Paper III

Deforestation or not – how reliable are image based mortality estimates?

Deforestation or not – how reliable are image based mortality estimates?

Gidske L. Andersen*

Department of Biology, University of Bergen
P.O.Box 7800, N-5020 Bergen, Norway

Abstract

Results from image based change analyses can give insights into ecological processes and population dynamics. Imagery from the US photo-reconnaissance satellite CORONA have great potential for change analyses since it offers high resolution data, allowing individual trees to be spotted, as early as from the period between 1960 and 1972. However, population estimates derived from change analyses can only be fully assessed if historical reference data exist. As a rule this is not the case, leaving estimates derived, in particular mortality, exposed to image interpretation errors.

Mortality is estimated from a change analysis, combining historical CORONA imagery (1965) with field maps of individual trees (2003). The reliability of the mortality estimate is tested by a method assessing the consistency of the image interpretation based on a Linear Mixed Effects Model and Fishers' exact test.

The image interpretation is consistent for the majority of sites studied. Inconsistencies detected in image interpretation suggest at which sites mortality estimates should be interpreted with care.

In regions where data are scarce, results from image analyses can improve the understanding of ecosystem processes and supply essential information for decision making and management strategies. Survival, mortality and recruitment estimates derived from change analyses are examples of such results. However, management relying on image based estimates can only become successful if the estimates themselves are reliable.

Synthesis and applications: Landscapes worldwide are changing. Remotely sensed data are increasingly important for global change studies and in particular for monitoring environmental and vegetation change. A recurrent problem in image based change analyses is to assess derived population estimates, in particular mortality, when historical reference data is lacking. The method presented here detects inconsistencies in the image interpretation and therefore suggests which estimates should be interpreted with particular care. This is essential additional information for users of image based population estimates.

Key-words: accuracy assessment, arid land ecology, CORONA KH-4a, Eastern Desert of Egypt, monitoring environmental change, high resolution imagery, linear mixed effects models

*E-mail address: Gidske.Andersen@bio.uib.no, fax: +47 55 58 96 67

Introduction

A main application of remotely sensed data is monitoring of landscapes and vegetation over time (Linderholm, 2006; Pettorelli et al., 2005; Tucker, Dregne & Newcomb, 1991) The extensive time series of historical earth observation data that provide detailed information about the ecological status of major parts of the globe presently cover a period of several decades. In particular in remote or poorly developed areas where data often are scarce such historical datasets are of great potential value for monitoring ecosystem variation and change over time. Imagery from photo-reconnaissance satellites such as the US CORONA is unique in this sense because it offers high resolution data (down to 1.8 m resolution) from as early as the period between 1960 and 1972 (McDonald, 1995). In arid areas where vegetation cover is sparse, this imagery offers great potential to study and monitor individual trees (Andersen, 2006). In combination with more recent data, e.g. field observations, insights into ongoing change processes and population dynamics of dominant tree species can be gained. Survival, mortality and recruitment estimates derived from such analyses are essential for a better understanding of these ecosystems and are needed if appropriate management decisions are to be taken. However, a prerequisite for the successful application of the results is that the derived estimates are reliable.

The reliability of image based estimates depends on the consistency and accuracy of the image interpretation. Although it does not always happen, it is widely accepted that generalized remote sensing map products should be accompanied by data stating their quality, and for this standard methods exist (Congalton, 1988, 1991; Foody, 2002; Stehman & Czaplewski, 1998). However, to assess the image interpretation requires contemporaneous reference data, which is almost always lacking for historical datasets (Okeke & Karnieli, 2006). This is also the case here, but a partial assessment of errors of omission, i.e. existing trees wrongly left out in the image interpretation, can be made using field observations as reference data (Andersen, 2006). However, an assessment of errors of commission, i.e. image structures that are falsely assumed to represent trees, cannot be made. This leaves the mortality estimate exposed to interpretation errors.

In the current study we present a method for assessing the consistency of the image interpretation in order to infer the reliability of the mortality estimate derived from a change analysis, combining historical CORONA imagery (1965) with field maps of individual trees (2003).

Methods and Material

During fieldwork in 2003 each individual tree on 19 different sites in the Eastern Desert of Egypt (ED) was registered and mapped by a handheld Global Positioning System receiver (Fig. 1). *Acacia tortilis* (Forssk.) Hayne and *Balanites aegyptiaca* (L.) Del. are dominant tree species growing in the dry river valleys (*wadis*) in this mountainous hyper-arid area (mean annual precipitation < 30 mm, coefficient of rainfall variation > 200 %; Andersen, 1999; Ayyad & Ghabbour, 1985).

Mortality is derived from the results of a change analysis comparing observed trees (2003) with points, i.e. possible trees, visually interpreted and digitized from the oldest (Dec. 1965) and best resolution (2.7 m at nadir) CORONA imagery available from the ED (KH-4a, mission 1027-1, frames 142-148). See Andersen (2006) for information about imagery, preprocessing and interpretation of vegetation content. A size and distance threshold is used to categorize interpreted points as belonging to one of two Change Categories (CC), either *surviving* (interpreted from 1965 imagery and observed in 2003) or *dead* (interpreted from 1965 imagery but not observed in 2003).

In a visually based process each interpretation will always be made with different degrees of certainty, introducing a bias into the final product which will be unknown to its users (Thierry & Lowell, 2001). Therefore, each interpreted point is also placed in an Interpretation Class (IC) based on an image analysis made independently of field observations. The frequency and characteristics of *surviving* and *dead* points within each IC are used to generate hypotheses testing the consistency of the image interpretation. This process is outlined in Fig. 2, and the result is presented in Fig. 3. Two ICs are applied:

- *Accepted* (IC1): an identifiable structure with good contrast which resembles a structure generated by a tree, and
- *Doubted* (IC0): an identifiable structure resembling a structure generated by a tree,

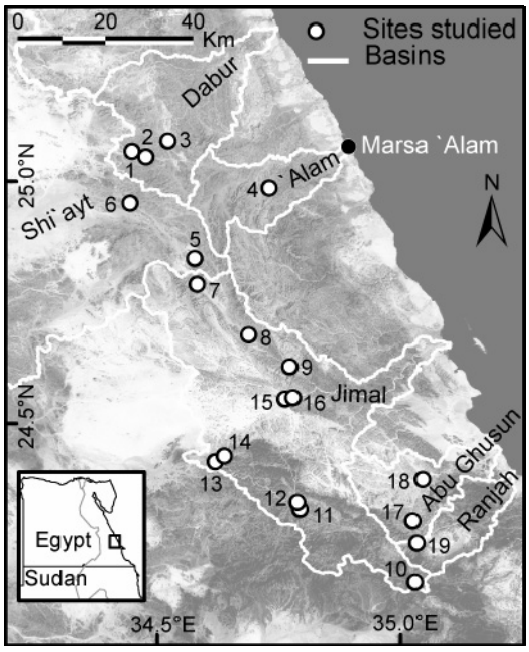

Figure 1: Study area. Studied sites are seen in relation to their basins. Each site is labelled with its ID (see Table 1). A Landsat TM image is displayed in the background.

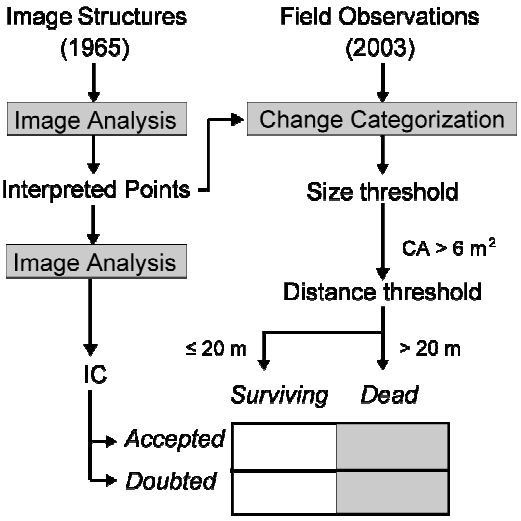

Figure 2: Flowchart. Image structures resembling trees are interpreted and represented as points. Together with trees recorded in the field, interpreted points are categorized as belonging to one of two Change Categories (CC): *surviving* and *dead*. In this process size (Canopy Area of 6 m^2) and distance (20 m) thresholds were applied to compensate for GPS inaccuracies and limited image resolution (Andersen, 2006). Interpretation classes (IC) of *accepted* and *doubted* are arrived at in a second image analysis process. ICs are applied together with *surviving* and *dead* to assess the reliability of the mortality estimated.

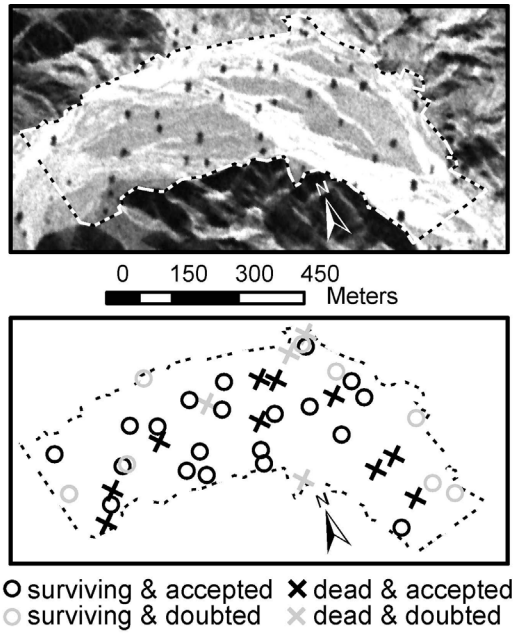

Figure 3: Interpreted points according to Interpretation Class and Change Category in Nuqrus. The CORONA image is seen in the upper panel.

but *doubted* either because a) it overlaps an *accepted* interpreted point (shape is strongly irregular, indicating that two trees are adjacent), or because b) it is indistinct due to poor contrast.

Contrast is quantified as "photocontrast", i.e. the ratio between the background pixels (the mean of marginal pixels in a window of size 7 x 7) and the central pixel of the interpreted image structure. Mean canopy diameter for *surviving* trees is 8 m.

In particular, interpreted points categorized as *doubted & dead* may contribute to an overestimation of tree mortality. On the underlying assumption that interpretations are correct, a set of hypotheses is formulated:

- $H0_1$: As for the interpreted points, there is no difference in the frequency between *doubted* points for *dead* trees and for *surviving* trees. If *doubted & dead* trees are overrepresented, the interpretation may have been too liberal, i.e. may have included false points.

$H0_1$ is tested for each site separately using the 2x2 contingency table defined by ICs and CCs applying Fisher's exact test (Crawley, 2002).

- $H0_2$: There is no difference in photocontrast between *doubted* and *accepted* points. Rejection of $H0_2$ would confirm the assumption of the visual interpretation, i.e. *doubted* points have a different photocontrast from *accepted* points.

- $H0_3$: *Doubted & dead* points do not differ in photocontrast from *doubted & surviving* points. Rejection of $H0_3$ would suggest a systematic trend in the distinctness of points across CC and IC.

To test $H0_2$ and $H0_3$ the clustered data of sites and interpreted points is analyzed using a linear mixed-effects model (Pinheiro, 2000). *Doubted* interpretations adjacent to *accepted* interpretations (IC0a) are excluded from the model because doubt in these cases is a function of shape, not photocontrast. The model selection is based on Akaike Information Criteria and Bayesian Information Criteria. In all models CC and IC are modeled as fixed effects, while the site factor is included as a random effect. The response variable (photocontrast) is assumed to follow a normal distribution.

A subset including only *doubted* interpretations was analyzed to assess the site-specific variability of CC. Model coefficients for individual sites (fixed effect + site-specific random effect) indicate which sites deviate most from the expectation shared by all sites (only fixed effects). All analyses were done in R, version 2.1.1 (http://www.r-project.org).

Results

$H0_1$ is rejected only for Jimal I (Table 1). At the confidence level of $\alpha=0.05$ we anticipated that one site might show significance even if everything were random, but Jimal I is significant also after a Bonferroni correction.

Three models with different variance structures for within-group errors were tested under $H0_2$. The best fit model includes different variance for each combination of CC and IC. The coefficients for this model (Table 2) show that *accepted* interpretations have a generally higher photocontrast (*dead*=2.26 and *surviving*=2.17) than *doubted* interpretations (*dead* =1.49 and *surviving* =1.61). Note that for *accepted* interpretations *dead* trees have a higher photocontrast than *surviving* trees (2.26 vs. 2.17), while the opposite is true for *doubted* interpretations. The significance of the interaction term of CC and IC (Table 2) warrants the rejection of $H0_3$.

A new model for the subset of *doubted* interpretations shows that the site-specific random variability in CC is considerable. This means that the effect of CC on photocontrast differs considerably among sites. The large site-specific random effect is, however, driven mainly by one site, i.e. Nuqrus (Table 3).

Table 1: Number of interpreted points per site studied for each combination of Interpretation Class (*doubted* vs. *accepted*) and Change Category (*dead* (D) vs. *surviving* (S)) and results from Fisher's exact test (FET). $H0_1$ was rejected only for Jimal I. Of *doubted* interpretations adjacent to *accepted* (IC0a) 19 were *absent* and 20 *present*. Site ID refers to Fig. 1.

		Doubted		Accepted		FET
Site-ID	Site	D	S	D	S	greater
1	Dabur I	2	9	2	20	0.407
2	Dabur II	5	8	3	18	0.116
3	Dabur III	4	6	3	13	0.230
4	Sukkari	6	5	9	11	0.447
5	Hanjaliyyah	8	7	22	15	0.763
6	Kharrasha	8	5	4	7	0.207
7	Nuqrus U	6	2	10	14	0.110
8	Nuqrus M	17	7	25	9	0.702
9	Nuqrus	4	8	10	20	0.635
10	Hulus U	1	2	5	10	0.730
11	Hulus M	1	6	4	12	0.870
12	Hulus M II	3	7	13	9	0.973
13	Gaetri	18	6	26	23	0.060
14	Hulus L	4	3	4	8	0.297
15	Jimal I	42	5	45	27	0.001
16	Jimal II	12	6	20	8	0.750
17	Sartut I	3	5	5	18	0.331
18	Abu Ghusun	4	5	12	21	0.471
19	Hulayfi	3	4	11	13	0.713
	Σ	151	106	233	276	

Table 2: Summary statistics of the best fit model for $H0_2$. Photocontrasts for the different combinations of Interpretation Class (IC) and Change Category (CC) are calculated from "Value". Intercept is the mean photocontrast for *doubted* & *dead* trees, CC is the difference in photocontrast between *doubted* & *dead* trees and *doubted* & *surviving* trees, df degrees of freedom, SE is Standard Error and SD is Standard Deviation. Treatment contrast is applied.

	Value	SE	df	t-value	p-value	SD
Intercept	1.49	0.06	705	25.90	< 0.001	0.22
CC	0.12	0.04	705	3.00	0.003	
IC	0.78	0.09	705	8.42	< 0.001	0.32
CC:IC	-0.21	0.07	705	-2.86	0.004	
Residual						0.24

Discussion

Most sites studied admit of consistent image interpretation, i.e. mortality estimates are reliable; but a few sites need further discussion:

At Jimal I *doubted* & *dead* interpreted points are overrepresented (Table 1); but since these points are as distinct as *doubted* & *surviving* ones (Table 3), the argument for higher mortality among *doubted* & *dead* individuals is strengthened. Since soils are light and homogeneous and therefore cause no disturbance in image interpretation, doubt is probably correlated with smaller canopy sizes (Andersen, 2006). In relation to water conditions, smaller trees (poorly developed roots) are more susceptible to both uprooting and water shortage. Uprooting was observed in Wadi Jimal, but it does not necessarily kill individuals because of their high capability to resprout (GL. Andersen, unpublished data). In the change algorithm applied, however, interpreted points categorized as *dead* increase mortality even if replaced by resprouts. This is a consequence of the size threshold applied (Fig. 2). As for water shortage, they are less likely in W. Jimal given the dominance of *B. aegyptiaca* which indicates generally good water conditions (Abdel Moneim, 2005; Hall,

1992). In relation to anthropogenic disturbance, e.g. charcoal production, which seems to be a main cause of tree mortality, respect for larger trees can explain why smaller ones show higher mortality (Hobbs, 1989; GL. Andersen, unpublished data).

Alternatively, *doubted* and *dead* individuals could be shrubs, e.g. *Leptadenia pyrotechnica* and *Ochradenus baccatus*, mistaken for (small) trees. They are perennial but less drought persistent than trees. However, some individuals were identified both in the field (2003) and in CORONA imagery (1965) at Jimal II. This indicates persistency and a potential to grow to sizes detectable in imagery. Such confusion of shrubs with trees would overestimate tree mortality; but in the field these shrub species were only recorded at the Jimal sites.

Nuqrus is the site driving the inconsistency in the image interpretation (cf. Results). This leaves 4 of 14 interpreted points categorized as *dead* unreliable (Table 1). In the worst case this reduces the mortality estimate at Nuqrus from 25 % to 21 %, and consequently affects population size and also other population estimates. In this specific study, however, there are reasons to believe that an underestimation of mortality caused by omission errors is a greater problem than a possible overestimation caused by commission errors. (Andersen, 2006) found that several larger trees (CA > 6 m^2) present in 2003, could not be recognized in imagery because they were close to features reducing image interpretability. If mortality among such undetectable trees is similar to that in the remaining population, mortality is probably underestimated at most sites. This could amount to an increase from 0.3 % (Kharrasha) to 5.1 % (Hulayfi).

Table 3: The coefficients for the individual sites, including both fixed and site-specific random effects. Intercept is the estimated mean photocontrast for each site for *doubted* & *dead* trees, and CC (change category) is the photocontrast value to be added to reach the level for *doubted* & *surviving* trees. Diff is the site-specific difference between dead and surviving points. This shows the variability in the fixed effect of CC (0.12) and the CC as estimates for the individual sites. The difference among sites is close to significant p=0.053

Site	Intercept	CC	Diff
Dabur I	1.52	0.15	0.04
Dabur II	1.67	0.17	0.05
Dabur III	1.56	0.21	0.10
Sukkari	1.42	0.07	-0.05
Hanjaliyyah	1.43	0.13	0.02
Kharrasha	1.33	0.04	-0.08
Nuqrus U	1.57	0.16	0.04
Nuqrus M	1.56	0.13	0.01
Nuqrus	1.83	0.43	0.31
Hulus U	1.36	0.00	-0.12
Hulus M	1.38	0.04	-0.08
Hulus M II	1.65	0.22	0.11
Gaetri	1.55	0.11	-0.01
Hulus L	1.65	0.21	0.09
Jimal I	1.41	0.01	-0.10
Jimal II	1.38	0.02	-0.09
Sartut I	1.35	0.03	-0.09
Abu Ghusun	1.53	0.11	-0.01
Hulayfi	1.34	0.00	-0.12

The data applied here, CORONA, are particularly useful in areas where tree or shrub populations are of interest and vegetation patterns permit solitary individuals to be detected, e.g. in sparsely vegetated areas. Ongoing changes in arid lands are a major topic for which this imagery supplies exactly the high resolution and the historical data required for a fuller

understanding of this much discussed problem (Ezcurra, 2006; Helldén, 1991; Hutchinson et al., 2005; Thomas & Middleton, 1994). However, the results should be quality controlled, and therefore accuracy assessment is considered as an integral part of any remote sensing application (Foody, 2002). In the application of historical data, however, the lack of contemporaneous reference data prevents proper accuracy assessment of the derived products (Okeke & Karnieli, 2006). An additional challenge for visual photo interpretation is the changing level of certainty of the interpretations which is not conveyed to the user (Thierry & Lowell, 2001). Both these challenges are addressed by the suggested method which will supply essential information about the reliability of the results, thus assuring that appropriate management measures can be taken.

Acknowledgements

The Norwegian Research Council (project no. 149181/730), the Faculty of Mathematics and Natural Sciences, University of Bergen, Olaf Grolle Olsens Legat and L. Meltzers Høyskolefond funded this research. Thanks to Knut Krzywinski and Richard H. Pierce for valuable contributions in discussions and in finalizing the manuscript and to Einar Heegaard for essential statistical help.

References

Abdel Moneim, A.A. (2005) Overview of the geomorphological and hydrogeological characteristics of the Eastern Desert of Egypt. Hydrogeology Journal, 13(2), 416-25.

Andersen, G.L. (1999) Change and variation in a hyper-arid cultural landscape: A methodological approach using remote sensing timeseries (LANDSAT MSS and TM, 1973-1996) from the wadi vegetation of the Eastern Desert of Egypt. cand. scient. thesis, University of Bergen/NERSC, Bergen.

Andersen, G.L. (2006) How to detect desert trees using CORONA images: Discovering historical ecological data. Journal of Arid Environments, 65(3), 491-511.

Ayyad, M.A. & Ghabbour, S.I. (1985). Hot deserts of Egypt and Sudan. In Hot desert and arid shrublands, B (eds M. Evenari, I. Noy-Meir & D.W. Goodall), Vol. 12B, pp. 149-202. Elsevier.

Congalton, R.G. (1988) A Comparison of Sampling Schemes Used in Generating Error Matrices for Assessing the Accuracy of Maps Generated from Remotely Sensed Data. Photogrammetric Engineering and Remote Sensing, 54(5), 593-600.

Congalton, R.G. (1991) A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. Remote Sensing of the Environment, 37(1), 35-46.

Crawley, M.J. (2002) Statistical computing: an introduction to data analysis using S-Plus Wiley, Chichester.

Ezcurra, E., ed. (2006) Global deserts outlook UNEP, Kenya.

Foody, G.M. (2002) Status of land cover classification accuracy assessment. Remote Sensing of Environment, 80(1), 185-201.

Hall, J.B. (1992) Ecology of a key African multipurpose tree species, Balanites aegyptiaca (Balanitaceae): the state-of-knowledge. Forest Ecology and management, 50, 1-30.

Helldén, U. (1991) Desertification - time for an assessment? Ambio, 20, 372-83.

Hobbs, J.J. (1989) Bedouin life in the egyptian wilderness University of Texas Press, Austin, Texas.

Hutchinson, C.F., Herrmann, S.M., Maukonen, T. & Weber, J. (2005) Introduction: The "Greening" of the Sahel. Journal of Arid Environments, 63(3), 535-37.

Linderholm, H.W. (2006) Growing season changes in the last century. Agricultural and Forest Meteorology, 137(1-2), 1-14.

McDonald, R.A. (1995) CORONA - Success for Space Reconnaissance, a Look into the Cold-War, and a Revolution for Intelligence. Photogrammetric Engineering and Remote Sensing, 61(6), 689-720.

Okeke, F. & Karnieli, A. (2006) Methods for fuzzy classification and accuracy assessment of historical aerial photographs for vegetation change analyses. Part I: Algorithm development. International Journal of Remote Sensing, 27(1), 153-76.

Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.M., Tucker, C.J. & Stenseth, N.C. (2005) Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends in Ecology & Evolution, 20(9), 503-10.

Pinheiro, J.C. (2000) Mixed-effects models in S and S-PLUS Springer, New York.

Stehman, S.V. & Czaplewski, R.L. (1998) Design and analysis for thematic map accuracy assessment: Fundamental principles. Remote Sensing of Environment, 64(3), 331-44.

Thierry, B. & Lowell, K. (2001) An uncertainty-based method of photointerpretation. Photogrammetric Engineering and Remote Sensing, 67(1), 65-72.

Thomas, D.S.G. & Middleton, N.J. (1994) Desertification: Exploding the myth Wiley, Chichester.

Tucker, C.J., Dregne, H.E. & Newcomb, W.W. (1991) Expansion and Contraction of the Sahara Desert from 1980 to 1990. Science, 253, 299-301.