcels from either group (high or low scores) with a third group of parcels that had never been protected. This estimation strategy is similar to that of Nelson and Chomitz (2011),

who compared the effectiveness of different protected-area categories (strict protection, sustainable use, indigenous lands) in reducing fire occurrence. However, distributions of key covariates of forest parcels differed considerably between areas with high and low METT scores (Supporting Information), and this reduced the extent to which differences in estimated effects on fire occurrences could be ascribed to differences in METT scores as opposed to differences in other characteristics. We therefore conducted 2 additional comparisons to achieve better covariate balance between forest parcels from protected areas with high and low METT scores. We matched forest parcels from protected areas with high METT scores to forest parcels from protected areas with low METT scores and compared differences in fire occurrences between the two. Then we repeated this process for parcels from protected areas with low versus high METT scores. We expected these two comparisons to generate different groups of forest parcels and different effect estimates (see Supporting Information for elaboration).

Our matching-based estimates of the effects of protected areas on fire occurrence required the assumption that the probabilities of forest fires were not autocorrelated. Recent findings cast some doubt on this assumption. Areas burned by wildfires are power-law distributed in a variety of different eco-regions, including the Amazon (Malamud et al. 2005; Pueyo et al. 2010), which introduces potential spatial dependence in the likelihood of fire occurrence on neighboring forest parcels. Although our method did not allow us to explicitly control for spatial autocorrelation, we reduced the risk of spatial dependence by randomly sampling a small percentage of forest parcels (2%) from the entire population of forest parcels and conducting our analyses on these samples (Andam et al. 2008; Joppa & Pfaff 2011; Nelson & Chomitz 2011).

We used Sekhon's (2007) matching library to match forest parcels in R. We used Mahalanobis distance NNM with replacement and bias adjustment. Average slope, elevation, rainfall, distance to forest edge, and travel time were covariates. Control parcels were all located in the same country (Peru, Bolivia) or state (in Brazil) as the treatment parcels (exact matching). We dropped treatment parcels for which no nearest neighbor could be found within 1 SD of each covariate (calipers). We repeated the process of random sampling and matching 30 times and averaged the estimated differences between treatment and artificial control groups. As in earlier matching studies (Andam et al. 2008; Joppa & Pfaff 2011), we used the Abadie-Imbens variance formula and reported the average standard error as the square root of the mean variance across 30 runs. We also tested whether our results were sensitive to the size of our calipers (Supporting Information).

## Effects and METT Scores of Individual Protected Areas

We examined the associations between METT scores and individual effects of protected areas on fire occurrence. We developed protected-area-level estimates of the probability of fire occurrence in the absence of protection by matching forest parcels from each protected area to unprotected forest parcels and averaging fire occurrence on the latter. We used the same data, covariates, and matching parameters as in our analyses of protected-area groups, but we applied calipers of 0.5 SD. We sampled 5% of the forest parcels from each protected area and matched them to a 5% sample of unprotected forest parcels. We repeated the sampling and matching process 30 times and averaged the resulting estimates of fire probability. Estimates were computed for each WDPA-reported protected area in Bolivia, Brazil, and Peru that was designated in or prior to 2002, contained at least 500 forest parcels, and did not overlap with other protected areas. After excluding protected areas for which matching was considered unrepresentative (protected areas for which >40% parcels had been dropped due to calipers), our final sample contained 182 protected areas, 29 of which had METT scores.

The distribution of our estimates of the probability of fire occurrence in the absence of protection was strongly skewed. Although 9.8% of protected areas were estimated to have >50% of their forest parcels exposed to fires in the absence of protection, 44% of protected areas were estimated to have <1% of forest parcels affected by fire if unprotected (Fig. 1). Such variation posed challenges to defining which of these protected areas were more effective in reducing fire occurrence. Where fire probabilities in the absence of protection are high, protected areas may have had a considerable effect on fire occurrence, although fires may still be frequent relative to protected areas with low fire probabilities in the absence of protection. The latter, in turn, may not have been exposed to forest fires at all but cannot be considered to have had major effects on fire occurrence either.

We asked what actual fire occurrence one would expect to observe in a protected area given the estimated probability of fire occurrence in the absence of protection. We estimated expected fire occurrence in protected areas, conditional on estimated fire probabilities, non-parametrically by fitting a locally weighted scatter plot smoothing function (LOESS, span = 0.75) to the complete sample of 182 protected areas. We divided the sample into 2 groups. Protected areas with fire occurrences below and above the threshold defined by the LOESS function (Fig. 1) were defined as belonging to the more effective or less effective category, respectively. In addition, we considered 3 alternative definitions of the relative effectiveness of protected areas in reducing fire occurrence (Table 2).

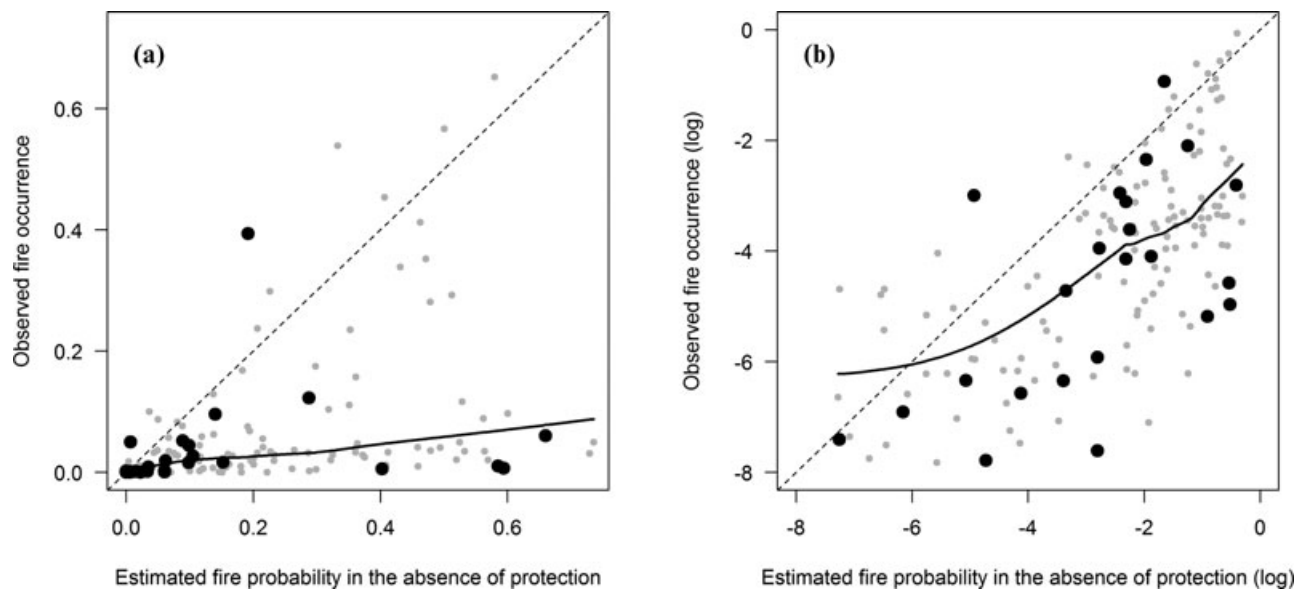


Figure 1. Relation between estimated probability ([a] original values and [b] log probability) of fire occurrence in protected areas in the absence of protection and observed fire occurrence in 182 forested areas in Peru, Bolivia, and Brazil with (black dots) and without (grey dots) data from the Management Effectiveness Tracking Tool (solid line, estimated fire occurrence within protected areas conditional on estimated probability of fire occurrence in the absence of protection, estimated via locally weighted scatterplot smoothing [LOESS]; dashed line, diagonal [ $y = x$ ]).

Given our small sample size, we used a simple 2-step approach to explore relations between METT scores and effects of individual protected areas on fire occurrence. First, we compared the differences in composite METT scores between more effective and less effective groups of protected areas. Second, we compared differences in selected METT indicators between more effective versus less effective protected areas. For each indicator, we discarded protected areas for which no score was provided. We used 2-tailed  $t$  tests to test for significant differences in score means between groups. With 18 score indicators and 29 observations, our objective in doing this was not to claim that differences in scores were causally related

to relative effectiveness, but to explore patterns within the data that could direct further research.

## Results

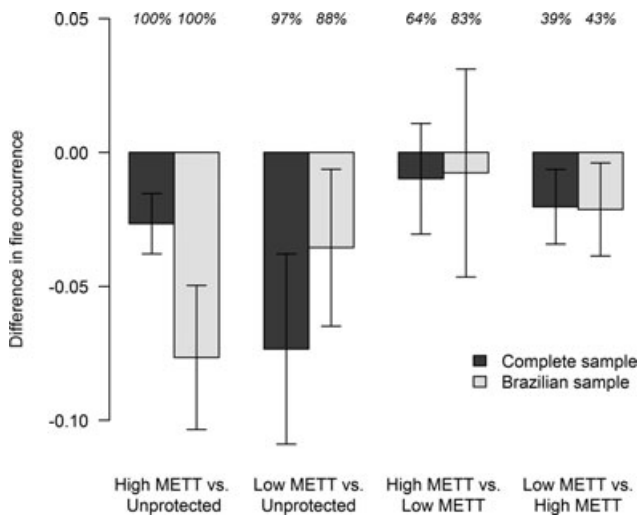
Group-matching results suggest that protected areas with high and low METT scores reduced the occurrence of fires within their boundaries relative to similar unprotected areas (Fig. 2). However, our results did not offer clear evidence that fire occurrence was lower in areas with high scores than in areas with low scores. In the complete sample, the estimated effect of protection on

Table 2. Definitions of 4 measures of the relative effectiveness of protected areas in reducing the occurrence of forest fires.

Measure	Definition of effectiveness	Threshold variable	Threshold for groups ( $n_{\text{less}} \cdot n_{\text{more}}$ ) <sup>a</sup>
Fire occurrence	Fire occurrence lower than in other protected areas	Fire occurrence	Median of total sample (11:18)
Fire reduction	Absolute effect of fire reduction on fire occurrence higher than in other protected areas	Fire probability in the absence of protection-fire occurrence	Median of total sample (18:11)
Percent fire reduction	Relative effect of fire reduction on fire occurrence is higher than in other protected areas	Fire occurrence or fire probability in the absence of protection	Median of total sample (14:14)
LOESS <sup>b</sup>	Fire occurrence is lower than in protected areas with comparable fire probability in the absence of protection	Fire occurrence	LOESS curve (9:20)

<sup>a</sup>Key:  $n_{\text{less}}$ , number of protected areas in the less effective group;  $n_{\text{more}}$ , number of protected areas in the more effective group.

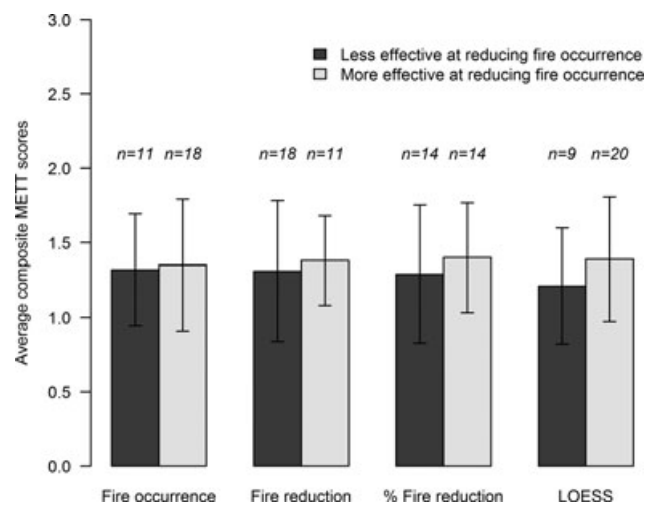
<sup>b</sup>Locally weighted scatterplot smoothing.



**Figure 2.** Estimated difference in fire occurrence among protected-area groups with high and low composite scores on the Management Effectiveness Tracking Tool (METT) (the threshold between high and low composite scores was 1.33 and 1.22 for the full sample and the Brazilian sample, respectively). Percentages above bars indicate statistical support for each estimator (average percentage of retained treated forest parcels across 30 matching estimates). Error bars indicate 95% confidence intervals (average  $SE \times 1.96$ ).

fire occurrence in areas with high METT scores was considerably smaller than for areas with low scores (−2.6% vs. −7.3%, respectively). The opposite pattern occurred with the Brazilian sample (−7.6% vs. 3.5%, respectively). The discrepancy seems to be due to 3 Brazilian protected areas with strong effects on fire occurrence (i.e., high estimated fire probabilities in the absence of protection, but low actual fire occurrence) switching groups as the METT threshold decreased from 1.33 to 1.22. Comparing only the upper with the lower-METT-score quartiles resulted in roughly similar estimates of effects of both protected-area groups on fire occurrence (complete sample, −2.7% vs. −2%; Brazilian sample, −3.3% vs. −3.6%) (Supporting Information).

Matching forest parcels in protected areas with high METT scores to those with low METT scores retained only 64% (full sample) and 83% (Brazilian sample) of the observations because matched pairs were dropped if they were not sufficiently similar (Fig. 2). After matching, differences in fire occurrence between parcels in protected areas with high versus low METT scores fell below 1%. Matching parcels in areas with low to high METT scores suggested that the former have had an even greater effect on fire occurrence than the latter (2.0% and 2.1%), although this estimate was based on less than half of the parcels in protected areas with low METT scores.



**Figure 3.** Composite scores of the Management Effectiveness Tracking Tool (METT) for protected-area groups considered more and less effective according to our definitions of relative effectiveness (Table 2) in reducing fire occurrence (error bars, SD of scores for each group; LOESS, locally weighted scatterplot smoothing).

Results from our analyses of individual protected areas also indicated weak relations between METT scores and effects on forest fires. Although METT composite scores varied substantially among protected areas (mean [SD] = 1.33 [0.41]), differences in METT composite scores between more effective and less effective protected areas were small (Fig. 3). The highest absolute difference in average composite METT scores (1.21 vs. 1.39) resulted when we used our preferred definition of the relative effectiveness of protected areas in reducing fire occurrence (LOESS). However, the difference was not statistically significant even at the 0.25 confidence level.

Although composite METT scores were not significantly related to the effectiveness of protected areas in reducing fire occurrence, several individual indicators in the METT were (Table 3). Given the ratio between indicators and observations, much of this variance could be the result of random variation. However, more effective areas in our sample tended to have higher subjective ecological-condition scores (indicator 27). Decision makers in such protected areas were more likely to cooperate closely with neighboring official and commercial land users (indicator 21), but were less likely to allow for input into management decisions from local communities (indicator 23). More effective protected areas were also likely to have more research activities (indicator 10) and better access control (indicator 28).

Many individual METT indicators did not exhibit observable or consistent differences between more effective and less effective groups. This was particularly true for management aspects traditionally assumed to be

**Table 3. Relation of indicators in the Management Effectiveness Tracking Tool to relative effectiveness of efforts to reduce fire occurrence in Amazonian rainforests.**

Indicator category <sup>a</sup>	Number and name of indicator on questionnaire	Scoring question	Measure of relative effectiveness <sup>b</sup>			
			Fire occurrence	Fire reduction	% Fire reduction	LOESS
Resources	12: staff numbers	Are there enough people employed to manage the protected area?	(-)			
	15: current budget	Is the current budget sufficient?				
	16: security of budget 18: equipment	Is the budget secure? Are there adequate equipment and facilities?	(+)	-	+	(+)
Institutions	2: regulations	Are inappropriate land uses and activities controlled?	(-)			
	3: law enforcement	Can staff enforce protected area rules well enough?		(+)		
	6: boundary demarcation 28: access assessment	Is the boundary known and demarcated? Is access/resource use sufficiently controlled?		(+)	+	(+)
Information	30: monitoring and evaluation	Are management activities monitored against performance?				
	10: research	Is there a program of management-oriented survey and research work?			(+)	++
	9: resource inventory	Do you have enough information to manage the area?				(+)
Planning	7: management plan	Is there a management plan and is it being implemented?				
	8: regular work plan	Is there an annual work plan?				
	17: budget management	Is the budget managed to meet critical management needs?		++	+	(+)
Stakeholder relationships	21: state and commercial neighbors	Is there cooperation with adjacent land users?	++		(+)	++
	22: indigenous people	Do indigenous and traditional peoples resident in or regularly using the protected area have input into management decisions?				
	23: local communities	Do local communities resident in or regularly using the protected area have input into management decisions?			--	(-)
Outcomes	27: condition assessment	What is the condition of the important values of the protected area as compared to when it was first designated?	(+)	(+)	+	+

<sup>a</sup>Heuristic categories were created to improve the readability of the table.

<sup>b</sup>Significance of differences between individual scores for protected-area groups that are more effective and less effective (2-tailed *t* test) in reducing forest fires. Positive relations: +++,  $p < 0.01$ ; ++,  $p < 0.05$ ; +,  $p < 0.1$ ; (+),  $p < 0.25$ . Negative relations: --,  $p < 0.05$ ; -,  $p < 0.1$ ; (-),  $p < 0.25$ . LOESS is locally weighted scatter plot smoothing. See Supporting Information for full scores.



closely related to protected-area effectiveness and thus to be classical targets of conservation investments, including adequacy (indicator 15) and security of budget (indicator 16), number of staff members (indicator 12), management plans (indicator 7), and boundary demarcation (indicator 6).

The behavior of institutional variables was inconsistent: Although controlling access or use of the protected area (indicator 28) was positively associated with effectiveness, other variables such as mechanisms for controlling inappropriate land use and activities (indicator 2) and capacities and resources of staff to enforce regulations and legislation (indicator 3) had weak or negative associations with the relative effectiveness of protected areas in reducing fire occurrence.

## Discussion

Our analyses of the associations between PAME scores and the effectiveness of protected areas in reducing fire occurrence was motivated by the goal to refine and improve existing strategies to measure and track management effectiveness. However, our use of fire data as a proxy for deforestation in the Amazon basin did not yield strong relations between the composite METT scores used for reporting and our multiple definitions of the relative effectiveness of protected areas in reducing fire occurrence. At least for the protected areas in our sample, METT scores seemed to fall short of their potential as an indicator for the capacity of a protected area to reduce the extent of undesired land-use changes.

Certainly, such correlation is not causality. Our failure to observe significant differences in 2005 METT scores among protected areas that were more versus less effective in reducing fire occurrence between 2000 and 2010 could be a result of conservation actors adapting their support strategies as a function of protected-area success. For example, it is plausible that support for protected areas has systematically targeted underperforming protected areas within our study period.

Although our findings do not allow us to establish claims of causality, the lack of observed associations between management-effectiveness indicators of protected areas and their effectiveness in reducing forest fires is illuminating. Developed by experienced protected-area experts, METT has been endorsed by major conservation donors as a mandatory evaluation tool, which makes METT a *de facto* standard for assessing PAME. If METT scores do not actually serve as a good proxy for the capacity of a protected area to reduce undesired land-use change, results of project evaluations that rely on METT scores may be biased. A greater concern is that METT and similar management-characteristic-based evaluations may create incentives for project implementers to invest in activities that improve effectiveness scores without nec-

essarily making a protected area more effective in terms of conservation outcomes.

Given the widespread use of PAME scores in conservation projects and policy worldwide, it seems necessary to direct further efforts into understanding the relation between protected area management, protected area effectiveness, and the indicators used to measure both. We suggest future studies examine the strength of associations between PAME indicators and effectiveness estimates of protected areas in other ecoregions and apply data from other widespread PAME methodologies (e.g., Rapid Assessment and Prioritization of Protected Areas, Parks in Peril Site Consolidation Scorecard) (Table 3) (Ervin 2003; Martin & Rieger 2003). Insights into relations between indicators and effectiveness of protected areas would allow evaluators to learn which indicators are more closely associated with effectiveness and adapt existing evaluation methods accordingly. The widespread use of PAME scores for accountability purposes also justifies a renewed quest for indicators that are cheap to verify, costly to fake (Ferraro 2008), and possibly more objective than the existing judgments of adequacy (Table 3), which can differ considerably among respondents and protected areas and over time.

However, to understand why some areas are effective and what type of support makes them effective, future analyses will need to examine causation rather than correlation. Do protected areas that received a specific type of support reduce undesired land-use changes to a larger extent than those that do not receive the same support—even if support allocation is influenced by expected effects? The large number of protected areas and support projects around the world make it increasingly possible to construct such counterfactual evidence for a number of management interventions, an approach that promises to provide strong evidence for the relative effectiveness of such investments.

Our methods offer a new way to improve the utility of matching methods in estimating the relative effectiveness of protected areas. We found that studies comparing protection-effect estimates of different protected-area groups versus unprotected groups (e.g., Nelson & Chomitz 2011) can conflate potential differences of the effectiveness of protected-area groups in reducing undesired land-use changes with differences in the probability of the occurrence of such land-use change in the absence of protection. Between-group matching allows one to single out these two estimates and thus to provide a better estimate of differences in the relative effectiveness between groups of interest. In addition, our approach to computing effectiveness estimates at the protected-area level allows for comparisons that assign the same weight to each protected area (and METT score) and are thus less vulnerable to differences in the size of protected areas. Our results suggest that although matching is not a methodological panacea, it can, if carefully designed,

be a useful tool with which to examine effectiveness of protected areas.

## Acknowledgments

We thank L. Coad, J. Geldmann, F. Leverington, and countless individuals who collected, compiled, and provided METT data over the years. We are grateful to P. Ferraro, E. Fleishman, M. Hanauer, K. Shedden, M. Zimsky, and 2 anonymous reviewers for insightful comments on earlier versions of this manuscript. Financial support came from the School of Natural Resources and Environment and the Rackham Graduate School at the University of Michigan, the National Science Foundation through (grant CNH0709545) and the Gordon and Betty Moore Foundation (grant 3006).

## Supporting Information

Methodological details, including summary statistics, balance tables, and sensitivity tests (Appendix S1) and the METT instrument (Appendix S2) are available online. The authors are solely responsible for the content and functionality of these materials. Queries (other than absence of the material) should be directed to the corresponding author.

## Literature Cited

- Andam, K. S., P. J. Ferraro, A. Pfaff, G. A. Sanchez-Azofeifa, and J. A. Robalino. 2008. Measuring the effectiveness of protected area networks in reducing deforestation. *Proceedings of the National Academy of Sciences* **105**:16089–16094.
- Aragão, L. E. O. C., Y. Malhi, N. Barbier, A. A. Lima, Y. Shimabukuro, L. Anderson, and S. Saatchi. 2008. Interactions between rainfall, deforestation and fires during recent years in the Brazilian Amazonia. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **363**:1779–1785.
- Bartholome, E., and A. S. Belward. 2005. GLC2000: a new approach to global land cover mapping from earth observation data. *International Journal of Remote Sensing* **26**:1959–1977.
- Brooks, T. M., R. A. Mittermeier, G. A. B. da Fonseca, J. Gerlach, M. Hoffmann, J. F. Lamoreux, C. G. Mittermeier, J. D. Pilgrim, and A. S. L. Rodrigues. 2006. Global biodiversity conservation priorities. *Science* **313**:58–61.
- CBD (Convention on Biological Diversity). 2004. Programme of work on protected areas. Convention on Biological Diversity, Montreal.
- Coad, L., F. Leverington, C. Nolte, and M. Hockings. 2011. Management effectiveness tracking tool: the global database. University of Oxford, Oxford, United Kingdom.
- Cochrane, M. A. 2011. The past, present, and future importance of fire in tropical rainforests. Pages 213–240 in M. Bush, editor. *Tropical rainforest responses to climate change*. Springer, Berlin.
- Dudley, N., A. Belokurov, L. Higgins-Zogib, M. Hockings, S. Stolton, and N. Burgess. 2007. Tracking progress in managing protected areas around the world. *World Wildlife Fund*, Gland, Switzerland.
- Ervin, J. 2003. WWF: rapid assessment and prioritization of protected area management (RAPPAM) methodology. *World Wildlife Fund*, Gland, Switzerland.
- Ferraro, P. J. 2008. Asymmetric information and contract design for payments for environmental services. *Ecological Economics* **65**:810–821.
- Ferraro, P. J. 2009. Counterfactual thinking and impact evaluation in environmental policy. Pages 75–84 in M. Birnbaum and P. Mickwitz, editors. *Environmental program and policy evaluation: addressing methodological challenges. New directions for evaluation*. American Evaluation Association, Fairhaven, Massachusetts.
- Ferraro, P. J., and S. K. Pattanayak. 2006. Money for nothing? A call for empirical evaluation of biodiversity conservation investments. *Public Library of Science Biology* **4** DOI:10.1371/journal.pbio.0040105.
- Fischer, G., F. Nachtergaele, S. Prieler, H. T. van Velthuisen, L. Verelst, and D. Wiberg. 2007. Global agro-ecological zones assessment for agriculture. *International Institute for Applied Systems Analysis*, Laxenburg, Austria.
- Fuller, R. A., E. McDonald-Madden, K. A. Wilson, J. Carwardine, H. S. Grantham, J. E. M. Watson, C. J. Klein, D. C. Green, and H. P. Possingham. 2010. Replacing underperforming protected areas achieves better conservation outcomes. *Nature* **466**:365–367.
- Global Environment Facility (GEF). 2009. GEF focal area: biodiversity. GEF, Washington, D.C.
- Gordon and Betty Moore Foundation (GBMF). 2006. Andes-Amazon initiative – overview. GBMF, Palo Alto, California.
- Hansen, M. C., R. S. Defries, J. R. Townshend, M. Carroll, C. Dimiceli, and R. A. Sohlberg. 2003. Global percent tree cover at a spatial resolution of 500 meters: first results of the MODIS Vegetation Continuous Fields algorithm. *Earth Interactions* **7**:1–15.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climated surfaces for global land areas. *International Journal of Climatology* **25**:1965–1978.
- Hockings, M., S. Stolton, F. Leverington, N. Dudley, J. Courrau, P. Valentine, and S. Editor. 2006. Evaluating effectiveness: a framework for assessing management effectiveness of protected areas. *International Union for Conservation of Nature*, Gland, Switzerland.
- Joppa, L. N., and A. Pfaff. 2011. Global protected area impacts. *Proceedings of the Royal Society B* **278**:1633–1638.
- Justice, C. 2002. The MODIS fire products. *Remote Sensing of Environment* **83**:244–262.
- Kasperek, M., R. Mack, C. Mayer, and K. Riha. 2010. Biodiversity in German development cooperation 2010. *Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH*, Eschborn, Germany.
- Leverington, F., K. L. Costa, J. Courrau, H. Pavese, C. Nolte, M. Marr, L. Coad, N. Burgess, B. Bomhard, and M. Hockings. 2010. Management effectiveness evaluation in protected areas – a global study. 2nd edition. University of Queensland, Brisbane.
- Malamud, B. D., J. D. A. Millington, and G. L. W. Perry. 2005. Characterizing wildfire regimes in the United States. *Proceedings of the National Academy of Sciences* **102**:4694–4699.
- Martin, A. S., and J. F. Rieger. 2003. The parks in peril site consolidation scorecard. *The Nature Conservancy*, Arlington, Virginia.
- Mascia, M. B., and S. Pailler. 2011. Protected area downgrading, downsizing, and degazettement (PADDD) and its conservation implications. *Conservation Letters* **4**:9–20.
- Morton, D. C., R. S. Defries, J. T. Randerson, L. Giglio, W. Schroeder, and G. R. Van Der Werf. 2008. Agricultural intensification increases deforestation fire activity in Amazonia. *Global Change Biology* **14**:2262–2275.
- NIMA (National Imagery and Mapping Agency). 2000. Vector MAP LEVEL 0 (Digital chart of the world). 5th edition. NIMA, Reston, Virginia.
- Nelson, A. 2008. Travel time to major cities: a global map of accessibility. *Global Environment Monitoring Unit, Joint Research Centre of the European Commission*, Ispra, Italy.
- Nelson, A., and K. M. Chomitz. 2011. Effectiveness of strict vs. multiple use protected areas in reducing tropical forest fires: a global analysis using matching methods. *Public Library of Science ONE* **6** DOI:10.1371/journal.pone.0022722.

- Nolte, C., F. Leverington, A. Kettner, M. Marr, G. Nielsen, B. Bomhard, S. Stolton, S. Stoll-Kleemann, and M. Hockings. 2010. Protected area management effectiveness assessments in Europe. A review of application, methods and results. Bundesamt für Naturschutz (BfN), Bonn.
- Olson, D. M. et al. 2001. Terrestrial ecosystems of the world: a new map of life on Earth. *BioScience* **51**:933–938.
- Pivello, V. R. 2011. The use of fire in the Cerrado and Amazonian rainforests of Brazil: past and present. *Fire Ecology* **7**:24–39.
- Pueyo, S., P. M. L. de Alencastro Graça, R. I. Barbosa, R. Cots, E. Cardona, and P. M. Fearnside. 2010. Testing for criticality in ecosystem dynamics: the case of Amazonian rainforest and savanna fire. *Ecology Letters* **13**:793–802.
- Saatchi, S. S., R. A. Houghton, R. C. Dos Santos Alvalá, J. V. Soares, and Y. Yu. 2007. Distribution of aboveground live biomass in the Amazon basin. *Global Change Biology* **13**:816–837.
- Sekhon, J. 2007. Multivariate and propensity score matching software with automated balance optimization: the Matching package for R. *Journal of Statistical Software* **42**:1–51.
- Soares-Filho, B. et al. 2010. Role of Brazilian Amazon protected areas in climate change mitigation. *Proceedings of the National Academy of Sciences* **107**:10821–10826.
- Stolton, S., M. Hockings, N. Dudley, K. MacKinnon, T. Whitten, and F. Leverington. 2007. Management effectiveness tracking tool. Reporting progress at protected area sites. 2nd edition. World Wildlife Fund, Gland, Switzerland.
- Trumper, K., M. Bertzky, B. Dickson, G. van der Heijden, M. Jenkins, and P. Manning. 2009. The natural fix? The role of ecosystems in climate mitigation. World Conservation Monitoring Center, United Nations Environment Programme, Cambridge, United Kingdom.
- World Conservation Monitoring Centre (WCMC). 2011. World database of protected areas. WCMC, Cambridge, United Kingdom.
- Zimsky, M., P. Ferraro, F. Mupemo, J. Robinson, and N. Sekhran. 2010. Results of the GEF biodiversity portfolio monitoring and learning review mission, Zambia. Enhancing outcomes and impact through improved understanding of protected area management effectiveness. Global Environment Facility, Washington, D.C.

