Detecting abundance trends under uncertainty: the influence of budget, observation error and environmental change

A. Nuno1,3, E. J. Milner-Gulland1 & N. Bunnefeld2

1 Department of Life Sciences, Imperial College London, Ascot, UK
2 Biological and Environmental Sciences, School of Natural Sciences, University of Stirling, Stirling, UK
3 College of Life and Environmental Sciences, University of Exeter Cornwall Campus, Penryn, Cornwall, UK

Keywords
abundance trends; monitoring decisions; observation error; population monitoring; savannah ungulates; threat inference; uncertainty; virtual ecologist.

Abstract
Population monitoring must robustly detect trends over time in a cost-effective manner. However, several underlying ecological changes driving population trends may interact differently with observation uncertainty to produce abundance trends that are more or less detectable for a given budget and over a given time period. Errors in detecting these trends include failing to detect declines when they exist (type II), detecting them when they do not exist (type I), detecting trends in one direction when they are actually in another direction (type III) and incorrectly estimating the shape of the trend. Robust monitoring should be able to avoid each of these error types. Using monitoring of two contrasting ungulate species and multiple scenarios of population change (poaching, climate change and road development) in the Serengeti ecosystem as a case study, we used a ‘virtual ecologist’ approach to investigate monitoring effectiveness under uncertainty. We explored how the prevalence of different types of error varies depending on budgetary, observational and environmental conditions. Higher observation error and conducting surveys less frequently increased the likelihood of not detecting trends and misclassifying the shape of the trend. As monitoring period and frequency increased, observation uncertainty was more important in explaining effectiveness. Types I and III errors had low prevalence for both ungulate species. Greater investment in monitoring considerably decreased the likelihood of failing to detect significant trends (type II errors). Our results suggest that it is important to understand the effects of monitoring conditions on perceived trends before making inferences about underlying processes. The impacts of specific threats on population abundance and structure feed through into monitoring effectiveness; hence, monitoring programmes must be designed with the underlying processes to be detected in mind. Here we provide an integrated modelling framework that can produce advice on robust monitoring strategies under uncertainty.

Introduction
Monitoring is an essential tool used to trigger interventions, inform decisions, measure success against stated objectives and learn about the system (Yoccoz, Nichols & Boulinier, 2001; Lindenmayer et al., 2012). Monitoring aims to draw inferences about changes in the observed system over time (Yoccoz et al., 2001) and must be able to detect true trends over time while taking costs into consideration (Kinahan & Bunnefeld, 2012). In some cases, time, budget and observational constraints may even mean that managers may be better off allocating resources to other interventions instead of monitoring (McDonald-Madden et al., 2010). Monitoring effectiveness (ability to detect trends) and efficiency (ability to do this at low cost) are thus key considerations when planning conservation interventions.

The importance of detecting changes at appropriate spatial and temporal scales and with adequate confidence levels has often been emphasized (Field et al., 2004; Jones, 2011), but a number of factors affect monitoring effectiveness, ultimately affecting management decisions and their robustness to uncertainty. For example, the time frame over which change can reliably be detected might not match that required for management (Maxwell & Jennings, 2005), monitoring effort may not be enough or appropriately targeted to detect trends (Katzner, Milner-Gulland & Bragin, 2007) and sampling design may not be optimal (Blanchard, Maxwell & Jennings, 2008). The degree of environmental
and demographic stochasticity also affects the reliability of monitoring data (Hauser, Pople & Possingham, 2006).

In the face of limited resources in conservation, monitoring is generally limited by budgets and varies with the manager’s willingness to accept different error types (Field et al., 2004; Rhodes et al., 2006; Lindenmayer et al., 2012). For example, type I errors (α; rejecting the null hypothesis when it is true, such as when a species is reported to be declining but is actually stable) may cause unnecessary restrictions and waste resources, while type II errors (β; failing to detect a difference that is present, such as concluding that a species is stable when it is actually declining) could mean failing to implement required management interventions and potentially allowing the species to go extinct (Brosi & Biber, 2009). Other potential types of error, rarely considered when planning and evaluating monitoring programmes, are type III errors (correctly rejecting the null hypothesis but incorrectly inferring the direction of the effect; Morrison, 2007), and misidentifying the shape of the population trajectory (e.g. by fitting linear models when trends are non-linear) despite the potential use of shapes of trends to identify threatening processes (Di Fonzo, Collen & Mace, 2013).

Long-term research in the Serengeti includes monitoring of a range of species, with wildlife censuses having been conducted since the 1950s (Sinclair et al., 2007). Poaching (Loibooki et al., 2002), climate change (Ritchie, 2008) and development of infrastructures, such as a commercial highway (Holdo et al., 2011), have been suggested as current or potential threats to this system. Poaching by local communities and environmental variability have been described as major sources of uncertainty in the system (Pascual & Hilborn, 1995; Nuno et al., 2013b). Observation error during the monitoring process further affects wildlife abundance estimates (Nuno, Bunnefeld & Milner-Gulland, 2013a). Using monitoring of two contrasting ungulate species and multiple scenarios of population change in the Serengeti ecosystem as a case study, we used a ‘virtual ecologist’ approach (Tyre, Possingham & Lindenmayer, 2001) to investigate monitoring effectiveness under uncertainty. Specifically, we explored how the prevalence of different types of error (I, II, III and shape) varies depending on budgetary, observational and environmental conditions.

Methods

Study area and species

We chose two ungulate species to investigate contrasting issues influencing the effectiveness and efficiency of monitoring in savannah ecosystems. The migratory wildebeest population *Connochaetes taurinus*, currently numbering around 1.3 million animals (Hilborn & Sinclair, 2010), has been extensively studied over the last 60 years (Mduma, Sinclair & Hilborn, 1999). The importance of the wildebeest migration has often been demonstrated, both for its ecological significance and as a source of tourism revenue (Norton-Griffiths, 2007; Holdo et al., 2011). The resident population of impala *Aepyceros melampus* found in the Grumeti-Ikorong Game Reserve (Supporting Information Appendix S1) has received considerably less attention but represents a suite of resident species important both for local livelihoods as bushmeat, and as constituents of the Serengeti mammal fauna. Currently, there are around 12 000 impala in the game reserve (Grumeti Fund, 2012).

Methodological framework

The modelling framework was divided into four main components (Fig. 1): (1) an ‘operating model’ that produced the simulated ‘true’ population dynamics under different scenarios of population change; (2) an ‘observation model’ that simulated monitoring of wildlife populations over time; (3) the ‘assessment model’ that simulated a manager’s estimation of trends in wildlife abundance based on the simulated monitoring data; and (4) an evaluation of monitoring effectiveness, in which discrepancies between ‘true’ and ‘observed’ trends and their drivers were investigated.

Operating model

Ungulate population dynamics

We used post-breeding, age-structured, two-sex matrix models to represent ungulate population dynamics (Caswell, 2001). The models include juveniles (<1 year old), yearlings (second year), adults (>2 years old) and senescent adults (impala: ≥8 years; wildebeest: ≥14 years). The matrix model was parameterized using vital rates from studies on wildebeest (Mduma, Hilborn & Sinclair, 1998; Mduma et al., 1999; Owen-Smith, 2006), impala (Jarman & Jarman, 1973; Fairall, 1983; Owen-Smith, Mason & Ogutu, 2005) and general ungulate life history (Gaillard et al., 2000).
The models account for polygynous mating behaviour (Caswell, 2001) and the effects of dry season rainfall and density dependence on ungulate mortality (Gaillard, Festa-Bianchet & Yoccoz, 1998; Mduma et al., 1999; Gaillard et al., 2000; Owen-Smith, 2006). The structure and parameterization of these models is fully described in Supporting Information Appendix S2.

**Drivers of change**

To investigate the ability of monitoring to detect population trends under a number of types of threat, we considered that both ungulate populations were potentially affected by three types of processes. We used simplified scenarios to illustrate realistic conditions of change under which monitoring may be conducted.

**Poaching**

Illegal hunting occurs throughout the Serengeti (Loibooki et al., 2002) but its prevalence is highly uncertain (Nuno et al., 2013b). Male bias in wildebeest offtake due to behavioural factors has often been suggested, with reported male selectivity ranging from 138 to 148% (i.e. the proportion of males in the harvest offtake is 38–48% higher than in the population; Georgiadiis, 1988; Hofer, East & Campbell, 1993; Holmern et al., 2006). The estimates of annual wildebeest offtake in the Serengeti range from 40 000 to 129 000 animals (Mduma et al., 1999; Rentsch & Packer, 2014), corresponding to 3–10% of the current population size.

In the poaching scenario, we assumed a 10% harvest rate for wildebeest, to be precautionary, and a 5% rate for impala, which is less heavily targeted by poachers (Rentsch & Packer, 2014). We assumed a rate of 143% male selectivity in wildebeest offtake, the median of the published estimates.

**Climate change affecting rainfall trends and variability**

In sub-Saharan Africa, the predicted primary effect of global climate change is on precipitation, but there still remains much debate as to which areas will receive more or less rainfall (Hulme et al., 2001; Cooper et al., 2013). Global climate models predict that annual rainfall will increase in East Africa but several studies have suggested that there will be a great deal of regional variation (Ogutu et al., 2008; Mango et al., 2011; Dessu & Melesse, 2012). Ritchie (2008) suggested that the Serengeti will experience decreased and less variable rainfall and that overall rainfall has decreased by approximately 25% over the past 50 years.

In the climate change scenario, the dry season rainfall mean (148 mm) and the dry season rainfall variability, expressed by its standard deviation (69 mm), were assumed to decrease exponentially (rate of annual change: −0.006), resulting in a cumulative 26% decrease over the 50 years of the simulation, similarly to ongoing trends (Ritchie, 2008). These changes subsequently affected rainfall-related processes in our simulated biological model, such as density-dependent survival and mortality (further details provided in Supporting Information Appendix S2).

**Potential effects of road construction**

Vital rates may be affected by a number of processes, such as encroachment and habitat fragmentation. For example, landscape fragmentation can lead to reduced population growth and a lower carrying capacity for migratory ungulates (Hobbs et al., 2008) and the proposed commercial highway in the Serengeti could affect the ability of migratory animals to effectively track high-quality forage resources across the landscape (Holdo et al., 2011). However, these processes often occur unexpectedly and their effects are poorly understood.

In this scenario, we considered potential impacts of a proposed road crossing the Serengeti (Holdo et al., 2011). We assumed consecutive declines in juvenile survival, yearling fecundity, adult fecundity and adult survival, following general patterns of stage-specific vulnerability to changes in environmental stressors among ungulate species (Gaillard et al., 2000), which started 3 years apart and then continued for the rest of the simulation at exponentially increasing rates (annual rate of change = −0.002), resulting in an approximately 10% decrease in vital rates over 50 years.

Changes to the parameters were applied after the initial transient dynamics in the baseline scenario (without any threats) and we ran 10 000 replicates of the operating model for each of the scenarios, producing estimates for ‘true’ trends of population abundance, and their associated uncertainties, under the three different sets of conditions. Five pre-threat and 50 post-threat years of each simulation and iteration were used as outputs.

**Observation model**

We simulated the monitoring of the ‘true’ wildlife abundance obtained from the operating model. In the Serengeti ecosystem, migratory wildebeest are monitored through aerial point sampling and impala using systematic reconnaissance flight surveys (see Nuno et al., 2013a for a description of the monitoring procedure and wildlife observation model). Monitoring was assumed to be carried out using the current methods, and we simulated the effects of low- and high-monitoring budgets as defined in Nuno et al. (2013a).

Unstandardized estimates of precision (measured as the coefficient of variation, CV) and accuracy (percentage discrepancy between the mean-estimated population size and the simulated known population size) were taken from Nuno et al. (2013a: Tables S2 and S3). These estimates were obtained by fitting generalized linear models to simulated precision and accuracy as a function of multiple sources of observation uncertainty for wildebeest and impala monitoring, such as sampling effort, ecological features and animal detectability (see Nuno et al., 2013a for a full
description). Values of accuracy and CV were then applied to the ‘true’ wildlife estimates to obtain ‘observed’ abundance from ‘true’ abundance for each simulation and iteration.

Monitoring scenarios
Monitoring was simulated under different conditions of survey frequency (every 1, 3 or 5 years), monitoring length (5, 10, 25 or 50 years), observation error (none, low or high as produced by monitoring budgets defined in Nuno et al. (2013a) and starting point (how long before or after the threat started did monitoring begin; 5 years before, at the same time as the threat started, or 5 years after).

To minimize the influence of simulation variability on any comparisons between different monitoring options, we generated complete datasets under maximum monitoring frequency (yearly) and length (50 years) for each simulated scenario. All monitoring designs were then applied by subsetting the complete dataset under specific conditions. We assumed that at least three data points would be needed for trend assessment so monitoring was annual if it was conducted only for 5 years, and conducted annually or every 3 years if it was conducted for 10 years.

Assessment model
The assessment model simulated the process of trend estimation from wildlife abundance data. Generalized additive models with a normal error distribution and identity link were fitted to both the observed and ‘true’ data, smoothing the time series of abundance using the package mgcv version 1.7–22 in R v.2.15.2 (R Core Team, 2012). We modelled the year effect as a cubic smoothing spline with 3 degrees of freedom (d.f.; given the length of the time series and our interest in trends instead of short-term fluctuations), as a linear term or as a constant (null model). Gamma was set to 1.4 to include a penalty for each additional degree of freedom within the model and prevent model overfitting (Wood, 2006). Selection of the most parsimonious model was performed using the Akaike information criterion corrected for small sample size (AICc). We considered that non-null models would be acceptable instead of null models, and non-linear instead of linear, only if ΔAIC ≥4; ΔAIC ≥4 indicates considerably less support for the alternative model (Burnham & Anderson, 2002). In order to explore the potential effects of ΔAIC on our results, we also ran the assessment model for ΔAIC = 2 (Supporting Information Appendix S5). We averaged model weights for each trend type over all the iterations and, based on the most parsimonious models, quantified how many of the 10 000 replicates showed decreasing or increasing trends for each trend type.

To identify the direction of the trend, we used the sign of the slope if year was fitted as a linear term, or the sign of the mean annual change in smoothed population size if year was fitted as a smoothing factor (Collen et al., 2011).

Analysis of monitoring effectiveness
We investigated differences between ‘true’ and estimated trends as a function of different ecological and monitoring conditions by quantifying different types of error for each scenario. Type I errors (α) were quantified as the percentage of the 10 000 replicates in which a negative or positive trend was detected in the ‘observed’ data but the trend from the ‘true’ data was actually stable (i.e. the null model was the most parsimonious model). Type II errors (β) were quantified as the percentage for which no significant trend was detected in the ‘observed’ data, although this was present in the ‘true’ data. A subset of the type II error (β2) represented the worst case in which negative trends were not detected, despite their presence. Type III errors (γ) were quantified as the percentage of cases in which a trend in the ‘observed’ data was identified in the opposite direction to that in the ‘true’ data. ‘Shape errors’ were quantified as the percentage of non-null cases in which we identified a linear trend as non-linear and vice versa.

To investigate the effect of monitoring conditions on the prevalence of each type of error, we used quasi-likelihood to fit a generalized linear model with a variance–mean relationship derived from the binomial distribution (to account for overdispersion) and a logit link to the simulation results (i.e. the number of times a certain error type occurred out of 10 000 simulations was modelled as a function of species, monitoring length, frequency, starting point, observation error and underlying threat). Relevant two-way interactions between monitoring length, frequency, starting point and observation error were included.

The monitoring budgets were calculated by multiplying current unitary costs from itemized monitoring expenses in the study area for wildlife surveys (J.G.C. Hopcraft, unpubl. data). Inflation, technological advancements and discount rates are expected to affect future expenses but are generally unknown; thus, we kept current costs to simulate into the future. The total costs for each monitoring scenario were expressed relative to the baseline scenario.

Results
Baseline scenarios: ‘true’ population trends under different threat conditions
Under the ‘no threat’ scenario and the baseline parameterization of the biological models (Table 2.1 in Supporting Information Appendix S2), wildebeest and impala generally stabilized at around 1.4 million animals and 14 000 animals, respectively. Other studies in the Serengeti have indicated similar carrying capacities for wildebeest (1.2–1.5 million; Mduma et al., 1999; Holdo et al., 2011) and the impala population in the game reserve has been stabilizing at 12 000 animals (Gruneti Fund, 2012), suggesting that our biological model produces relatively realistic carrying capacities.
Declines were greatest in the scenarios of poaching and road construction, with both species declining, on average, by 43–69% in 50 years (Fig. 4.1 in Supporting Information Appendix S4). On average, non-linear models had greater support than linear and null models for all the scenarios, but impala and wildebeest populations showed differences in the prevalence of the shape and direction of abundance trends depending on the threat type (Table 1). The wildebeest populations generally declined non-linearly in response to all threats, although 37% of populations declined linearly in response to potential effects of road construction. Most impala populations declined non-linearly in response to the potential effects of road construction. In response to poaching, about half of the impala simulations showed non-linear declines and half showed linear declines. The shape and direction of the effects of climate change on impala were more uncertain; 15% of populations remained stable or increased while the others decreased, on average by 31% over the 50 years.

The prevalence of different error types according to threat scenario and species

Under the best monitoring conditions (i.e. 50 years of data collected annually with high-monitoring budgets resulting in CVs around 0.15 for wildebeest and 0.23 for impala), the prevalence of different error types was affected by the specific threat conditions and their impacts on population abundance, structure and trajectory (Table 2).

Under the simulated conditions, the occurrence of type I errors, when a negative or positive trend is detected in the ‘observed’ data but the trend in the ‘true’ data is actually stable, was very low for all threat scenarios ($\alpha \leq 0.02$, Table 2). Similarly, type III errors (identifying a trend in the ‘observed’ data with an opposite direction to that in the ‘true’ data) were low ($\gamma \leq 0.03$) for both the impala and wildebeest populations, although 82% of type III errors related to the more serious situation in which a negative trend was observed as positive (i.e. the remaining 18% of type III errors were positive trends identified as negative).

Type II errors, failing to find a significant trend in the ‘observed’ data although this was present in the ‘true’ data, were relatively low for wildebeest ($\beta \leq 0.34$) but higher for

### Table 1
Average Akaike model weights per trend type (N: null; L: linear; NL: non-linear), prevalence (percentage of 10 000 simulations) of best-fit models and trend direction (increasing ↑ or decreasing ↓), and average total change (%) per trend direction over 50 years for each threat scenario, for the ‘true’ abundance of wildebeest and impala

<table>
<thead>
<tr>
<th>Threat scenario</th>
<th>Ungulate species</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wildebeest</td>
<td>Impala</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average weights</td>
<td>Prevalence (%)</td>
<td>Average total change</td>
<td>Average weights</td>
<td>Prevalence (%)</td>
</tr>
<tr>
<td>None</td>
<td>N: 0.22</td>
<td>N: 57</td>
<td></td>
<td>N: 0.03</td>
<td>N: 7</td>
</tr>
<tr>
<td></td>
<td>L: 0.29</td>
<td>L: ↓9, ↑23</td>
<td>0</td>
<td>L: 0.11</td>
<td>L: ↑11, ↓11</td>
</tr>
<tr>
<td></td>
<td>NL: 0.49</td>
<td>NL: ↑9, ↓3</td>
<td></td>
<td>NL: 0.86</td>
<td>NL: ↑48, ↓22</td>
</tr>
<tr>
<td>Poaching</td>
<td>N: 0</td>
<td>N: 0</td>
<td></td>
<td>N: 0</td>
<td>N: 1</td>
</tr>
<tr>
<td></td>
<td>L: 0</td>
<td>L: ↑0, ↓0</td>
<td>↓50%</td>
<td>L: 0.22</td>
<td>L: ↑0, ↓50</td>
</tr>
<tr>
<td></td>
<td>NL: 1</td>
<td>NL: ↑0, ↓100</td>
<td></td>
<td>NL: 0.78</td>
<td>NL: ↑0, ↓49</td>
</tr>
<tr>
<td>Climate change</td>
<td>N: 0</td>
<td>N: 0</td>
<td></td>
<td>N: 0.02</td>
<td>N: 6</td>
</tr>
<tr>
<td></td>
<td>L: 0</td>
<td>L: ↑0, ↓0</td>
<td>↓40%</td>
<td>L: 0.15</td>
<td>L: ↑4, ↓28</td>
</tr>
<tr>
<td></td>
<td>NL: 1</td>
<td>NL: ↑0, ↓100</td>
<td></td>
<td>NL: 0.83</td>
<td>NL: ↑5, ↓57</td>
</tr>
<tr>
<td>Potential effects of road</td>
<td>N: 0</td>
<td>N: 0</td>
<td></td>
<td>N: 0</td>
<td>N: 0</td>
</tr>
<tr>
<td>construction</td>
<td>L: 0.15</td>
<td>L: ↑90, ↓37</td>
<td>↓63%</td>
<td>L: 0.03</td>
<td>L: ↑0, ↓7</td>
</tr>
<tr>
<td></td>
<td>NL: 0.85</td>
<td>NL: ↑0, ↓63</td>
<td></td>
<td>NL: 0.97</td>
<td>NL: ↑0, ↓93</td>
</tr>
</tbody>
</table>

### Table 2
Prevalence of different error types (out of 10 000 simulations) for each threat scenario from the ‘observed’ wildebeest and impala data, monitored annually over 50 years with a high-monitoring budget

<table>
<thead>
<tr>
<th>Threat scenario</th>
<th>Ungulate species</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wildebeest</td>
<td>Impala</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha$: 0.02</td>
<td>$\alpha$: 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta$: 0.34</td>
<td>$\beta$: 0.76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta_i$: 0.20</td>
<td>$\beta_i$: 0.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\gamma$: 0.01</td>
<td>$\gamma$: 0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>‘shape’: 0.16</td>
<td>‘shape’: 0.51</td>
<td></td>
</tr>
<tr>
<td>Poaching</td>
<td>$\alpha$: 0</td>
<td>$\alpha$: 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta$: 0</td>
<td>$\beta$: 0.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta_i$: 0</td>
<td>$\beta_i$: 0.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\gamma$: 0</td>
<td>$\gamma$: 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>‘shape’: 0.55</td>
<td>‘shape’: 0.50</td>
<td></td>
</tr>
<tr>
<td>Climate change</td>
<td>$\alpha$: 0</td>
<td>$\alpha$: 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta$: 0</td>
<td>$\beta$: 0.66</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta_i$: 0</td>
<td>$\beta_i$: 0.59</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\gamma$: 0.03</td>
<td>$\gamma$: 0.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>‘shape’: 0.15</td>
<td>‘shape’: 0.58</td>
<td></td>
</tr>
<tr>
<td>Potential effects of road</td>
<td>$\alpha$: 0</td>
<td>$\alpha$: 0</td>
<td></td>
</tr>
<tr>
<td>construction</td>
<td>$\beta$: 0</td>
<td>$\beta$: 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta_i$: 0</td>
<td>$\beta_i$: 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\gamma$: 0</td>
<td>$\gamma$: 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>‘shape’: 0.52</td>
<td>‘shape’: 0.68</td>
<td></td>
</tr>
</tbody>
</table>

$\alpha$, prevalence of type I errors; $\beta$, prevalence of type II errors; $\beta_i$, prevalence of ‘worst case scenario’ type II errors; $\gamma$, prevalence of type III errors; shape, prevalence of ‘shape errors’.
impala (β ≤ 0.76), except in the scenario of potential effects of road construction in which both species had similar low levels (Table 2). Moreover, 90% of the type II errors involved negative trends being undetected (i.e. only 10% of undetected trends were positive). Reporting the wrong trajectory shape, that is, identifying a linear trend as non-linear and vice versa, was common for all threat scenarios; on average, 46% of the non-null trends were misclassified, 96% of which were identified as linear but were actually non-linear.

The effect of monitoring conditions on the prevalence of different error types

The occurrence of type I and type III errors was unaffected by any of the monitoring conditions (frequency, length, observation error and starting point), given the threat and assessment scenarios considered in this study, generally remaining at very low levels (Table 4.1 in Supporting Information Appendix S4). Increasing monitoring length did not significantly affect the occurrence of different types of errors; changing the monitoring length tended to change the shape, direction and magnitude of the true trends, offsetting the expected benefit of increasing monitoring length. For example, if monitoring was conducted for only 5 years after the threat, virtually no errors were found because the ‘actual’ trend, to which observed trends were compared, was identified as stable. Therefore, increasing monitoring length was likely both to increase the possibility of errors occurring as well as the possibility of detecting them.

Type II and shape errors were more likely to occur when surveys were conducted with observation error or less frequently. The effects of the level of observation error were, however, strongly conditioned on survey frequency and length of the monitoring period (Table 4.1 in Supporting Information Appendix S4): as surveys were conducted more frequently or monitoring length increased, the importance of observation error increased in determining the ability of monitoring to detect trends correctly. For example, in order to detect true negative trends in wildebeest numbers more than 80% of the time over a 50-year period, one would have to monitor with no observation error every 3 years or with low error every 2 years (Fig. 2). Starting monitoring 5 years before or after the actual threat started only affected the probability of occurrence of type II errors; fewer negative trends went undetected when monitoring started 5 years before, although this effect was less important as monitoring length increased.

Characteristics related to threat type and species explained some of the differences in the likelihood of type II and shape errors. Impala populations were 1–5% more likely to present these errors than those of wildebeest, while keeping all the other variables constant. Threat scenarios were only 4–5% more likely to have shape errors than the ‘no threat’ scenario but the likelihood of failing to detect a negative trend (subset of the type II error; β2) was 7–15% higher in threat scenarios than in the no threat scenario.

Trade-offs between monitoring budget and effectiveness

Negative trends would go undetected 32 or 2% of the time in impala and wildebeest populations, respectively, if conducting annual surveys over 50 years with a low observation error (Fig. 3). A reduction in budget leading to reduced survey frequency and higher observation error would increase the likelihood of not detecting negative trends and misclassifying the shape of trends (Table 4.1 in Supporting Information Appendix S4). For example, when compared with the total costs of conducting annual surveys for 50 years with low observation error, conducting surveys only every 5 years and with higher levels of observation error would save up to 90% of the budget, but negative trends would go undetected more than 80% of the time (Fig. 3).

Discussion

Using a ‘virtual ecologist’ approach, we have linked changes in population abundance and structure caused by simulated realistic conservation threats to specific monitoring
Different factors might affect the monitoring of different wildlife species. In the Serengeti, monitoring of highly aggregated species, such as wildebeest, is improved by increasing survey precision, while for random or slightly aggregated species, such as impala, accuracy is the key factor, being most sensitive to observer effects (Nuno et al., 2013a). By linking population models with a realistic representation of the observation and assessment processes, we can provide monitoring advice that makes explicit the trade-offs between monitoring budgets and power to detect changes. For example, if monitoring in the Serengeti is conducted with the current survey frequency (approximately every 3 years; TAWIRI, 2010) and low observation error, negative trends in wildebeest and impala populations might go undetected approximately 23 or 30% of times, respectively. A reduction in monitoring budgets by 66%, leading to higher observation error and surveys being conducted less frequently, could increase this likelihood to 50%. Our results also suggest that the likelihood of not detecting negative trends would be particularly high in scenarios of climate change for impala. These findings can be used to interpret the data on ungulate population abundance being currently collected in the study area and to aid decisions on budget allocation to make monitoring fit for purpose. This is particularly relevant given the internationally expressed importance of identifying robust and reliable monitoring targets, such as CBD Aichi targets, that can be used to infer declines in specific populations and biodiversity in general (Collen et al., 2009; Rhodes et al., 2011; Porszt et al., 2012). Moreover, the feasibility of acting upon different drivers of change (e.g. rainfall changes or poaching) should be considered when making management decisions and planning their implementation. By providing a relatively intuitive single framework that links population dynamics, scenarios of change and monitoring effectiveness, our approach may ultimately contribute to more robust decisions in applied settings.

Although type II errors, failing to detect effects, may result in serious consequences to the ecosystem, type I errors, incorrectly rejecting the null hypothesis, would result in unnecessary restrictions and waste of resources. Much more attention is given to type I errors (Field et al., 2004; Brosi & Biber, 2009), but our results suggest that for this case study the type I error rate is low and unaffected by most forms of uncertainty under the considered assessment conditions (such as choice of ΔAIC). This means that, under the simulated conditions, reports of population decline in the system are very unlikely to be wrong, suggesting that this information should be promptly used to inform management decisions.

Linear models are commonly applied to population trend assessment (Thomas, 1996) but most populations naturally exhibit non-linear dynamics (Clutton-Brock et al., 1997). Our study shows that making a correct distinction between linear and non-linear trends may be challenging, although these might have different implications for management decision making (e.g. population under ongoing decline or stabilizing at lower carrying capacity). While in
conservation more attention is generally given to detecting declines, our results demonstrate that it is also relatively common for positive trends to be wrongly identified as declines; this could be particularly relevant, for example, when dealing with overabundant populations (Chee & Wintle, 2010). Our approach could be extended to include populations that are increasing; the expectation is that these populations would grow in a non-linear way, and hence that more complex models would better reflect their true trend. In addition, our framework could be useful for exploring how modelling assumptions in both the operating model (e.g. more complex population dynamics models; Cornulier et al., 2013) and assessment model (e.g. distinguishing several types of non-linear shapes; Di Fonzo et al., 2013) might affect the robustness of decisions based on monitoring data. Additionally, as choice of $\Delta$AIC implies trade-offs between type I and II errors (Supporting Information Appendix S5), a ‘strength of evidence’ approach could be adopted instead of defining AIC thresholds that might affect robustness of scientific results (Johnson, 2013).

Applying the concept of ‘learning by doing’ from adaptive management (Keith et al., 2011), our virtual monitoring tool can allow managers to learn from past monitoring data and to narrow down the range of possible processes that could be producing the shapes and trends concerned. A process of adaptive monitoring in which multiple monitoring strategies are implemented and adapted in response to data collected could also be developed (Lindenmayer & Likens, 2009). Most biological surveys are constrained by observational and economic constraints that affect the way resources can be allocated (Field, Tyre & Possingham, 2005). The implications of monitoring uncertainty are often unknown and given little consideration in the design of monitoring programmes worldwide. As shown in this study, the likelihood of not detecting negative trends and misclassifying shapes may be too high to be ignored. Uncertainty mitigation efforts must be focused on the kinds of information that are most valuable and make a meaningful difference to our understanding of processes, and to the way we manage threats (Wintle, Runge & Bekessy, 2010; Runge, Converse & Lyons, 2011; Running, Wilson & Rhodes, 2013). Decision-theoretic approaches that incorporate uncertainties and trade-offs are essential to providing clear and transparent advice for conservation decision making (Chee & Wintle, 2010; McDonald-Madden et al., 2010), and ultimately promoting the efficient use of scarce conservation resources (Mackenzie, 2009).

Acknowledgements

We thank Grant Hopcraft for sharing unpublished data. A.N. was supported by the Portuguese Foundation for Science and Technology (FCT; doctoral grant SFRH/BD/43186/2008). N.B. and E.J.M.-G. were supported by the European Commission under the HUNT project of the 7th Framework Programme for Research and Technological Development. Jonathan Rhodes, Andrew Knight, Nigel Yoccoz and an anonymous reviewer gave valuable comments on the manuscript. This paper is a contribution to Imperial College’s Grand Challenges in Ecosystems and the Environment initiative.

References


Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

Appendix S1. Map of study area.

Appendix S2. Ungulate population dynamics.

Appendix S3. Parameter determination for wildebeest rainfall- and density-dependence.

Appendix S4. Supplementary figures and tables supporting results.

Appendix S5. Exploratory analysis of the potential effects of AICc.