MIKE analysis for Africa - Summary

Carl Schwarz
Statistician: MIKE-ETIS Technical Advisory Group (TAG)

Contents

1 Introduction

In this document, we summarize how yearly-trends in the Proportion of Illegally Killed Elephants (PIKE) from MIKE (Monitoring Illegally Killed Elephants) monitoring sites are found. These are first found at the site level and then aggregated to the continental and sub-regional level.

1.1 Survey protocol

Briefly, MIKE data is collected on an annual basis in designated MIKE sites by law enforcement and ranger patrols in the field and through other means. When an elephant carcass is found, site personnel try to establish the cause of death and other details, such as sex and age of the animal, status of ivory, and stage of decomposition of the carcass. This information is recorded in standardized carcass forms, details of which are then submitted to the MIKE Programme. As expected, different sites report widely different numbers of carcasses, as encountered carcass numbers are a function of: population abundance; natural mortality rates; the detection probabilities of elephant carcasses in different habitats; differential carcass decay rates; levels of illegal killing; and levels of search effort and site coverage. Because of these features of the survey data, the number of carcasses found is unlikely to be proportional to the total mortality and trends in observed numbers of illegally killed elephants may not be informative of the underlying trends in poaching pressure.

1.2 Definition of ***PIKE***

Consequently, the observed proportion of illegally killed elephants (PIKE) as an index of poaching levels has been used in the MIKE analysis in an attempt to account for differences in patrol effort between sites and over time:

$$PIKE_{sy} = \frac{Number\ of\ illegally\ killed\ elephants_{sy}}{Total\ Carcasses\ Examined_{sy}}$$

where the subscripts *sy* refer to site and year respectively.

Computing a continent-wide PIKE is challenging for several reasons, including as mentioned above:

- Detection probabilities of elephant carcasses in various habitats differ.
- Levels of search effort and site coverage differ between sites.
- Not all sites report in all years.
- Number of carcasses in both categories varies considerably across space and time.

In past years, a simple linear model on the PIKE values was computed as described in https://cites.org/sites/default/files/notif/E-Notif-2019-046.pdf. This is denoted the *LSMeans* approach.

The PIKE trend is calculated using estimated marginal means of a linear model weighted by the total number of observed carcasses. The continental PIKE trend is estimated based on a model with subregion and year as factors, while the subregional trends are estimated from a model using country and year as factors.

As indicated in the MIKE report to the 18th meeting of the Conference of Parties to CITES available at https://cites.org/sites/default/files/eng/cop/18/doc/E-CoP18-069-02.pdf, the CITES Secretariat, in collaboration with the MIKE-ETIS TAG statisticians and an independent statistician, initiated a process to review the MIKE analytical methodology to determine whether it could be refined or its scientific robustness improved, and to further enhance the analytical basis for MIKE. The approach included a review of the current methodology, and consideration of new statistical developments and, therefore, alternative methods or models for *PIKE* trend analysis, while taking into consideration the imbalances and inconsistencies inherent in the data. Burn, Underwood and Blanc (2011) used a Bayesian hierarchical model based on a generalized linear mixed model (logistic regression with random effects) to resolve many of these issues.

Issues with the collection of the data such as:

- search patterns are not random and often directed to illegally killed elephants;
- misclassification of animals as unknown causes of mortality rather than illegally or natural mortality based on the preponderance of evidence;
- errors in the data

are not considered here, but the CITES MIKE CCU is preparing discussion papers that explains these issues in more detail.

This document is a brief summary of the result of using a GLMM to estimate trends in *PIKE*. A technical document is available with detailed results, model assessment, and a sensitivity analysis.

1.3 Why change from the LSMeans approach

The *LSMeans* approach has been used for the *PIKE* trend analysis in the reports to previous meetings of the Conference of the Parties (CoP16, Bangkok, 2013 in document CoP16 Doc. 53.1; and CoP17, Johannesburg, 2016, in document CoP17 Doc. 57.5); and to meetings of the Standing Committee (SC62, Geneva, July 2012, in document SC62 Doc. 46.1 (Rev. 1); SC65, Geneva, July 2014, in document SC65 Doc. 42.1; SC66, Geneva, January 2016, in document SC66 Doc. 47.1; SC69, Geneva, November 2017, in document SC69 Doc. 51.1; and SC70, Sochi, October 2018, in document SC70 Doc. 49.1). However, there are a number of issues with this approach

- Predicted *PIKE* values could be less than 0 or greater than 1
- Each country is given equal weight regardless of the number of *MIKE* sites or the abundance of elephants in each site at the sub-regional level. For example, in Eastern Africa, both Kenya (4 *MIKE* sites) and Eritrea (1 *MIKE* site) are given equal weight in computing the subregional PIKE. But at the continental level, countries are ignored and data pooled over all sites in a sub-region.
- The *LSMeans* is not consistent in how to aggregate to larger levels. The mean of the *LSMeans* estimates of *PIKE* at the sub-regional level from the individual sub-regional models, will not match the estimated *PIKE* at the continental level estimated using the continental model because country is or is not included in the two models.
- It is difficult to apply different weighting, e.g. by elephant populations at the *MIKE* or country level.
- The *LSMeans* approach imputes *PIKE* for missing country-year combinations at the sub-regional level, but the individual *MIKE* sites could differ in the patterns of missingness across years. Consequently, the country may appear to have data for every year, but the country-wide aggregate data is based on a different combination of *MIKE* sites across the years. A similar problem occurs in the analysis at the continental level.
- Not all sources of variation are included in estimates of uncertainty. For example, binomial variation (Section 5.2 of the technical document) at the site-year level is not included, i.e. if the actual *PIKE* of all elephants at the *MIKE* site is .20, then the observed *PIKE* in the sample of carcasses examined will vary around 0.20.

The impact of giving equal weight at the country level or at the *MIKE* site level is explored in more detail in Appendix 3 where this has a noticeable impact on the sub-regional *PIKE* for East Africa.

1.4 Advantages of GLMM

At the recent MIKE-ETIS TAG meeting (September 2019, Nairobi), the use of the generalized linear mixed model (GLMM) was recommended going forward. This document provides an analysis of the *PIKE* data using a GLMM, compares the results to those from the previous analyses, and explores the impacts of various assumptions on the estimated *PIKE*.

The advantages of the GLMM approach are:

- The new model fully accounts for the binomial structure at the site-year level, i.e. of n carcasses observed, x are illegally killed.
- The new model fully accounts for different sample sizes, i.e. a *PIKE* based on observing 1 illegally killed elephants out of 2 elephant carcasses is given different weight than a *PIKE* based on observing 20 illegally killed elephants out of 40 elephant carcasses.
- Imputation of missing data takes place at the MIKE site level based on the relationship of *PIKE* in this site to the *PIKE* at other sites across time.
- Each *MIKE* site is given equal weight when computing the continental or sub-regional *PIKE*. Consequently, countries with more *MIKE* sites will automatically be given more weight in the aggregate *PIKE* estimates.
- It is easy to apply other weightings, e.g. by the elephant population abundances at each *MIKE* site when computing an aggregate *PIKE*
- Multiple sources of variation are automatically included, e.g. binomial variation at the site-year level, site-year interactions representing the changes in *PIKE* over time at a particular *MIKE* site, and

site-to-site variation. This variation is automatically included in the uncertainty of the aggregate *PIKE* estimates.
- Uncertainty of the estimated *PIKE* at the continental or sub-regional level can be computed assuming that the current *MIKE* sites are index sites (and fixed) or random sample of potential *MIKE* sites. The current design is somewhat between these two extremes, and so the uncertainty reported under these two ways of viewing the current set of *MIKE* sites represents a lower and upper bound of uncertainty.
- It is easy to include estimates of uncertainty in elephant abundances when weighting *MIKE* sites by elephant abundances. This is currently not yet done because of issues in determining the uncertainty in these abundance estimates (see Section 8.7 of the technical document).
- The current model implicitly accounts for spatial autocorrelation in the *PIKE* among *MIKE* sites that are geographically close. The estimated site-level effects are similar for sites that are geographically close and have similar levels of governance and poaching.
- It will be possible to extend this model to account explicitly for spatial and spatial-temporal autocorrelation (Zuur, 2019)

1.5 *PIKE* as an index of poaching pressure

The value of *PIKE* computed by *LSMeans* and the *GLMM* approach should be considered as an INDEX of poaching pressure. We hope that trends in the index reflect trends in the actual levels of poaching. Converting the value of *PIKE* into a measure of actual poaching mortality is complicated due to the following:

- Effects of changing natural mortality over time are confounded with changes in *PIKE*.
- MIKE sites are not selected at random and so the marginal mean *PIKE* may not be representative of the actual marginal mean *PIKE*.
- Poaching at MIKE sites may not be representative of poaching in areas outside of the MIKE site.
- The unweighted marginal mean *PIKE* gives equal weight to every MIKE site. We have investigated in the technical report the impact of weighting sites by the population abundance of elephants associated with that site, but have not included estimates of uncertainty in the population abundance estimates.
- There are a number of issues with the collection of the data such as:
  - Management related deaths are included in the number of carcasses examined (the denominator of *PIKE*) at this point in time;
  - search patterns are not random and often directed to illegally killed elephants;
  - misclassification of animals as unknown causes of mortality rather than illegally or natural mortality based on the preponderance of evidence;
  - errors in the data

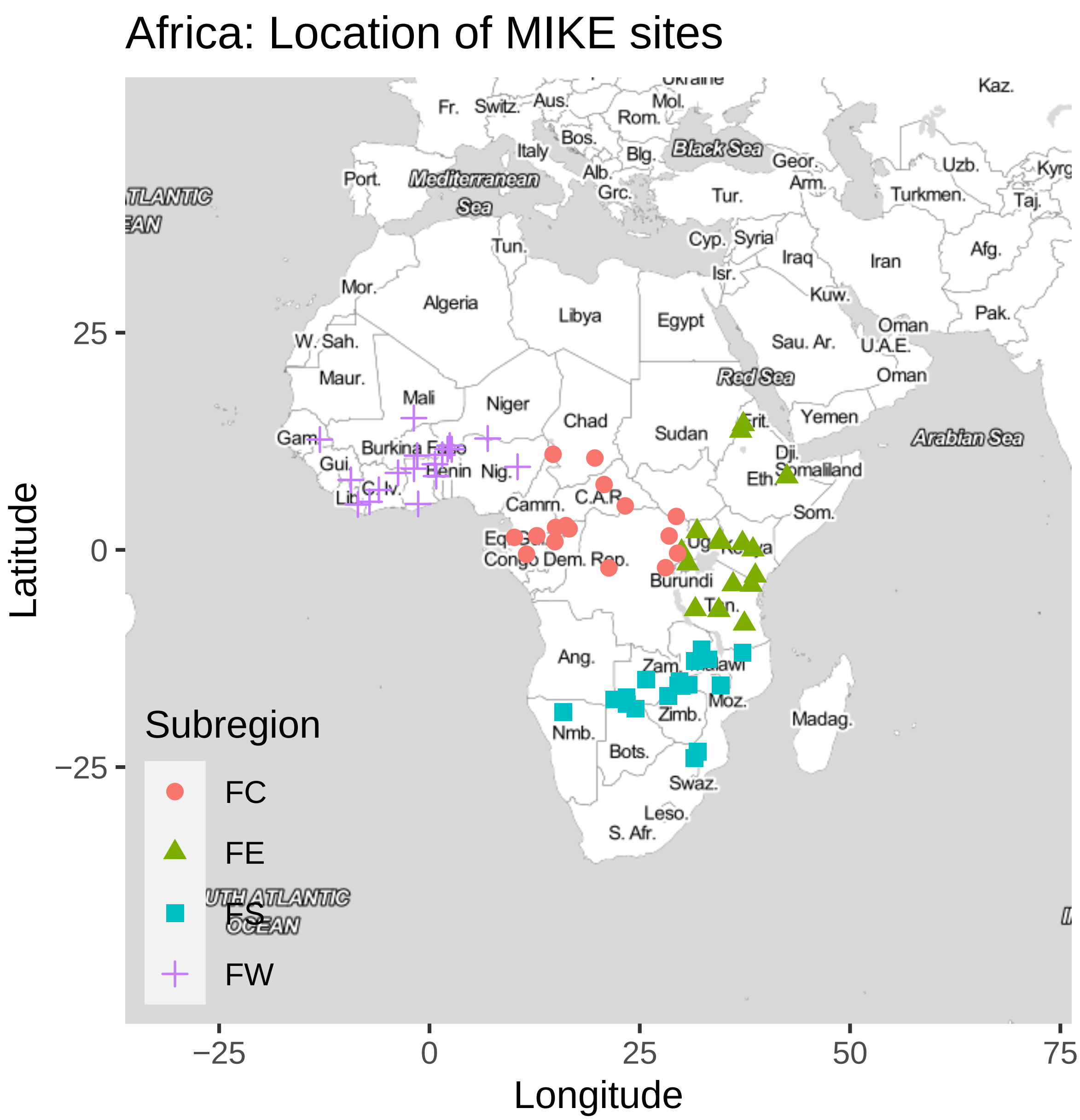
The CITES MIKE CCU is preparing a discussion paper that explores data issues in more detail.

For these reasons, great care should to be taken and assumptions should be well documented when converting the estimated *PIKE* to actual levels of poaching mortality.

2 Exploration of PIKE data

2.1 Location of MIKE sites

There are 68 MIKE sites in Africa broken into 4 regions. Data from only 61 MIKE sites across the sub-regions are used in the analysis because some sites have never submitted data or are currently lacking population abundance data.

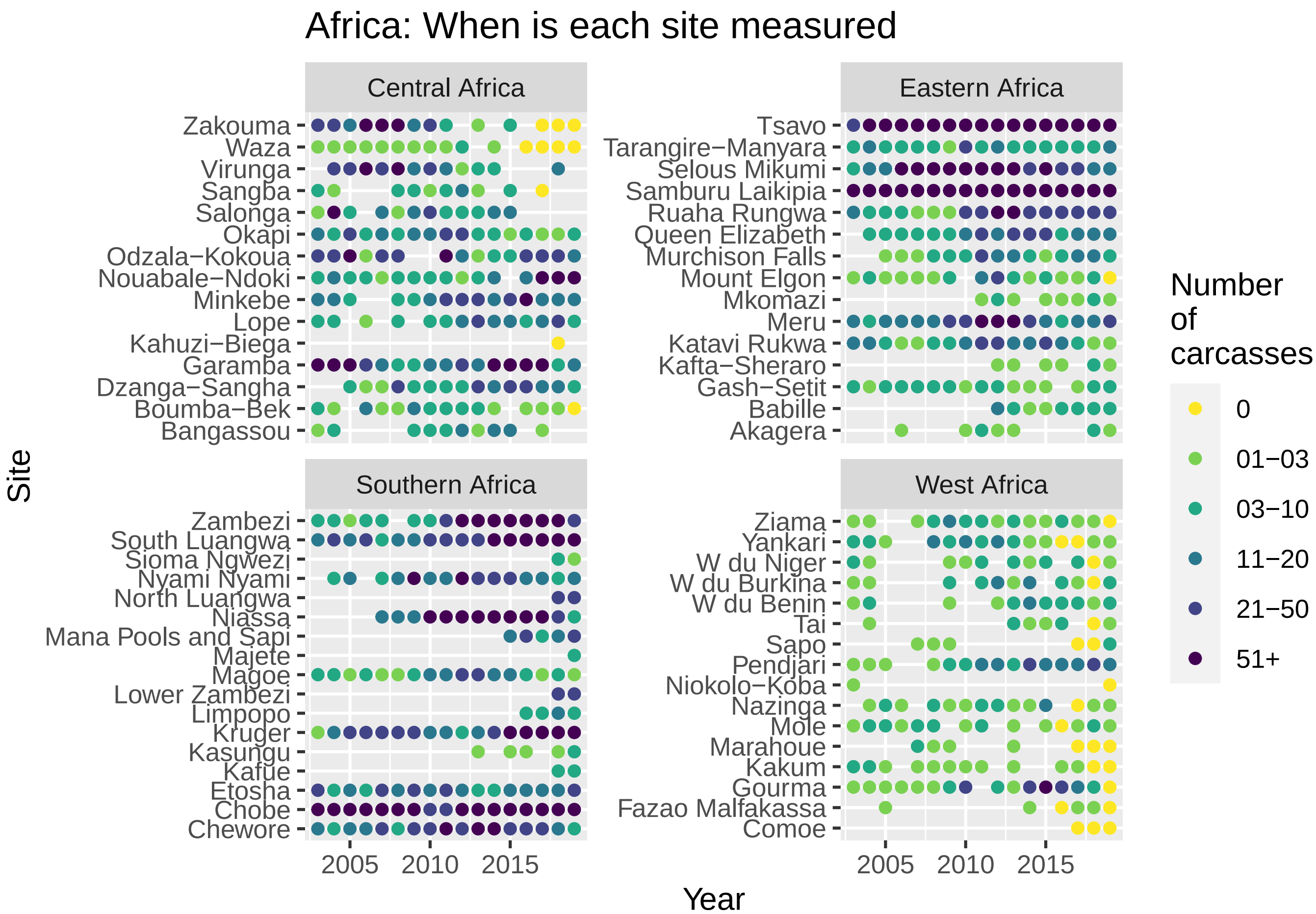


2.2 MIKE sites with *PIKE* data

In addition to the above, some sites have reported 0 (zero) carcasses detected in some years.

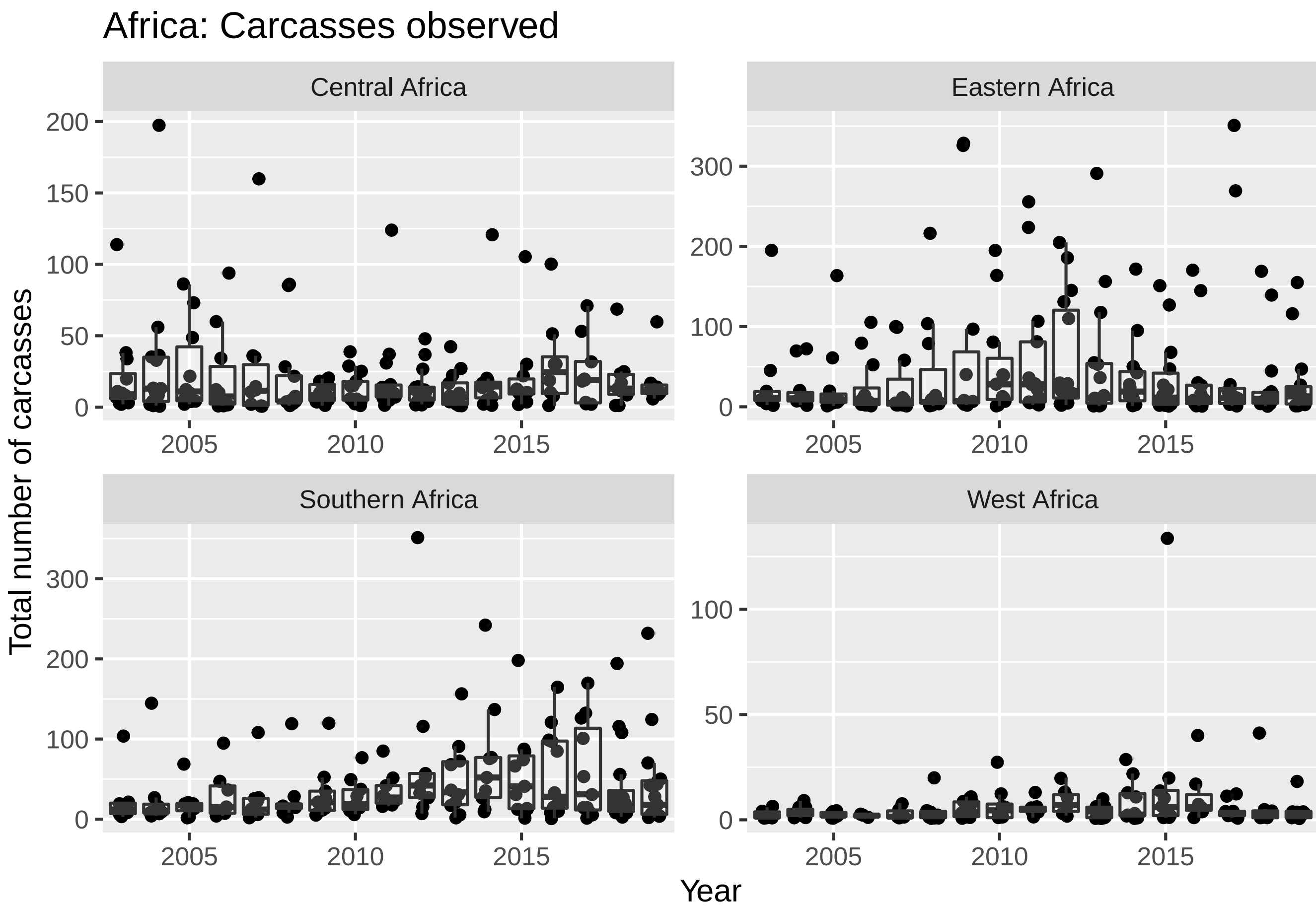
The current analysis treats site that did not report on any carcasses in a year (no patrol effort) and a site that reports 0 carcasses examined in year (patrol effort but no carcasses found) in the same way. This is because information on patrol effort is not currently used in the analysis and only the number of carcasses examined and the number of illegally killed elephants in the sample of carcasses is used. In the latter case, 0 illegal carcasses out of 0 carcasses examined gives a *PIKE* for that site-year of 0/0 which is indeterminate and cannot be used in any mathematical analysis of *PIKE*.

The following plot shows that there are some sites that have reported data for at least one carcass in as little as one year, but other sites have reported data for at least one carcass in almost every year.



In total, there are 711 unique site-years in Africa since 2003 where data has been reported (and the number of reported carcasses > 0).

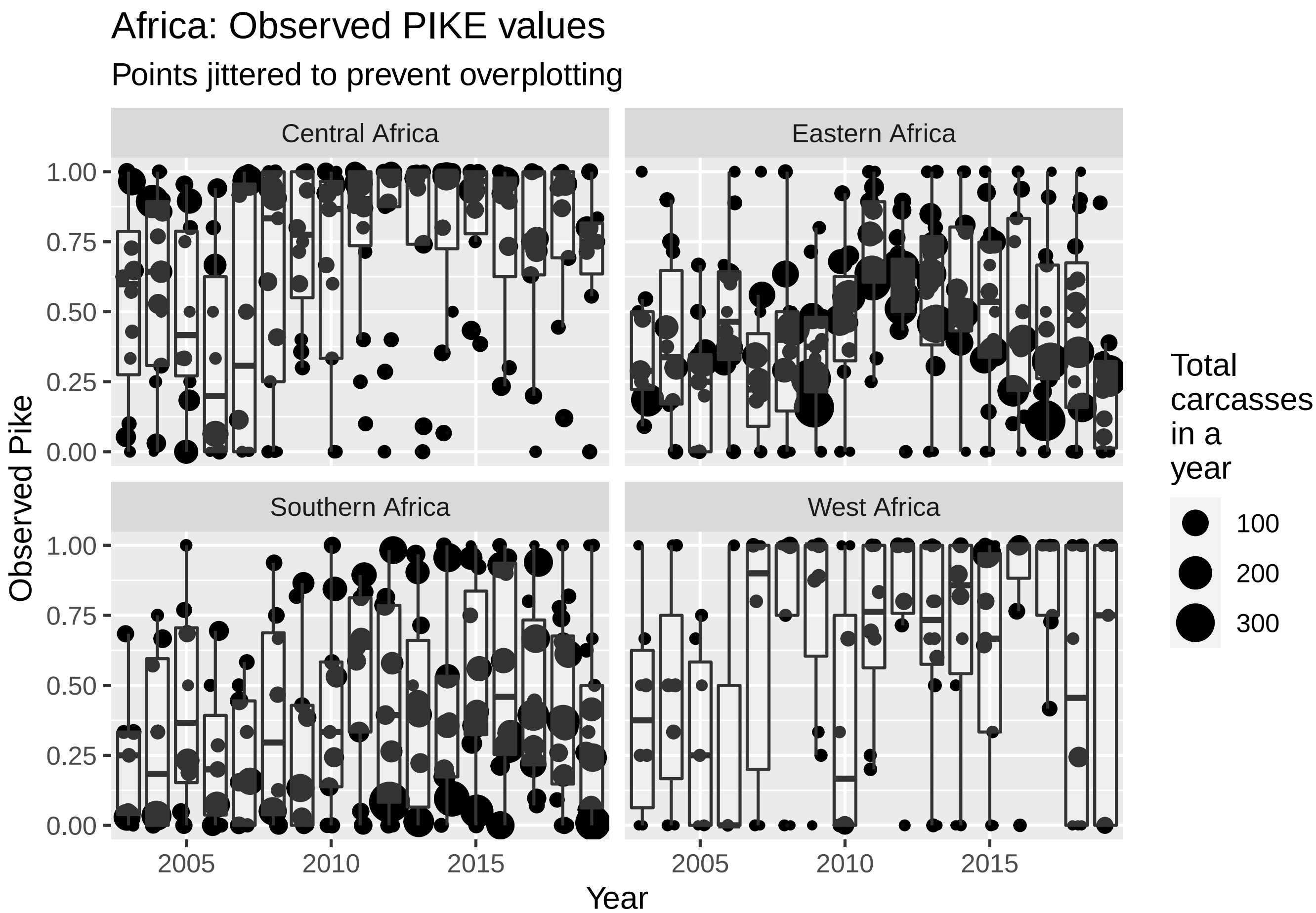
The number of carcasses reported in each site-year since 2003 varies enormously from 1 to 351 carcasses.



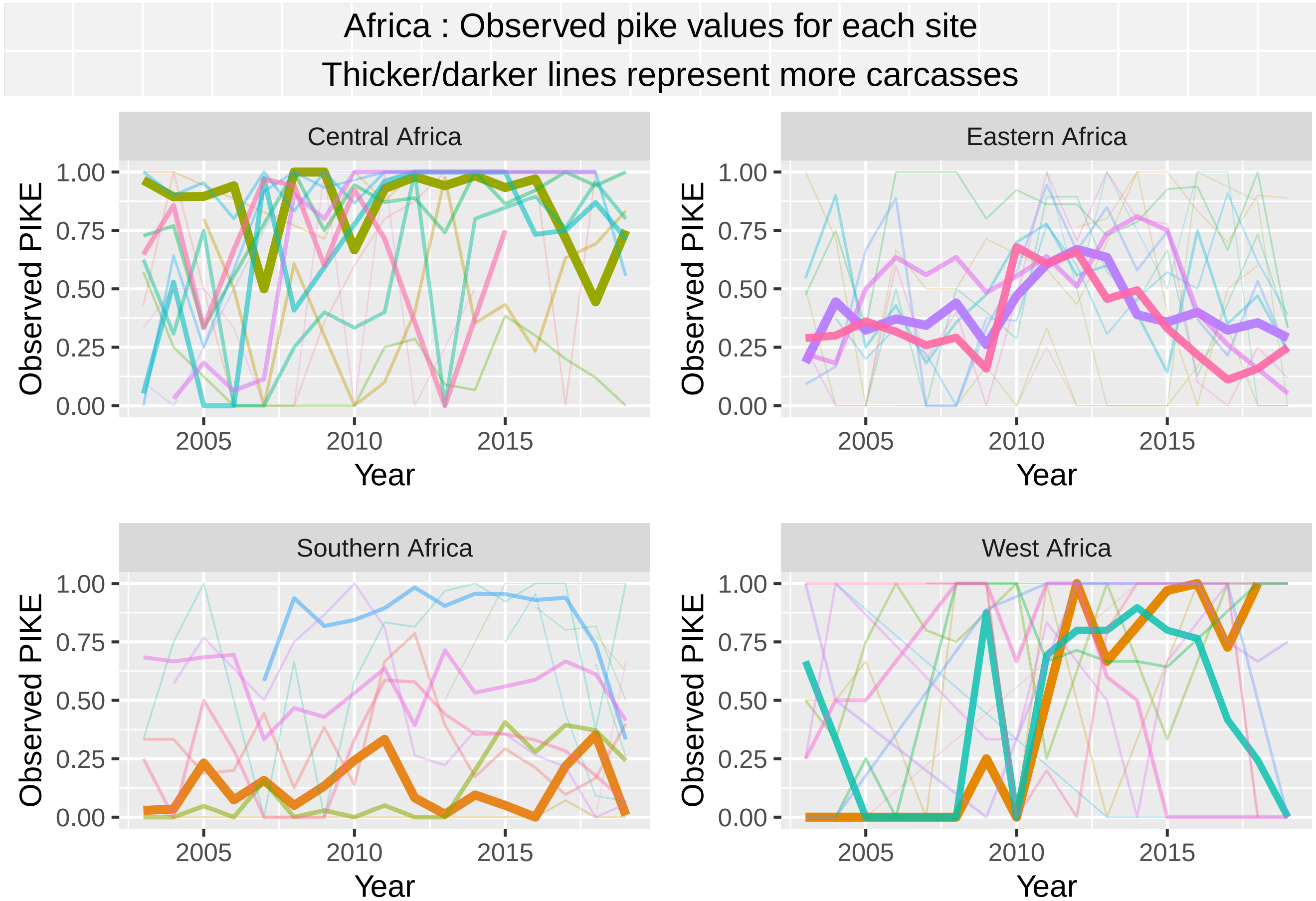
The unusual data point for West Africa where one site reported a large number of carcasses in one year is correct and corresponds to the MIKE site Gourma (GOU) with total number of carcasses equal to 134, of which 130 were poached by armed groups who entered the area.

2.3 Observed *PIKE*

A plot of the observed *PIKE* values from each site-year shows a wide range in the observed *PIKE* values, but many of the observed *PIKE* values close to 0 or 1 occur in sites with only a small number of carcasses examined in a year:

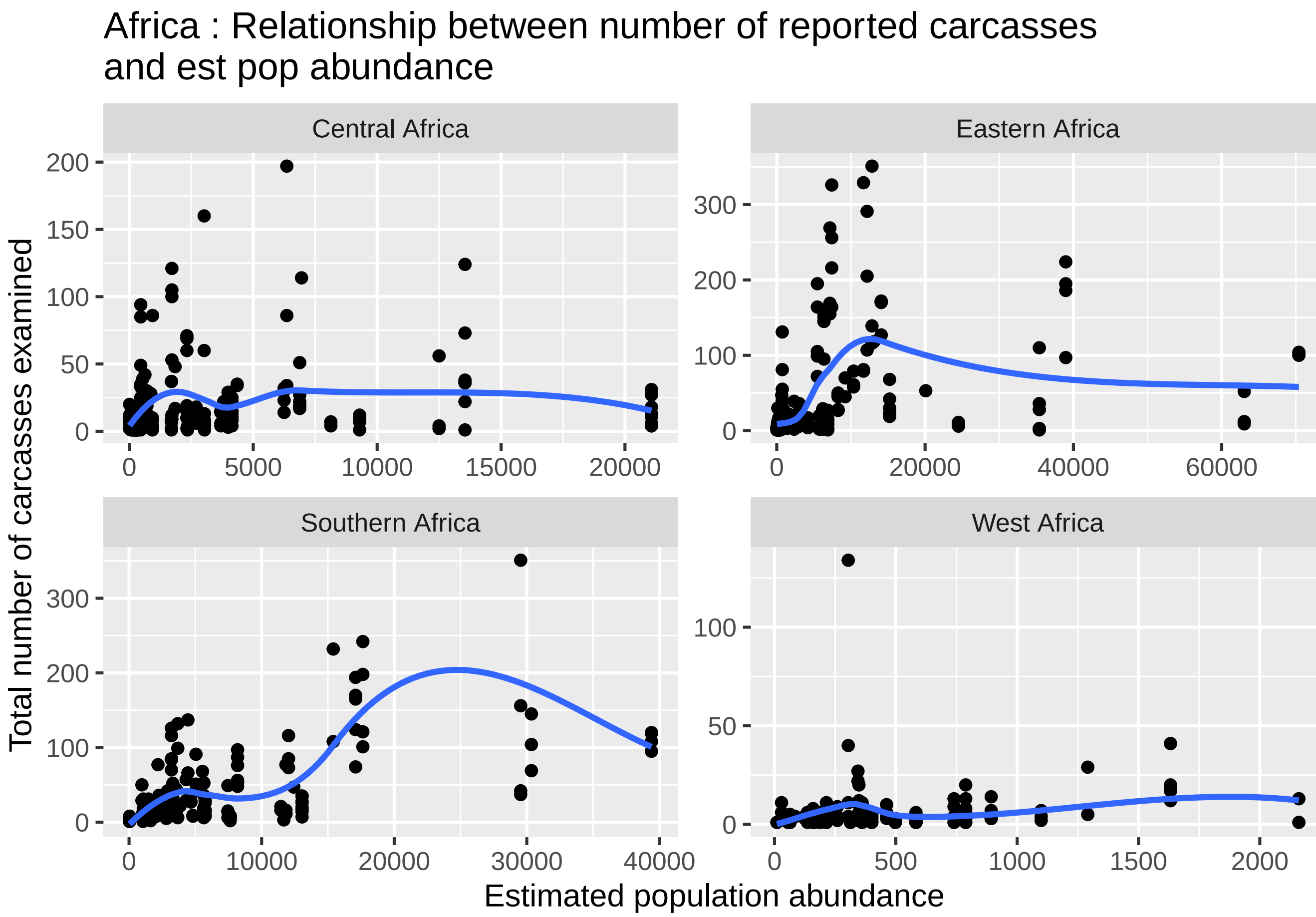


The trend in the observed *PIKE* values for each site is:



Note that with a small number of carcasses reported (e.g. 0 or 1) it is quite common for the reported *PIKE* to be 0 or 1 because either none or all of the carcasses have been illegally killed. Consequently, the trends are difficult to interpret for many sites with only a few carcasses reported.

Patrol effort varies considerably among sites, so the number of examined carcasses only bears a weak relationship to estimated population abundances around the MIKE sites:



Consequently, the total number of carcasses found is unlikely to be a consistent fraction of the population abundance and the number of illegally killed elephant carcasses is not a good index for poaching pressure. Because the *PIKE* index does not depend on effort, it is thought to be a better indicator of poaching pressure than the observed number of carcasses of illegally killed elephants.

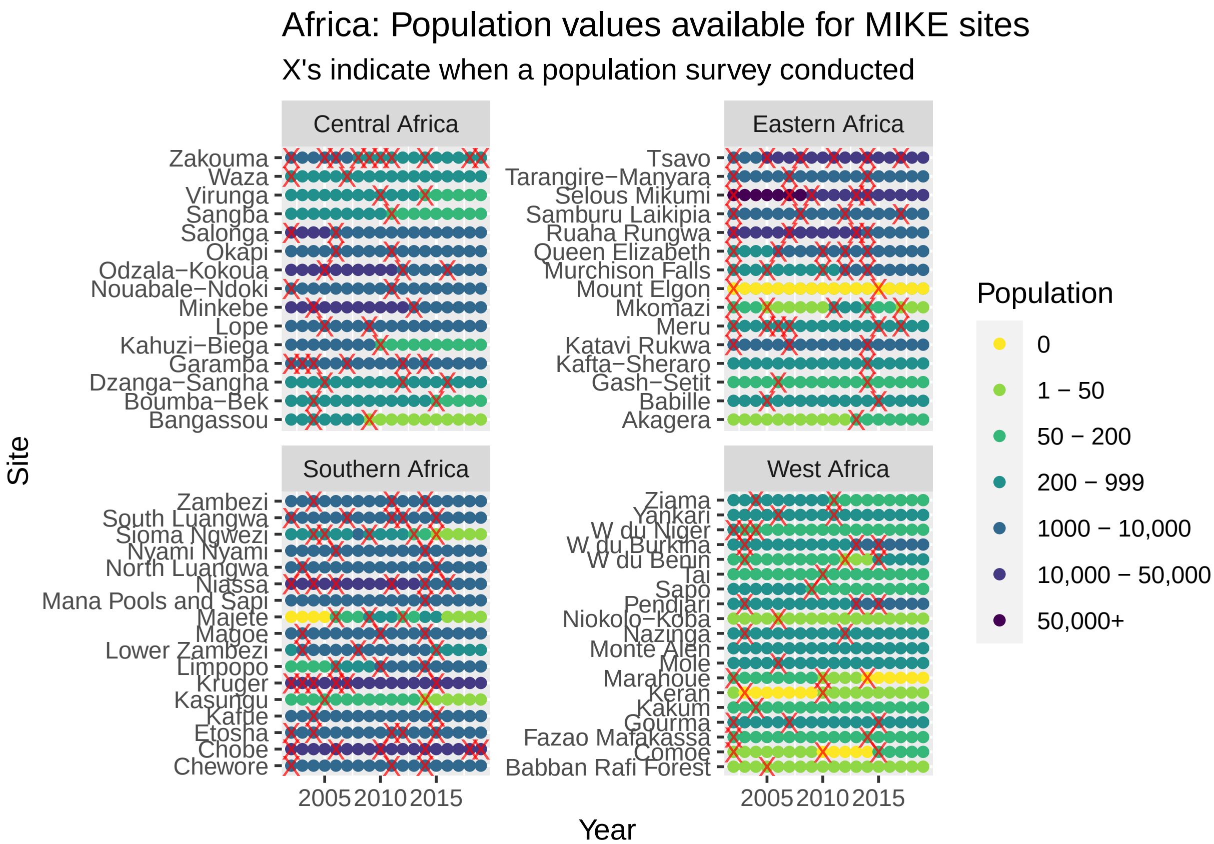
2.4 MIKE sites with population data

Population data is available for 67 sites and has been extracted from Thouless et al (2016). Population surveys are not done each year and population numbers are not updated between population surveys. In these cases, for purposes of illustration, these missing values were imputed using the most recent abundance value. For example, if the estimated population abundances from a survey conducted in 2010 was 500 elephants and from a survey conducted in 2015 was 400 elephants, then the imputed population abundances for 2011, 2012, 2013, and 2014 is also 500 elephants.

For some sites, no population surveys have been done, and population estimates are "best guesses".

These population values are NOT the number of elephants in the MIKE site. Rather, the MIKE sites were chosen to be representative of a broader population and the population values here refer to this broader population. **The use of population values to compute a population-weighted PIKE is currently experimental and being evaluated and the actual implementation may change in the future.**

A plot of when an estimate of the population abundance is available by MIKE site is:



2.5 Final dataset used

The final dataset is an amalgamation of the MIKE population data and observed *PIKE* data. Only those sites for which population data is available and for which at least one carcass has been reported over the period of interest are used.

The following table identifies the MIKE sites not included in the analysis based on data from one of the data sources being absent:

Table 1: Summary of MIKE sites where data not present in all sources

MIKEsiteID	Population values available	PIKE reported	MIKE Center Available
ALE	TRUE	FALSE	TRUE
BBR	TRUE	FALSE	TRUE
COM	TRUE	FALSE	TRUE
EGU	TRUE	FALSE	TRUE
KER	TRUE	FALSE	TRUE
KHB	TRUE	FALSE	TRUE
LGL	FALSE	FALSE	TRUE

The final data set consists of 61 MIKE sites from 2003 to 2019 over the subregions as shown below:

Table 2: Summary of MIKE sites used in analysis

Subregion Name	Number of sites	# Site-Years	Mean # carcasses reported per year	Site IDs
Central Africa	14	188	20.4	BBK, BGS, DZA, GAR, LOP, MKB, NDK, ODZ, OKP, SAL, SGB, VIR, WAZ, ZAK
Eastern Africa	15	210	41.7	AKG, BBL, EGK, GSH, KSH, KTV, MCH, MKZ, MRU, QEZ, RHR, SBR, SEL, TGR, TSV
Southern Africa	17	169	42.8	CHE, CHO, ETO, KFE, KRU, KSG, LPP, LZN, MAG, MAN, MJT, NIA, NLW, NYA, SLW, SMN, ZBZ
West Africa	15	144	6.2	FAZ, GOU, KAK, MAR, MOL, NAZ, NKK, PDJ, SAP, TAI, WBF, WBJ, WNE, YKR, ZIA

3 Generalized Linear Mixed Model (GLMM)

3.1 Model description

The Generalized Linear Mixed Model (GLMM) is a generalization of the simple regression model currently used (the *LSMeans* approach). The GLMM begins by modeling the multiple sources of variation present in the data

- Binomial variation in the number of illegally killed elephants in the number of carcasses detected. For example, the true underlying *PIKE* at a MIKE site may be 30%. However, the rangers will not detect all carcasses at the MIKE site. Suppose that 10 carcasses are detected. Statistical theory says that if the underlying *PIKE* is 30%, then the number of illegally killed elephants is likely to range from 1 to 6 in the sample of 10 selected due to random chance.
- Year-to-year changes in the continental *PIKE* due to factors that change from year-to-year. These are called the **year effects**.
- Some MIKE sites tend to have higher than average *PIKE* year-after-year. These **site-effects** could be due to factors that are consistent for a site across years.
- Within a MIKE site, the *PIKE* will vary from year-to-year within the MIKE site due to factors that vary locally from year to year. These are known as **site-year** effects.

The model will estimate the year, site, and site-year effects, and then use them to impute the *PIKE* in years where a MIKE site has not reported any carcass returns. This imputation is based on the year trends in other sites; the relation of *PIKE* among sites (i.e. some sites tend to always have higher *PIKE* values); and a range in site-year effect based on what is observed in sites with data. This imputation step then gives a *PIKE* value for every site in every year. A range of values is generated for each imputed-value to account for the uncertainty in each of the individual effects.

Once the *PIKE* is found for every site in every year (some site-years are observed; site-years with missing data have *PIKE* imputed), the marginal mean *PIKE* is found in two ways:

- Unweighted marginal mean *PIKE* as the simple average of the *PIKE* in a year over all sites.
- Weighted marginal mean *PIKE* where the weights are the estimated population abundance of elephants represented by the MIKE site in each year.

The population abundances are allowed to vary within a MIKE site over time. Uncertainty in the estimated population abundances is not incorporated into this analysis.

The use of population values to compute a population-weighted PIKE is currently experimental and being evaluated and the actual implementation may change in the future.

3.2 Model fitting

There are many ways to fit a model to data. Simple regression models use least squares where the sum of the squared vertical deviations between the fitted line and the observed data is minimized. The least squares approach is a special case of a more general approach called maximum likelihood. However, maximum likelihood methods often perform poorly (e.g. failure to converge) with models with multiple levels (as in our model) and it is difficult to create custom summaries (such as marginal means) using likelihood methods.

For this reason, we have implemented our model in a Bayesian context using a method called Markov Chain Monte Carlo (MCMC) sampling that deals with many of the technical difficulties found in a likelihood fit without a great deal of effort. Details are available in the technical document.

A Bayesian model combines information from the prior beliefs about the values of certain parameters and the information about these parameters from the data (through a likelihood function). If there is a large amount of data, the information about the parameters would usually overwhelm the information from the prior beliefs, but in cases of sparse data, the prior beliefs may be more important. The end product of a Bayesian analysis is the posterior distribution of belief about a parameter. In the context of this *PIKE* analysis, we have used vague priors with little information about the parameters so the final result is driven almost entirely by the data from the MIKE program.

The concept of posterior belief, while technically challenging to compute, has an intuitive understanding that many people practice. For example, suppose you are waiting to be picked up by a friend. But the friend is late. *A priori* (prior belief), you could place different weights on two hypotheses of why your friend is late - traffic is bad and so your friend is delayed in traffic, or your friend forgot. So if your friend is generally reliable, you may place a higher prior weight on the hypothesis that traffic was bad. While waiting, you overhear from someone else about an accident on the road that has caused a large traffic jam. In light of this new information (data), you update your prior beliefs about the two hypotheses and would now place even more weight on the traffic jam hypothesis than the "forgot" hypothesis. The updated beliefs are your posterior beliefs about the hypotheses which is a combination of prior beliefs in the two hypotheses (prior distribution) and information (data, or likelihood). It would be sensible to make statements such as "I have a 75% belief that my friend is stuck in a traffic jam". Note that your friend either is or is not stuck in traffic – there is no probability associated with the actual state. As more and more data are received about the size of the traffic jam, your posterior beliefs about the two hypotheses will shift.

Similarly, you may have a prior belief about a persons age who you have never met. In this case, the prior belief is not two discrete hypotheses, but a continuum, i.e. you can picture a "normal"-like distribution of prior belief with a peak, say at age 50, but ranging from 40 to 60. As you get more information (data) such as the fact that person witnessed the fall of the Berlin Wall as a young adult, you would update your prior belief by shifting the peak of the prior age distribution and/or shrinking the range of possible ages. This is now your posterior belief summarized by a posterior distribution. It would be sensible to make statements such as "I have a 40% posterior belief that this person is more than 55 years of age". Note that the actual age of the person does NOT have a distribution – the distribution refers to your knowledge of the age based on a combination of prior information and actual data. If you find the year of birth of the person, the posterior distribution becomes extremely concentrated because you now know the age of the person to within a year.

In an analogous fashion, a Bayesian analysis summarizes the posterior belief about a parameter using the posterior distribution. This distribution could be discrete (e.g. two competing hypotheses) or could be continuous (e.g. a range of values summarized by a distribution with a peak and spread.)

We summarize the posterior distribution in this report in several ways:

- Mean of the posterior distribution the (mathematical) expectation of the belief distribution representing our point estimate about the parameter value.
- Standard deviation of the posterior distribution representing a measure of uncertainty about our belief in the parameter value.
- A credible interval (typically a 95% credible interval) representing uncertainty in the belief. A 95% credible interval would contain 95% of our belief.
- Total belief that the parameter is above/below a certain value found as the portion of the posterior distribution that is above/below the specified value. As seen below, we are interested in our posterior belief that the slope in marginal *PIKE* was less than zero in the last five years representing our posterior belief that the marginal *PIKE* has declined in the last five years.

A rough correspondence exists between the results of a least-squares or likelihood model and a Bayesian model. The maximum likelihood estimate roughly corresponds to the mean of the posterior distribution; the standard error for a parameter from a likelihood fit roughly corresponds to the standard deviation of the posterior distribution; a confidence interval from a likelihood fit roughly corresponds to a credible interval from a Bayesian model. However, the correspondence is not "exact" and there are fundamental differences in the technical definitions of each measure that make them not commensurable. In particular, the total belief that a parameter is above/below a specific value has no correspondence in the likelihood context.

3.3 Choice between the weighted and unweighted marginal mean *PIKE*

The simple combined *PIKE* (i.e. total illegally killed elephants / total observed carcasses) may be a suitable estimator for the marginal mean *PIKE* if carcasses reported were proportional to the population abundances (if you are interested in a weighted *PIKE*) or equal across sites (if you are interested in an unweighted *PIKE*) and if all sites reported in all years. The number of carcasses reported is not equal across sites, bears only a weak relationship to the population abundances, and there is much missing data so the simple combined *PIKE* is difficult to interpret and seldom of interest.

The choice between the weighted and unweighted marginal mean *PIKE* is more complex.

The unweighted marginal mean *PIKE* should be thought of as an index of poaching pressure that gives each site the same weight regardless of the underlying population abundance. We hope that trends in the index are broadly informative of changes in poaching pressure at the continental scale.

The weighted marginal mean *PIKE* may reflect actual trends in poaching at the population level under several assumptions (e.g. that the observed *PIKE* in a MIKE site is reflective of the poaching in the larger population; that the population abundance estimate for a MIKE site is reflective of the actual population abundance; that all populations are monitored (i.e. every population has a MIKE site). Consequently, the weighted marginal mean *PIKE* looks appealing but also must make many assumptions before reflecting the poaching pressure on a population level.

The use of population values to compute a population-weighted PIKE is currently experimental and being evaluated and the actual implementation may change in the future.

Given the very weak relationship between the number of carcasses and the population abundances, the simple combined *PIKE* will tend to track the weighted marginal mean *PIKE* closer than the unweighted marginal mean *PIKE* values.

4 Continental level trends in *PIKE*

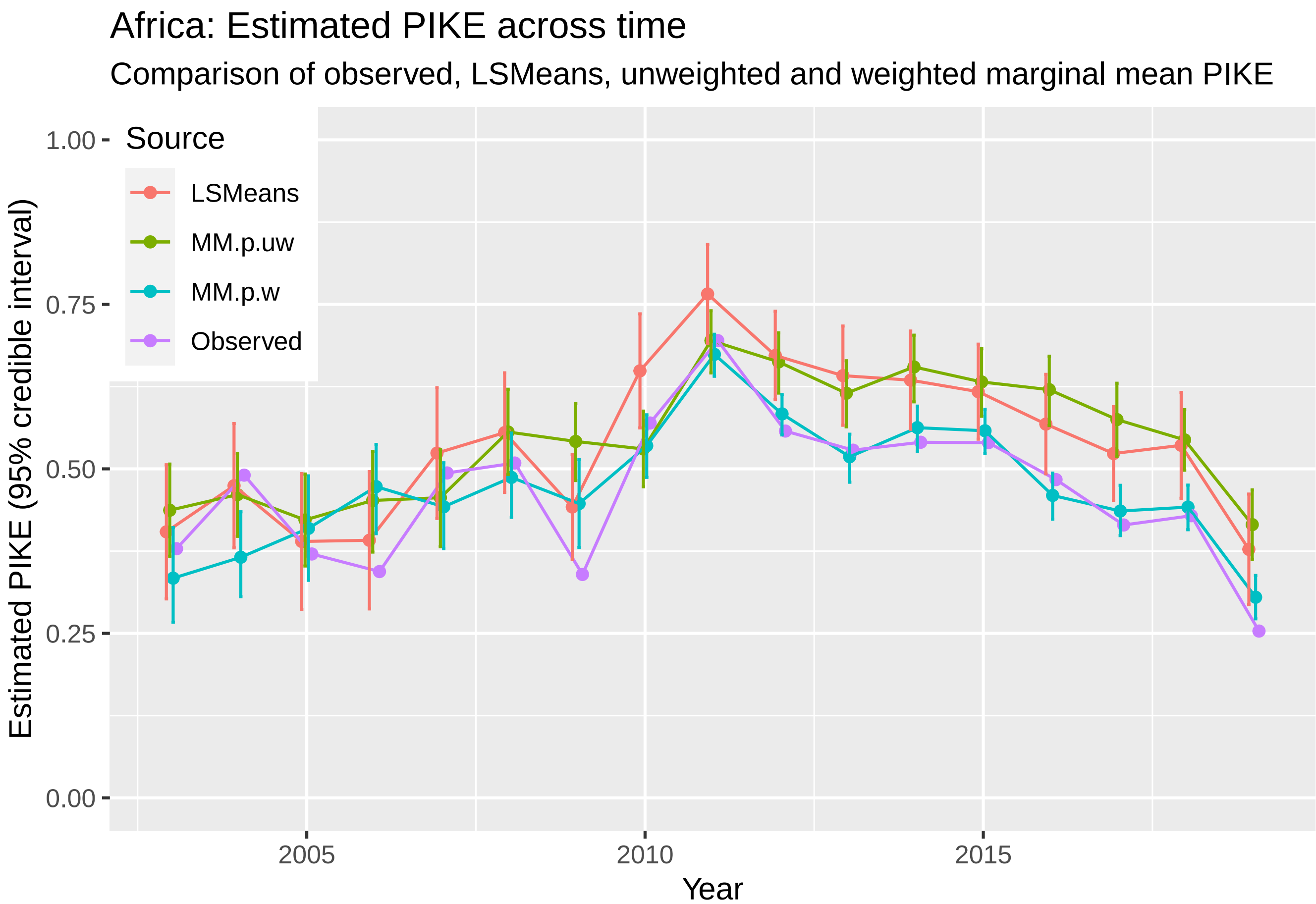
The GLMM model was coded using *BUGS* (Lunn et al, 2012), a common way to specify Bayesian models and run using *JAGS* (Plummer, 2003) within *R* (R Core Team, 2020).

The estimated *LSMeans*, unweighted and weighted marginal mean *PIKE* are:

Table 3: Estimated marginal mean PIKE from the GLMM

	Unweighted		Weighted	
Year	Mean	SD	Mean	SD
2003	0.44	0.04	0.33	0.04
2004	0.46	0.03	0.37	0.03
2005	0.42	0.04	0.41	0.04
2006	0.45	0.04	0.47	0.03
2007	0.46	0.04	0.44	0.03
2008	0.56	0.03	0.49	0.03
2009	0.54	0.03	0.45	0.03
2010	0.53	0.03	0.54	0.02
2011	0.69	0.02	0.67	0.02
2012	0.66	0.02	0.58	0.02
2013	0.61	0.03	0.52	0.02
2014	0.65	0.03	0.56	0.02
2015	0.63	0.03	0.56	0.02
2016	0.62	0.03	0.46	0.02
2017	0.57	0.03	0.44	0.02
2018	0.54	0.02	0.44	0.02
2019	0.42	0.03	0.30	0.02

The following plot compares the $PIKE$ computed using the $LSMeans$ approach (the current approach), the unweighted ($MM.p.uw$) and weighted marginal mean ($MM.p.w$) computed using the Bayesian model



The new proposed unweighted marginal mean ($MM.p.uw$) tracks the previously computed $LSMeans$ fairly well except for 2009-2011. The fitted trend lines are consistently higher than the observed *PIKE* values because the latter gives more weight to sites with more carcasses whereas the $MM.p.uw$ give equal weight to each site regardless of number of carcasses (or underlying population abundance). The weighted marginal mean ($MM.p.w$) tracks the observed *PIKE* closely after 2010 because sites with large number of carcasses (and larger elephant populations) will dominate both in a similar way.

Once the sample from the posterior is available, it is relatively easy to estimate the posterior belief that the trend is negative in the last five years. This is done by estimating the slope in the last five years for each sample from the posterior and then the posterior belief that the trend is negative is the proportion of fitted slopes that are less than zero. The posterior belief that the slope in PIKE is negative in the last five years is 1.00, i.e. we have a high posterior belief that the trend in PIKE in the last five years is negative (i.e., is declining).

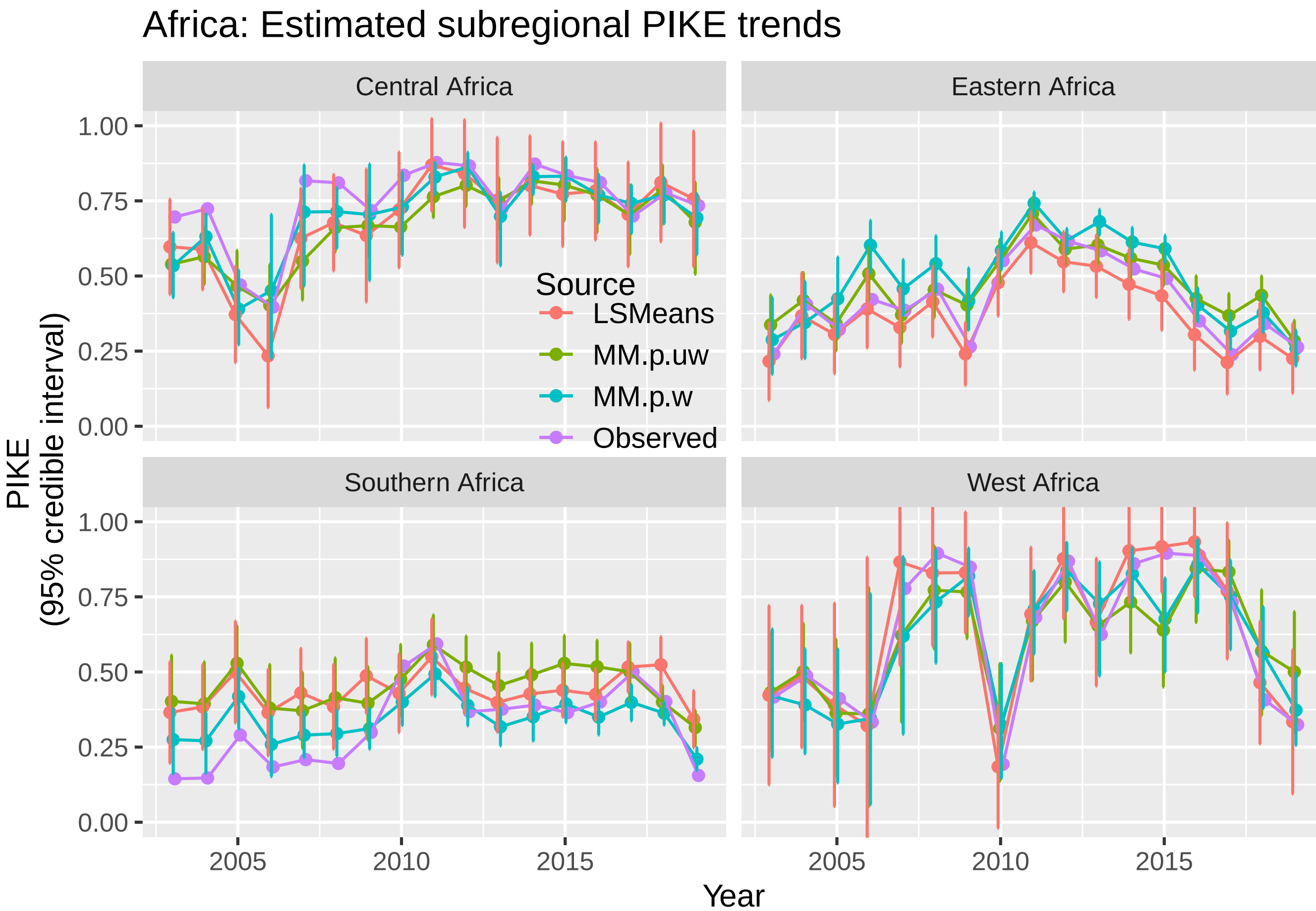
5 Subregional trends in *PIKE*

The above analyses were repeated at the sub-regional level. Only the data from each sub-region was used in each analysis, i.e., completely separate analyses were performed for each sub-region.

A more complex model where the four sub-regions were modelled as part of a larger model could be attempted. This larger model would be useful in cases where a sub-region has sparse data but has a consistent relationship in *PIKE* trends with other sub-regions. This is similar to the case where sparse data for a particular MIKE site is improved if there is a consistent relationship between the *PIKE* in several MIKE sites. In these cases, the consistency in the relationship allows information from other sub-regions to improve the estimates of trend in the sub-region with sparse data.

This more complex model is examined in the technical document, but because there does not appear to be a great deal of similarity in trends among sub-regions, we do not expect this more complex model to provide much of an improvement over individual models for each sub-region.

The following plots compare the mean *PIKE* computed using the *LSMeans* method, the unweighted marginal mean *PIKE* (*MM.p.uw*) and the weighted marginal mean *PIKE* (*MM.p.w*).



The estimates at each subregion have wider confidence intervals because the amount of data is smaller in each subregion compared to the continental results.

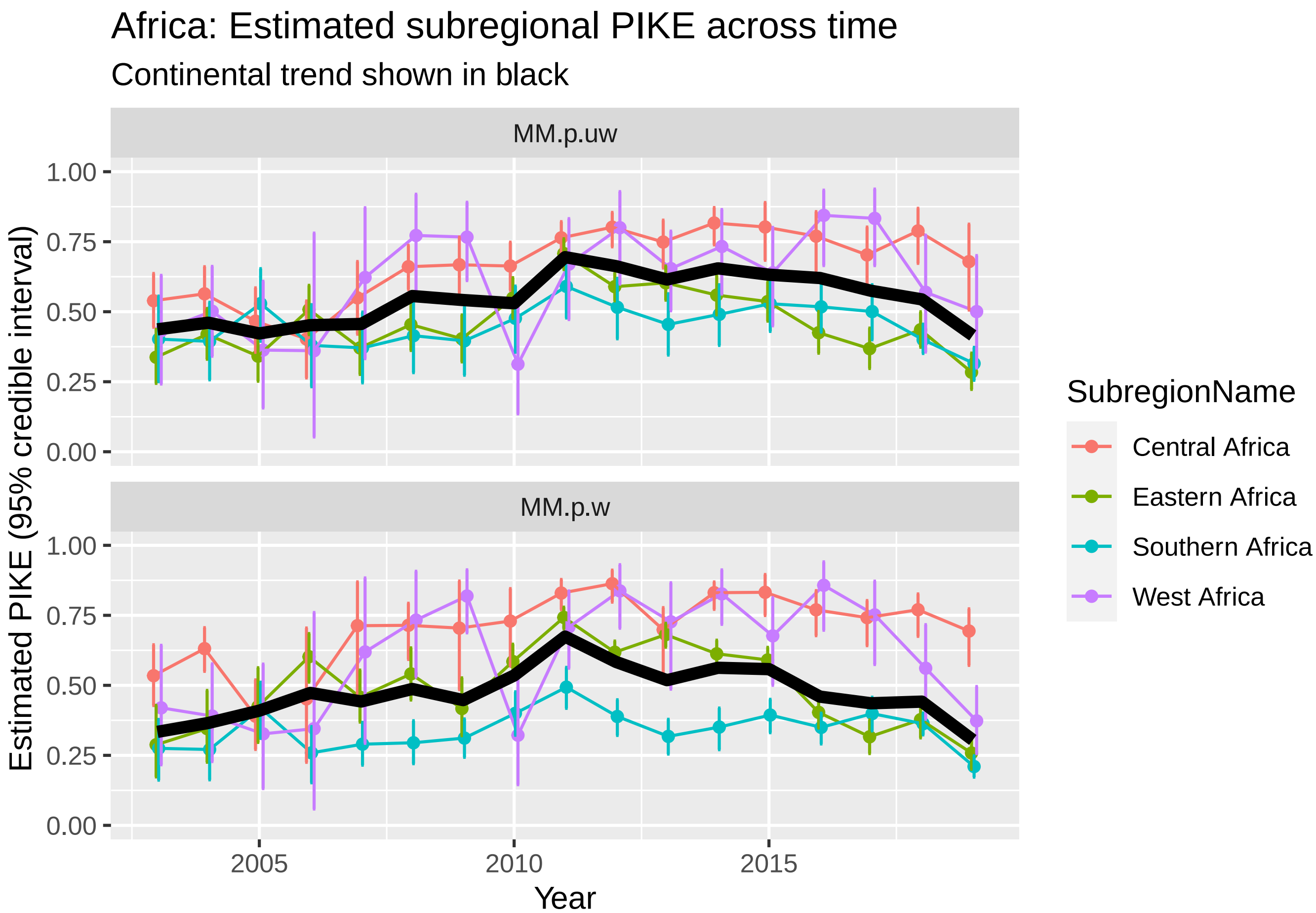
It is hard to make a general statement comparing the estimates computed using *LSMeans* and the unweighted marginal means (*MM.p.uw*) because of the wide confidence intervals, but the respective trend lines mostly track each other. The apparent consistent difference seen for Eastern Africa and Southern Africa are artefacts of how data are aggregated to the sub-regional level in the *LSMeans* and Bayesian GLMM models (see Appendix 3 of the technical document for details).

In Central Africa, the weighted marginal mean *PIKE* is consistently higher than the unweighted marginal mean *PIKE* but the credible intervals overlap); in Eastern and West Africa, they track each other closely; while in Southern Africa, the unweighted marginal mean *PIKE* is consistently higher than the weighted mean *PIKE*. This "inconsistency" is a function of the relationship between the number of carcasses examined and the population abundance examined earlier in this document.

But regardless if the weighted or unweighted marginal mean *PIKE* are examined, trends are similar.

The use of population values to compute a population-weighted PIKE is currently experimental and being evaluated and the actual implementation may change in the future.

It is interesting to compare the regional trends with the continental trends



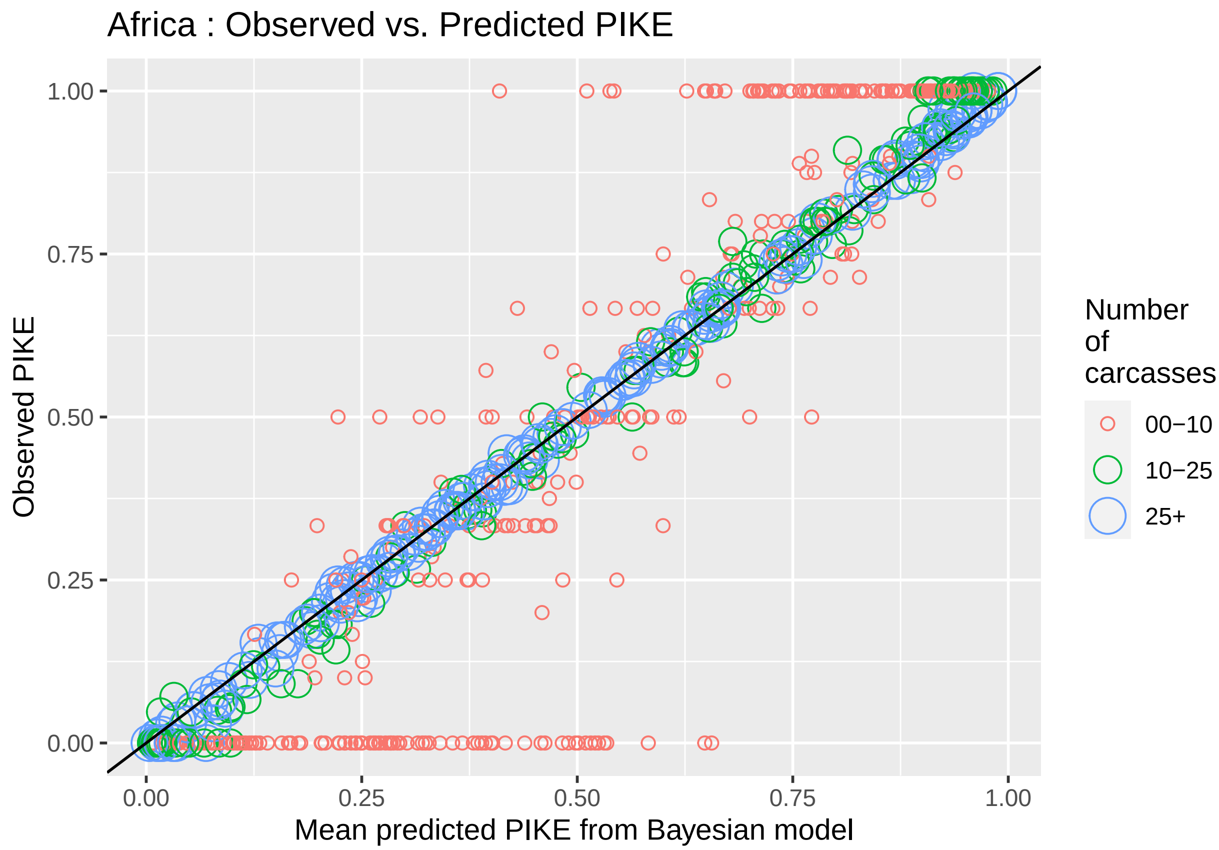
We see that *PIKE* in Southern Africa is consistently lower than the continental *PIKE*, while *PIKE* in Central Africa is consistently higher.

6 Model assessment

Whenever a new method is used, it is important to determine if the model fits the data well. We performed detailed model assessments (see the technical document) and found no evidence of problems except for a residual spatial autocorrelation in the estimated site effects.

6.1 Observed vs. predicted *PIKE*

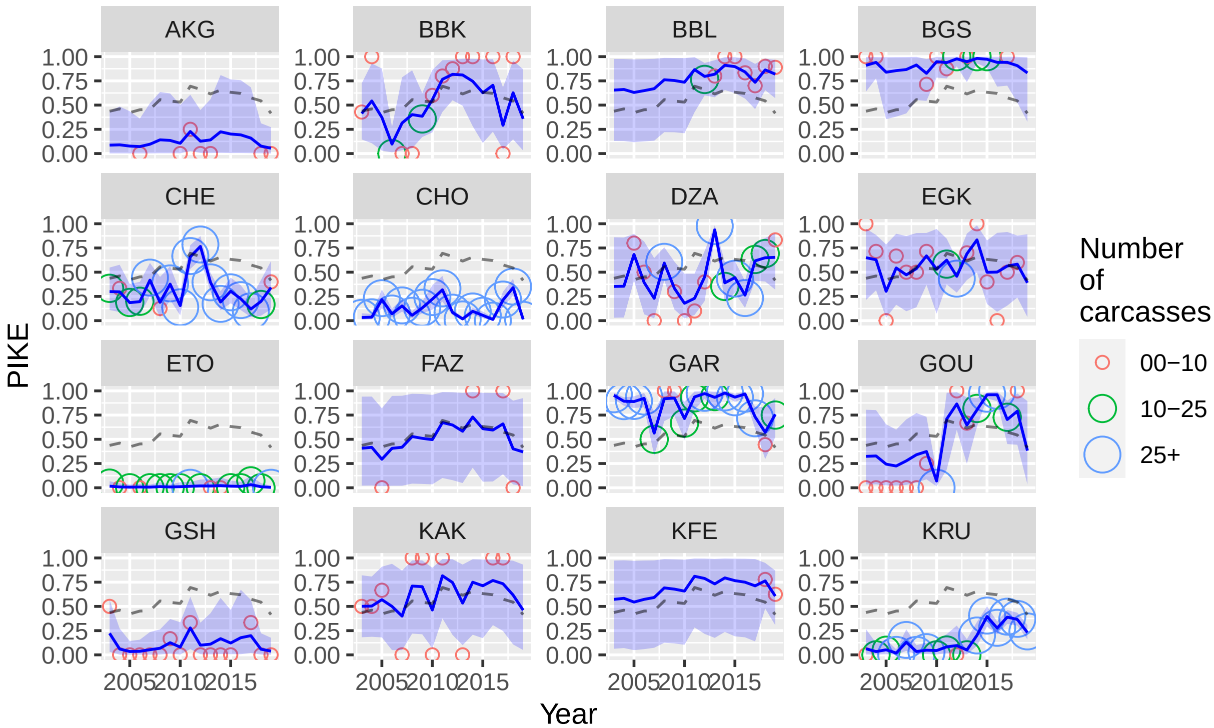
A plot of observed *PIKE* in each year.site vs. the predicted *PIKE* is:



The fit is generally very good. For site-years where the number of carcasses was very small (< 10) and the observed *PIKE* was 0 or 1, the estimated *PIKE* is pulled towards the yearly average for that year. For site-years with large number of carcasses (> 25) the estimated *PIKE* matches closely with the observed *PIKE*. For site-years with intermediate number of carcasses, the estimates are shrunk slightly towards the mean for that year.

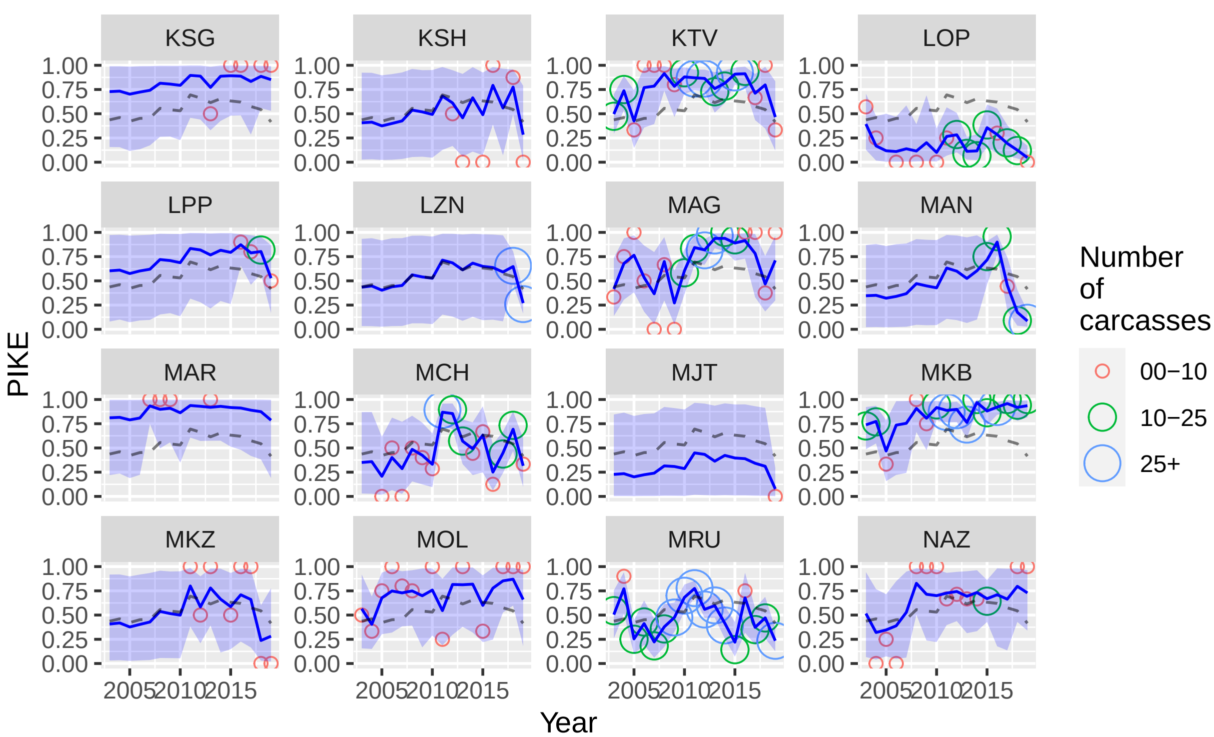
This can also be seen in the plots of observed and fitted *PIKE* for the individual MIKE sites:

Africa : Observed and predicted PIKE for individual MIKE sites

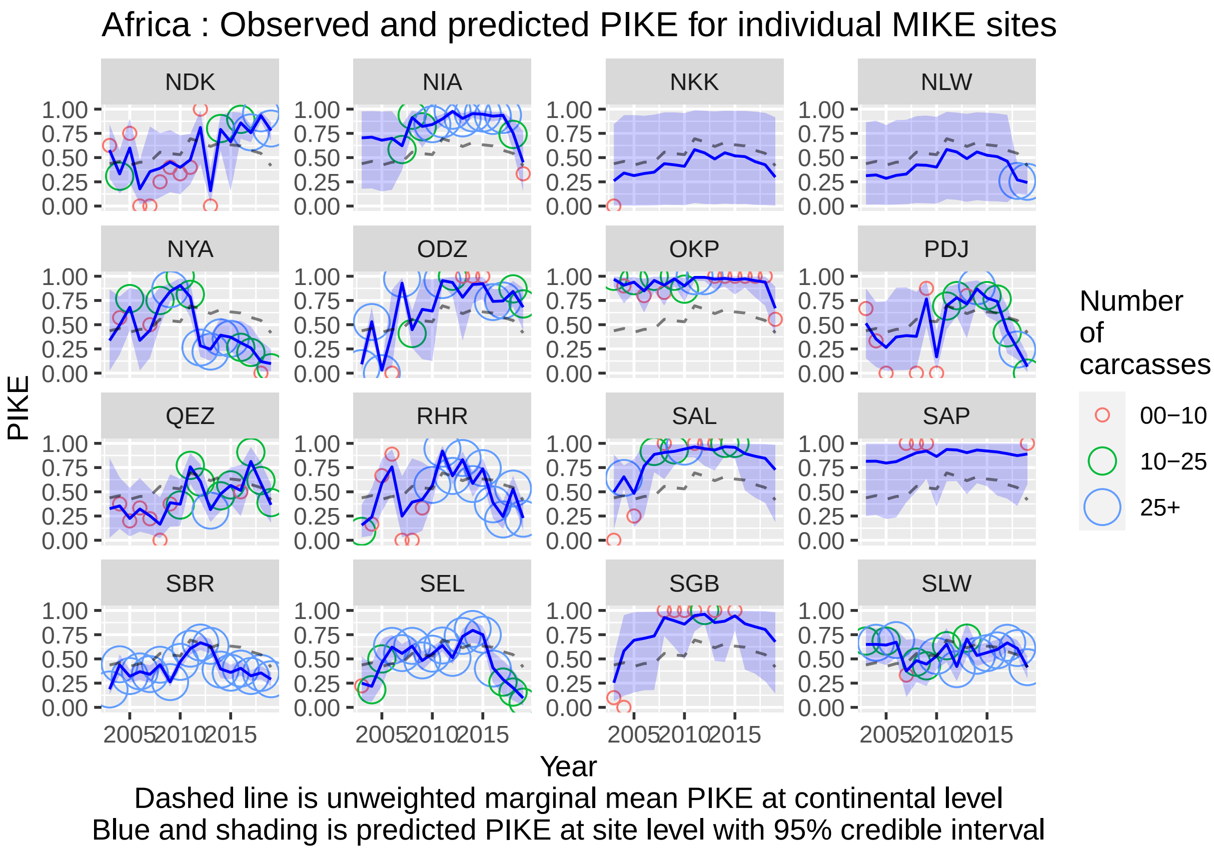


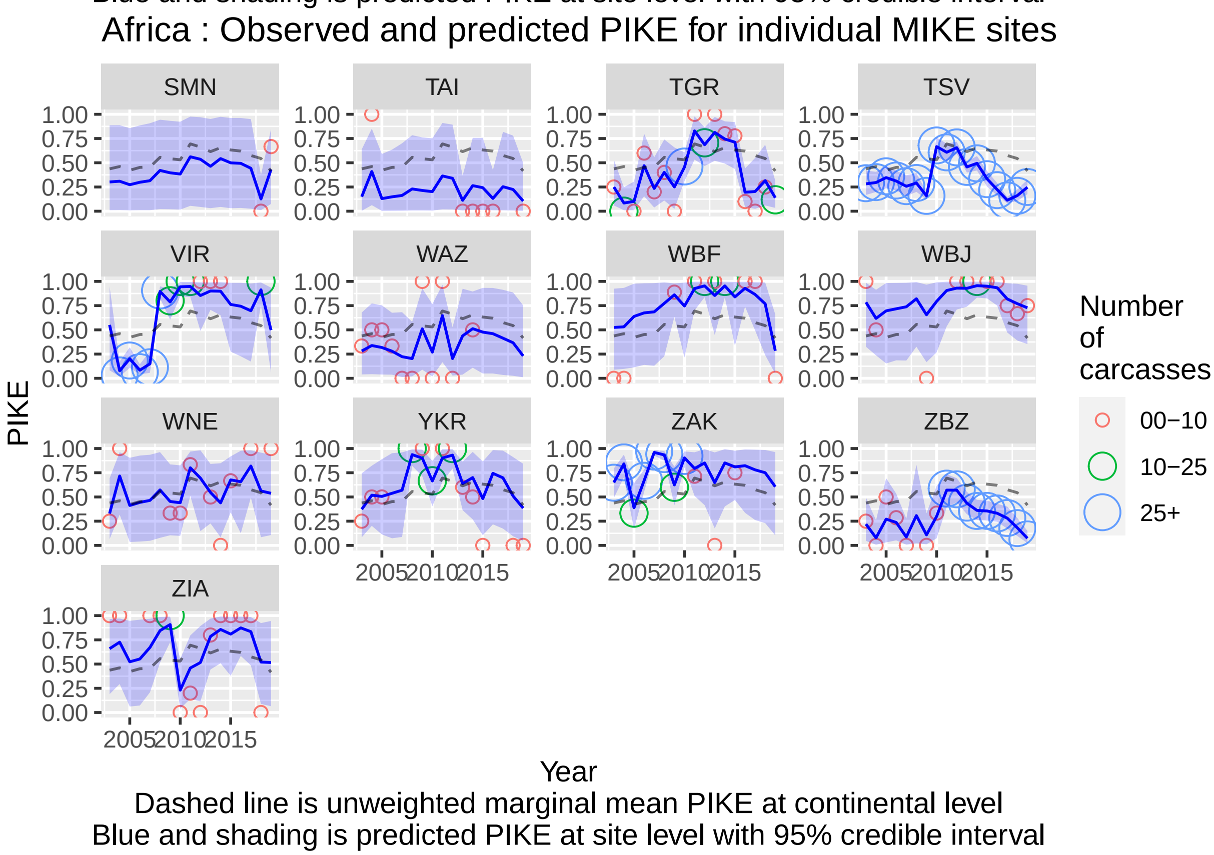
Dashed line is unweighted marginal mean PIKE at continental level
Blue and shading is predicted PIKE at site level with 95% credible interval

Africa : Observed and predicted PIKE for individual MIKE sites



Dashed line is unweighted marginal mean PIKE at continental level
Blue and shading is predicted PIKE at site level with 95% credible interval


Africa : Observed and predicted PIKE for individual MIKE sites
NDK
NIA
NKK
NLW
NYA
ODZ
OKP
PDJ
QEZ
RHR
SAL
SAP
SBR
SEL
SGB
SLW
PIKE
Year
Number of carcasses
00−10
10−25
25+
Dashed line is unweighted marginal mean PIKE at continental level
Blue and shading is predicted PIKE at site level with 95% credible interval


Africa : Observed and predicted PIKE for individual MIKE sites
SMN
TAI
TGR
TSV
VIR
WAZ
WBF
WBJ
WNE
YKR
ZAK
ZBZ
ZIA
PIKE
Year
Number of carcasses
00−10
10−25
25+
Dashed line is unweighted marginal mean PIKE at continental level
Blue and shading is predicted PIKE at site level with 95% credible interval

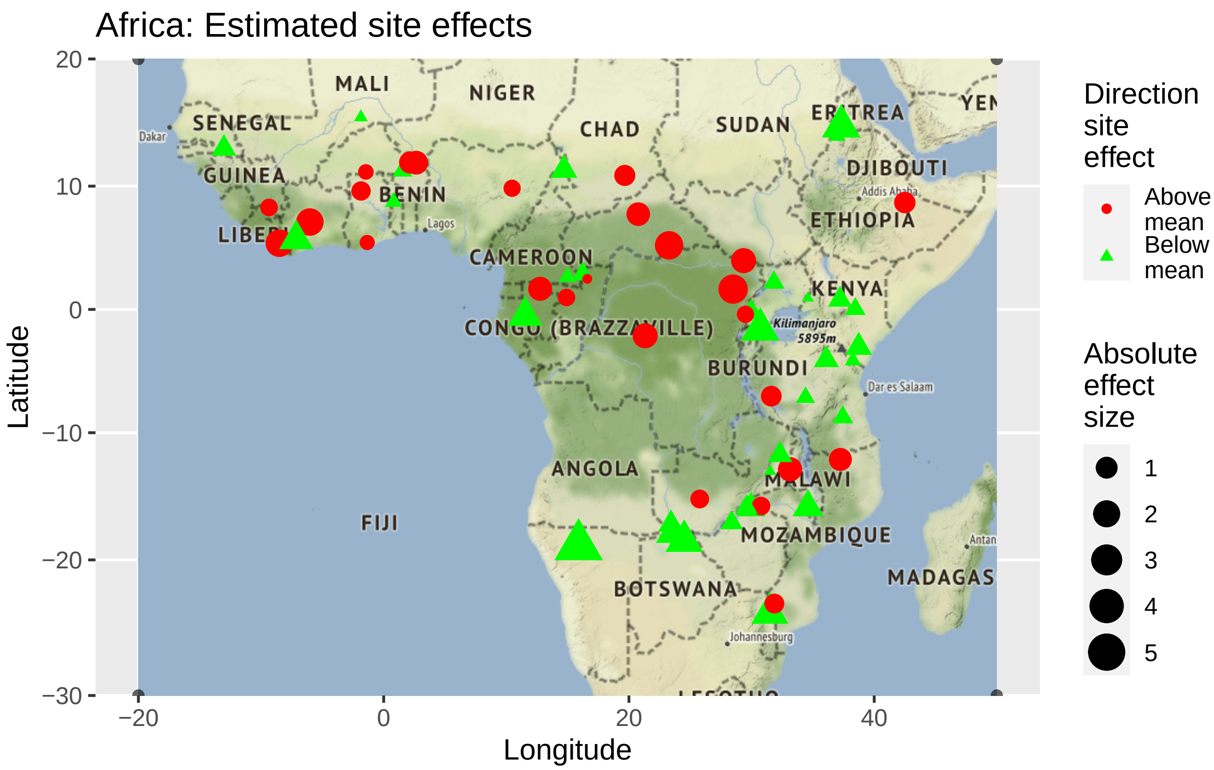
There are several interesting patterns that illustrate the features of the model.

- Site *CHO*. This site reports nearly every year with a large number of carcasses (large blue circles). The estimated yearly site-level *PIKE* closely follows the observed data (as expected).
- Site *BBK*. In some years, this site reports a large number of carcasses and the estimated yearly site-level *PIKE* for these years matches the observed *PIKE*. In some years, the observed *PIKE* based on large number of carcasses is above the continental marginal mean (e.g. 2003) and in some some years it is below the continental marginal mean (e.g. 2006). On average, this site tracks the continental trend. So in years where this site only reports on a few carcasses (small red circles) and the observed *PIKE* is mostly 0 or 1, the estimated *PIKE* is close to the continental trend. For example, with a small number of carcasses examined, a value of 2 illegally killed elephants from 2 carcasses examined (observed *PIKE* of 1) is consistent with an estimated *PIKE* closer to the continental marginal *PIKE*. Notice that in years with a small number of carcasses reported, the credible interval for the estimated site-level *PIKE* is very wide.
- Site *AKG*. This site mostly has smallish sample sizes, but the observed *PIKE* is consistently close to 0. The estimated site-level *PIKE* is then also close to 0 in years with no reports, but notice the wide credible intervals.
- Site *KFE*. This is a new MIKE site with data only starting in 2018. In this year, the observed *PIKE* was slightly higher than the continental marginal mean *PIKE*, and so the estimated site-level *PIKE* for earlier years is slightly above the continental trend, but notice the wide credible intervals.

In summary, in years with many carcasses reported, the estimated site-year *PIKE* will closely match the observed site-year *PIKE*. In years with few carcasses reported, the estimated site-year *PIKE* will be pulled towards the continental trend after accounting for the observed relationship between this sites *PIKE* and the continental trend.

6.2 Spatial correlation in site effects

The (random) site effects have been modelled as independent random effects without explicitly accounting for the spatial structure of the data. However, we find that sites that are close geographically have similar site effects.



Sites that have *PIKE* consistently above the continental average are labelled as *Above the mean*; sites that have *PIKE* consistently below the continental average are labelled as *Below the mean*.

We notice that sites that are close geographically tend to have similar site effects (size of dot) and in the same direction (above or below the mean, color of dots). This implies there is a spatial correlation among the site effects that has not been directly accounted for in the analysis.

The current analysis is still valid, but inefficient because it has not used the spatial correlation to improve inference. If spatial autocorrelation is explicitly modelled, then information is shared among sites that are geographically close, i.e., if the *PIKE* increases in one site, then spatial autocorrelation would imply that it would tend to also increase in a nearby site. Of course, if the sites are in different countries with different levels of enforcement or other covariates that impact *PIKE*, an explicit spatial autocorrelation could introduce a spurious relationship between the *PIKE* in the two sites unless these other factors (law enforcement etc.) are also modelled. The explicit spatial autocorrelation models rapidly become more complex to account for these features.

Because the current analysis treats all sites as independent (rather than spatially correlated), the uncertainty in the overall yearly *PIKE* is slightly smaller than from a model with explicity spatial autocorrelation because the effective number of sites used in computing the overall yearly *PIKE* is smaller when autocorrelation is explicitly modelled. This in turn, implies that the uncertainty of a trend (e.g. trend in the last five years) in the currently model may be slightly understated as well and the posterior belief in a trend will be higher in the current model compared to the model with an explicity spatial autocorrelation. We believe such effects are minor given the spase data at many sites, the large amount of missing site.years and the potential breaking of spatial autocorrelation across country borders.

A potential improvement to the current analysis may be to add another level of random effects (country effects) so that points from the same country that have related site effects then experience a common country effect. This model is currently under investigation.

6.3 Sensitivity analysis

We also examined in the technical document, the sensitivity of the new model to adding new MIKE sites, to dropping MIKE sites (perhaps due to local extirpation), increasing monitoring effort at MIKE sites, and the impact of the prior beliefs as part of the Bayesian model.

Adding a new MIKE site may have impacts on previous estimated *PIKE* values. Presumably the MIKE site had elephants and potential illegal killings prior to being added to the programme. The current model will impute the *PIKE* for this site in the years before the site was added based on the yearly trend in *PIKE* in other sites and the relationship between the new site's *PIKE* and other sites' *PIKE*. For example, if the new MIKE site had a *PIKE* value that was 10 percentage points above the mean *PIKE*, then a *PIKE* will be imputed for this site for all past years that is also 10 percentage points above the yearly mean.

The imputation will be very coarse in the first few years after a MIKE site is added until enough yearly *PIKE* data has been collected to estimate reasonably well the relationship between this site's *PIKE* and the overall mean (i.e., several years will be needed to estimate the (random) *site* effect well).

Fortunately, the (unweighted) marginal mean *PIKE* is currently computed using over 50 MIKE sites and so adding a new MIKE site is expected to only have minimal impact on the yearly mean *PIKE* values unless the new site has an extreme *PIKE*. The weighted marginal mean *PIKE* could be influenced more if the new MIKE site represented a large elephant population that was not previously monitored.

In most cases, the effect of dropping individual sites is slight except when *PIKE* is computed by weighting by population abundance. In these cases, the yearly mean *PIKE* values can change substantially when a site with a large underlying population is dropped. This influence could push the year mean *PIKE* up or down depending if the particular site has a site-specific *PIKE* that was larger or smaller than the overall yearly mean *PIKE*.

Increasing monitoring effort at MIKE sites will shrink the width of credible intervals, but missing site-years is the limiting factor in determining the uncertainty of the marginal mean *PIKE* because imputation must be performed for sites with missing values and increasing the sample size at other MIKE sites has little impact on the uncertainty in the imputation.

The data is rich enough (more than 20 years with more than 50 sites) so there prior belief distributions have minimal impact on the results.

We therefore conclude that the proposed GLMM model is robust.

7 Conclusions

The GLMM method is an improvement over the current method (the *LSMeans* method) by accounting for the binomial structure of the data, the interrelationship between *PIKE* among sites, and for local variations in *PIKE* at the site level over time. The results from the proposed model are similar to the current results and conclusions about trends in *PIKE* are the same under both models. The GLMM model can also be easily extended to finding weighted marginal mean *PIKE* (e.g. weighted by population abundances) and posterior beliefs about features of the data (e.g. trend in the last five years) that are impractical to obtain from the *LSMeans* approach. The GLMM model is also robust to changes in the data (i.e. not sensitive to data changes).

8 References

Burn, R.W., Underwood, F.M., Blanc J. (2011). Global Trends and Factors Associated with the Illegal Killing of Elephants: A Hierarchical Bayesian Analysis of Carcass Encounter Data. PLoS ONE 6(9): e24165.

Lunn, D., Jackson, C., Best, N., Thomas, A. and Spiegelhalter, D. (2012). The BUGS Book – A practical introduction to Bayesian Analysis. Chapman and Hall/CRC Press.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), March 20–22, Vienna, Austria. ISSN 1609-395X.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.