SADC REGIONAL PROGRAMME FOR RHINO CONSERVATION

A Review of the RMG Black Rhino Carrying Capacity Model Version 1

Kevin M. Dunham and Raoul F. du Toit

Improving and standardising methods of black rhino carrying capacity assessment II: review

Semester 7 Task 4.2-2.2

COOPERAZIONE
ITALIANA
ITALIAN COOPERATION
AID 5064

CESVI
cooperaziene e sviluppe

SADC

IUCN
The World Conservation Union

WWF

SPECIES SURVIVAL COMMISSION
AFRICAN RHINO SPECIALIST GROUP

PUBLICATION CREDITS:

Title: SADC RMG Black Rhino Carrying Capacity Model Version 1.0: A Review

Author: Kevin Dunham & Raoul du Toit

Date: April 2003

Special acknowledgements:

This report is an output from a task of the **SADC Regional Programme for Rhino Conservation**

ABOUT the SADC Regional Programme for Rhino Conservation:

The Programme is funded by the Italian Ministry of Foreign Affairs, Directorate General for Development Cooperation (Project AID 5064).

The Programme is contracted to CESVI and implemented through a regional consortium which comprises:

- The Secretariat of the Southern Africa Development Community (SADC)
- IUCN-ROSA (The World Conservation Union - Regional Office for Southern Africa)
- The IUCN African Rhino Specialist Group
- WWF-SARPO - (World Wide Fund for Nature - Southern Africa Regional Programme Office)
- CESVI (Cooperazione e Sviluppo)

The ***Programme goal*** is to contribute to maintain viable and well distributed metapopulations of Southern African rhino taxa as flagship species for biodiversity conservation within the SADC region.

The ***Programme objective*** is to implement a pragmatic regional rhino strategy within the SADC region following the acquisition of sound information on, firstly, the constraints and opportunities for rhino conservation within each range state and secondly, the constraints and opportunities for rhino metapopulation management at the regional level.

DISCLAIMER

CONTACT DETAILS FOR THE PROGRAMME:

Dr. Rob Brett
SADC Regional Rhino Programme Coordinator
IUCN-ROSA
6 Lanark Road Belgravia PO Box 745 Harare, Zimbabwe
Tel: 263-4-728266 Fax: 263-4-720738
Email: robb@iucnrosa.org.zw

SADC Secretariat
Directorate of Food, Agriculture and Natural Resources
P O Box 0095 Gaborone Botswana

CESVI Zimbabwe Office
9 Northwood Rise, Mt Pleasant, Harare, Zimbabwe
TEL/ FAX: 263-4-882243, 884492, 850898
EMAIL: cesvi@africaonline.co.zw

Ministero degli Affari Esteri
Direzione Generale per la Cooperazione allo Sviluppo,
Unità Tecnica Centrale
Viale Contarini (angolo Viale Farnesina)
00194 ROMA ITALY

PROGRAMME WEB SITE: We invite you to vist our web site : www.rhino-sadc.org

SUMMARY

The RMG black rhino carrying capacity model (version 1) was examined and recommendations are made to improve the presentation and user-friendliness of its user guide and spreadsheets, and the statistical development of a revised model.

The model represents a bold, almost courageous, plunge into a very complicated set of issues. Considerable work has gone into preparation of the user guide and model, but the former still needs to be ‘polished’, (to achieve a clearer presentation of theory, and to sort out editing glitches) with a typical user (not rhino experts) in mind. The number of spreadsheets (currently three) should be reduced by combining worksheets into one spreadsheet. Worksheets that are ‘under development’, those for data that are not used in the present model, and those that are simply lists should be removed from a revised spreadsheet. The remaining worksheets should be revised to simplify their use, with particular attention paid to differentiating between cells used for data entry, automatic calculations, and outputs.

The model uses four linear regression equations (which are similar in terms of their independent variables) to predict carrying capacity and then uses the mean of the four predicted values as the model’s predicted carrying capacity. This approach raises question marks: why use four similar equations? Why not three, or five? Is this not something of a salvo of potentially inaccurate shots fired through a set of similar muskets to try to bracket a target, rather than a single shot using a more precise rifle? Such an approach lacks statistical elegance.

The regression equations predict that (all other things being equal) carrying capacity will increase as rainfall concentration increases (i.e. as the duration of the rainy season declines and the duration of the dry season increases). The biological basis for such a relationship is unclear.

The ‘prior’ expert estimates of carrying capacity in the baseline areas are not the same in the user guide’s diagrams and in the data set used to develop the model. The obvious question: ‘prior’ to what?

The accuracy of the carrying capacity predictions made by the model was tested by jack knifing (leave-one-out). During this process, 15 revised models were prepared, each time dropping a different one of the 15 baseline areas from the data set and then preparing new models that were otherwise identical to the RMG model. For each of the 15 versions of the model, the values of the independent variables for the site that was left out of the data set were entered into the revised model, and the predicted carrying capacity for that area compared with the expert estimate. This process produced, for each of 15 areas, five carrying capacity predictions, one from each of four regression equations and the mean of these four.

None of the four regression equations was significantly better or worse than the others at predicting carrying capacity accurately.

For eight of the 15 baseline areas, the expert estimate of carrying capacity was outside the range of the four predicted values.

In absolute terms (rhinos per square kilometre) the error (i.e. the difference between the model prediction and the expert estimate) was small (< 0.1 rhinos km^{-2}) in 13 out of 15 cases. In percentage terms, however, the error was often not small (< 25 % in just six cases and > 50 % in five cases).

For one baseline area, the prediction error was large (predicted carrying capacity = 0.54 rhinos km^{-2}; expert estimate = 0.2 rhinos km^{-2}). This area was the only one with high mean annual rainfall and infertile soils, and linear regression equations did not capture the approximately polynomial relationship between rhino density and rainfall on infertile soils (East 1984).

Independent rhino experts should define the degree of accuracy that they expect from a prediction model.

The model must not only be reasonably accurate, but it must also be robust to inter-observer differences in assessing qualitative variables such as soil fertility and browse suitability.

Although these are initially assessed on a 9-point scale, they are entered into the equations as continuous variables with a precision one or two orders of magnitude greater than this. Using categorical variables – which more accurately reflect the state of knowledge – might address this problem.

General additive models and general linear models are probably more appropriate (than multiple linear regressions) for modelling complex multivariate relationships. An alternative approach would be to develop an expert system that uses both qualitative data and simple quantitative relationships.

The development of a reliable model will require more baseline areas and, if the model is to be useful within the SADC region, it must be applicable to sites outside South Africa and Namibia, and in particular must be relevant to lowveld mopane areas and habitats with miombo elements (which are said to be outside the model's predictive range at present). On the other hand, before a great deal of effort is made to collect model-building data for other areas, it may be desirable to refine and test the model (and especially to prove its robustness when applied independently by various users) within the existing range of baseline sites.

Finally, it must be emphasised that this is no more than a preliminary look at the model and its theoretical justification; it is not a complete review of all aspects.

TABLE OF CONTENTS

INTRODUCTION

Adcock (Undated) has produced a model that predicts the ecological carrying capacity of black rhinos for an area. The model is a set of four multiple regression equations. Each equation predicts a carrying capacity and then the model uses the mean of these four values as the predicted ecological carrying capacity for the area.

The regression equations were developed from data for 15 baseline areas (13 in South Africa, 2 in Namibia; see Table 2) for which prior estimates, produced by experts, of black rhino carrying capacity were available. In each regression equation, the dependent variable is the expert estimate of carrying capacity and there are six independent variables representing:

1. mean annual rainfall;
2. rainfall concentration (monthly spread);
3. likelihood of frost;
4. soil fertility;
5. browse biomass; and
6. browse suitability, or browse quality.

In addition, the likely maximum number of adult males that any given area can tolerate socially is predicted from the size of the area, and from a regression between the mean range size of adult males and carrying capacity.

Our review of the model is by no means comprehensive, although both of us considerably exceeded our time allocations for this exercise, and the completion of this review was therefore delayed because the exercise could not be undertaken within the time frame that was originally scheduled. The issues are complex, and their presentation is sometimes confusing, hence it would take considerable additional reviewing effort to get to grips with all aspects of the background theory, statistical methods and presentation of the carrying capacity model. Thus this review constitutes little more than first impressions of the approach, from two ecologists with relevant experience, but with insufficient time to digest the full range of issues that arise. We have sometimes had to give examples of problem areas, rather than identifying each and every instance where that type of problem arises.

It is reasonable to assume that a rhino expert is unlikely to use the model, because he or she already has – in their mind – their own expert system. Therefore, to be valuable, the model must be useable by people who are not rhino experts (although the user guide notes that any user must have, at least, an "appropriate ecological” background). During this review, the easy-of-use of the model was considered from the viewpoint of an appropriately experienced ecologist or park manager (but not a rhino expert) who wants to reintroduce black rhinos to some area and who, for financial, logistical and other reasons, wants to estimate the size of a sanctuary large enough to accommodate a founder population of 20 rhinos and their offspring.

In this review, recommendations for improvements to the presentation and design of the model and its user guide are presented in italics.

FIRST IMPRESSIONS – THE USER GUIDE

The user guide was received as two rich-text-format files (the first containing the guide text and the second containing the references cited in this) and 14 gif (picture format) files of the diagrams referred to in the guide. When the rich-text-format files were opened in MS WORD, they were in ‘Technical’ font (e.g. SADC Regional programme for Rhino Conservation). It is not clear if this font was chosen by the author, or whether WORD lost something during file conversion, but the result was a font that was difficult to read. Combining all 16 files into a single, easy-to-use file was a slow, time-wasting procedure, and showed that the "product" that had been achieved during the previous SADC task had not been satisfactorily "packaged". The consolidated document comprised 59 pages, with 35 pages of text – including 3½ pages of acknowledgements - plus appendices and references.

The user guide is comprehensive in discussing the factors that do or might influence the carrying capacity of black rhinos, and the construction of the model. It is clear that considerable work went into preparing the user guide, but this effort is undermined by some editing problems, for example this quote from the second paragraph of section 1: "…15 "baseline" black rhino areas (12 in South Africa, 2 in Namibia)*…"*. Elsewhere (e.g. "…Dougall and Glover ###…") the text gives the impression that the user guide is a document still in preparation, rather than one that is complete and ready for distribution.

On Figure 2.3.2 of the user guide (for example, but the same applies to other Figures), the prior estimate of carrying capacity for Umfolozi Park is plotted as 0.34 rhinos km^{-2}. But in Table A3.2 of the user guide, the prior estimate of carrying capacity for Umfolozi Park is given as 0.43 rhinos km^{-2}. And this is not an isolated occurrence. There are obvious discrepancies between the prior estimates of carrying capacity as plotted on Figure 2.3.2 and as given in Table A3.2 for Addo Elephant NP, Andries-Vosloo, Eastern Shores (Tewati), Hluhluwe, uMkhuze and Vaalbos NP, and suspected, but small, discrepancies for several other areas. At best, these discrepancies suggest editing mistakes (but see below).

Apart from simple editing problems, there are some more fundamental points of confusion that arise in the presentation of ecological theory and which tend to derail the understanding that the reader of the user guide is supposed to be gaining from it. For instance, the author states:

> "Productivity curves for large mammals are skewed towards carrying capacity, and are not near ½ of carrying capacity as with smaller mammals. This is due to their typically high adult survivorship, long gestation periods and relatively old ages at first calving, which limit the range over which life-history parameters can change in response to changing density or food supply."

This kind of statement has to be re-read several times, and even then the logic remains unclear. When the author writes about 'productivity curves', presumably she is referring to the parabolic relationship between sustained yield and density for a population growing logistically (Caughley 1977).

The presentation of theory on carrying capacity (which is indeed a tricky concept) becomes even more confusing with a later statement:

> "In areas with a high degree of annual variation in browse production (e.g., high variation in annual rainfall), rhino densities may never attain the "average" carrying capacity density, but can reach densities which are high enough to inflict long-term damage on vital browse resources during times of low browse supply (e.g. drought)".

What is the "average" carrying capacity here? How can rhino densities reach figures that are "high enough to inflict long-term damage", yet still remain below "average" carrying capacity? Average carrying capacity is surely exactly that, as demonstrated by long-term population performance in the face of whatever ecological feedback mechanisms act on rhino density in a given area, with the population density over time going as much above this line as it goes below the line.

Some recommendations:

- *Although considerable effort has been made to pull together relevant theory on herbivore feeding ecology, equivalent effort needs to be made to present the key concepts in a way that builds understanding, to the extent required by a person who is not a rhino expert, but who has some ecological background and seeks to understand the principles upon which the model is based, before blindly plugging in data.*

- *All text, references and diagrams comprising the user guide should be combined in a single file. Readability of the user guide should be improved by using a conventional font instead of one that mimics handwriting.*
- *The manual needs to be compartmentalised more cleanly into a section that serves as a step-by-step guide to the use of the model and a section on background theory; at present there is quite of lot of theoretical comment embedded in Section 3, which is supposed to be the step-by-step data collection guide.*
- *The text needs to be finished to a standard similar to that which the editor of a scientific journal would expect of a paper submitted for publication.*
- *Sections that are 'under development' should be removed from the guide until such time as their development is complete.*

THE PRIOR ESTIMATES OF CARRYING CAPACITY

The word prior, as in 'prior estimates' of carrying capacity, is not unambiguously defined in the user guide, but the implication of section 2.4 of the user guide ("The model should predict as closely as possible the prior, "expert" estimates of ecological carrying capacities of the 15 baseline areas.") is that the carrying capacities were determined by experts prior to (i.e. before) development of the model.

There are obvious discrepancies between the prior estimates of carrying capacity as plotted on Figure 2.3.2 of the user guide and as given in Table A3.2 for Addo Elephant NP, Andries-Vosloo, Eastern Shores (Tewati), Hluhluwe, uMkhuze, Umfolozi Park and Vaalbos NP and possibly other areas. Are these discrepancies are the result of simple transcription errors? Or have they arisen because the 'prior' estimates were changed at some point during or after model construction? There may have been a good reason for changing the 'prior estimates': for example, the experts may have refined their estimates completely independently of the process of model development. However, by apparently changing the 'prior' estimates after the start of model development and failing to explain why, the author is open to accusations that the data used to develop the model have been massaged.

Box 1

Exactly how was the prior estimate of carrying capacity in Umfolozi (0.43 rhinos km^{-2}) derived from these data? (From Appendix 3 of the user guide)

In Umfolozi, black rhino were rare in the late '50's. Numbers built up during the 80's and early '90's up to *c*. March '94 where they peaked at *c*. 321. Removals averaged 12.4 rhino per year, or 4.1% of the population from '89 to '98 (but with more removals in later years). From end '94 to the end of '98, numbers have varied from *c*. 319 to *c*. 283, and annual increment accounting for removals has averaged around 3.5%.

	Density range from April '94 to Dec. '98
Masinda	0.473 to 0.586 rhino/km^2
Mbhuzane	0.346 to 0.442 rhino/km^2
Makhamisa	0.52 to 0.613 rhino/km^2

The average densities in Umfolozi as a whole have been around 0.458 rhino/km^2 since March '94.

Documented Umfolozi carrying capacity estimates for '90's conditions have been:

Masinda rhino/km^2	Mbhuzane rhino/km^2	Makhamisa rhino/km^2	Reference
0.5	0.5	0.5	(Emslie & Adcock 1993)
0.46-0.5	0.45	0.44	(Goodman *et al*. 1996 NPB meeting)
0.4-0.55	0.4-0.55	0.37-0.5	(Brooks & Adcock 1997)

Although Appendix 3 of the user guide summaries what is known about black rhino densities in the 15 baseline areas, there is often no explanation of how each expert estimate was derived from these data (if in fact it was) (see Box 1 for an example).

THE MODEL

Model Design

The model consists of three EXCEL spreadsheets (BrCCSupport, BrCCDataInput and BrCCModel).

Spreadsheet BrCCSupport contains five worksheets:

1. Available Cover Scale;
2. Geology Info;
3. Soil Info;
4. Diet Info; and
5. Suitability Profile.

The worksheet Available Plant Cover was obviously easier to prepare as a spreadsheet, but it must be in paper form to be useful to the model user. The next three worksheets detail information about the geology, soils and rhino diet in some of the baseline areas used to develop the model, but the advantage to the model user of having this information in spreadsheet format, rather than in table form within the user guide, is not obvious. The worksheet Suitability Profile is an example of a table that uses data on the rhino diet and species composition of the various vegetation types within a study area to facilitate the determination of browse suitability ratings. However, the instructions at the top of this worksheet are inadequate: for example, there is no explanation of the numbers in the light green and dark green areas of the worksheet (i.e. within the columns labelled with vegetation type code numbers). Plant densities perhaps? Also the model user has to guess the meanings of the abbreviations in the preference rating column.

Spreadsheet BrCCDataInput contains six worksheets:

1. Final Variable Set;
2. Vegetation;
3. Soil Geology;
4. Rainfall;
5. Min July Temp; and
6. Auxilliary Browsers.

The worksheet Auxilliary Browsers is not used to predict rhino carrying capacity during the current version of the model, but the user is expected to enter data about vegetation, soil or geology, rainfall and temperature into the relevant worksheets. The Final Variable Set worksheet automatically provides the values of the independent variables for use in the model's regression equations.

The multicoloured worksheets are not easy to use immediately. Sometimes the sections are in reverse numerical order (e.g. in the Rainfall worksheet, section 1 is at the bottom, section 4 is at the top – see Fig. 1a). It is seldom immediately obvious which cells the user must enter data into and which conduct calculations automatically – it often necessary to examine the contents of a cell to discover its purpose. Many parts within these worksheets are cell-protected, to prevent the user inadvertently altering formulae.

Spreadsheet BrCCModel contains four worksheets:

1. Predict ECC of an Area;
2. Predict Avg Male Range;
3. Original variables; and

4. Model and Variables.

The user has to transfer the values of the variables from the Final Variable Set worksheet of the BrCCDataInput spreadsheet to the Predict ECC of an Area worksheet of the BrCCModel spreadsheet. Then, the predicted carrying capacity is automatically calculated. The user must then enter this number – and the size of the user's study area – into the Predict Avg Male Range worksheet, which then automatically calculates the likely number (to one decimal place!) of dominant males that can be supported by the study area.

There is considerable scope for improving the design and layout of the worksheets to make them easier to use; for removing from spreadsheets those worksheets that do not require relevant (to the current model) data entry; and for placing the remaining worksheets into a single spreadsheet.

- *The worksheets should be revised to simplify their use. Fig. 1 shows the current Rainfall worksheet and a suggested improvement.*
- *Cell colourings, and font colour and size should be used consistently to indicate different cell types, e.g.:*
  - *cells that contain instructions;*
  - *cells into which the user needs to enter data;*
  - *cells that calculate automatically (i.e. the internal workings);*
  - *cells that contain the final parameters; and*
  - *cells that contain error or warning messages.*
- *Section 1 should be at the top of a worksheet and other sections labelled numerically working down the worksheet.*
- *Cells checks (e.g. that Sum = 1 in cell L12 of the Soil Geology worksheet) should be conducted automatically and an error message displayed if necessary (e.g. if Sum < or >1).*
- *Numbers shown as percentages should be described as such and not described (incorrectly) as proportions, e.g. the worksheet Original Variables.*
- *The worksheets should be arranged in the order in which they should be completed, e.g. the Final Variable Set should be after the worksheets for vegetation, soil/geology, rainfall and temperature.*
- *Worksheets (e.g. Auxilliary Browsers) that require data that are not (yet) used by the model should be removed during the current version of the model. Alternatively, if these worksheets are retained, the user guide should contain explicit instructions on how the user should utilise this information.*
- *The spreadsheets BrCCDataInput and BrCCModel should be combined into one spreadsheet. The worksheet Final Variable Set is then unnecessary. Other worksheets, such as Original variables and Model and Variables, are not needed by the model user and can be removed.*
- *The model could be reduced to two spreadsheets. The main one should contain only those worksheets that are currently needed by the model user. The second one should contain the other worksheets that are required for an 'evaluation' version of the model. Support tables and diagrams should be in the user guide.*

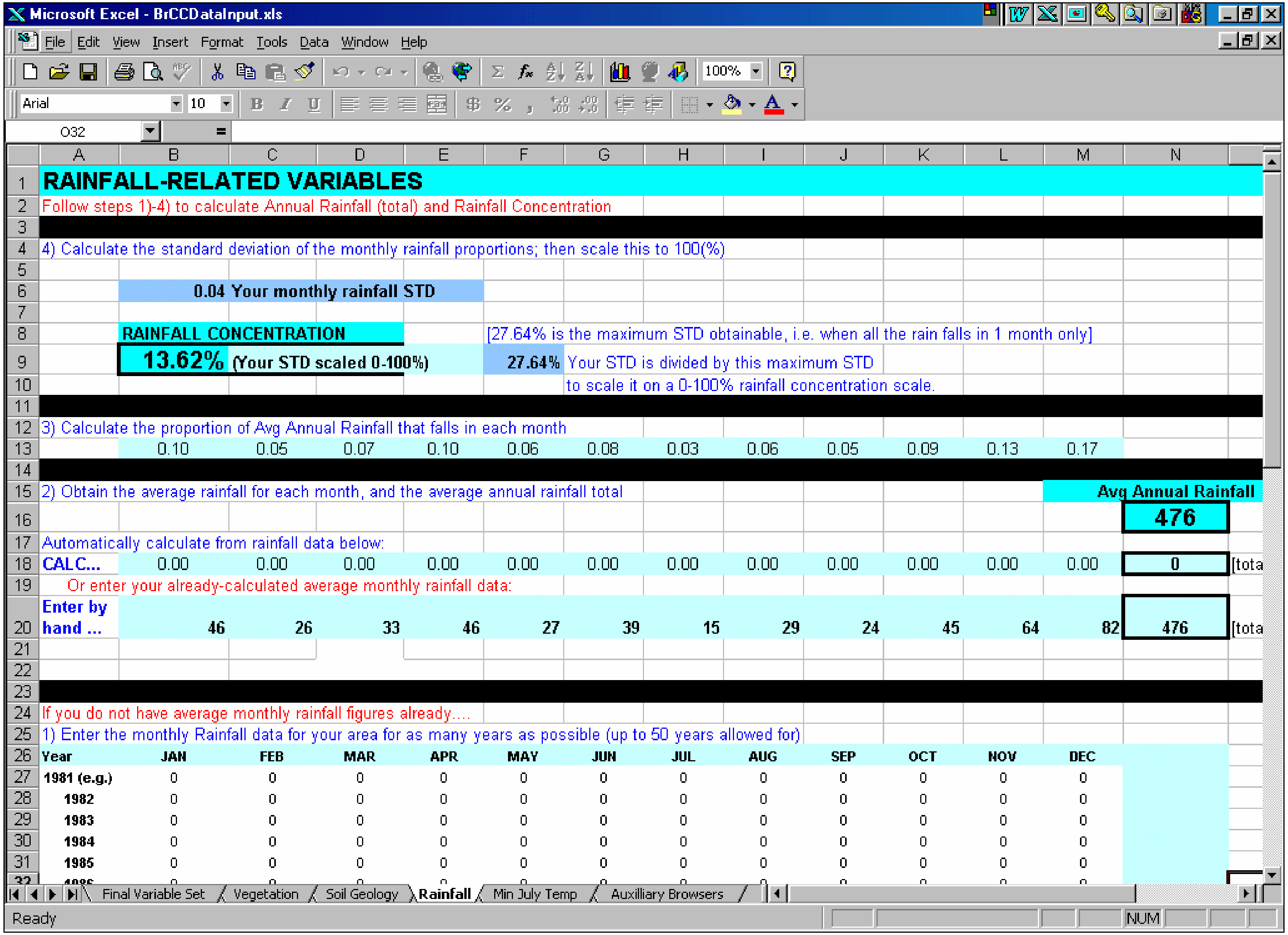

Figure 1 (a). The original worksheet Rainfall in the spreadsheet BrCCDataInput

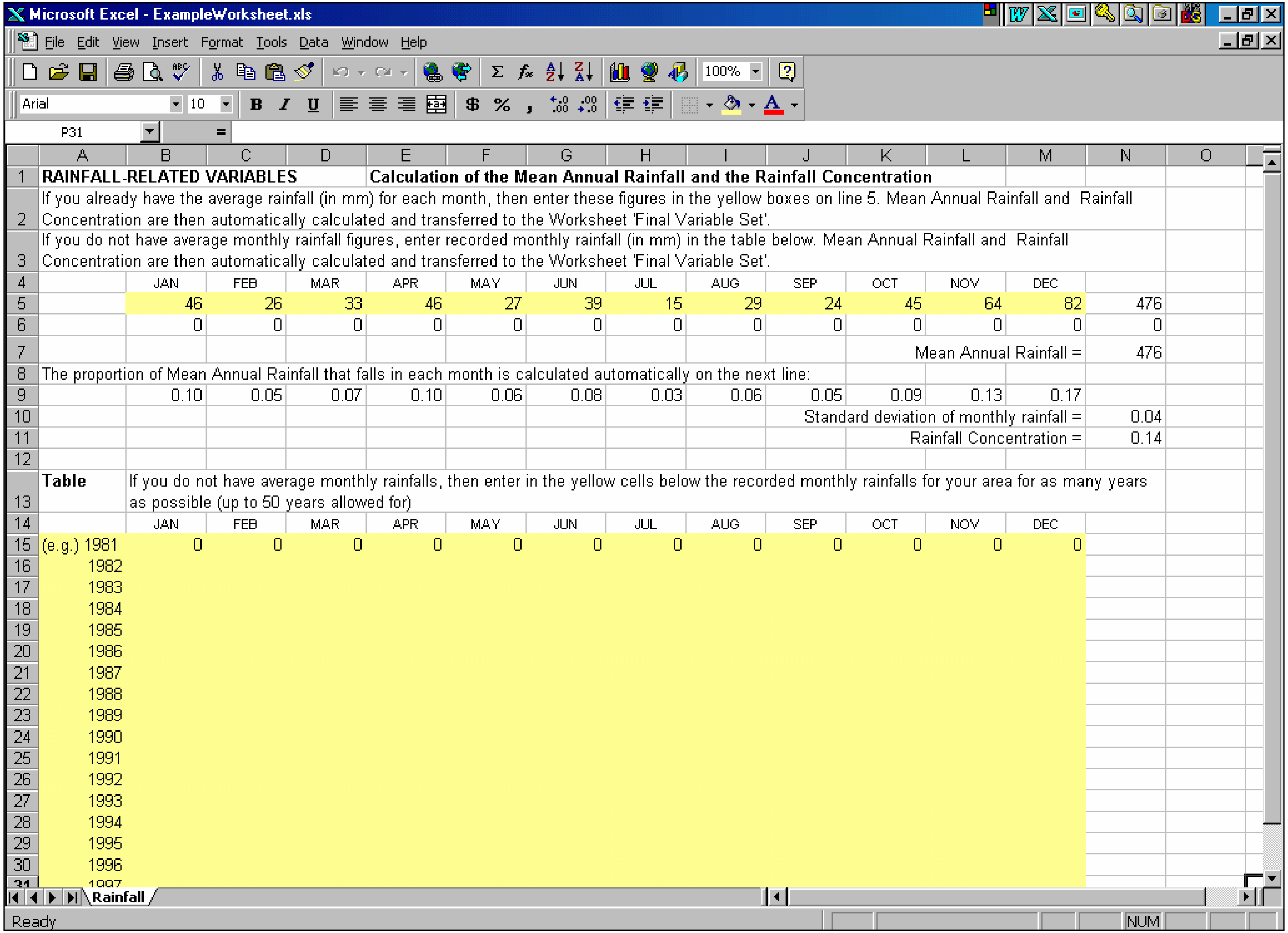

Figure 1 (b). A simplified and improved version of the worksheet Rainfall

Predicting Carrying Capacity with Linear Regression Equations

Linear Regression Analysis

The use of linear regression equations requires several assumptions about the data used, including that:

1. the independent variables were measured without error;
2. the relationship between the dependent and independent variables is a straight line;
3. for any given value of an independent variable, the dependent variable is independently and normally distributed; and
4. there is homogeneity of variances.

In the user guide, the author acknowledges that not all these assumptions were met:

a) most probably there were errors in the measurement of some independent variables, particularly the non-climatic ones (violation of assumption 1);
b) despite the use of data transformations, relationships were not always linear (violation of assumption 2); and
c) greater coefficients of variation when mean annual rainfall is low (compared with when it is high) imply potentially greater fluctuations in carrying capacity in low rainfall areas (violation of assumption 4).

In the user guide, it is stated that proportions were arcsine transformed before inclusion in the linear regression analyses. (It is standard practise during regression analysis to transform proportions in this manner.) But in fact the variable Proportion of Study Area with Low Browse Suitability was not transformed before inclusion in the multiple regression equations.

Rainfall concentration is also a proportion that has not been arcsine transformed before inclusion in the regression analyses. The soil fertility and browse suitability indices are variables which, like proportions, have lower and upper bounds (for a proportion these are 0 and 1, for the soil fertility index they are 1 and 9, and for the browse suitability index they are 10 and 90).

- *On theoretical grounds, the soil fertility index and the browse suitability index should be arcsine transformed before inclusion in the regression equations.*

The four regression equations used in the model all included six independent variables, namely mean annual rainfall, rainfall concentration, mean minimum July temperature, soil fertility, browse availability and one of two measures of browse quality (Table 1). The independent variables are included because the author believed that "the most critical error in developing theoretically-based models is the omission of one or more key predictive variables (i.e. specification error). The author felt that on a theoretical basis, all 6 variables deserved a place in the models." However, as a consequence of this belief, some independent variables were included in the equations even though their inclusion was not statistically justified. In other words, the predictive power of the regressions was not increased by their inclusion. This is not surprising, because the three climatic variables were strongly correlated with each other (at least for the baseline areas used to develop the model). The ideal option is probably to use several equations (all with a sound statistical basis) that have little or no overlap in terms of their independent variables.

- *The regression equations used in the model should contain only independent variables whose inclusion is statistically significant.*

The use of four regression equations that are so similar – in terms of independent variables - is of doubtful value. If the four equations predict the same (or similar) carrying capacities, this would not be surprising, because the independent variables are so similar. However, if the predictions differ greatly, the entire process of model development might be considered doubtful.

Table 1. The variables included (+) in the four linear regression equations that are used to predict black rhino carrying capacity.

Equation	Dependent variable		Independent variables						
	Carrying capacity	Natural logarithm of carrying capacity	Mean annual rainfall	Rainfall concentration	Mean minimum July temperature	Soil fertility	Browse availability	Proportion of area with browse of low suitability	Browse suitability
1	+		+	+	+	+	+		+
2	+		+	+	+	+	+	+	
3		+	+	+	+	+	+		+
4		+	+	+	+	+	+	+	

Measurement of the Independent Variables

Three independent variables - mean annual rainfall, rainfall concentration and mean minimum July temperature - are based on climate data that are easy to collect, with minimum error. Given the same dataset for a study area, different observers would produce, if not the same, then extremely similar values for these variables. In other words, for these variables, repeatability is high.

Soil fertility, browse availability and browse suitability are all assessed on a 9-point scale for each of the vegetation types within the study area and a weighted mean is then calculated. For example, the soil fertility index for North Luangwa NP (see below) was stated as 3.44. Thus, in effect, soil fertility was entered into the model using an 801-point scale (i.e. a scale ranging from 1.00 to 9.00, with 0.01 intervals). This degree of precision is clearly unrealistic. Furthermore, its repeatability between observers would be very low. The author is clearly aware of the problems involved in assessing soil fertility accurately and suggests that 3 or 5-point scales can be used in place of the 9-point one. Nonetheless, soil fertility is treated as a continuous variable in the regression equations. As a consequence, differing soil fertility indices will produce different predictions of carrying capacity, even though the differing soil fertility indices reflect differences in assessment and not differences in fertility.

The robustness of the model (i.e. the repeatability of its predictions when used by different, independent observers assessing the same data set) is just as important as its accuracy. Our reading of the user guide suggests that, although model accuracy was a high priority during model development and design, its robustness was either a low priority or not considered. The robustness of the model was not tested during this review, because this would have required several people with appropriate ecological backgrounds to make field visits and conduct the required assessments (including use of new browse availability assessment guides that we do not possess) at sites within the ecological range of the baseline sites (i.e. outside Zimbabwe).

- *One way to reduce inter-observer variability in the assessment of indices, such as the soil fertility index, might be to use the modal average (not the weighted mean) as the index for the entire study area.*

Predicting Maximum Number of Adult Males

The maximum number of adult male black rhinos that can be accommodated in an area without significant, fighting-induced mortality of adult males is predicted within the model using the size of the study area, and a regression that predicts, from the carrying capacity, the mean size of the home range of adult males.

The regression equation used in the model is one that relates mean male range size and the *predicted* carrying capacity for the 15 baseline areas. This is unsatisfactory, because any error in predicting this carrying capacity produces an additional error in the predicted male range size.

- *Mean male range size should be estimated in the model by substituting the predicted carrying capacity into a regression equation that relates mean male range size and the expert estimate of carrying capacity*.

Not surprisingly, this equation is similar to that used in the current version of the model, but the coefficient of determination is actually slightly greater for this equation than for that which relates male range size and the predicted carrying capacity (R^2_{adj} = 0.840 *vs* 0.823).

SENSITIVITY ANALYSIS OF THE MODEL

A sensitivity analysis was conducted to determine the consequences of an error in the value of one of the variables entered into the model. Adcock has used the model to predict black rhino carrying capacity in a sanctuary in North Luangwa NP (NLNP) where it is proposed to reintroduce black rhinos. NLNP was used in this review for two simple reasons: 1) Adcock's original values of the regression variables for NLNP were available; and 2) NLNP was not one of the baseline areas used to develop the model.

Using the original variable values, the carrying capacity was predicted to be 0.33 rhinos km^{-2}. The original value of just one of the independent variables was increased by 20 % and decreased by 20%, and the predicted carrying capacity recalculated after each adjustment. Only one variable was altered during each run of the model. The full extent of the proposed reintroduction site was regarded by Adcock as low suitability habitat and so this proportion could not be increased (Fig. 2).

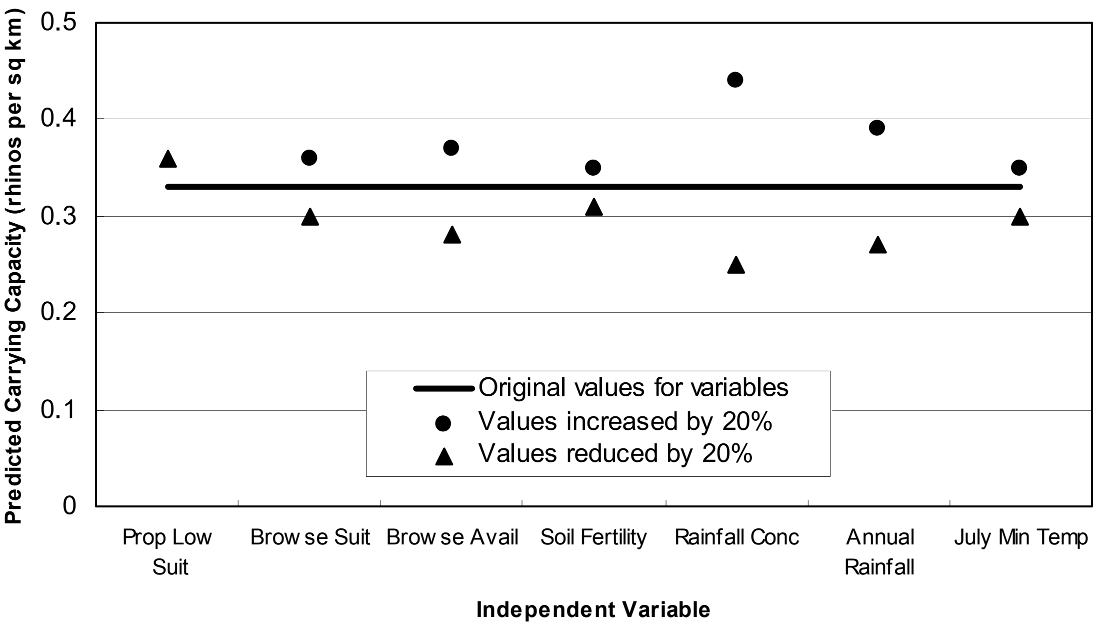

Figure 2. The results of a sensitivity analysis to determine the effect of a 20% increase or decrease in one of the variables included in the model to predict black rhino carrying capacity.

(Incidentally, this designation as 'low density habitat' would, in a Zimbabwean or Zambian context, be queried: it is at least 'medium density' relative to the range of rhino habitat types that are available, hence the definitions of what is 'high', 'medium' or 'low' density would have to be revised if the model acquires a more regional context – habitats like Addo are 'exceptionally high'.)

The predicted carrying capacity was most sensitive to changes in rainfall concentration, annual rainfall and browse availability. It was relatively little affected by changes in the other independent variables (Fig. 2). For all independent variables except rainfall concentration, a 20% change in the value of a variable caused a change of less than 20% in the predicted carrying capacity. But changes in the value for rainfall concentration caused proportionally greater changes in the predicted density.

IS THE MODEL BIOLOGICALLY SENSIBLE?

There are a total of seven independent variables in the model to predict carrying capacity. Each of the four regression equations contains the same five variables. In addition, two of the equations include the proportion of the study area that is low suitability habitat as a variable, and the other two equations contain browse suitability instead. The regression coefficient for the proportion of low suitability habitat is negative, but all other regression coefficients are positive. In the case of browse suitability, July mean minimum temperature, soil fertility and annual rainfall, this is to be expected (although note the interaction between annual rainfall and soil fertility – see below).

Rainfall concentration (i.e. whether rainfall is "concentrated in a few months or spread over most months of the year") was included in the model because it was believed that the durations of the wet and dry seasons would be important determinants of carrying capacity, because of their effects on seasonal changes in the food supply for browsers. However, the user guide does not explicitly state whether the relationship between carrying capacity and rainfall concentration is expected to be negative or positive.

In the regression equations, the coefficient for rainfall concentration is positive, which means that (all other things being equal) predicted carrying capacity increases as rainfall concentration increases. In other words, the model predicts that the shorter the rainy season – and thus the longer the dry season – the greater the carrying capacity. But an ecologist might surmise that a high rainfall concentration index (i.e. a short rainy season and a long dry season) would cause relatively great seasonal changes in browse availability and, consequently, a low carrying capacity for a browser such as the black rhino. Thus, further explanation of the reason for the positive relationship between rainfall concentration and rhino carrying capacity appears to be required.

At least for the 15 baseline study areas, rainfall concentration was closely correlated with two other climatic variables (annual rainfall and mean minimum July temperature) that were also included in the model. Hence, the inclusion of rainfall concentration – at least in the present version of the model – must be questioned, especially given that its inclusion is statistically unjustified for at least two of the regression equations (P = 0.201, 0.478, 0.088, 0.099 for the equations 1 to 4 respectively).

PREDICTIVE POWER OF THE MODEL

Any test of the predictive power of the model requires that its predictions of black rhino carrying capacity are compared with expert-derived estimates of carrying capacity that have not been used during construction of the same model. The information required to construct Adcock's model was available for only 15 study areas. Hence, jack knifing is the most suitable method of testing the model's predictive powers (Sokal & Rohlf 1981).

Jack knifing involved constructing a new model and new multiple regression equations using the data for just 14 study areas. The new equations contain the same independent variables as used by Adcock, transformed in the same ways. Regression analyses were undertaken using SYSTAT v.7.0.1 (Wilkinson 1997). Data (i.e. values for the independent variables) for the 15th study area were then substituted into the new regressions and model. The expert-estimated carrying capacity was then compared with the carrying capacities predicted by the new regression equations 1, 2, 3 and 4 and by the new model.

This process was then repeated 14 times, each time dropping data for a different study area from the data set but retaining data for the other 14 study areas to construct 14 sets of regression equations and 14 new models.

At the end of this process, there are 15 study areas for which one has an expert-derived estimate of black rhino carrying capacity and a prediction of carrying capacity from a model. Most importantly, that model was constructed using data from sites other than the area for which carrying capacity is predicted. For each study area, there are in fact five predicted carrying capacities: those predicted by multiple regression equations 1, 2, 3 and 4 and that predicted by the model, which is the mean of the first four (Table 2).

Table 2. A comparison between expert estimates of black rhino carrying capacity and model predictions. For each study area, carrying capacity was predicted with a model constructed using data from the other 14 areas. Each model uses four multiple regression equations to predict carrying capacity and the mean of these predictions is the model-predicted carrying capacity. Bold figures indicate the 'best' estimate and italics indicate the 'worst' estimate. Best and worst are defined as the predictions closest to and furthest from the expert estimate.

Study area	Area code	Black rhino carrying capacity (rhinos km^{-2})					
		Expert estimate	Predicted				
			Regression 1	Regression 2	Regression 3	Regression 4	Model
Addo Elephant NP	ADD	0.600	0.442	*0.435*	0.660	**0.593**	0.532
Augrabies Falls NP	AUG	0.050	0.007	*-0.008*	**0.017**	0.015	0.008
Andries-Vosloo	AVS	0.350	*0.502*	0.463	0.479	**0.379**	0.456
Eastern Shores – Tewati	ESH	0.200	**0.204**	0.302	0.665	*0.977*	0.537
Hluhluwe	HLU	0.430	**0.414**	0.451	0.454	*0.668*	0.497
Ithala GR	ITA	0.210	0.259	**0.174**	0.254	*0.120*	0.202
Kunene West	K-W	0.020	**0.017**	0.040	0.052	*0.056*	0.041
Lapalala NR	LAP	0.100	0.084	**0.098**	*0.077*	0.085	0.086
uMkhuze	MKU	0.270	0.408	*0.427*	**0.247**	0.345	0.357
Ndumo GR	NDU	0.360	*0.383*	0.364	**0.360**	0.366	0.368
Pilanesberg NP	PIL	0.100	0.142	*0.187*	**0.112**	0.138	0.145
Umfolozi	UMF	0.430	0.340	0.303	**0.345**	*0.231*	0.305
Vaalbos NP	VAA	0.060	*0.110*	**0.082**	0.106	0.086	0.096
Weenen NR	WEE	0.290	0.264	**0.299**	*0.212*	0.360	0.284
Waterberg Plateau Park	WPP	0.085	*-0.021*	-0.005	0.047	**0.052**	0.018

If it is assumed that the difference between the predicted carrying capacity and the expert estimate is a measure of the accuracy of a regression, then it is clear that none of the four regression equations was greatly better or worse than the others, although equation 3 was marginally better than the others (Table 2). Each equation gave the best prediction in 3-5 of 15 cases and the worst in 2-5 cases. Equation 3 gave the best prediction in 5 cases and the worst in 2 cases (Table 2 and Figs 3, 4, 5, 6 and 7).

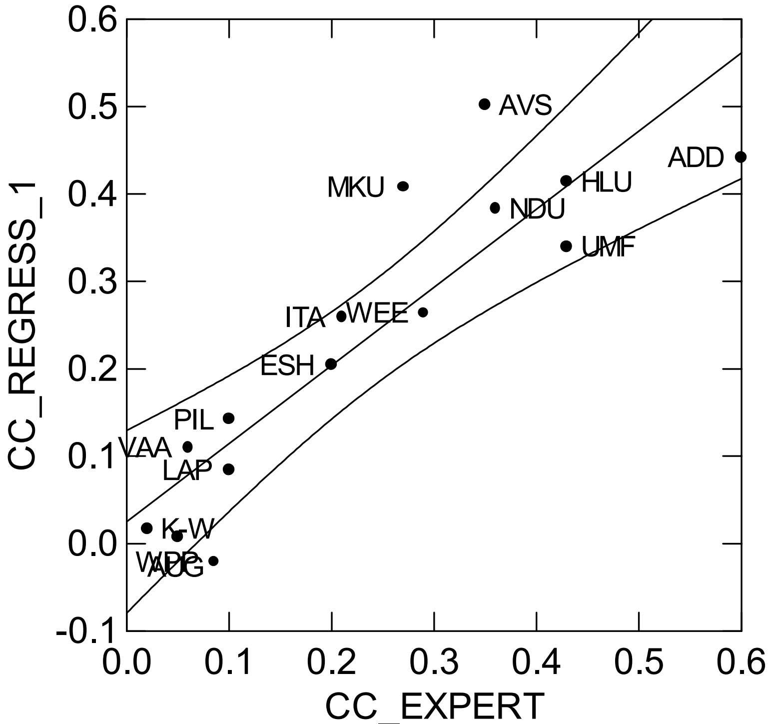

Figure 3. The carrying capacity of black rhinos at 15 study areas, as predicted by regression equation 1 (CC Regress 1), compared with the experts' estimate of carrying capacity in the same areas (CC Expert). Carrying capacities in rhinos km^{-2}. For each area, carrying capacity was predicted using a multiple regression equation derived from data for the other 14 areas. The regression line relating the two carrying capacities and its upper and lower 95 % confidence limits are shown. Study area abbreviations as in Table 2.

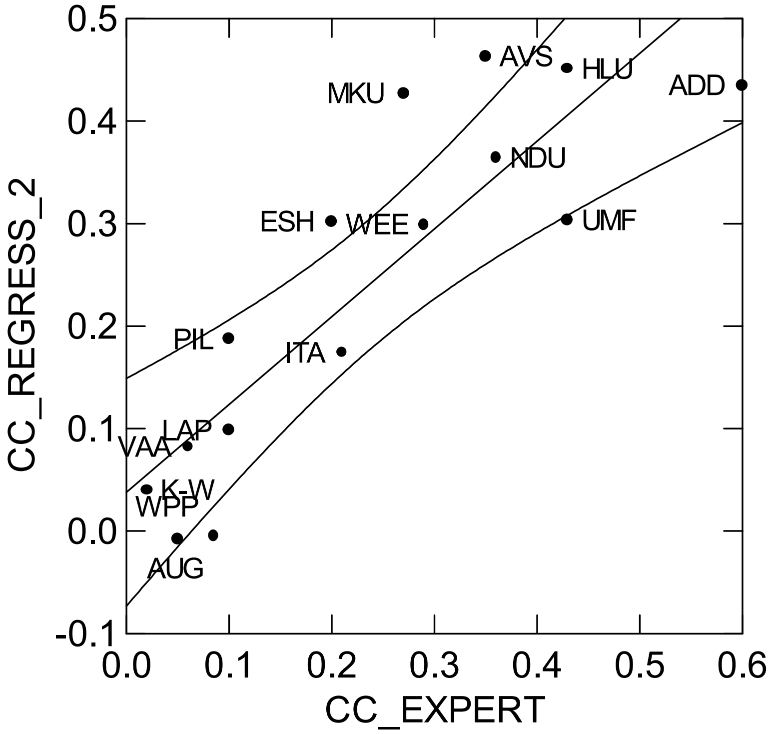

Figure 4. The carrying capacity of black rhinos at 15 study areas, as predicted by regression equation 2 (CC Regress 2), compared with the experts' estimate of carrying capacity in the same areas (CC Expert). Carrying capacities in rhinos km^{-2}. For each area, carrying capacity was predicted using a multiple regression equation derived from data for the other 14 areas. The regression line relating the two carrying capacities and its upper and lower 95 % confidence limits are shown. Study area abbreviations as in Table 2.

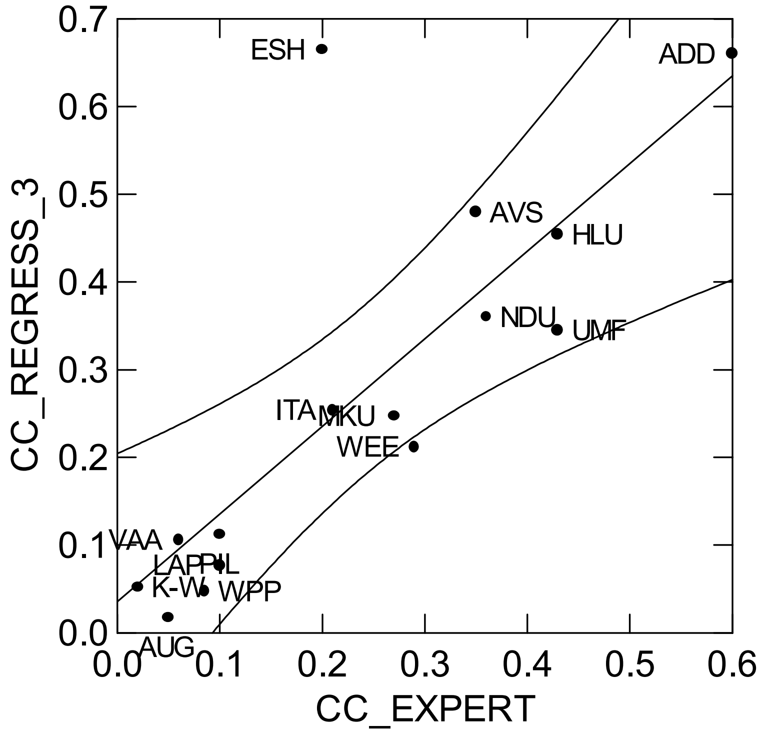

Figure 5. The carrying capacity of black rhinos at 15 study areas, as predicted by regression equation 3 (CC Regress 3), compared with the experts' estimate of carrying capacity in the same areas (CC Expert). Carrying capacities in rhinos km^{-2}. For each area, carrying capacity was predicted using a multiple regression equation derived from data for the other 14 areas. The regression line relating the two carrying capacities and its upper and lower 95 % confidence limits are shown. Study area abbreviations as in Table 2.

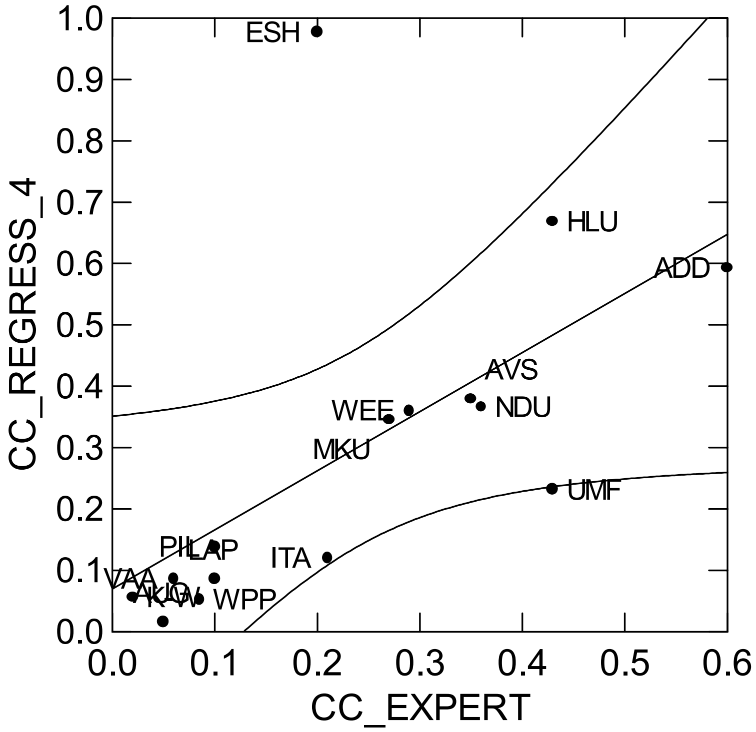

Figure 6. The carrying capacity of black rhinos at 15 study areas, as predicted by regression equation 4 (CC Regress 4), compared with the experts' estimate of carrying capacity in the same areas (CC Expert). Carrying capacities in rhinos km^{-2}. For each area, carrying capacity was predicted using a multiple regression equation derived from data for the other 14 areas. The regression line relating the two carrying capacities and its upper and lower 95 % confidence limits are shown. Study area abbreviations as in Table 2.

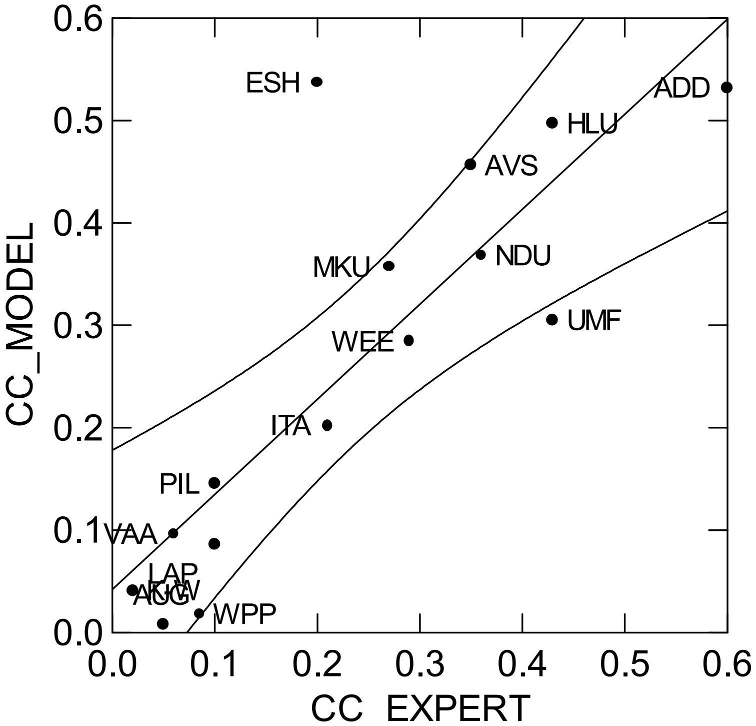

Figure 7. The carrying capacity (in rhinos km$^{-2)}$ of black rhinos at 15 study areas as predicted by Adcock's model (CC Model) compared with the experts' estimate of carrying capacity in the same areas (CC Expert). For each area, carrying capacity was predicted using a variation of Adcock's model that averaged the carrying capacities predicted by four multiple regression equations derived from data for the other 14 sites. The regression line relating the model-predicted and expert-estimated carrying capacities and its upper and lower 95 % confidence limits are shown. Study area abbreviations as in Table 2.

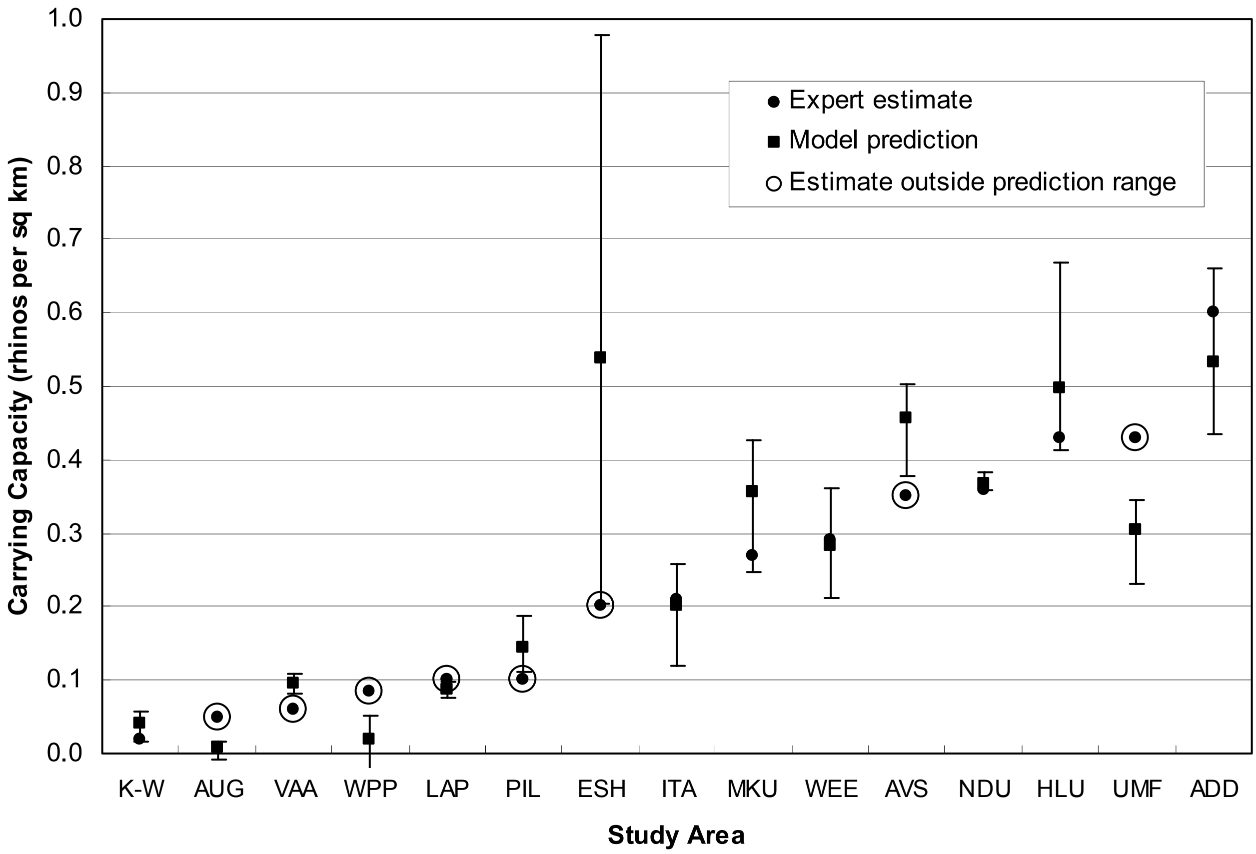

Figure 8. Comparison between the black rhino carrying capacity estimated by experts and the carrying capacity predicted by the model for 15 study areas. The model prediction is the mean of the carrying capacities predicted by four multiple regression equations: vertical lines indicate the range of values predicted by these equations. Expert estimates that lie outside the range of predicted values are highlighted.

For eight of the 15 baseline areas, the expert estimate of carrying capacity was outside the range of the four estimates of carrying capacity produced by the four regression equations (Fig. 8). For four areas, the model overestimated carrying capacity and for four areas it underestimated it. Taken at face value, these results suggest that the predictive power of the model is poor.

The seriousness of these errors depends on one's expectations of the model. When the magnitude of the absolute error (i.e. the difference – in rhinos km^{-2} – between the carrying capacity predicted by the model and the expert estimate) is examined, it is found that the error is less than 0.05 km^{-2} for 8 study areas and less than 0.1 km^{-2} for 13 areas (Fig. 9a). When the percentage error is considered, it is less that 25 % for 6 areas, less than 50 % for 10 areas, but more than 100 % for 2 areas (Fig. 9b).

- *That the model is intended to provide a prediction of carrying capacity that is in the correct 'ballpark' could be emphasised by expressing the model prediction as a range of carrying capacities, rather than as a single figure. For the same reason, care must be taken in deciding the appropriate precision (i.e. number of significant figures) to use when expressing the predicted carrying capacity.*

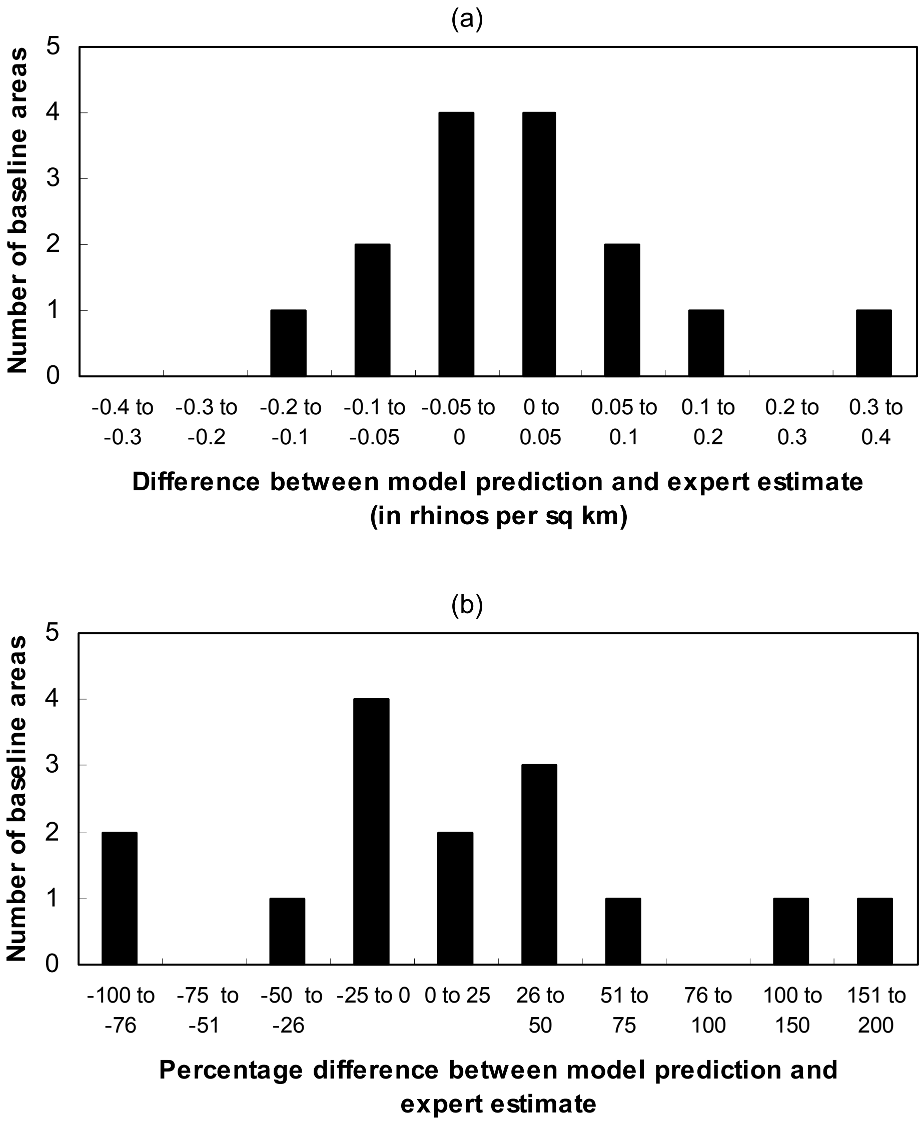

Figure 9. The frequency distribution of errors in the model's prediction of black rhino carrying capacity. The error is defined as the difference between the model's prediction and the expert estimate: (a) error in terms of rhinos km^{-2}; and (b) error as a percentage of the expert estimate.

By any criterion, the difference between the model's prediction of carrying capacity (0.54 rhinos km^{-2}) at Eastern Shores (Tewati Wilderness section) and the expert estimate (0.2 rhinos km^{-2}) is probably unacceptable. This area was the only baseline area where mean annual rainfall exceeded 800 mm, and was the only baseline area with high annual rainfall and low soil fertility (Fig. 10).

In areas with soils that are of medium and high fertility status, black rhino density is known to increase as mean annual rainfall increases (East 1982, Fig. 11). But in areas where soils are infertile, rhino density increases as rainfall increases only up to an annual rainfall of about 700 mm. When annual rainfall exceeds approximately 700 mm, rhino density declines, probably because the high rainfall leaches nutrients from the soil. Hence, the relationship between rhino density and annual rainfall in areas with infertile soil is not linear even when density and rainfall are plotted on a logarithmic scale (Fig. 12). The equation that describes the relationship between density and rainfall (even after data transformation) will never be the linear relationship that was assumed when linear regression equations were used in the carrying capacity model. The relationship is closer to a quadratic one and such relationships can be dealt with during linear regression analysis by including both x and x^2, i.e. in this case (Mean annual rainfall) and (Mean annual rainfall)^2 as independent variables.

That the relationship between rhino density and mean annual rainfall varies depending on the degree of soil fertility illustrates an interaction between rainfall and fertility. The regression equations that form the carrying capacity model do not include any interaction terms and statistical interactions are not mentioned in the user guide.

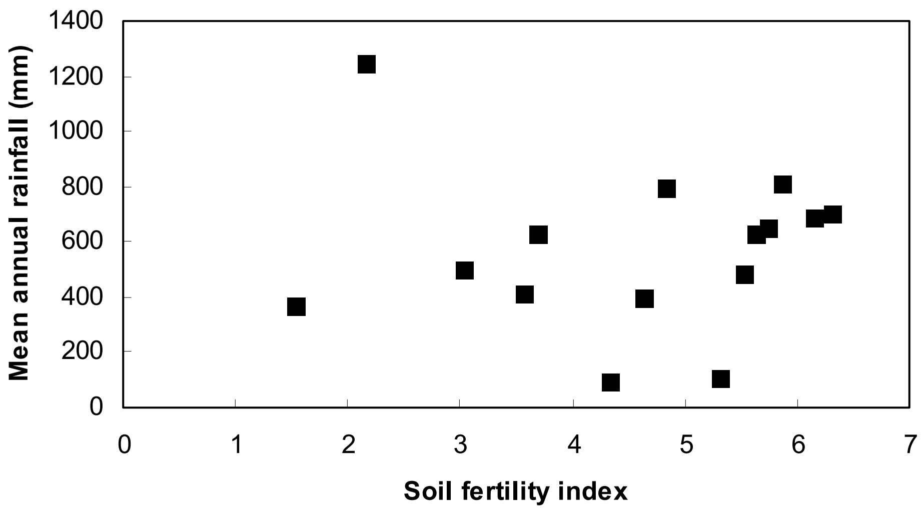

Figure 10. The relationship between annual rainfall and soil fertility at the 15 baseline areas. The point for Eastern Shores (Tewati Wilderness section) is at the top of the graph.

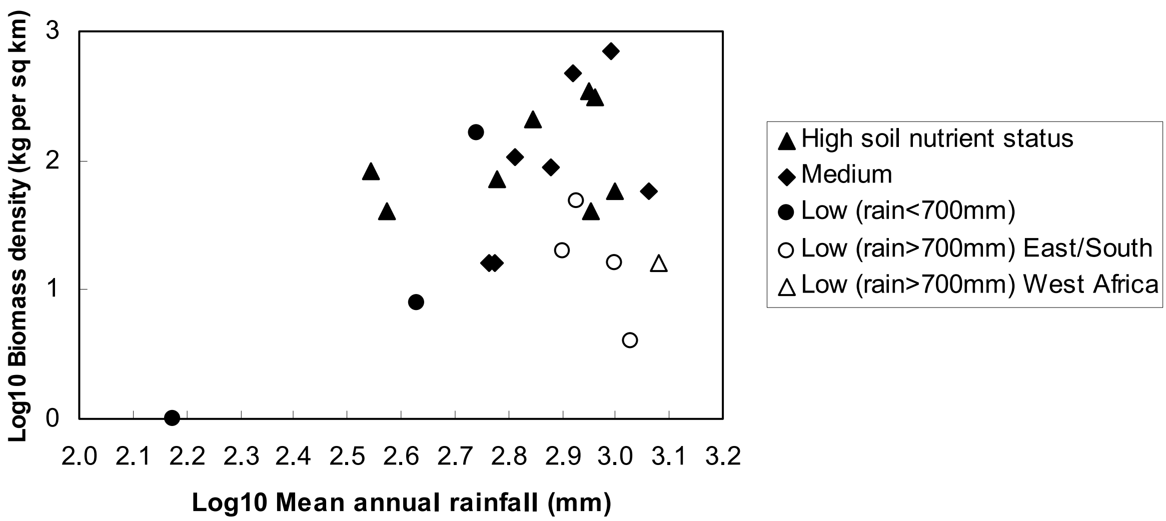

Figure 11. Relationship between the (former) biomass density of black rhinos at various sites in East, South and West Africa, and the mean annual rainfall and soil nutrient status at these sites (from East 1984). Soil nutrient status recorded as high, medium or low. At sites with soils of high or medium nutrient status, biomass density increased with annual rainfall. At sites with soils with low nutrient status, biomass density increased with annual rainfall until rainfall was about 700 mm annually, but then declined as annual rainfall increased further. (700 mm rainfall = 2.85 on $\log_{10}$ scale).

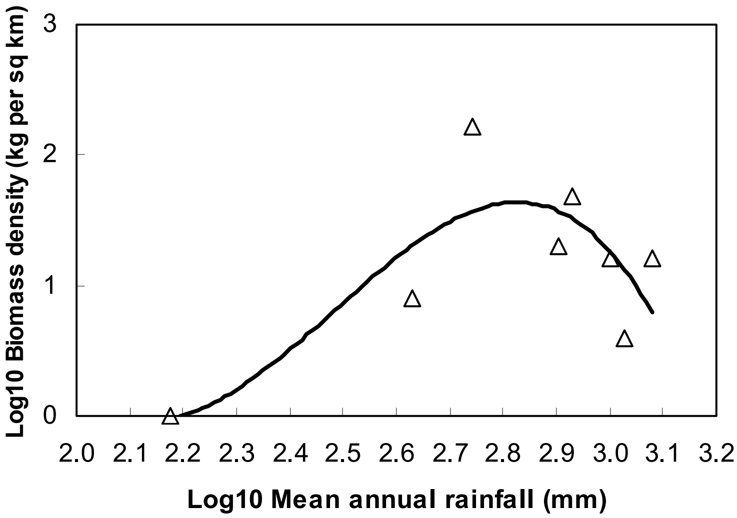

Figure 12. The approximate relationship between rhino density and annual rainfall in areas with infertile soils (data points from East 1984; line shows approximate trend).

THE WAY FORWARD

1. The current model is designed to provide a 'ballpark' estimate of black rhino carrying capacity for an area. The size of this ballpark needs to be defined, because without this information, the adequacy of the model cannot be fully assessed.

    - *A number of independent rhino experts should be asked to state what level of accuracy they expect from a model such as this. In other words, what is the largest difference (between the expert estimate of carrying capacity and the model prediction) that they would find acceptable? Does the model always have to predict within this range, or is it good enough to be accurate on just 90 % or 95 % of occasions? The maximum permissible error must be clearly expressed in absolute terms (e.g. rhinos km^{-2}), or in relative terms (i.e. as a percentage), because perceptions of model adequacy can vary according to which measure is used (Fig. 9).*

2. The model was developed with data from just 15 baseline sites and, as the author is well aware, this is a small sample size. Most of these sites were in a single country.

    - *If a revised version of this model is going to be prepared, the number of baseline sites should be increased, preferably to at least 30.*
    - *Ideally, the additional sites should cover a larger geographical range than the present baseline sites. This is particularly important if the model is to be useful within the SADC region as a whole.*

3. The effect of the interaction between rainfall and soil fertility, and the polynomial relationship of rhino density with rainfall on infertile soils, are inadequately addressed by linear regression analysis.

    - *Generalised Linear Models (GLMs) are often used in place of linear regression analysis to model non-linear relationships, but Generalised Additive Models (GAMs) may be more appropriate. One advantage of GAMs is that, by spline smoothing, they use the data themselves to dictate the form of a relationship and so more accurately reflect a species' response to the environment. GAMs cannot reliably predict beyond the range of the data however and, when data are few, one approach is to use GAMs to determine the shape of a response curve and then to use GLMs to model that form.*

4. A good model will not only predict carrying capacity accurately, it should also be robust. In other words, different people, independently assessing the same data set for a study area, should be able to use the model to predict similar carrying capacities.
   - *The robustness of a model can be improved during model design by:*
     - *maximising the use of independent variables that can be measured with little error and a high degree of repeatability between observers;*
     - *minimising the use of independent variables that can be measured or assessed with only a relatively low degree of accuracy, or for which inter-observer repeatability is low; and*
     - *entering some variables in the model as categorical ones rather than as continuous ones. For example, soil fertility could be modelled as 'low', 'medium' or 'high', rather than as a value ranging from 1.00 to 9.00. This recommendation might reduce the accuracy of the model (although that depends on the size of the 'ballpark'), but it would increase its robustness (because there is likely to be a high level of agreement between observers about which fertility category any given study area is in).*
5. Greater use of categories for variables that are difficult to assess accurately, and consistently between observers, might make it easier to include (as baseline areas during model revision) some areas where rhinos no longer occur, but for which historical data on rhinos and their habitats exist, for example Leader-Williams' studies (1985a, 1985b, 1988) in the core area of South Luangwa NP.
6. An alternative to compiling complex multivariate models for predicting rhino carrying capacity would be to construct an expert system that combines qualitative information and simple quantitative relationships. For example, the alternative to attempting to use a polynomial to model the relationship between (log) carrying capacity and (log) rainfall on infertile soils would be as follows. First decide if soil fertility in the study area is generally high, medium or low. If it is low, find out if mean annual rainfall is greater or less than 700 mm. Either answer might lead to the use of a simple linear regression equation (positively relating (log) carrying capacity to (log) rainfall in the first case, negatively in the second) to provide a first approximation of carrying capacity. This approximation is then modified in the light of answers to further questions, to produce the final predicted carrying capacity.

   The advantages of this method is that: (a) it can replicates the approach that experts might use to predict carrying capacity; and (b) it should allow a much wider range of sites to be considered during model development, because sites can be used even if the information available for them is incomplete.

REFERENCES

Adcock, K. Undated. *User guide to the RMG black rhino carrying capacity model. Version 1. Evaluation Draft*. SADC Regional Programme for Rhino Conservation & Wild Solutions.

Caughley, G. 1977. *Analysis of Vertebrate Populations*. J. Wiley & Sons, Chichester.

East, R. 1984. Rainfall, soil nutrient status and biomass of large African savanna mammals. *African Journal of Ecology* **22**, 245-270.

Leader-Williams, N. (1985a). Black rhino in South Luangwa National Park: their distribution and future protection. *Oryx* **19**, 27-33.

Leader-Williams, N. (1985b). *Problems associated with telemetry of black rhinos in Luangwa Valley and results from a preliminary study of home range size*. Unpublished report submitted to National Parks & Wildlife Service & Save the Rhino Trust.

Leader-Williams, N. (1988). Patterns of depletion in a black rhinoceros population in Luangwa Valley, Zambia. *African Journal of Ecology* **26**, 181-187.

Sokal, R.R. & Rohlf, F.J. 1981. *Biometry. The Principles and Practice of Statistics in Biological Research*. 2nd edition. W.H. Freeman & Co., San Francisco.

Wilkinson, L. 1997. *SYSTAT 7.0: New Statistics*. SPSS Inc., Chicago.