This is a preprint of an article published in Ecological Indicators 2014 (36): 524-531.
For the final version please go to: http://dx.doi.org/10.1016/j.ecolind.2013.09.008

Convergence between ANPP estimation methods in grasslands – A practical solution to the comparability dilemma

Jan Christian Ruppert[a,b], Anja Linstädter[a]

[a] Range Ecology and Range Management Group, Botanical Institute, University of Cologne, Zülpicher Straße 47b, D-50674 Cologne, Germany

[b] Crop Science Group, INRES, University Bonn, Katzenburgweg 5, D-53115 Bonn, Germany

Corresponding author:

Jan Christian Ruppert

Email: jan.ruppert@uni-koeln.de

Tel.: +49 (0) 221 - 470 7906

Abstract

Aboveground net primary production (ANPP) is a key ecosystem characteristic and of fundamental importance for essentially all aspects of matter and energy fluxes in terrestrial ecosystems. Various methods for estimating ANPP are available and despite partial consensus on 'best practice methods' important methodological issues remain unresolved: ANPP data obtained with different methods differ in their magnitude, variability and their tendency to over- or underestimate primary production. Paradoxically, despite the large number of published ANPP data, the limited comparability of ANPP estimates across studies leads de facto to a scarcity of ANPP data for assembled large-scale studies. We aimed to overcome these problems by establishing conversion rates between the most commonly used ANPP methods, thus making the large body of published ANPP data more comparable and thus useful for assembled large-scale studies.

Using seasonal biomass dynamics from 89 sites representing various biomes and climata, we established linear conversions for all 21 combinations between the seven most common ANPP estimation algorithms in grass-dominated vegetation. We also checked for confounding effects of environmental factors such as biome, management and climatic aridity. Aridity was the only factor with a clear influence on ANPP conversions, and in six cases we thus calculated separate relationships for dry and humid conditions. In these cases, dryland ANPP was systematically underestimated by the respective methods. As these methods are insensitive to turn-over processes from live to senescent biomass, we assume this underestimation is related to climate-induced differences in biomass turn-over rates, with more arid sites having higher rates.

The majority of the resulting 27 conversions had high (pseudo) R^2 values (≥ 0.65; full range: 0.31 - 0.92), indicating clear linear relationships between most ANPP estimation methods. Given the large size of the dataset and the accuracy of statistical models, we assume that most conversion formulae are generally valid. We classified conversions with respect to their R^2 values and their methodological comparability, and concluded that 16 conversions can be fully recommended. For those cases where a recalculation of ANPP on basis of original biomass data is not possible, our conversion formulae offer an easy and practical approach to synchronize ANPP estimates from divergent algorithms and sources.

Keywords: Aboveground Net Primary Production, Grasslands, Global ANPP dataset, ANPP estimation, Ecosystem services

Abbreviations: (A)NPP – (Aboveground) Net Primary Production, ORNL DAAC – Oak Ridge National Laboratory Distributed Active Archive Center

1. Introduction

Aboveground net primary production (ANPP) is a key ecosystem characteristic and of fundamental importance for essentially all aspects of matter and energy fluxes in terrestrial ecosystems. It is a prominent core ecological currency and one of the best documented quantitative estimate for several ecosystem services such as forage or lumber (Scurlock et al., 2002). However, as it represents a concept rather than a precise physical quantity or attribute, ANPP can only be estimated by surrogate measurements and not measured directly (Lauenroth et al., 2006).

Many different procedures and methods for estimating ANPP have been developed. Particularly in grass-dominated ecosystems, a wide variety of different estimation protocols have been developed within recent decades. The most common methods to estimate ANPP (hereafter simply 'ANPP methods') have been thoroughly evaluated and compared in literature (Lauenroth et al., 2006; McNaughton et al., 1996; Milner and Hughes, 1968; Sala and Austin, 2000; Scurlock et al., 2002; Singh et al., 1975). However, despite a partial consensus on 'best practice methods', discussion regarding various methodological issues is still ongoing, and as a result, numerous ANPP estimation methods are in use and compete up until today. Generally, ANPP methods can be sub-divided into complex elaborated methods and simple, less elaborated ones. Elaborated methods, which account for dynamics in live, senescent, and moribund tissue simultaneously throughout the growing season, have often been recommended (Singh et al., 1975; Scurlock et al., 2002). However, these methods are far more labor-intense and costly than other 'simple' estimations (e.g. *Peak standing crop*, or *Peak live biomass*) which have a tendency to underestimate production. Unsurprisingly, less elaborate methods are far more often applied, as they are faster and cheaper. Unfortunately, different ANPP methods differ not only in their general accuracy (i.e. their tendency to over- or underestimate ANPP), but also with respect to magnitude, variability and uncertainty (Scurlock et al., 2002; Lauenroth et al., 2006). These differences render estimates based on different methods more or less incomparable. Scurlock et al. (2002) have shown that ANPP estimates at one site and date may vary up to more than 6-fold depending on the computational method used. Examples from our own dataset show even more extreme differences of up to 10- to 15-fold in certain cases (data not shown).

In the past, simple methods like *Peak standing crop* were sufficient for common questions in vegetation and rangeland ecology. They give robust estimates which are sufficient for determining carrying capacity, assessing the influence of climatic characteristics, or comparing the effects of contrasting management strategies at local scale (e.g. Blaisdell, 1958; Dye and Spear, 1982; Smoliak, 1986)). However, in recent years there is a growing demand for both more accurate and better comparable ANPP data across larger scales. In fact the lack of large-scale ANPP data has been stated as one of the most crucial data gaps in ecology in recent times (Ni, 2004; Scurlock et al., 2002; Scurlock and Olson, 2002). Paradoxically, despite the large number of studies presenting ANPP data on field and site scale, the limited comparability of ANPP data across sites, regions and studies de facto leads to a scarcity of ANPP data for supra-regional or large-scale studies.

In the light of the climate and land-use change debate, the need for reliable and adequately scaled large-scale and global ANPP datasets is urgent, as each of cross-system analyses, meta-analyses, as well as land-use, climate and vegetation models imminently require them. Since adequate biomass and ANPP monitoring is not only time consuming but also costly, numerous scientists rely on assembling ANPP datasets from published data (Hsu et al., 2012; Lauenroth and Sala, 1992; Ni, 2004; Ruppert et al., 2012). However, due to differences between ANPP estimation methods, this pragmatic solution is not without its pitfalls. Surprisingly, only a small proportion of studies discuss the issue of comparability of ANPP data assembled from various sources, and based on different estimation and/or computation methods (see 3.1 Results). To date, authors of large-scale studies and meta-analyses either had to neglect major proportions of published data for the sake of comparability or accept the limited and unknown comparability, a true *'comparability dilemma'*.

Still, little is known about the incidence and frequency of ANPP comparability issues in assembled datasets.

Being confronted with this comparability dilemma ourselves (Ruppert et al., 2012; Ruppert et al. in prep.), we aimed to overcome these problems by searching for conversions rates between common ANPP methods. We found that Singh et al. (1975) presented conversions for a set of different ANPP method combinations, developed on the basis of ten short-term datasets form North American grasslands. Surprisingly, practically no use was made of these conversions thereafter. A review (see 2.1 Materials and methods) of all 165 studies citing Singh et al. (source: Google Scholar) revealed that only two studies used the conversions, both by authors of the original paper (Lauenroth and Whitman, 1977; Singh et al., 1983). This poor adoption may be explained by various reasons including: (1) the paper was largely a detailed review, and the conversions were not mentioned in the abstract limiting their visibility; (2) the strong interest in large and global scale ANPP datasets was not as virulent in the 1970s as it is today; and (3) perhaps most critically, the study was based on a restricted dataset and did not test whether conversions were applicable to data from other regions or ecosystems.

We believe that the attempt by Singh et al. (1975) was simply ahead of its time and that it offers a starting point to assess the comparability for future assembled studies. However, the problems and shortcomings of Singh's study, as mentioned under point (3) above, can be overcome by using a large global dataset allowing a more systematic assessment of the comparability of the most common ANPP methods. This is the scope of the present study.

We aim to establish simple conversion formulae between the most common ANPP estimation methods for grass-dominated vegetation. Our study is based on data from 89 sites with more than 850 years of biomass data.

2. Materials and methods

2.1 Literature reviews

Two literature reviews were carried out for this study: (1) A review of the 165 studies citing Singh et al. (1975) to determine whether or not they made use of the presented ANPP conversions (see 1. Introduction). (2) We reviewed the 150 most recent studies presenting field measured ANPP data, and noted the ANPP estimation method(s) employed. We only selected papers from peer-reviewed journals, and excluded ANPP data which was derived from modeling or remote sensing indices. In detail, we searched the term 'ANPP' in the years 2012 and 2011 and selected the 150 most recent papers (written in English, French, German or Spanish). ANPP estimation methods were classified into twelve groups (see Table 1), generally based on the nomenclature of Scurlock et al. (2002) but slightly extended (see Table 1 and below). All literature reviews were carried out using Google Scholar in December 2012, as this source gives more complete results compared to other platforms (Beckmann and von Wehrden, 2012).

2.2 Dataset

Our ANPP dataset combines established datasets with data obtained from complementary literature reviews. It only comprises datasets which allow the calculation of at least two common ANPP estimation methods. All methods considered in this study are given and described in Table 1, their selection and nomenclature follows Scurlock et al. (2002).

One of the two main sources for ANPP data is the *Net Primary Production Dataset* distributed by the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC, http://daac.ornl.gov). The second major source is a self-assembled ANPP dataset comprising long-term monitoring data from arid and semi-arid ecosystems. The principal data search and acquisition methods are described in Ruppert et al. (2012), but the current dataset has been considerably updated and extended compared to that presented therein. Furthermore, suitable ANPP datasets which were found during the above described literature reviews (see 2.1) were added. Table S1 in the supplementary material presents a complete overview on sources and references for all 89 datasets included in analyses.

2.3 Data analysis

2.3.1 ANPP estimation methods

Estimating ANPP is a two-step procedure, starting with the measurement (or estimation) of biomass, followed by the computational processing of these measurements. Here we will focus on the latter aspect of calculation algorithms only, and will concentrate on those algorithms most commonly used in recent studies. Generally two groups of estimation methods can be distinguished: (1) *'Peak methods'*, using single biomass measurements at peak biomass conditions to estimate ANPP and (2) *'Incremental methods'*, which sum the incremental accumulation of biomass on a seasonal or annual basis.

The seven (to eight) most common methods – their calculation, inherent assumptions and possible pitfalls – have been comprehensively described by Scurlock et al. (2002). We generally followed their nomenclature but split Method 2 *'Peak standing crop'* into two sub-methods (Table 1). Method 2a is the original *Peak standing crop* method (as described in Scurlock et al., 2002), which uses the maximum amount of live plus recent (current year's) dead material as estimate of ANPP. We found several studies which also included previous year's dead material (and sometimes even non-standing, de-attached litter), and labeled this approach as Method 2b. We chose to distinguish between these sub-methods for two reasons: Firstly, Method 2b is of limited applicability only, since it can be biased by the previous year's production. Secondly, lumping both methods together would have introduced considerable variability into 'Peak standing crop' data.

Since only one site reported sufficient data to calculate ANPP via Method 7 (Sum of positive increments in live and dead biomass with an adjustment for decomposition), we excluded this method from our analyses.

2.3.2 Statistical analyses – Regressions and conversion formulae

Data exploration to avoid common statistical problems (e.g. with respect to outliers, normal distribution and homogeneity of variances) was performed visually as proposed by Zuur et al. (2010). Due to several cases of a violation of the homoscedasticity assumption in least squares regression, we used generalized least squares regression (GLS). By implementing flexible variance structures of the covariate, GLS allows to correct for heteroscedasticity (Zuur, 2009). For each conversion model we tested, five (generalized) least squares models were derived, reflecting different common variance structures of the covariate for ecological data (no variance structure, fixed variance structure, power of the covariate variance structure, exponential variance structure, and constant plus power of the variance structure, see Zuur, 2009). We used Akaike's information criterion (AIC) to select the best-fitting model and checked again for homoscedasticity.

For some method combinations we had indications that systematic differences between data from drylands (arid and semi-arid) and humid areas existed, based on either methodological issues or visual observation of the regressions. We thus used ANCOVAs to test the influence of climate regime on the respective regression models. For six method combinations we found a significant influence of the climate regime and therefore split the data accordingly to establish climate-specific conversion formula (see Table 2 and Figure 1).

Established conversion formulae were classified on the basis of their pseudo R^2 values into three groups (highly reliable, reliable, and unreliable), representing their reliability and usability as conversion models. Class borders were set at pseudo $R^2 \leq 0.5$ for unreliable, > 0.5 and < 0.7 for reliable, and ≥ 0.7 for highly reliable, respectively. Pseudo R^2 calculation was based on the generic definition of the coefficient of determination and was calculated as: *1 – residual sum of squares / total sum of squares.* If the final selected model was based on standard least squares regression, pseudo R^2 values were thus equivalent to standard R^2 values.

We also assessed the comparability of each method combination. Comparability between *Peak methods* (Method 1, 2a & 2b) was assumed to be moderate (labeled as "+ -" in Table 2): While all methods are based on single observations during peak biomass conditions, they refer to different estimates of biomass. Comparability between *Peak methods* and *Incremental methods* ranged from poor (- -) to moderate (+ -), depending on the type of biomass used for the estimation. If both methods were based on the same type of biomass (live biomass, live plus recent dead, etc.; e.g. Method 1 : Method 3) their comparability was rated as moderate; if not, comparability was rated as poor (e.g. Method 1 : Method 6). The comparability between *Incremental methods* ranged from moderate (+ -) to good (+ +). Comparability was rated as good if both methods were based on the same type of biomass (e.g. Method 3 : Method 4) and as moderate if not (e.g. Method 3 : Method 5). This assessment of the methodological and ecological comparability adds some information about the applicability of conversions, in addition to the statistical classification based on pseudo R^2 values.

All statistical calculations were performed in R, version 2.15.2 (R Development Core Team, 2012). The *rms* package (version 3.6-3) and the *nlme* package (version 3.1-105) were used to calculate and visualize GLS models.

Table 1.

Group / Method for ANPP estimation[a]	Description	%		
Method 1	Peak live biomass	12.7	Peak methods: 50.0%	
Method 2a[b]	Peak standing crop (live plus recent dead)	18.7		
Method 2b[b]	Peak standing crop (live plus recent and old dead)	18.7		
Method 3	Maximum minus minimum live biomass	1.3	Incremental methods: 15.3%	Incremental + Other incremental methods : 20.7%
Method 4	Sum of positive increments in live biomass	12.0		
Method 5	Sum of positive increments in live and recent dead (Smalley's Method)	1.3		
Method 6	Sum of positive increments in live and total dead (recent plus old dead)	0.0		
Method 7[c]	Sum of positive increments in live and dead biomass with an adjustment for decomposition	0.7		
Other ANPP methods	ANPP methods which could not be sorted into the above.	12.6		
	Other – incremental methods	(5.3)		
	Other – sum methods	(4.0)		
	Other – unspecified	(3.3)		
Assembled ANPP studies	Studies which assembled ANPP datasets from more than one source of ANPP data (supposedly) comprising more than one estimation method for ANPP.	5.3		
Misleading (or wrong)	Abbreviation ANPP was used in a misleading (or wrong) way. In most cases daily productivity data was presented.	4.0	Wrong or no info: 16.7%	
No information	No information on ANPP estimation methodology was given.	12.7		

[a] Nomenclature follows Scurlock et al. 2002.

[b] Differing from Scurlock et al. 2002 the 'peak standing crop' method was split into two subgroups.

[c] Note that we had to skip Method 7 from analyses due to insufficient data.

3. Results

3.1 Literature reviews

The most recent 150 publications presenting ANPP data showed that *Peak biomass* estimates (Methods 1, 2a & 2b) dominated with 50 % of all studies using them. *Incremental methods* (Methods 3-7) followed with 15.3 %. A smaller proportion of 12.7 % of studies used very specific ANPP estimation methods, which could not be assigned to one of the common methods, and therefore were allotted in 'Other ANPP methods'. Within this group, the largest share (representing 5.3% of all studies) were other, 'non-canonical', incremental methods, followed by methods calculating ANPP as the sum of several cuts throughout a season or year (4% of studies). Combining the canonical ANPP methods (Methods 3-7, 15.3 %) and these specific non-canonical methods (5.3 %), increased the total share of incremental methods to 20.7% over all studies.

In total 5.3% of all studies (8 studies of 150) presented *Assembled ANPP datasets* with more than one source of ANPP data. These studies often combined several methods in one dataset. Another 4% of all studies used the term ANPP in a misleading way. In most cases, authors presented aboveground net primary productivity, which is production per time (e.g. g m^{-2} d^{-1}). The remaining 12.7 % gave no information, on how ANPP was estimated.

The group of *Peak biomass* estimates was dominated by the two varieties of *Peak standing crop*, Method 2a and Method 2b, with 18.7 % each, as compared to *Peak live biomass* (Method 1) with 12.7 %. *Incremental methods* are dominated by Method 4 (Sum of positive increments in live biomass) with 12.0 %. All other methods were rarely used. Method 3 (Maximum minus minimum in live biomass) and Method 5 (Sum of positive increments in live and recent dead, aka Smalley's Method) have been used in 1.3 % of all cases each (2 in 150 each), Method 7 (Sum of positive increments in live and dead biomass with an adjustment for decomposition) were used in 0.7 % of all cases (1 in 150), and Method 6 (Sum of positive increments in live and total dead) was not used in recent publications.

Table 2. Overview on the established conversion formulae.

	Statistical reliability class	& comparability	Conversion formulae						Std. Err. slope	n	Pseudo R^2
Recommended	***Highly reliable***	**+ +**	*Method 3*	=	*0.89*	*x*	*Method 4*	*+ 6*	0.02	255	0.91
		+ +	*Method 5*	=	*0.9*	*x*	*Method 6*		0.04	38	0.78
		+ -	*Method 1*	=	*0.69*	*x*	*Method 2a*		0.02	227	0.82
		+ -	*Method 1*	=	*1.05*	*x*	*Method 3*	*+ 29*	0.02	384	0.92
		+ -	*Method 1*	=	*0.97*	*x*	*Method 4*	*+ 32*	0.02	679	0.89
		+ -	*Method 2a*	=	*0.56*	*x*	*Method 2b*	*+ 57*	0.06	29	0.71
		+ -	*Method 2a*	=	*0.73*	*x*	*Method 6*	*+ 92*	0.06	30	0.71
		+ -	*Method 2b*	=	*0.81*	*x*	*Method 6*	*+ 176*	0.10	18	0.80*
		+ -	*Method 3_{arid}*	=	*0.34*	*x*	*Method 6_{arid}*		0.03	29	0.73
		+ -	*Method 4_{arid}*	=	*0.39*	*x*	*Method 6_{arid}*	*+ 11*	0.03	29	0.71
		- -	*Method 1_{arid}*	=	*0.35*	*x*	*Method 6_{arid}*	*+ 50*	0.03	29	0.81*
	Reliable	**+ -**	*Method 3_{humid}*	=	*0.49*	*x*	*Method 5_{humid}*	*+ 85*	0.06	47	0.60
		+ -	*Method 3_{humid}*	=	*0.44*	*x*	*Method 6_{humid}*	*+ 103*	0.09	24	0.51*
		+ -	*Method 4_{arid}*	=	*0.53*	*x*	*Method 5_{arid}*	*+ 19*	0.05	39	0.65
		+ -	*Method 4_{humid}*	=	*0.64*	*x*	*Method 5_{humid}*		0.05	44	0.66
		+ -	*Method 4_{humid}*	=	*0.72*	*x*	*Method 6_{humid}*		0.07	24	0.62
Not recommended	***Unreliable***	**+ -**	*Method 2a*	=	*0.83*	*x*	*Method 5*	*+ 96*	0.06	70	0.60
		+ -	*Method 2b*	=	*0.81*	*x*	*Method 5*	*+ 188*	0.13	39	0.52*
		- -	*Method 2a*	=	*1.23*	*x*	*Method 3*	*+ 87*	0.08	79	0.67
		- -	*Method 2a*	=	*1.13*	*x*	*Method 4*	*+ 96*	0.08	79	0.63
		+ -	*Method 1*	=	*0.24*	*x*	*Method 2b*	*+ 96*	0.05	52	0.33*
		+ -	*Method 3_{arid}*	=	*0.41*	*x*	*Method 5_{arid}*	*+ 28*	0.05	39	0.50
		- -	*Method 1_{arid}*	=	*0.35*	*x*	*Method 5_{arid}*	*+ 82*	0.06	39	0.50*
		- -	*Method 1_{humid}*	=	*0.58*	*x*	*Method 5_{humid}*	*+ 94*	0.06	47	0.50
		- -	*Method 1_{humid}*	=	*0.69*	*x*	*Method 6_{humid}*	*+ 43*	0.04	24	0.31
		- -	*Method 2b*	=	*1.27*	*x*	*Method 3*	*+ 264*	0.28	47	0.31*
		- -	*Method 2b*	=	*1.25*	*x*	*Method 4*	*+ 245*	0.27	46	0.33*

All regression parameters were significant on $p \leq 0.001$ (slopes) or on $p \leq 0.05$ (intercepts). Pseudo R^2 values marked with an asterisk are standard R^2 values. Here model selection selected non-GLS models (= least squares regression). Statistical reliability class borders were set according to (pseudo) R^2 values: ≤ 0.5 poor, > 0.5 and < 0.7 moderate, ≥ 0.7 good. Classification of comparability classes (+ +, + -, and - -) is described in 2.3.2 Materials and Methods. For full model descriptions please refer to Table S3.

In the group of *Assembled ANPP studies* only three out of eight studies gave information on the respective ANPP estimation method for all datasets and addressed issues of comparability (Adler et al., 2011; Robinson et al., 2012; Ruppert et al., 2012). The other studies either mentioned the most commonly used methodologies only (Hsu et al., 2012; Yahdjian et al., 2011), simply stated that datasets were comparable (Hector et al., 2011), or did not comment on the nature of ANPP data at all (Eldridge et al., 2011; Evans et al., 2011). It should be mentioned that Eldridge et al. (2011) and Yahdjian et al. (2011) only presented ANPP response ratios (treated vs. non-treated), therefore differences in ANPP estimation algorithms should be of minor concern.

3.2 Established conversions between ANPP estimations methods

Using the statistical protocol described above (see 2.3.2 Materials and Methods), we analyzed all 21 possible (one-way) combinations between the seven considered ANPP estimation methods (Method 1, 2a, 2b, 3, 4, 5, and 6). Since six of these combinations exhibited systematic influences of climate (dryland vs. humid), we established a total of 27 conversion formulae (Table 2). Based on their coefficients of determination, eleven models were classified as rendering highly reliable conversions, nine as reliable and seven as unreliable. The assessment of method comparability generally mirrored the statistical classification. The class of *highly reliable models* included the only two method combinations which were rated as highly comparable (Method 3 : Method 4, and Method 5 : Method 6). Furthermore, this class only includes one method combination which has been rated as poorly comparable (Method 1_{arid} : Method 6_{arid}), the remaining eight combinations were rated as moderately comparable. The class of *reliable models* mostly contains combinations which were rated as moderately comparable, and only two poorly comparable combinations. The majority of poorly comparable method combinations are found in the *unreliable class*, which apart from these combinations only includes two moderately comparable combinations.

Table 2 presents all established conversions formulae in a standardized linear model format ($y = mx + b$). Furthermore, the standard error of the slope, the number of observations for the respective model, and the pseudo R^2 is given. Figure 1 gives a graphical representation of selected conversions. It presents nine method combinations and their eleven respective conversion models together with their confidence intervals. These method combinations represent the most frequently used ANPP methods according to our literature review (Methods 1, 2a, 2b and 4; see Table 1). In addition, we have included Method 5 as an example for an often recommended elaborate method (Singh et al., 1975, Scurlock et al., 2002). The selection in Figure 1 also gives examples for all statistical reliability classes: highly reliable (Figure 1A, B, D), reliable (Figure 1E, F, H, I), and unreliable (Figure 1C, G). An overview of all other established conversion formulae can be found in Figure S1 in the supplementary material.

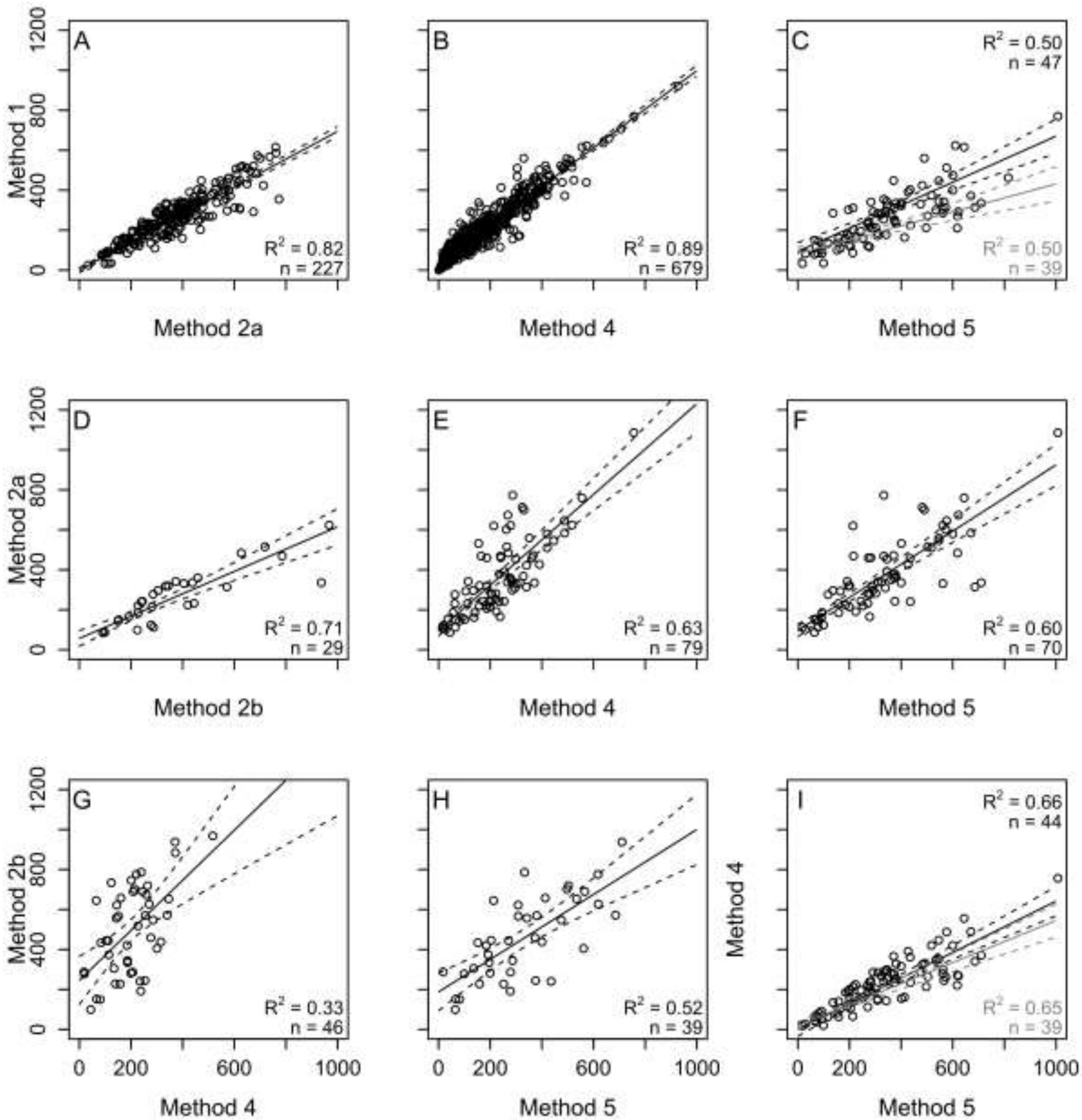

Figure 1. Selection of conversion models (GLS regressions) between common ANPP estimation methods together with corresponding number of observations (n) and (pseudo) R^2. Linear regressions are given as solid black lines. Where regressions were calculated separately for humid and dry sites (see 2.3.2 Material and Methods), black line represent the humid model. Solid grey lines represent the arid model, where applicable. Broken lines indicate the .95 confidence interval. Note: Selection of models comprises recommended and not recommended conversions models (see 2.3.2 Materials and Methods). Models in A, B, D, and I are recommended. See also Figure S1 for a complete graphical overview on all conversions models.

4. Discussion

The aim of this study was to establish conversions between the most common ANPP estimation methods, to improve comparability between ANPP estimates derived from different methods, and thus provide better access to the large body of published ANPP data. This was mainly motivated by the growing demand for large- or global-scale ANPP datasets which has evolved as a direct consequence of the climate and land-use change debate.

We were able to establish linear conversion formulae between the seven most commonly used ANPP estimation methods for grass-dominated biomes, and to assess their reliability and usability with statistical and methodological means.

4.1 Faster, simple methods are more often used than elaborate but labor-intense methods

The review on the use of ANPP in recent literature revealed that the simple and fast methods of the *Peak biomass* group were most frequently applied. Every second publication in our review used one of these methods. The frequency of use of the three sub-methods in this group was nearly identical. The more elaborate, but also more time- and labor-intense, *Incremental methods* were used less often. Only one in five publications used one of these methods; when only the canonical methods are considered, this frequency further drops to one in six to seven. While this general trend is not surprising and consistent with the dataset structure in Scurlock et al. (2002), it is surprising that recommendations to use the more elaborate algorithms, accounting for dynamics of live and dead plant matter (Method 5, 6 and 7), have not been adopted by the scientific community. Indeed, only 3 of 150 publications used one of these methods (Table 1). However, far more concerning is that 12.7 % of the studies did not provide information on which ANPP method was used.

Given this use frequency of common ANPP estimation algorithms, scientists who seek to compile large-scale ANPP datasets from various sources face the *'comparability dilemma'* described above (see 1. Introduction). To make matters worse, the rare data derived from elaborate and supposedly more accurate algorithms would be the first to be dropped for the sake of comparability.

4.2 Using recommended conversion formulae to overcome the 'comparability dilemma'

Our main impetus for the study was to overcome the above described 'comparability dilemma' by mitigating the trade-off between the demand for large datasets and data comparability. Motivated by the compilation of a global ANPP dataset for drylands (Ruppert et al., 2012, Ruppert et al., in prep), and inspired by Singh et al. (1975), we found linear conversion formulae to be a simple, versatile, and straight-forward approach to convert between different ANPP estimation algorithms.

Based on seasonal biomass dynamics from 89 sites from various grass-dominated biomes and climate regimes, we deduced conversion formulae for all method combinations representing the most commonly used ANPP estimation algorithms (Scurlock et al., 2002). Six out of all 21 method combinations showed a significant influence of climate regime (dry vs. humid), thus leading to a total of 27 conversions formulae (see 4.3 Influence of climate regime on conversions formulae and ANPP methods). Even though we were able to deduce statistically sound and significant regressions for all model combinations, not all conversions can be fully recommended.

Generally, all models which were rated as *highly reliable* in terms of statistical criteria can be recommended for use without exceptions. In contrast, formulae classified as *unreliable* cannot be recommended and should be avoided. Even though conversion models in the latter group are highly significant, the underlying data exhibit considerable variance, which is also reflected in the pseudo R^2 values. Therefore, products derived from these models would involve considerable uncertainty. The line separating recommendable and non-recommendable conversions runs through the group of statistically *reliable models*. Our decision to classify the conversions between Method 2a and Method 3, 4 and 5, as well as conversions between Method 2b and Method 5 as not recommended is based on the visual assessment of the respective scatterplots (Figure S1-4, and Figure 1E, F, H respectively). For all combinations, a high spread of relatively equally spaced datapoints can be observed. For most cases, the spread also shows a tendency to increase with higher ANPP values, indicating heteroscedasticity. Therefore, derived conversion products would largely suffer from uncertainty. However, these conversion formulae might still be applicable for ANPP data from less productive sites (e.g. from drylands) with respective input estimates up to circa 200 g m^{-2}. For this range in ANPP data, the spread in the data is rather small, particularly for the conversions between Method 2a and Method 3, 4 and 5.

4.3 Influence of climate regime on conversions formulae and ANPP methods

The six possible combinations between Methods 1, 3 and 4 on the one hand and Methods 5 and 6 on the other (and only these six) showed a significant influence of climate regime (arid vs. humid) and were split into climate-specific conversion formula (see Figure 1, S1 and Table 2).

Notably, in all six cases, the slope of the dry climate model is less steep as compared to the humid model. If we assume Methods 5 and 6 to be the best proxy to 'real' ANPP (as they are 'best practice' methods), Methods 1, 3 and 4 underestimate ANPP in drylands more strongly than in humid ecosystems.

We assume that this systematic error could be ecologically explained by the higher turn-over rate from live to senescent biomass in drylands due to increased tissue senescence rate in response to water stress (Coughenour and Chen, 1997). While Methods 5 and 6 are sensitive to changes in live, senescent and moribund material and thus account for all biomass turn-over processes, Methods 1, 3 and 4 only assess live biomass. Thus, the latter three methods have specific ways of neglecting turn-over processes. Method 1 registers only live biomass at peak conditions, neglecting all produced live biomass which already turned senescent before peak. Methods 3 and 4 miss all live biomass which has turned over between minimum and maximum live biomass, or between sampling intervals, respectively. Thus these methods are inherently prone to differences in turn-over rates between different climates or ecoregions.

4.4 Applicability and generality of the conversion formulae

Given the clear patterns in the conversion models (Fig. 1 & S1) and considering the large underlying dataset, we expect the conversion formulae to be generally valid. Furthermore, despite the importance of climate regime for some conversions, we found no evidence for systematic influences of other factors (e.g. biome or long-term management). The generality of conversions is also supported by a comparison to those presented in Singh et al. (1975). Although the selection of ANPP estimation methods differs between the two studies, a subset of six conversions can be compared. The conversions between Method 1 and Method 4 are discussed as an example.

Based on our data we established the conversion formula:

Method 1 = 0.97 x Method 4 + 32

(n = 679)

Singh and colleagues (1975) found a very similar conversion formula (the formula has been converted to fit our format, see fourth formula in Table IV, Singh et al., 1975):

Method 1 = 1.06 x Method 4
(n = 33)

The slightly higher slope in Singh's formula can be explained by the fact that all linear conversions were forced through the origin. An overview of the remarkable consistency between our results and those of Singh et al. (1975) and other published data (Linthurst and Reimold, 1978) is presented in the Supplementary Material (Table S2 and Figure S2).

Some authors have assumed that differences between ANPP methods might be site-specific (Linthurst and Reimold, 1978; Long et al., 1989; Scurlock et al., 2002). They based this assumption on their observation that ranking sites according to their production, using several ANPP estimation methods, yielded varying outcomes. Interpreted towards the use of the conversion models this means that the respective proportion of under- or overestimating ANPP by applying a respective conversion is site-specific. However, this source of uncertainty is a general feature of predictions based on regression models.

Our analysis clearly shows that there are strong systematic relationships between several ANPP estimation algorithms. This underlines the usability of our conversion models, especially those which have been labeled as recommended on the basis of statistical and methodological criteria.

4.5 Uncertainties in estimating ANPP

Lauenroth et al. (2006) raised the issue of uncertainty in estimating (A)NPP and hypothesized that estimation algorithms differ not only with respect to magnitude and accuracy (over- or underestimation) but also with respect to uncertainty. They analyzed the amount of uncertainty which is mathematically introduced in ANPP estimates based on different estimation algorithms, as compared to the uncertainty in the input data (biomass estimates). Considering their findings we can assume that all estimation methods which we used for conversions should exhibit very low levels of uncertainty (i.e. corresponding to the level found in the biomass input data or even less). *Peak methods* simply transmit the uncertainty of the single biomass measurements on which they are based to the ANPP estimate. Since biomass can be measured or estimated with low uncertainty, these ANPP algorithms will exhibit the same low uncertainty. *Incremental methods* (Methods 3 to 6) are based on sums or differences over sequential biomass data. For these methods, the amount of uncertainty is even lower as compared to the average uncertainty of the input data. Only algorithms which contain product terms (i.e. Method 7) might increase (or also decrease) uncertainty as compared to the input data (biomass), but these methods have not been used in this study (see 2.3.1 Material and Methods).

Hence, we assume that possible interference, caused by divergent uncertainty in the ANPP methods when converting between different methods, can be neglected for the conversion formulae presented here.

4.6 Conclusions and recommendations

The conversions formulae established within this study offer an easy and practical approach to recalculate and compare between ANPP estimates derived by divergent estimation algorithms. Authors who assemble large-scale ANPP datasets, or generally wish to combine ANPP data from various sources, can surely benefit from our approach, since it allows generating comparably scaled ANPP estimates based on published data.

Though we found statistically significant models for all combinations of the most common ANPP estimates in grass-dominated biomes, not all conversions can be recommended. The combined classification via statistical (pseudo R^2) and methodological attributes (comparability of ANPP estimation algorithms) offered a sound basis for recommendations (Table 2). Based on these statistical and methodological criteria, we rated 16 out of 27 conversions formulae as recommendable. The remaining 11 conversions are afflicted with high statistical or methodological uncertainty and should only be used with care, if at all.

In this context another important outcome was that we found an ecological explanation for the phenomenon that certain ANPP methods differ in their tendency to underestimate ANPP across ecoregions (Singh et al., 1975; Scurlock et al., 2002). We assume that this tendency is related to differences in plants' turn-over rates from live to senescent biomass as a function of climatic aridity. We conclude that those methods which are highly sensitive to this turn-over (Methods 1, 3, and 4) should not be used in warm xeric environments where biomass turn-over rates appear to be particularly high.

Note that this study does not advocate relying on conversion options only. Even the best conversion formula is still second best to a recalculation of ANPP which can be done by applying the desired algorithm to the original biomass data. Our approach offers a practical solution for those cases where this option is not possible or feasible, and is superior to previous attempts to solve the comparability dilemma (i.e. combining incomparably scaled ANPP data or skip available published data).

We are confident that a prudent use of conversion formulae, will promote the compilation of assembled ANPP datasets, and that our conversions will greatly facilitate the usability of published ANPP data in assembled regional or global studies.

Acknowledgements

We thank Marcelo Sternberg from Tel Aviv University and Zalmen Henkin from the Agriculture Research Organization for providing unpublished biomass data from the Karei Deshe Experimental Farm, Israel. We thank the editor and two anonymous referees for their insightful comments on the manuscript. Furthermore, we thank Roelof Oomen for fruitful discussions during the development of this study and Heidi Webber for improving the English. The research of Jan C. Ruppert was funded by the Foundation of German Business (Stiftung der Deutschen Wirtschaft, sdw) and by the German Science Foundation (Deutsche Forschungsgemeinschaft, DFG) through a grant to the Research Unit (FOR 1501). Research of Anja Linstädter was supported by the DFG through FOR 1501 and the German Federal Ministry of Education and Research (BMBF) via the WASCAL initiative (West African Science Service Center on Climate Change and Adapted Land Use). Data from Jornada Basin, Konza Prairie, and Sevilleta was provided by the Long Term Ecological Research (LTER) Program which is significantly funded by the U.S. National Science Foundation Long Term Ecological Research program (NSF Grant numbers BSR-8811906, DEB-0080529, DEB-0217774, DEB-0236154, DEB-0618210, DEB-0823341, DEB-0832652, DEB-0936498, DEB-9411976, DEB-9634135). Finally, we thank the many, many – often anonymous – researchers and research assistants who gathered the biomass data underlying our dataset.

References

Adler, P.B., Seabloom, E.W., Borer, E.T., Hillebrand, H., Hautier, Y., Hector, A., Harpole, W.S., O'Halloran, L.R., Grace, J.B., Anderson, T.M., Bakker, J.D., Biederman, L.A., Brown, C.S., Buckley, Y.M., Calabrese, L.B., Chu, C.J., Cleland, E.E., Collins, S.L., Cottingham, K.L., Crawley, M.J., Damschen, E.I., Davies, K.F., DeCrappeo, N.M., Fay, P.A., Firn, J., Frater, P., Gasarch, E.I., Gruner, D.S., Hagenah, N., Hille Ris Lambers, J., Humphries, H., Jin, V.L., Kay, A.D., Kirkman, K.P., Klein, J.A., Knops, J.M.H., La Pierre, K.J., Lambrinos, J.G., Li, W., MacDougall, A.S., McCulley, R.L., Melbourne, B.A., Mitchell, C.E., Moore, J.L., Morgan, J.W., Mortensen, B., Orrock, J.L., Prober, S.M., Pyke, D.A., Risch, A.C., Schuetz, M., Smith, M.D., Stevens, C.J., Sullivan, L.L., Wang, G., Wragg, P.D., Wright, J.P., Yang, L.H., 2011. Productivity is a poor predictor of plant species richness. Science 333, 1750-1753.

Beckmann, M., von Wehrden, H., 2012. Where you search is what you get: literature mining - Google Scholar versus Web of Science using a data set from a literature search in vegetation science. J Veg Sci 23, 1197-1199.

Blaisdell, J.P., 1958. Seasonal development and yield of native plants on the upper Snake River Plains and their relation to certain climatic factors. U.S. Dept. of Agriculture, Washington.

Coughenour, M.B., Chen, D.-X., 1997. Assessment of grassland ecosystem responses to atmospheric change using linked plant-soil process models. Ecol Appl 7, 802-827.

Dye, P.J., Spear, P.T., 1982. The Effects of Bush Clearing and Rainfall Variability on Grass Yield and Composition in Southwest Zimbabwe. Zimbabwe J Agr Res 20, 103-118.

Eldridge, D.J., Bowker, M.A., Maestre, F.T., Roger, E., Reynolds, J.F., Whitford, W.G., 2011. Impacts of shrub encroachment on ecosystem structure and functioning: towards a global synthesis. Ecol Lett 14, 709-722.

Evans, S.E., Burke, I.C., Lauenroth, W.K., 2011. Controls on soil organic carbon and nitrogen in Inner Mongolia, China: A cross-continental comparison of temperate grasslands. Global Biogeochem Cycles 25.

Hector, A., Bell, T., Hautier, Y., Isbell, F., Kery, M., Reich, P.B., van Ruijven, J., Schmid, B., 2011. BUGS in the Analysis of Biodiversity Experiments: Species Richness and Composition Are of Similar Importance for Grassland Productivity. Plos One 6.

Hsu, J.S., Powell, J., Adler, P.B., 2012. Sensitivity of mean annual primary production to precipitation. Global Change Biol 18, 2246-2255.

Lauenroth, W.K., Sala, O.E., 1992. Long-term forage production of North American shortgrass steppe. Ecol Appl 2, 397-403.

Lauenroth, W.K., Wade, A.A., Williamson, M.A., Ross, B.E., Kumar, S., Cariveau, D.P., 2006. Uncertainty in calculations of net primary production for grasslands. Ecosystems 9, 843-851.

Lauenroth, W.K., Whitman, W.C., 1977. Dynamics of Dry-Matter Production in a Mixed-Grass Prairie in Western North-Dakota. Oecologia 27, 339-351.

Linthurst, R.A., Reimold, R.J., 1978. Evaluation of Methods for Estimating the Net Aerial Primary Productivity of Estuarine Angiosperms. J Appl Ecol 15, 919-931.

Long, S.P., Moya, E.G., Imbamba, S.K., Kamnalrut, A., Piedade, M.T.F., Scurlock, J.M.O., Shen, Y.K., Hall, D.O., 1989. Primary Productivity of Natural Grass Ecosystems of the Tropics - a Reappraisal. Plant Soil 115, 155-166.

McNaughton, S.J., Milchunas, D.G., Frank, D.A., 1996. How can net primary productivity be measured in grazing ecosystems? Ecology 77, 974-977.

Milner, C., Hughes, R.E., 1968. Methods for the measurement of the primary production of grassland, IBP Handbook no. 6. Blackwell, Oxford, UK.

Ni, J., 2004. Estimating net primary productivity of grasslands from field biomass measurements in temperate northern China. Plant Ecol 174, 217-234.

R Development Core Team, 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Robinson, T.M.P., La Pierre, K.J., Vadeboncoeur, M.A., Byrne, K.M., Thomey, M.L., Colby, S.E., 2012. Seasonal, not annual precipitation drives community productivity across ecosystems. Oikos, 727-738.

Ruppert, J.C., Holm, A.M., Miehe, S., Muldavin, E., Snyman, H.A., Wesche, K., Linstädter, A., 2012. Meta-analysis of rain-use efficiency confirms indicative value for degradation and supports non-linear response along precipitation gradients in drylands. J Veg Sci 23, 1035-1050.

Sala, O.E., Austin, A.T., 2000. Methods of estimating aboveground net primary production, in: Sala, O.E., Jackson, R.B., Mooney, H.A., Howarth, R.H. (Eds.), Methods in Ecosystem Science. Springer, New York, pp. 31-43.

Scurlock, J.M.O., Johnson, K., Olson, R.J., 2002. Estimating net primary productivity from grassland biomass dynamics measurements. Global Change Biol 8, 736-753.

Scurlock, J.M.O., Olson, R.J., 2002. Terrestrial net primary productivity - A brief history and a new worldwide database. Environmental Reviews 10, 91-109.

Singh, J.S., Lauenroth, W.K., Heitschmidt, R.K., Dodd, J.L., 1983. Structural and Functional Attributes of the Vegetation of Northern Mixed Prairie of North-America. Bot Rev 49, 117-149.

Singh, J.S., Lauenroth, W.K., Steinhorst, R.K., 1975. Review and assessment of various techniques for estimating net aerial primary production in grasslands from harvest data. Bot Rev 41, 181-232.

Smoliak, S., 1986. Influence of Climatic Conditions on Production of Stipa-Bouteloua Prairie over a 50-Year Period. J Range Manage 39, 100-103.

Yahdjian, L., Gherardi, L., Sala, O.E., 2011. Nitrogen limitation in arid-subhumid ecosystems: A meta-analysis of fertilization studies. J Arid Environ 75, 675-680.

Zuur, A.F., 2009. Mixed effects models and extensions in ecology with R. Springer, New York, NY.

Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol Evol 1, 3-14.