Estimating biomass and phenology in African Rangelands using Earth Observation

Background: INES (Lars Eklundh and Jonas Ardö) recently got a contract for a project (RAMONA, staring in December 2021) with the European Space Agency (ESA). This project aims to develop and apply methodology to assess grazing resources in African Rangelands using (mainly) data from the Sentinels (Sentinel 1,2, and 3 in this case). The project includes assessment of rangeland type and extent, herbaceous biomass, biomass anomalies and rangeland phenology.

Thesis topics: As a party of this project we provide opportunity for 2-3 master thesis projects during 2022 (both spring an autumn) focusing on **1)** estimation of herbaceous rangeland biomass, **2)** assessment of rangeland phenology and **3)** quantification of biomass anomalies.

Below you will find descriptions of these three topics from the original project plan. Based on information below we will define suitable topics, subtopics, study regions, study periods etc. and develop testable hypotheses, in cooperation with interested students.

When: Spring semester 2022 (potentially autumn semester 2022 as well).

Contact: Jonas Ardö, jonas.ardo@nateko.lu.se, 072-2025028 or Lars Eklundh, Lars.Eklundh@nateko.lu.se for more information

1.1.3.3.2 Herbaceous rangeland biomass

Herbaceous rangeland biomass [t DM ha^{-1}] is the net accumulation of the photosynthetic gain of carbohydrates (gross primary productivity, GPP, [g C m^{-2} day^{-1}]), and losses through autotrophic respiration (Ra). Whereas GPP can be estimated by remote sensing in a fairly straight forward way, this is not the case with the respiratory losses. In limited geographical areas it may be possible to estimate grassland biomass productivity by linear models directly from GPP (Liu, Dahlgren et al. 2019), or by using simple scalars to estimate net primary productivity (NPP) from GPP as in the current Copernicus Global Land Service Dry Matter Productivity product where dry matter productivity is assumed to be static and calculated as 0.5 x gross dry matter production (Svinnen, Tote et al. 2020). However, in Africa, there is insufficient ground data to calibrate such a model. In this proposal, we will tackle the problem by combining remote sensing and ecosystem modelling. Using an ecosystem model to convert GPP to biomass provides better constraints to the relationships across this huge geographical area. The seasonal and spatial dynamic in the model is provided by GPP. GPP is estimated via a state-of-the-art implementation of the light use efficiency (LUE) methodology (Monteith 1972, Monteith 1977) based on time series of fAPAR processed from the fused Sentinel 2 and 3 data (**Error! Reference source not found.**). The general algorithm quantifies GPP [g C m^{-2}] as a product of the absorbed photosynthetic radiation (APAR=fAPAR x PARin, [MJ]) and the efficiency by which APAR is converted into carbohydrates, ε_{max} [g C MJ^{-1}] and environmental scalars that downregulates ε_{max} based on resource limitation or environmental constrains (Medlyn 1998):

$$GPP = \varepsilon_{max} \bullet fAPAR \bullet PAR_{in} \bullet scalars \qquad (1)$$

GPP [g C m^{-2} day^{-5}] will be calculated every five days and **NPP** [g C m^{-2} day^{-5}] will be calculated every five days as accumulated GPP – autotrophic respiration (Ra). **Biomass dry weight** [t DM ha^{-1}] = NPP/0.45 assuming a carbon content of 45% (Atjay, Ketner et al. 1979) will be produced every five days. ε_{max} denotes the maximum rate of conversion of light energy and varies with photosynthetic pathway (C_3/C_4) and among rangeland types and may vary spatially (Madani, Kimball et al. 2014). **Candidate** ε_{max} will be derived from; 1) current eddy covariance flux measurements from African rangelands sites, 2) historical flux measurements from the FLUXNET2015 data base (Pastorello, Trotta et al. 2020) and additional African sites (Sjöström 2012, Abdi 2017) as listed in section 1.1.3.6.1 and Tab. 1, 3) extracted from literature (Garbulsky, Penuelas et al. 2010, Garbulsky, Filella et al. 2014) and 4) extracted from existing LUE models (Running and Zhao 2015, Stocker, Wang et al. 2019).

fAPAR can be accurately estimated based on satellite data (Myneni and Williams 1994, Zhang, Xiao et al. 2005) and is a key variable for the methodology. **Candidate variables** include several vegetation indices (NDVI, EVI, kNDVI) and the S2ToolBox Level 2 fAPAR product. All candidates will be temporally gap-filled with the fused S1/S2/S3 data (section 1.1.3.1.1.2). TIMESAT-smoothed data from the S1/S2/S3 fusion provides regular input for the modelling. The NDVI has been used in many models to estimate fAPAR, and this relationship has been shown to be linear in the Sahel (Fensholt, Sandholt and Rasmussen 2004). However, across large areas the influence of several factors affects the relationship, particularly variations in soil background colour (Myneni and Williams 1994, Zhao et al. 2018). The relationship has even shown non-linear sensitivity to canopy architecture during vegetative stages (Gitelson, Peng and Huemmrich 2014). Several studies have found that the EVI has better relationship with fAPAR (Zhao et al. 2018, Ogutu and Dash 2013, Xiao et al. 2004b) than other vegetation indices. These good properties relate strongly to the broader dynamic range and lower sensitivity to soil background of EVI (Huete et al. 2002, Gao et al. 2000), but there is also increasing evidence that the efficiency of EVI is related to its response to

$fAPAR_{chlorophyll}$ while NDVI is more related to $fAPAR_{canopy}$ (Xiao et al. 2004b, Xiao et al. 2005, Zhang et al. 2005, Liu et al. 2017). These reasons help explaining why EVI has been successfully used for GPP estimation across a range of biomes (Xiao et al. 2004a, Sims et al. 2006, Schubert et al. 2010, Schubert et al. 2012, Cai et al. 2021), including grasslands (Noumonvi and Ferlan 2020), and has been applied by the applicants in African drylands (Sjöström et al. 2011, Ardö et al. 2018b, Abdi et al. 2019.).

PAR_{in}, the incoming photosynthetic radiation (Surface net solar radiation [J m^{-2}]) quantifies the energy available for conversion to biomass. PAR_{in} is normally derived as a fraction (0.48) of incoming shortwave radiation (Papaioannou, Papanikolaou et al. 1993) and depend on position, time, cloudiness and aerosols. **Candidate variables** include ERA5-Land reanalysis data set (Surface net solar radiation, hourly, 9x9 km resolution) (ECMWF 2021) and the Daily Downward Surface Shortwave Flux (daily, 4x4 km spatial resolution) provided by EUMETSAT and originating from the SEVIRI radiometer on the Meteosat Second Generation (MSG) platform (SAF 2012).

Scalars, quantifying environmental downregulation of the light use efficiency include, temperature and moisture availability. The scalars vary over space and time and within rangeland types and plant functional types. For down-regulation due to **temperature** we will mimic the MOD17 algorithm (Running and Zhao 2015) with further calibration from flux measurements. Candidate variables include ERA-5 Land, 2m temperature (hourly, 9x9 km resolution) (ECMWF 2021).

Water availability can be expressed in several ways. **Vapour pressure deficit** (VPD, [Pa]) is commonly used to assess moisture stress and has been applied successfully in African environments (Sjöström, Zhao et al. 2013) even if the advantage of using **soil moisture**, especially during drought conditions have been stressed (Stocker, Zscheischler et al. 2018, Stocker, Zscheischler et al. 2019). **Candidate variables:** VPD will be derived from temperature and relative humidity, available from ERA5-Land (hourly, 9x9 km resolution). Soil moisture available to plants can be obtained at 20 m spatial resolution through the modelling of soil evaporation and plant transpiration (evapotranspiration – ET) using optical and thermal data as successfully exemplified in ESA Sen-ET (https://www.esa-sen4et.org/) and ET4FAO (https://et4fao.dhigroup.com/#/) projects and as successfully applied for quantifying soil moisture stress on crops (Jaafar and Mourad 2021). Alternatively, soil moisture estimates are available from ERA-5-Land for four horizons (0-7 cm,7-28 cm, 28-100 cm and 100-200 cm) at 9 km resolution. Soil Water index global scale, 12.5 km resolution, based on the Metop ASCAT soil moisture observations (https://land.copernicus.eu/global/products/swi). Evaporative fraction (LE/(LE+H), where LE = latent heat and H = sensible heat, has been reported strongly correlated to GPP (Sjöström, Ardö et al. 2011*, Figure 1*) and, similarly to soil moisture, could be obtained through own modelling or derived from ERA5-Land.

NPP will be derived from GPP based on the carbon used efficiency (CUE, NPP/GPP) derived from a state-of-the-art dynamic vegetation model (LPJ-GUESS, (Sitch, Smith et al. 2003)) (*Figure 1)*. This provides a significant improvement to previous approaches, e.g. the current Copernicus Global Land Dry Matter Productivity, where dry matter production is assumed to be static and 0.5 x gross dry matter production. This is not always correct and CUE has been shown to vary substantially (Amthor, 2000). This approach will provide a temporally and spatially varying CUE (Ardö 2015, Ardö, Tagesson et al. 2018) per plant functional type. LPJ-GUESS, a dynamical global vegetation model partly developed in Lund, calculates *NPP* as 0.75 x (*GPP- maintenance respiration*) (Sitch, Smith et al. 2003). CUE will be derived from existing dynamic global vegetation model outputs originating from modelling exercises such as CMIP5, CMIP6 (Eyring, Bony et al. 2016) and available from the Earth System Grid Federation (ESFG), the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) and Multi-scale Synthesis and Terrestrial Model Intercomparison Project (Ms-TMIP).

Biomass [DM] is assumed to equal *NPP/carbon content* on monthly basis, where carbon content of vegetation is on average 0.45 (Martin et al. 2018). *A*dditional correction will be applied for shoot:root ratio and biomass turnover. The shoot:root ratio (Wilsey and Wayne Polley, 2006) quantify allocation to aboveground versus belowground biomass increment, and will be quantified as we aim to estimate aboveground biomass only. Biomass turnover is the NPP/mean assessed biomass (Long et al, 1989, Scurlock et al, 2002) i.e., the biomass increases not captured by standard field inventory methods which render a general underestimation of NPP in field-based methods. The shoot:root ratio and turnover will be estimated based on a combination of literature, trait data bases such as the TRY data base (Kattage et al 2020), available in-situ data, and modelling results.

Calibration of the LUE model to estimate herbaceous rangeland biomass will be performed using data collected at the test sites and additional available existing data. Candidate variables and their combinations will be evaluated at the test sites in order to find the most suitable and robust combination for continental upscaling of herbaceous rangeland biomass during phase 2. **Validation** to be carried out as outlined in section ***Error! Reference source not found.***

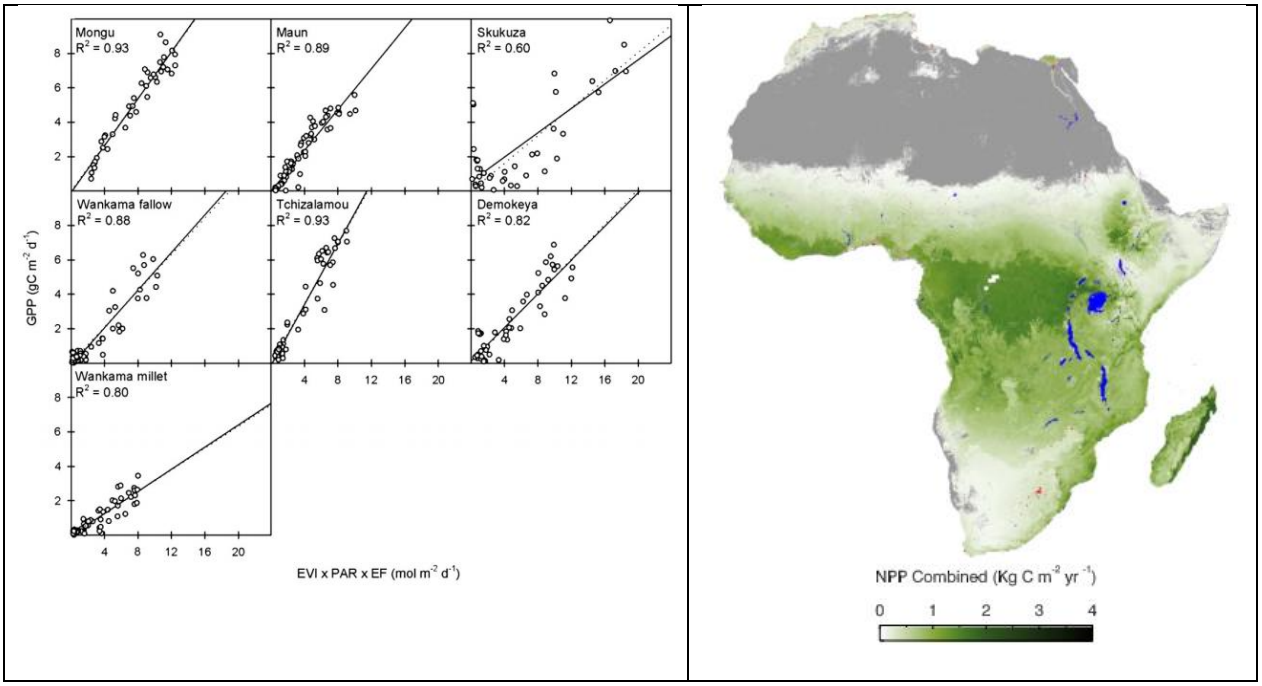

Figure 1 Steps towards estimating herbaceous rangeland biomass for Africa. To the left: The relationship of GPP vs EVI x PAR x EF for seven eddy covariance flux sites in Africa (Sjöström, Ardö et al 2011). To the right: Combining GPP estimated using EO data and the light use efficiency concept with the carbon use efficiency (NPP/GPP) from a dynamic vegetation model, incorporating eco-physiological skill and processed based representation of plant physiology and ecosystem biogeochemistry in the process chain towards biomass (Ardö, 2015).

1.1.3.4 Experimental products

1.1.3.4.1 Rangeland Herbaceous Biomass Anomalies

Biomass anomalies will be generated from standardized z-scores of fAPAR derived from medium-term Copernicus Sentinel-3/Proba-V at 300 m resolution as outlined by Meroni *et al* (2014). These data will supply a minimum of a 7-year baseline which is expected to be appropriate in the highly variable and rapidly changing drylands of Africa, at least for on-the-ground users more interested in comparing present conditions to the recent past than to conditions over past decades. If 7-year baselines turn out unsatisfactory it is possible to fall back on MODIS NPP data (MOD17A3hgfv061, 500 m resolution, going back 20 years). Harmonization of data to ensure consistency with the biomass originating from the fused S1/S2/S3 data will be performed in the pilot areas. The data will be seasonally and annually integrated (see section on Rangeland Phenology) and z-scores for each season/year computed per rangeland pixel. Harmonization between the Proba-V and Sentinel-3 time-series may be required to ensure consistency.

1.1.3.4.2 Rangeland Phenology

Phenology estimates serve as input to classify rangeland types but also as an experimental output product. For the first purpose, phenology data will be derived with TIMESAT during the smoothing of the fused S1/S2/S3 fAPAR as an intermediate product at five-day time step. For the second purpose, we will improve on standard phenology products based on vegetation indices and instead extract phenological metrics and parameters from the herbaceous biomass product. In contrast to traditional phenology products based on vegetation indices, biomass phenology has a clear ecological interpretation and is the key variable of interest for potential users interested in raising livestock or maintaining wildlife. Phenological parameters will be derived from smooth seasonal trajectories at five-day time step of grassland productivity described in section 1.1.3.3.2, and expressed in the same physical units [t DM ha^{-1}].

Extraction of seasonal phenology parameters

From the smooth seasonal trajectories several parameters are computed in TIMESAT. Potential users will be most interested in the dates of start of season and end of season and the seasonal integral, which summarises the total amount of herbaceous biomass produced over the season. Additional relevant parameters include length of the growing season, amplitude, slopes of greening and browning, and seasonal minimum and maximum (**Figure 2**).

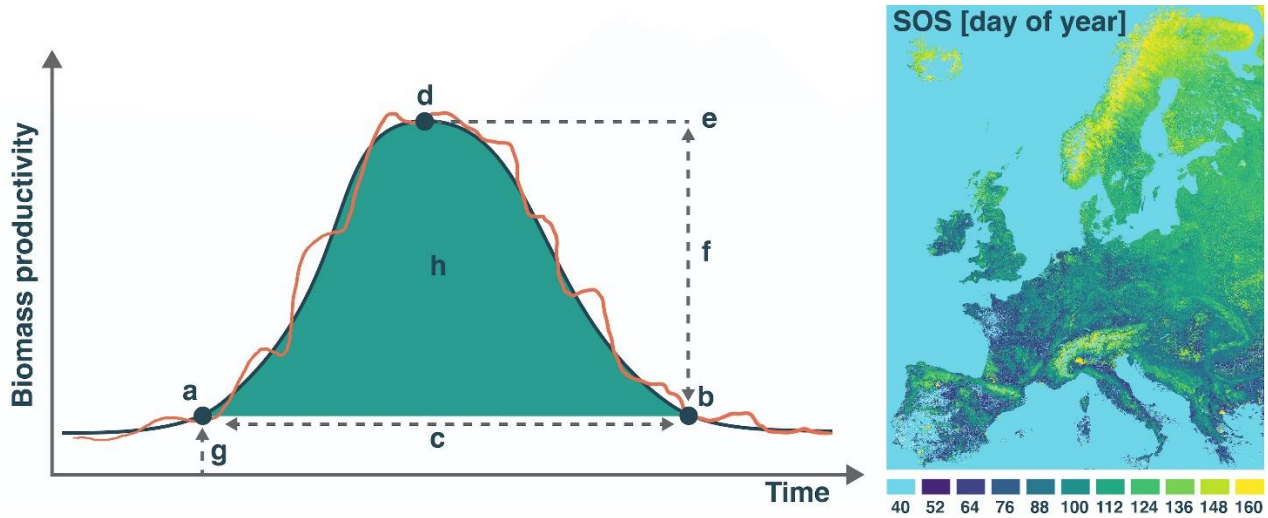

Figure 2: TIMESAT rangeland seasonality parameters (left panel): (a) start of season, (b) end of season, (c) length of season, (d) peak time, (e) peak value, (f) amplitude, (g) minimum, (h) seasonal productivity. Apart from these several points on the curve as well as greening and browning slopes can be extracted. The orange line is the input data and the smooth green line is the fitted TIMESAT function. Pan-European estimation of Start-of-Season (SOS) [day of year] using TIMESAT (Based on Tian, F., Cai, Z., Jin, H., Hufkens, …Ardö, J. & Eklundh, L. (2021). https://doi.org/10.1016/j.rse.2021.112456) (right panel).

Two growing seasons per year can be extracted, which is important as significant parts of East Africa have bimodal rainfall and growing seasons. Furthermore, seasons can traverse calendar years, which is crucial to accurately estimate phenological parameters in e.g. southern Africa. Based on calibration data from the pilot sites, including eddy covariance, rainfall data and field observations, we will analyse suitable amplitude thresholds (Tian et al. 2021) and derivatives (Buitenwerf et al. 2015) for determining start and end of seasons. In TIMESAT several quality metrics are calculated that are supplied with the final RAMONA products. These express the accuracy of the seasonal trajectories as a function of data density, and are propagated to the phenological parameters. A large number of phenological parameters can potentially be derived from the seasonal trajectories, and since many of these are intercorrelated we will analyse and select a subset of parameters that most accurately depict the seasonal and interannual variability in rangeland biomass productivity.

Relevance of the seasonal phenology parameters: Seasonal biomass productivity will provide gross forage production whereas the rate of increase/decrease and the seasonal amplitude provide indicators of forage quality (productivity and perennial to annual species). Shifts and variability in timing will provide data on vegetation response to climate drivers, potentially affecting productivity and, in the long term, land degradation.