# Final report

## Planet – Projecting Livestock, Agriculture, Nature, Ecology and Technologies, Australia Edition

Project code: P.PSH.1391

Prepared by: Peer Ederer
Global Food and Agribusiness Network

Date published: 14-Mar-2025

PUBLISHED BY
Meat & Livestock Australia Limited
PO Box 1961
NORTH SYDNEY NSW 2059

This is an MLA Donor Company funded project.

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication.

*Abstract*

This project was undertaken as part of an agreement between GOALSciences and MLA, with the overarching aim of determining global grazing capacity in both a worldwide context and on a country or region-specific level. By comparing estimated grazing capacity with livestock distribution data, it is possible to identify regions where additional livestock could be supported, as well as to determine areas that are overstocked and require supplementary feeding. This report summarizes the project's objectives, methods, and findings, indicating which goals were achieved and highlighting the limitations that must be addressed for more precise, localized applications.

# Executive summary

## Background

This project was undertaken as part of an agreement between GOALSciences and MLA, with the overarching aim of determining global grazing capacity in both a worldwide context and on a country or region-specific level. By comparing estimated grazing capacity with livestock distribution data, it is possible to identify regions where additional livestock could be supported, as well as to determine areas that are overstocked and require supplementary feeding. This report summarizes the project's objectives, methods, and findings, indicating which goals were achieved and highlighting the limitations that must be addressed for more precise, localized applications.

## Objectives

The principal objective was to quantify the world's overall grazing capacity, while simultaneously breaking this capacity down by individual countries or regions. A related goal was to compare these capacity estimates with current livestock numbers so that areas with the potential for expanded herds, as well as those necessitating supplementary feed, could be identified. A further objective was to understand what proportion of the global grazing capacity is already being utilized.

## Methodology

In order to accomplish these objectives, the project team began by consulting industry experts and reviewing existing scientific literature to identify the primary drivers of grazing capacity. While various parameters—such as breed, production system, land type, soil type, grass species, and nutrient density—were initially considered, global coverage for many of these factors was insufficient. Consequently, the analysis focused on parameters for which worldwide data are widely available. These included climate, rainfall, land-cover class, livestock numbers for key ruminant species (sheep, goats, cattle), dry matter productivity (DMP), and the Normalized Difference Vegetation Index (NDVI).

## Results/key findings

Overall, these findings highlight substantial surpluses of dry matter across all regions, suggesting opportunities for further livestock production or more efficient resource utilization. However, it also highlights the challenges in accurate grazing estimations as much of the dry matter availability may account for forests and tree or shrub species which are inedible yet cannot be accurately distinguished by the global monitoring tools, datasets and satellites such as described by the following limitations:

This study, while comprehensive in its approach, faced several limitations that may have influenced the outcomes and their applicability. Firstly, the availability and resolution of data were restricted in some cases. For instance, while NDVI and DMP data were available at relatively fine resolutions (300m x 300m), other critical datasets, such as rainfall and temperature, were only accessible at coarser resolutions (10km x 10km). This discrepancy in data granularity may have affected the precision of certain analyses, particularly in heterogeneous landscapes where finer-scale variability plays a significant role in livestock and vegetation interactions.

Furthermore, limited publicly available research and country-specific information on herd composition and utilisation ability posed challenges. Generalizations had to be made based on existing literature, which, while providing a foundation, may not fully capture regional or local nuances. This is particularly relevant in cases where livestock systems vary significantly across countries or agro-ecological zones, necessitating assumptions that may not align with specific conditions.

Another limitation lies in the reliance on generalized global datasets, such as SRTM for slope and elevation. While these datasets are valuable for broad-scale analyses, they may not account for microclimatic or localized topographic effects that influence vegetation and water availability. Similarly, the classification of land-cover types from global sources may not perfectly align with national classifications, leading to potential discrepancies in land-use assessments.

## Benefits to industry

By comparing estimated grazing capacity with livestock distribution data, it is possible to identify regions where additional livestock could be supported, as well as to determine areas that are overstocked and require supplementary feeding.

This project thus identifies a need for more local, ground-truth data collection on a country-basis in order to accurately estimate the availability of dry matter as well as the utilization of dry matter according to the unique needs of each country's production systems. Nevertheless, this project provides a stepping stone into gaining more insights on the role of ruminants on a global level and how they can be repurposed to covert non-arable land into productive pieces of land.

## Future research and recommendations

Future research should prioritize the integration of higher-resolution datasets where available, particularly for key variables such as rainfall, temperature, and land cover. Enhanced regional or country-specific datasets could significantly improve the accuracy and relevance of analyses. Additionally, targeted field validation of grazing capacity and herd composition assumptions would address gaps in empirical data, providing more precise insights into livestock-environment interactions.

Further exploration is also warranted into the impacts of climate variability on livestock distribution and vegetation dynamics. Developing predictive models that incorporate both biophysical and socio-economic factors could provide valuable tools for policymakers and farmers alike. Finally, fostering collaborations with local institutions and stakeholders would enable the inclusion of traditional knowledge and region-specific expertise, ensuring that future research outcomes are both scientifically robust and practically applicable.

While the utilisation factors applied in this research were designed to be broadly applicable on a large scale, they do not account for the species-specific variation of vegetation across countries and regions. For example, in areas with tree cover and livestock presence, the dry matter utilisation may be significantly overestimated. This is due to the general assumption that livestock browse on available vegetation in tree-covered areas, which may not always be accurate. In some cases, livestock may only graze and not utilise tree-related dry matter at all. Conversely, in high-quality grasslands under excellent management, dry matter utilisation could reach 60–80%, making the applied factor of 0.5 an underestimation. On the other hand, in poor-quality grasslands or areas with poor management, the assumed utilisation factor of 30% may be more realistic. These variations, which satellite and global datasets cannot capture, have led to potential overestimations or underestimations of dry matter productivity in certain contexts.

Future research will focus on refining these limitations by incorporating local utilisation data specific to land-cover classes where such information is available on a country level. More detailed investigations into country- and site-specific herd dynamics and livestock production systems will also be conducted as additional data becomes available. Efforts will be made to integrate detailed vegetation quality and species-specific information wherever possible, allowing for a more accurate representation of dry matter productivity and utilisation.

Furthermore, continuous validation against local data on grazing capacity will remain a priority. This process will involve collaboration with industry experts to regularly update and verify assumptions, ensuring the results are robust and reflective of actual field conditions. Future models will also aim to balance global-scale applicability with regional accuracy, integrating both high-resolution datasets and traditional knowledge to improve the precision and applicability of the findings.

# Table of contents

# 1. Background

## 1.1 Introduction

This project was undertaken as part of an agreement between GOALSciences and MLA, with the overarching aim of determining global grazing capacity in both a worldwide context and on a country or region-specific level. By comparing estimated grazing capacity with livestock distribution data, it is possible to identify regions where additional livestock could be supported, as well as to determine areas that are overstocked and require supplementary feeding. This report summarizes the project's objectives, methods, and findings, indicating which goals were achieved and highlighting the limitations that must be addressed for more precise, localized applications.

# 2. Objectives

The principal objective was to quantify the world's overall grazing capacity, while simultaneously breaking this capacity down by individual countries or regions. A related goal was to compare these capacity estimates with current livestock numbers so that areas with the potential for expanded herds, as well as those necessitating supplementary feed, could be identified. A further objective was to understand what proportion of the global grazing capacity is already being utilized.

# 3. Methodology

## 3.1 Industry & literature consultation

In order to accomplish these objectives, the project team began by consulting industry experts and reviewing existing scientific literature to identify the primary drivers of grazing capacity. While various parameters—such as breed, production system, land type, soil type, grass species, and nutrient density—were initially considered, global coverage for many of these factors was insufficient. Consequently, the analysis focused on parameters for which worldwide data are widely available. These included climate, rainfall, land-cover class, livestock numbers for key ruminant species (sheep, goats, cattle), dry matter productivity (DMP), and the Normalized Difference Vegetation Index (NDVI).

## 3.2 Verification of global datasets

After identifying suitable global parameters, the next step involved verifying these datasets using measurements from an experimental farm in Mpumalanga, South Africa. On-site information was compared with satellite-derived data to check for consistency in dry matter estimates and vegetation health indicators. This revealed the challenges of reconciling coarse-scale data—intended to cover vast geographic areas—with the realities of specific farm conditions. The large-scale data were found to be inadequate for precisely matching the test farm's true grazing capacity, which necessitated refining the methodology. It became clear that a broad-scale estimation focusing on general trends in dry matter availability and livestock suitability would be more realistic.

As a result, the approach was refined to focus on broader-scale approximations of dry matter availability and suitability for livestock. Rather than attempting to adapt dry matter utilization rates to diverse, location-specific production systems and breeds—an infeasible approach at a global scale— the research team developed a generalized model that estimates where livestock grazing is environmentally viable and how much dry matter is likely present. Under this refined strategy, the emphasis shifted from applying intricate, production-system-specific parameters to providing an estimation of how much dry matter is potentially available in each region and whether this area is suitable for grazing or not. Individual producers, or national agencies, can then adjust these estimates according to actual on-the-ground production systems.

## 3.3 Data Collection and preprocessing

The below data sources in Table 1 constituted the primary repositories for parameters such as DMP, NDVI, land-cover class, slope, elevation, water availability, and livestock distribution. Any additional information needed was obtained from country-specific national governmental websites, publicly accessible databases and repositories, or through searches on platforms such as ResearchGate and Google Scholar. For example, data related to herd composition or utilization capacity was gathered from these sources. However, scientific research at the continental, country-specific, or global level was limited, which necessitated the use of multiple generalizations based on the existing literature or industry consultation, as explained in subsequent sections.

**Table 1: Main data sources for the key variables used**

| Variable | Source | Description |
|---|---|---|
| Dry matter productivity | Copernicus land monitoring services | Dry Matter Productivity (kg/ha/day) tracks daily biomass accumulation by vegetation. Data was updated every 10 days at a 300 m resolution. |
| Normalised differentiation vegetation index | | NDVI measures vegetation health by comparing near-infrared reflectance with red light absorption. Higher values indicate healthier vegetation. Data was updated every 10 days at a 300 m resolution. |
| Climate | | Rainfall (mm) measures precipitation, and temperature (°C) refers to land surface temperature (LST). Both were updated every 10 days at a 10 × 10 km resolution. |
| Land cover | EarthMap | CCI/ESA (2022): Classifies surface types (e.g., forests, grasslands, croplands) |
| River network | | (WFF HydroSheds) (2022) |
| Total available water | | Estimates overall water resources (rainfall, surface water, groundwater) |
| Slope – SRTM | | Indicates terrain incline, affecting accessibility, erosion risk, and land-use decisions. |
| Elevation - SRTM | | Measures land height above sea level, influencing climate conditions and vegetation growth. |
| Livestock numbers | FAO | The GLW 4: Gridded Livestock Density (Global - 2020 - 10 km) converted to 1x1km |

**Predicted livestock**

Based on the assembled datasets, a prediction model was developed for livestock distribution. This model is intended to fill in gaps where FAO data are missing or incomplete, as well as to allow for more frequent updates should users need a more current picture of livestock numbers and their geographic spread. The prediction model provided a clearer understanding of which variables contribute to the distribution of livestock, such as climate, proximity to water, proximity to cropland and others. The below graphs in Figure 1A and 1B visualises the relationship of livestock distribution to the different variables by means of plotting the importance scores of each variable (Longitude, latitude, elevation, slope, water, land cover class, average DMP, maximum average temperature, minimum average temperature, average rainfall). An example of the values are visible in Supplementary Table S1.

The main variables identified were geographically related such as slope or elevation, longitude and latitude or temperatures, whereas the second most important variables were all captured by DMP (i.e. DMP and rainfall). It became evident that even if an area appears to have a large amount of dry matter, it may be unsuitable for livestock due to varying reasons such as extreme heat, lack of adequate water supplies, or unfavourable slopes. However, where livestock are present, DMP is a great contributor thereof, leading to the assumption that these areas have edible species. A detailed description of the methodology for the prediction model can be found in Supplementary file 2.

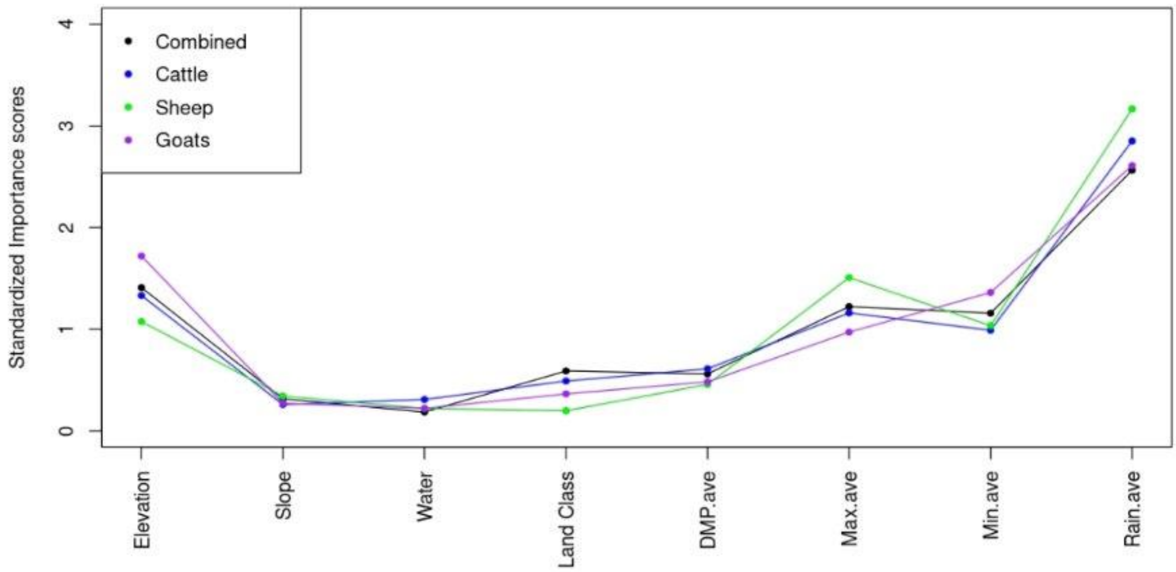



**Figure 1.A) South Africa variable importance for the prediction of livestock numbers**

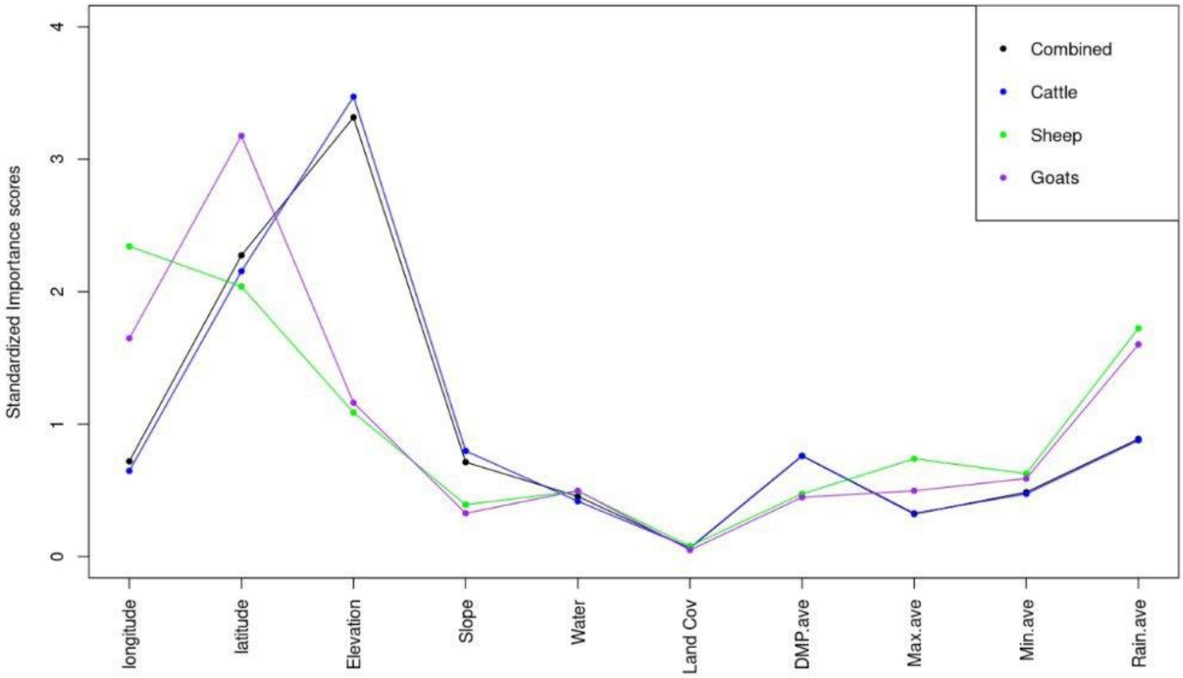



**Figure 1.B) Switzerland variable importance for the prediction of livestock numbers**

**Total dry matter and available utilizable dry matter**

Dry Matter Productivity (DMP, kg/ha/day) measurements were derived from satellite products that record vegetation growth at regular intervals (e.g., ten-day composites). The initial resolution of data is 300m, meaning the data reflects the average growth (kg/ha/day) for a 300m x 300m cell, which was then multiplied by 9 to obtain the total growth for the 300m x 300m cell. To covert this to a 1km x 1km cell, the total growth for all 300mx300m cells within a 1kmx1km cell were summed. Total DMP per 1km x 1km were then interpolated to obtain daily values and summed across 365 days to obtain the total amount of dry matter produced within a cell for a full year. Although actual grazing or defoliation at the time of satellite recording can reduce biomass during the time of measurement in ways that may not be captured by satellite sensors (thus potentially leading to underestimation of DMP), these losses are assumed to be small and offset by the interpolation process and by integrating data throughout the year, thus representing a full season of no growth, minimal growth and abundant growth.

Once annual dry matter (DM) values were established, a utilization factor was applied to account for realistic grazing scenarios in different land-cover classes. In the case where no livestock is present (according to FAO gridded livestock data and the prediction models), a 0-utilization factor was applied based on the assumption that climatic and environmental conditions or species of dry matter prevent livestock presence. In grassland regions, a factor of 0.5 was employed—a figure commonly cited in rangeland ecology literature as a sustainable utilization rate for extensive grazing systems. In shrub- and tree-dominated areas, utilization factors were low or zero, depending on the initial presence of livestock. Croplands were assumed to leave 30% residue after harvest (based on recommended residue) of which only half is recommended to be sustainably removed, leading to a total utilization factor of 15% or 0.15. Wetlands and bare areas were deemed unsuitable for ruminant grazing and were assigned a utilization factor of zero in most cases. However, water bodies often appear close to cropland or grassland thus although a water body is not grazable, a portion of the 1kmx1km cell may be grazable, thus indicting the presence of livestock. Hence, some unsuitable areas have livestock present and thus have higher than expected utilisation factor. The full list of land-cover classes and their respective utilisation rates can be found in Supplementary table S2. The following equation was thus used, reflecting the total available dry matter (to be used as grazing, fodder or hay) within a year:

Available utilizable DM = Total DMP (sum of 365 days) x utilization factor

## 3.4 Grazing capacity, requirement and supply

**Large Stock Unit conversion and Required Dry Matter**

Raw livestock numbers (head of animals per grid cell) were converted into Large Stock Units (LSUs) based on typical herd compositions and weight categories. One LSU was defined as a 450 kg cow; thus, bulls, cows, heifers, weaners, and steers were each assigned a weight-based factor to convert to LSUs, adjusted by the specific proportions observed in each region. For example, in South Africa, local herd structures may lead to a conversion factor of approximately 0.8 for beef cattle, based on the typical herd composition and weights of cattle. This is lower than the typical LSU, due to the majority of informal cattle herds and thus lower mature weights of animals. Australia, however, have larger bulls and cows at maturity, leading to a conversion factor of near 1. Similar calculations were done for dairy, sheep and goat, for which the herd composition for each has each own LSU conversion factor. These were implemented to the existing animal numbers, with cattle numbers being adjusted according to the percentage of dairy versus beef animals. Where no numbers were available for this split, 10% dairy and 90% beef were assumed. The supplementary tables (S3) reflect region-specific herd-data, a composite of multiple global and region-specific sources, which were used for calculations of each region's specie conversion factor. Where available, country-specific herd structures were applied. The following formula was thereafter applied to determine the large stock units (LSU) in each cell, based on the regional conversion factors. The LSU conversions are recorded in Table 2.

LSU = (Cattle numbers x beef % x beef conversion factor) + (Cattle numbers x dairy % x dairy conversion factor) + (Sheep numbers x sheep conversion factor) + (Goat numbers x goat conversion factor)

In keeping with standard intake assumptions, cattle were assumed to consume 3% of their body weight in dry matter per day, with calves at a lower rate (2%) owing to partial suckling. The grazing capacity in subsequent sections were calculated on the assumption that the full 3% is to be sourced from grazing. However, dairy cows, which require higher intakes and higher quality feed during lactation, were assumed to consume only 2% of body weight from grazing and an additional 2% from supplementary feeds. Due to the classification of land-cover classes, planted pastures may also be reflected as cropland, thus the 2% of supplementary feed may either be from high-quality planted pastures or cropland. Similarly, steers were assumed to consume 2% of body weight from grazing and 2% from supplementary feed, either in the form of cropland or planted pasture (both reflected as cropland in this case). Sheep and goat were calculated on a 2.5% of body weight intake, except for lambs and mature ewes which consumed an additional 0.5% from supplement due to either being finished for slaughter, or lactation and gestation demands. The full tables are visible in the supplementary material (Table S3). Similar to the LSU calculation, the intakes for each specie were weighted based on livestock numbers per country to present an intake requirement per LSU for each cell. The required intake from grazing was thus calculated as below for each grid cell (1km x 1km).

Step 1: Grazing intake requirement per LSU = (Weighted cattle numbers x beef % x beef intake factor) + (Weighted cattle numbers x dairy % x dairy intake factor) + (Weighted sheep numbers x sheep intake factor) + (Weighted goat numbers x goat intake factor)

Step 2: Required dry matter = LSU x grazing intake requirement per LSU x 365 days

**Table 2: Regional large stock unit conversion factors and intakes per large stock unit (LSU)**

| Region | Species | LSU conversion | Dry matter intake/LSU | Cropland intake LSU |
|---|---|---|---|---|
| **North America** | Beef | 1.020 | 12.379 | 1.029 |
| | Dairy | 0.952 | 10.314 | 5.174 |
| | Sheep | 0.130 | 10.904 | 1.641 |
| | Goat | 0.082 | 10.919 | 1.653 |
| **Europe** | Beef | 0.950 | 12.368 | 1.074 |
| | Dairy | 0.891 | 10.257 | 5.295 |
| | Sheep | 0.126 | 10.834 | 1.754 |
| | Goat | 0.083 | 10.890 | 1.723 |
| **Asia** | Beef | 0.783 | 12.191 | 1.213 |
| | Dairy | 0.728 | 10.453 | 4.718 |
| | Sheep | 0.111 | 10.799 | 1.830 |
| | Goat | 0.071 | 10.855 | 1.823 |
| **Africa** | Beef | 0.720 | 12.302 | 1.181 |
| | Dairy | 0.677 | 10.301 | 5.169 |
| | Sheep | 0.1.1 | 10.850 | 1.743 |
| | Goat | 0.066 | 10.877 | 1.744 |
| **Latin America** | Beef | 0.849 | 12.282 | 1.118 |
| | Dairy | 0.878 | 10.338 | 5.148 |
| | Sheep | 0.115 | 10.825 | 1.753 |
| | Goat | 0.082 | 10.867 | 1.751 |
| **Oceania** | Beef | 0.995 | 12.352 | 1.055 |
| | Dairy | 0.924 | 10.342 | 5.116 |
| | Sheep | 0.130 | 10.904 | 1.751 |
| | Goat | 0.088 | 10.924 | 1.753 |

**Grazing capacity, shortage and supplementation**

Based on the unique LSU's of each region and associated unique intakes per LSU, the below formula was applied to calculate the estimated grazing capacity (GC) of each cell and divided again by 100 to reflect GC per hectare. The grazing capacity is expressed as the number of LSU which can be sustained per cell or hectare, respectively.

GC = Available utilizable DM / 365 / Intake per LSU

The available utilizable dry matter was compared with the intake requirements of the estimated livestock numbers to establish whether there is a shortage or surplus in utilizable dry matter as follows, with a negative value indicating a shortage and a positive value indicating surplus.

Shortage or surplus = Available utilizable DM – Required DM

In areas where a shortage was reflected, the total supplementation required was calculated based on the shortage amount, plus the 2% body weight estimated to be already supplemented to steers and lactating dairy cows (as in Table 2). Thus, a supplement intake requirement was calculated based on the herd structures as portrayed earlier, totalled with the dry matter shortfall.

Step 1: Supplement intake requirement per LSU = (Weighted cattle numbers x beef % x beef intake factor) + (Weighted cattle numbers x dairy % x dairy intake factor) + (Weighted sheep numbers x sheep intake factor) + (Weighted goat numbers x goat intake factor)

Step 2: Required supplement = Dry matter shortage + ( LSU x supplement intake requirement per LSU x 365 days)

The available cropland was further adapted to determine the amount and percentage of current cropland production required to be fed by ruminants per country. Cropland production (kg dry matter) was multiplied by 0.7, based on the 30% residue left on land. No additional factor was applied for the portion of the harvested material which can be utilized since no data exists of how each cropland is utilised. In addition, harvest by-product can often be used by ruminants such as by making silage, incorporating soybean hulls into the diet, and others. Therefore, a zero wastage of harvested cropland was assumed.

Available cropland per country = Cropland production x 0.7

Percentage cropland required = (Supplement required / available cropland) * 100

# 4.0 Results

The results of the study provide a comprehensive overview of global dry matter (DM) availability, livestock requirements, and grazing capacity as summarised for six regions in Table 3. The results offer insights into regional grazing surpluses, deficits, and supplementary feeding needs.

**Table 3: Regional results for dry matter, grazing capacity, livestock numbers and supplementation required**

| Region | Total available DM | Total required DM | Surplus | Total predicted LSU | Total grazing capacity | Total available cropland | Total supplement required | Portion cropland required |
|---|---|---|---|---|---|---|---|---|
| | Mill. tons | Mill. tons | Mill. tons | Mill LSU | Mill LSU | Mill. tons | Mill. tons | % |
| Africa | 6332,81 | 1590,67 | 4742,14 | 386,03 | 1545,00 | 2689,11 | 792,63 | 29% |
| Asia | 5476,18 | 589,66 | 4886,53 | 416,58 | 4193,47 | 4176,30 | 182,14 | 4% |
| Europe | 6079,70 | 622,75 | 5456,95 | 227,03 | 2709,08 | 4126,63 | 140,49 | 3% |
| Latin America | 10033,04 | 1580,79 | 8452,25 | 469,61 | 3765,54 | 3456,07 | 220,18 | 6% |
| North America | 4003,86 | 674,28 | 3329,58 | 159,91 | 1095,19 | 2107,74 | 136,10 | 6% |
| Oceania | 1966,11 | 196,72 | 1769,39 | 79,17 | 1047,74 | 460,46 | 63,38 | 14% |

*DM = Dry Matter; LSU = Large stock unit

In Africa, the total available dry matter was estimated at 6.33 × $10^{12}$ kg, with a livestock DM requirement of 1.59 × $10^{12}$ kg. This leaves a surplus of 4.74 × $10^{12}$ kg, indicating significant underutilization of grazing resources. Africa supports an estimated 386 million large stock units (LSU), yet its grazing capacity is approximately four times higher at 1.54 billion LSU. This region has a total of 2.69 × $10^{12}$ kg of available cropland DM, and only 29.48% of this cropland would be required to meet livestock supplementation needs, suggesting an opportunity for optimized land use.

Asia demonstrates a similarly high surplus of 4.89 × $10^{12}$ kg from a total DM availability of 5.48 × $10^{12}$ kg, with a relatively low livestock requirement of 5.90 × $10^{11}$ kg. Despite supporting 416.6 million LSU, the continent's total grazing capacity is estimated at 4.19 billion LSU, underscoring an ample supply of grazing resources. With 4.18 × $10^{12}$ kg of cropland DM available, only 4.36% of this would be required for supplementary feeding, reflecting Asia's strong self-sufficiency in grazing resources.

Europe exhibits one of the highest DM surpluses at 5.46 × $10^{12}$ kg, with a total DM availability of 6.08 × $10^{12}$ kg and a livestock requirement of 6.23 × $10^{11}$ kg. The continent supports an estimated 227 million LSU, while its grazing capacity is significantly higher at 2.71 billion LSU. With 4.13 × $10^{12}$ kg of cropland DM available, only 3.40% of this resource is necessary to meet supplementary feeding requirements, making Europe the most efficient region in terms of land and feed resource use.

Latin America reported the highest DM availability at 1.00 × $10^{13}$ kg, with a surplus of 8.45 × $10^{12}$ kg after accounting for the 1.58 × $10^{12}$ kg required for livestock. The region supports 469.6 million LSU, with a grazing capacity of 3.77 billion LSU. Supplementary feeding would require just 6.37% of its total cropland DM availability of 3.46 × $10^{12}$ kg, illustrating a balance between livestock density and grazing resources.

North America, with a total DM availability of 4.00 × $10^{12}$ kg, exhibits a surplus of 3.33 × $10^{12}$ kg after accounting for a livestock DM requirement of 6.74 × $10^{11}$ kg. The region supports 159.9 million LSU, with a grazing capacity of 1.10 billion LSU. Supplementation needs would require 6.46% of its total cropland DM of 2.11 × $10^{12}$ kg, signifying a manageable demand for supplementary feeding relative to grazing capacity.

In Oceania, the total DM availability was estimated at 1.97 × $10^{12}$ kg, with a livestock requirement of 1.97 × $10^{11}$ kg. The surplus of 1.77 × $10^{12}$ kg reflects a strong potential for additional livestock support. The region currently supports 79.2 million LSU but has a grazing capacity of 1.05 billion LSU. However, Oceania exhibits the highest proportion of cropland DM required for supplementary feeding at 13.76%, indicating the need for careful management of cropland resources to sustain livestock populations.

# 5.0 Conclusion

Overall, these findings highlight substantial surpluses of dry matter across all regions, suggesting opportunities for further livestock production or more efficient resource utilization. However, it also highlights the challenges in accurate grazing estimations as much of the dry matter availability may account for forests and tree or shrub species which are inedible yet cannot be accurately distinguished by the global monitoring tools, datasets and satellites such as described by the following limitations:

This study, while comprehensive in its approach, faced several limitations that may have influenced the outcomes and their applicability. Firstly, the availability and resolution of data were restricted in some cases. For instance, while NDVI and DMP data were available at relatively fine resolutions (300m x 300m), other critical datasets, such as rainfall and temperature, were only accessible at coarser resolutions (10km x 10km). This discrepancy in data granularity may have affected the precision of certain analyses, particularly in heterogeneous landscapes where finer-scale variability plays a significant role in livestock and vegetation interactions.

Furthermore, limited publicly available research and country-specific information on herd composition and utilisation ability posed challenges. Generalizations had to be made based on existing literature, which, while providing a foundation, may not fully capture regional or local nuances. This is particularly relevant in cases where livestock systems vary significantly across countries or agro-ecological zones, necessitating assumptions that may not align with specific conditions.

Another limitation lies in the reliance on generalized global datasets, such as SRTM for slope and elevation. While these datasets are valuable for broad-scale analyses, they may not account for microclimatic or localized topographic effects that influence vegetation and water availability. Similarly, the classification of land-cover types from global sources may not perfectly align with national classifications, leading to potential discrepancies in land-use assessments.

# 6.0 Future research and recommendations

Future research should prioritize the integration of higher-resolution datasets where available, particularly for key variables such as rainfall, temperature, and land cover. Enhanced regional or country-specific datasets could significantly improve the accuracy and relevance of analyses. Additionally, targeted field validation of grazing capacity and herd composition assumptions would address gaps in empirical data, providing more precise insights into livestock-environment interactions.

Further exploration is also warranted into the impacts of climate variability on livestock distribution and vegetation dynamics. Developing predictive models that incorporate both biophysical and socio-economic factors could provide valuable tools for policymakers and farmers alike. Finally, fostering collaborations with local institutions and stakeholders would enable the inclusion of traditional knowledge and region-specific expertise, ensuring that future research outcomes are both scientifically robust and practically applicable.

While the utilisation factors applied in this research were designed to be broadly applicable on a large scale, they do not account for the species-specific variation of vegetation across countries and regions. For example, in areas with tree cover and livestock presence, the dry matter utilisation may be significantly overestimated. This is due to the general assumption that livestock browse on available vegetation in tree-covered areas, which may not always be accurate. In some cases, livestock may only graze and not utilise tree-related dry matter at all. Conversely, in high-quality grasslands under excellent management, dry matter utilisation could reach 60–80%, making the applied factor of 0.5 an underestimation. On the other hand, in poor-quality grasslands or areas with poor management, the assumed utilisation factor of 30% may be more realistic. These variations, which satellite and global datasets cannot capture, have led to potential overestimations or underestimations of dry matter productivity in certain contexts.

Future research will focus on refining these limitations by incorporating local utilisation data specific to land-cover classes where such information is available on a country level. More detailed investigations into country- and site-specific herd dynamics and livestock production systems will also be conducted as additional data becomes available. Efforts will be made to integrate detailed vegetation quality and species-specific information wherever possible, allowing for a more accurate representation of dry matter productivity and utilisation.

Furthermore, continuous validation against local data on grazing capacity will remain a priority. This process will involve collaboration with industry experts to regularly update and verify assumptions, ensuring the results are robust and reflective of actual field conditions. Future models will also aim to balance global-scale applicability with regional accuracy, integrating both high-resolution datasets and traditional knowledge to improve the precision and applicability of the findings.

This project thus identifies a need for more local, ground-truth data collection on a country-basis in order to accurately estimate the availability of dry matter as well as the utilization of dry matter according to the unique needs of each country’s production systems. Nevertheless, this project provides a stepping stone into gaining more insights on the role of ruminants on a global level and how they can be repurposed to covert non-arable land into productive pieces of land.

# 7.0 Appendix

## 7.1 Animal Production Module

**Data report:P.PSH.1391**

**Animal Production Module**

---

Table of contents

2. Animal Production Module

We will briefly discuss the various resources used in constructing the Animal Production Module, which includes the herd structure model and the associated nutrient requirements for pigs, chickens, and cattle, as well as the Animal Production System Evaluator (APSE) tool.

**Background**

The objective of the Animal Production Module is to display the movement of animals over the course of a year using a Sankey diagram, providing detailed information on animal sub-classes, yearly movements, losses, slaughter, and live animal trade across different production systems. Additionally, providing the associated nutrient requirements for the animals during the same period. Refer to section 2 for illustration examples.

**Resources**

**National statistics**

The statistics department of the Food and Agricultural Organization of the United Nations (FAO) forms the basis for national country numbers used in the various animal models.

- https://www.fao.org/faostat/en/#data

Data is sourced from various domains within the FAO statistics database. From the production domain, data is retrieved on animal inventory numbers, slaughter numbers, and meat, milk, and egg production over the course of a year. From the trade domain, data on live animal export and import numbers is utilized.

- https://www.fao.org/faostat/en/#data/QCL
- https://www.fao.org/faostat/en/#data/TCL

**Production systems**

The data for pig and chicken livestock production systems is derived from the Global Livestock Production Systems (GLPS) mapping done by FAO for base year 2010. Pigs are distributed into four production systems namely, industrial, intermediate, backyard and African backyard. While chickens are distributed into industrial and backyard systems. Then the industrial system for chickens is further divided into broiler and layer systems using data from the Climate change: Agri-food system emissions domain.

- https://www.fao.org/livestock-systems/production-systems/en/
- Gilbert, Marius; Conchedda, Giulia; Van Boeckel, Thomas P.; Cinardi, Giuseppina; Linard, Catherine; Nicolas, Gaëlle; Thanapongtharm, Weerapong; D'Aietti, Laura; Wint, G. R. William; Newman, Scott H.; Robinson, Timothy P., 2018, "Global distribution of chickens and pigs raised in extensive, semi-intensive and intensive systems in 2010 (5 minutes of arc)", https://doi.org/10.7910/DVN/A7GQXG , Harvard Dataverse, V2

- Gilbert M, Conchedda G, Van Boeckel TP, Cinardi G, Linard C, Nicolas G, et al. (2015) Income Disparities and the Global Distribution of Intensively Farmed Chicken and Pigs. PLoS ONE 10(7): e0133381. doi:10.1371/journal.pone.0133381, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0133381
- https://www.fao.org/faostat/en/#data/GLE


Defining cattle production systems on a global scale is challenging due to variations in research classifications. We have chosen to classify cattle production systems on three levels, namely dairy, beef and feedlot, based on country-specific research where available, and applying known profiles to countries with no available data.


- Robinson, Timothy P.; Thornton, Philip; Franceschini, Gianluca; Kruska, Russ; Chiozza, Federica; Notenbaert, An; Cecchi, Giuliano; Herrero, Mario; Epprecht, Michael; Fritz, Steffen; You, Liang; Conchedda, Giulia; See, Linda, 2018, "Global distribution of ruminant livestock production systems V5 (5 minutes of arc)", https://doi.org/10.7910/DVN/WPDSZE , Harvard Dataverse, V1
- Robinson, T.P., Thornton P.K., Franceschini, G., Kruska, R.L., Chiozza, F., Notenbaert, A., Cecchi, G., Herrero, M., Epprecht, M., Fritz, S., You, L., Conchedda, G. & See, L. 2011. Global livestock production systems. Rome, Food and Agriculture Organization of the United Nations (FAO) and International Livestock Research Institute (ILRI), 152 pp. isbn: 978-92-5-107033-8, https://www.fao.org/4/i2414e/i2414e.pdf
- https://www.fao.org/gleam/en/

**Production systems parameters**

The various production system parameters, such as fertility, piglets born, eggs hatched, culling rate, mortality, and others, were inferred from supply chain studies conducted by the FAO, the GLEAM model, and breeder guides.


- FAO. 2022a. Global Livestock Environmental Assessment Model, Model Description, Version 3.0. https://www.fao.org/fileadmin/user_upload/gleam/docs/GLEAM_3.0_Model_description.pdf
- FAO. 2022b. GLEAM 3 Dashboard. In: Shiny Apps. Cited [accessed 2023, 2024]. https://foodandagricultureorganization.shinyapps.io/GLEAMV3_Public/
- MacLeod, M., Gerber, P., Mottet, A., Tempio, G., Falcucci, A., Opio, C., Vellinga, T., Henderson, B. & Steinfeld, H. 2013. Greenhouse gas emissions from pig and chicken supply chains – A global life cycle assessment. Food and Agriculture Organization of the United Nations (FAO), Rome. https://www.fao.org/4/i3460e/i3460e.pdf
- Opio, C., Gerber, P., Mottet, A., Falcucci, A., Tempio, G., MacLeod, M., Vellinga, T., Henderson, B. & Steinfeld, H. 2013. Greenhouse gas emissions from ruminant supply chains – A global life cycle assessment. Food and Agriculture Organization of the United Nations (FAO), Rome. https://www.fao.org/4/i3461e/i3461e.pdf
- https://www.picrsa.co.za/manuals/
- Boehringer Ingelheim Animal Health GmbH. The Real Pig Handbook. 4th ed., SA, 2018. https://www.picrsa.co.za/wp-content/uploads/2019/09/PIC451-Real-Pig-Handbook-Web.pdf
- https://aviagen.com/tmea/tech-center/
- https://aviagen.com/assets/Tech_Center/Ross_Broiler/Ross-BroilerHandbook2018-EN.pdf
- https://aviagen.com/assets/Tech_Center/Ross_PS/Aviagen_Ross_PS_Handbook_2023_Interactive_EN.pdf
- https://aviagen.com/assets/Tech_Center/Ross_Broiler/RossxRoss308-BroilerPerformanceObjectives2022-EN.pdf
- https://layinghens.hendrix-genetics.com/en/technical-support/management/
- https://layinghens.hendrix-genetics.com/documents/980/Management_guide_commercial_cage_English_vs_L0260-6_.pdf

**Nutrient requirements**

The nutrient requirements for animals are derived from the Nutrient Requirements Collection (NRC) published by the National Academics of Science, Engineering and Medicine for pigs and cattle. While Aviagen and Hendrix-genetics guides are used for the chickens.


- National Research Council. 2012. Nutrient Requirements of Swine: Eleventh Revised Edition. Washington, DC: The National Academies Press. https://doi.org/10.17226/13298.
- National Academies of Sciences, Engineering, and Medicine. 2021. Nutrient Requirements of Dairy Cattle: Eighth Revised Edition. Washington, DC: The National Academies Press. https://doi.org/10.17226/25806
- National Research Council. 2001. Nutrient Requirements of Dairy Cattle: Seventh Revised Edition, 2001. Washington, DC: The National Academies Press. https://doi.org/10.17226/9825
- National Academies of Sciences, Engineering, and Medicine. 2016. Nutrient Requirements of Beef Cattle: Eighth Revised Edition. Washington, DC: The National Academies Press. https://doi.org/10.17226/19014.
- National Research Council. 2000. Nutrient Requirements of Beef Cattle: Seventh Revised Edition: Update 2000. Washington, DC: The National Academies Press. https://doi.org/10.17226/9791.
- https://aviagen.com/assets/Tech_Center/Ross_Broiler/Ross-BroilerNutritionSpecifications2022-EN.pdf
- https://aviagen.com/assets/Tech_Center/Ross_PS/Ross308-ParentStock-NutritionSpecifications-2021-EN.pdf
- https://aviagen.com/assets/Tech_Center/Ross_PS/Ross308-ParentStock-PerformanceObjectives-2021-EN.pdf
- https://layinghens.hendrix-genetics.com/en/technical-support/nutrition/
- https://layinghens.hendrix-genetics.com/documents/883/Nutrition_Guide_English_vs4.pdf

**Animal Production System Evaluator (APSE) tool**

The Animal Production System Evaluator (APSE) tool is designed to evaluate and compare livestock production systems, illustrating their contribution to national protein food baskets while considering environmental impact indicators. Drawing on data from FAO, GLEAM, and PLANET Food System Explorer, the APSE tool aims to bring a variety of data sources (production systems, land use, greenhouse-gas emissions, and protein production and consumption) together to offers a holistic assessment of livestock production systems. More information on the APSE tool can be found in an article to be published in the first edition of Animal Frontiers in 2025, titled “Shifting the Focus from Animal Species to Livestock Production Systems: An Interactive Tool for Evaluating Food Contributions Relative to Environmental Impacts.” This peer-reviewed article has been approved for publication.

**Conclusion**

Attachment 1: “GOAL Sciences _ Pig Herd Structure”, Attachment 2: “GOAL Sciences _ Cattle Herd Structure” and Attachment 3: “GOAL Sciences _ Chicken Flock Structure” serves as an example of how information is combined to achieve the outcomes of the Animal Production Module.

Each of the module can be accessed on the GOALSciences platform:

- https://goalsciences.org/food-system-explorer/herd-structure
- https://goalsciences.org/planet-food-system-explorer/animal-nutrition-requirements
- https://goalsciences.org/food-system-explorer/animal-production-system-evaluator

## 3. Annex

### Animal Herd Figures

### Pig module illustrations:

Figure 2.1 to 2.3 illustrates Sankey diagram of pig herds in Australia.

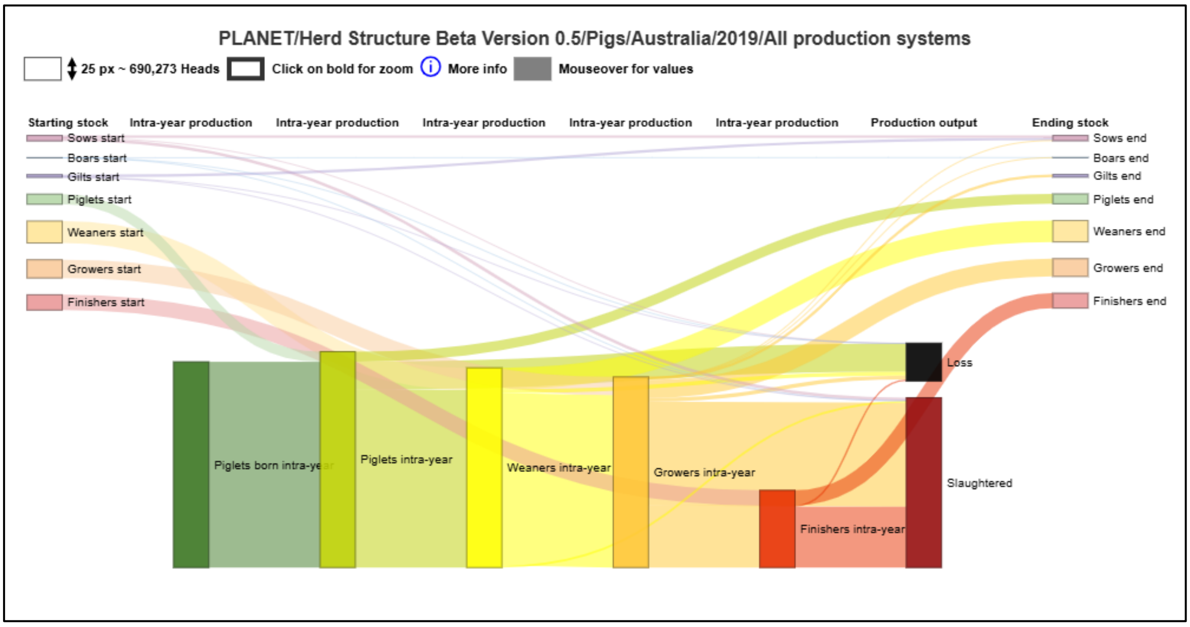



**Figure 2.1: Sankey illustration of pig herd flow in a year period: All production systems**

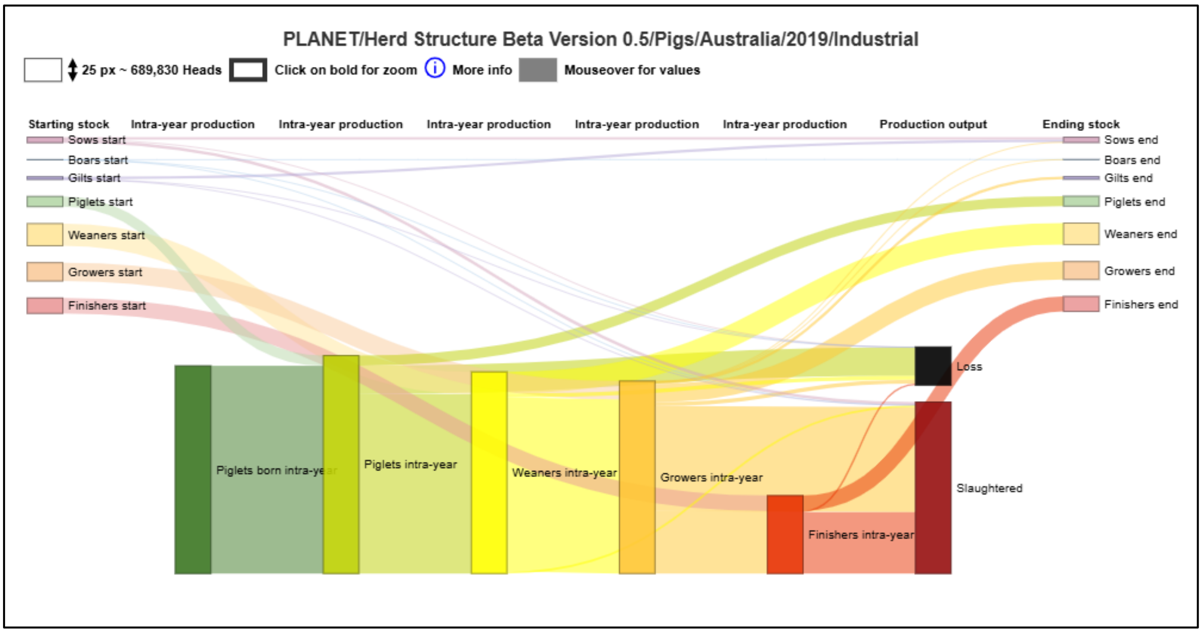



**Figure 2.2: Sankey illustration of pig herd flow in a year period: Industrial system**

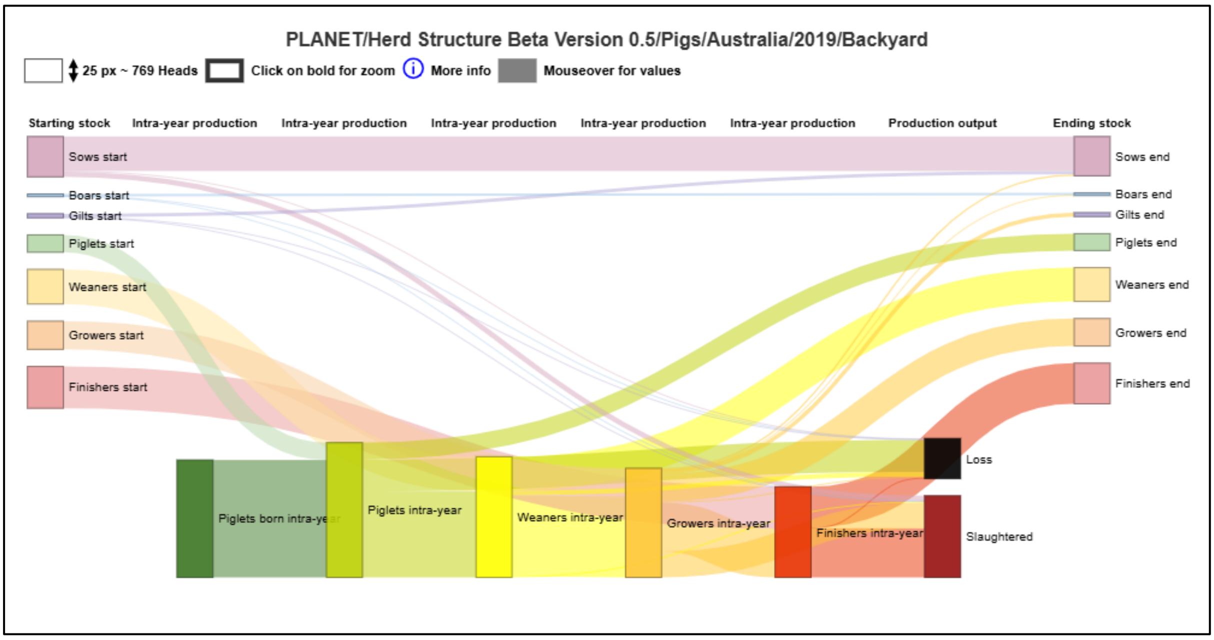



**Figure 2.3: Sankey illustration of pig herd flow in a year period: Backyard**

**Chicken module illustration:**

Figure 2.4 to 2.6 illustrates Sankey diagram of chicken flock in Australia.

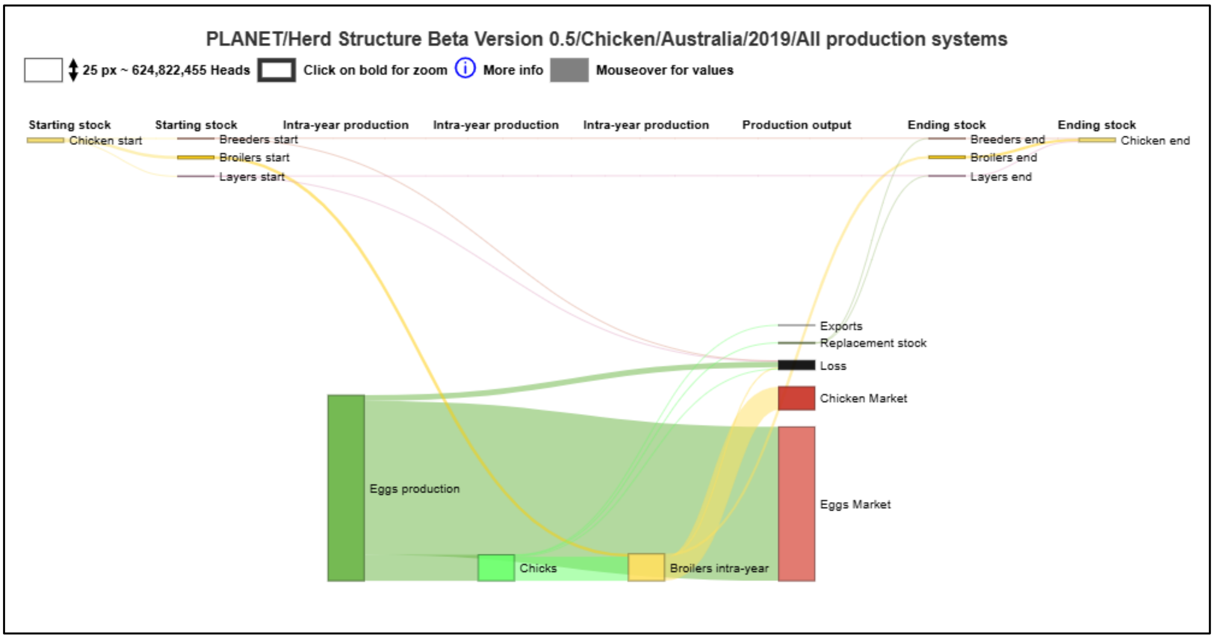



**Figure 2.3: Sankey illustration of chicken flock flow in a year period: All production systems**

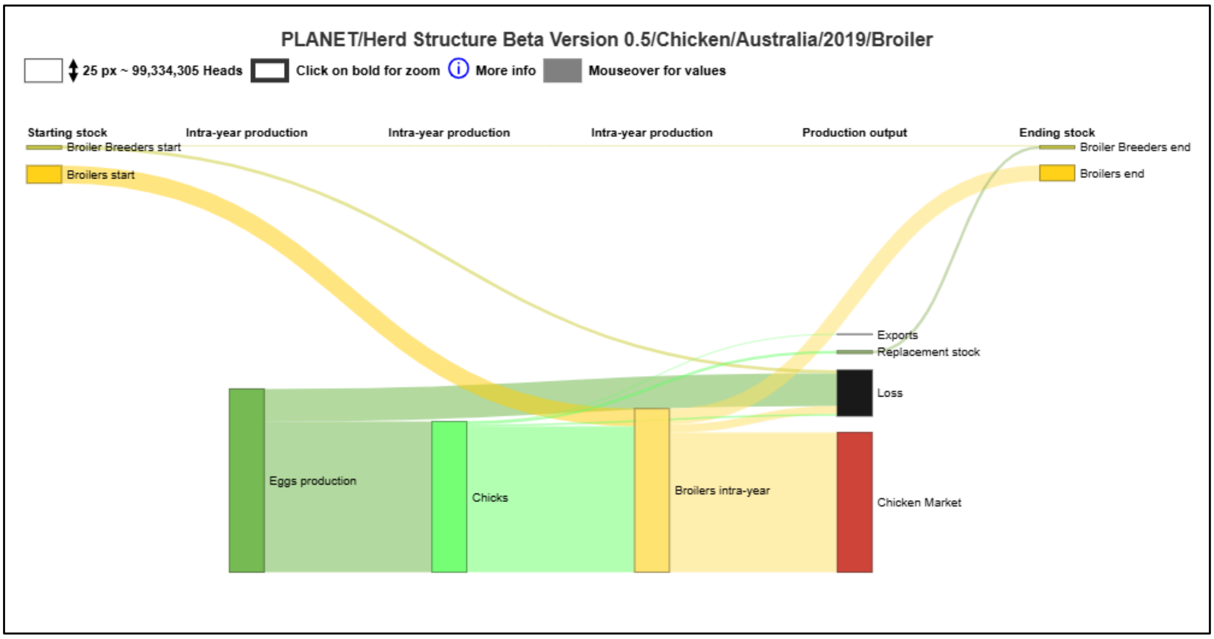



**Figure 2.4: Sankey illustration of chicken flock flow in a year period: Broiler system**

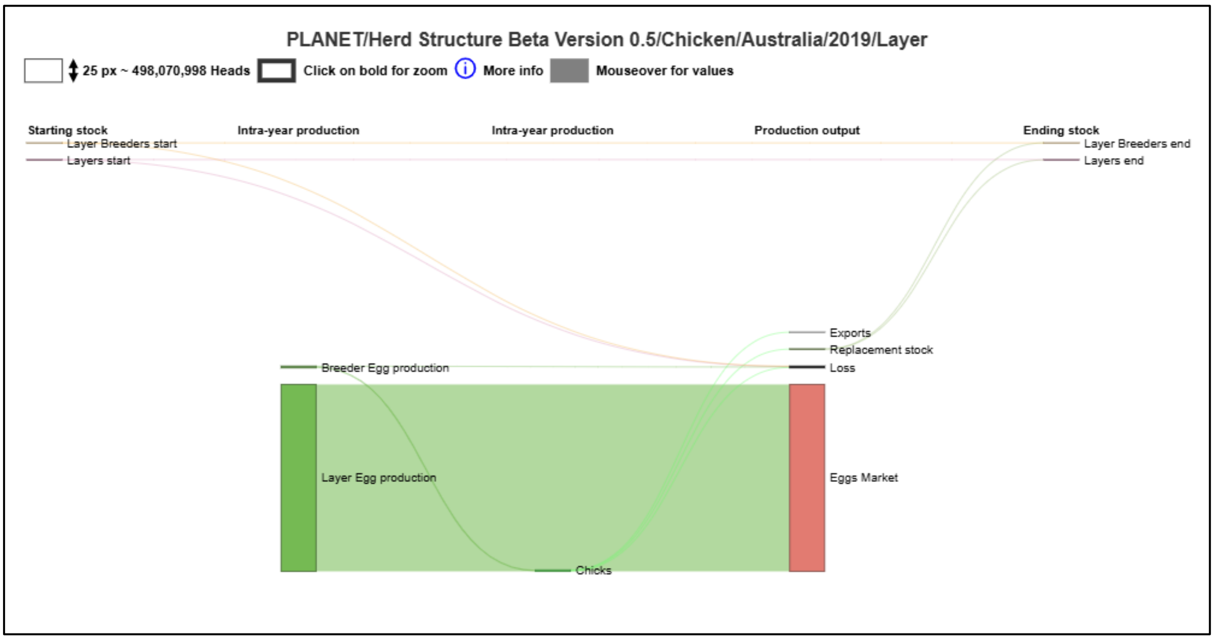



**Figure 2.4: Sankey illustration of chicken flock flow in a year period: Layer system**

**Cattle module illustrations:**

Figure 2.7 to 2.10 illustrates Sankey diagram of cattle herd in Australia.

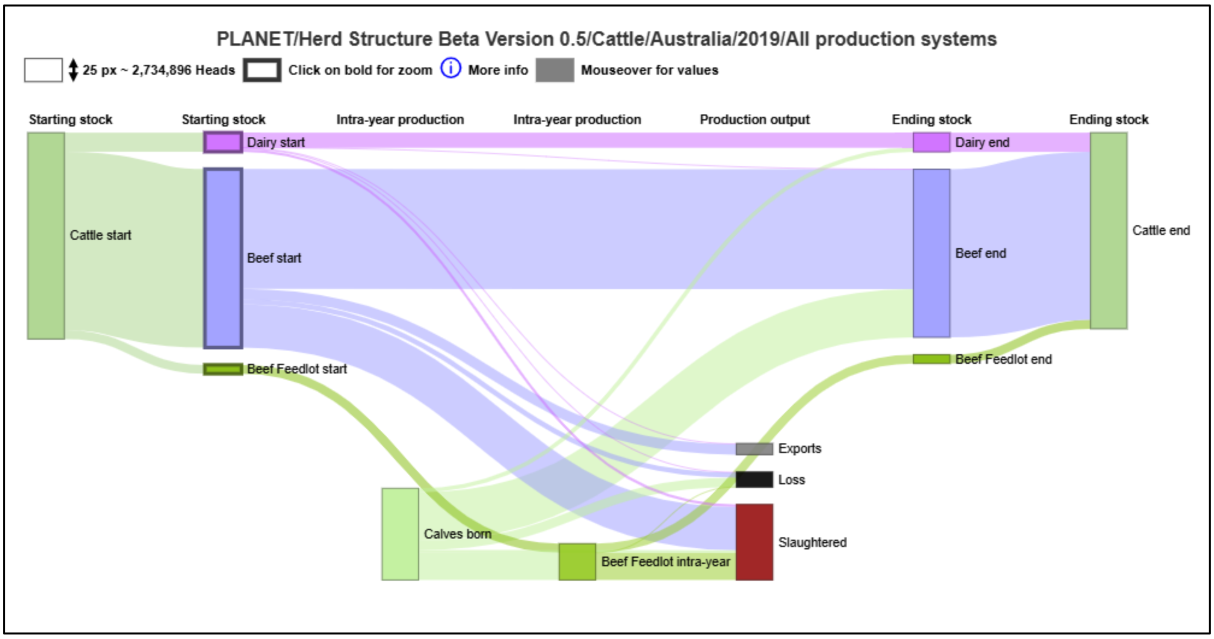



**Figure 2.7: Sankey illustration of cattle herd flow in a year period: Main screen**

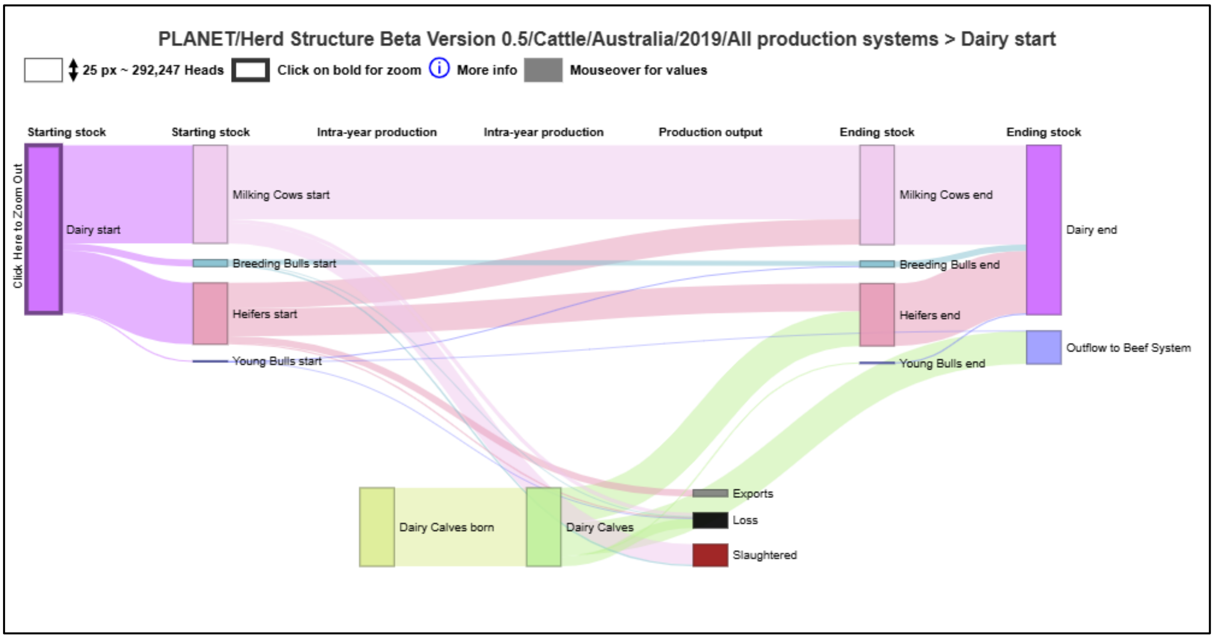



**Figure 2.8: Sankey illustration of cattle herd flow in a year period: Dairy**

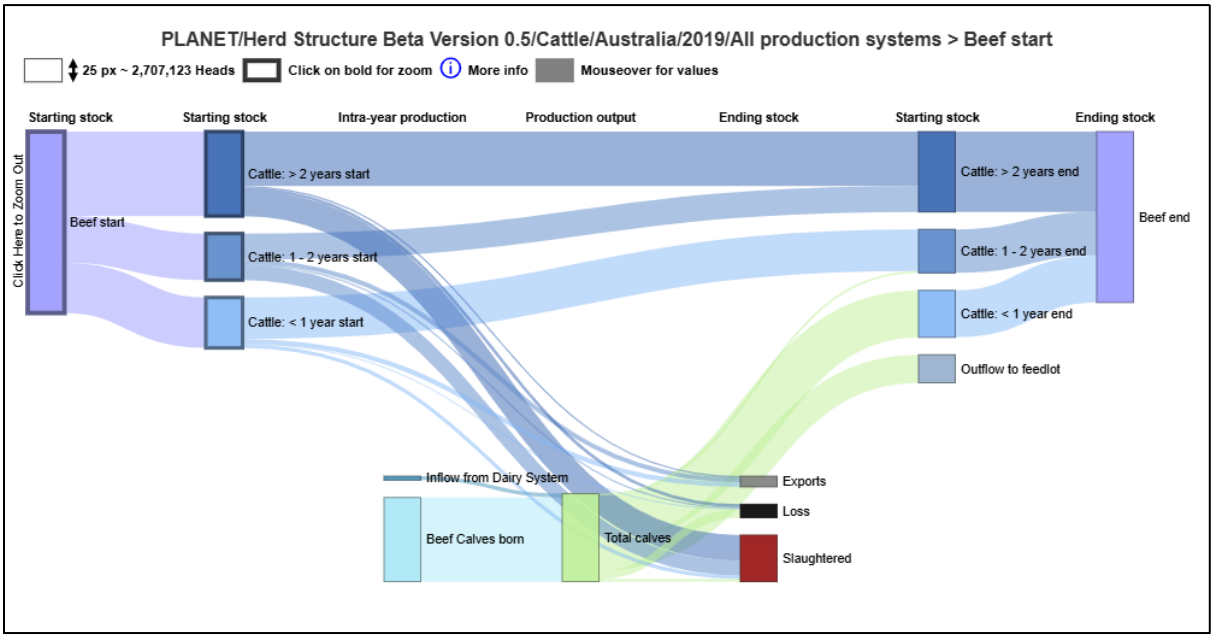



**Figure 2.9: Sankey illustration of cattle herd flow in a year period: Beef**

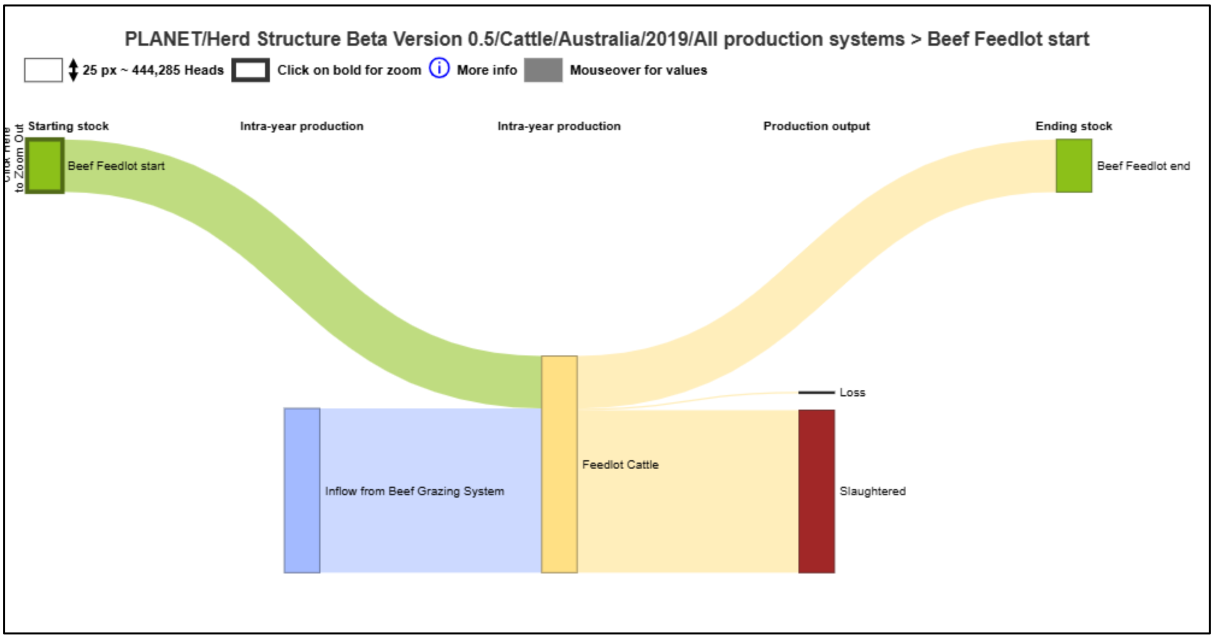



**Figure 2.10: Sankey illustration of cattle herd flow in a year period: Beef Feedlot**

## Animal Nutrient Figures

Figure 2.11 to 2.13 illustrates graphs displaying nutrient requirements of different animal classes in Australia.

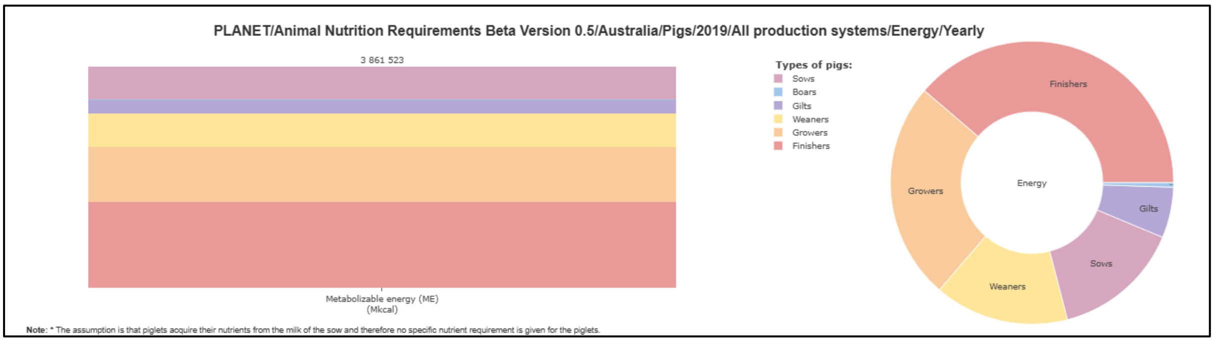



**Figure 2.11: Nutritional requirement displays of different pig classes: Energy**

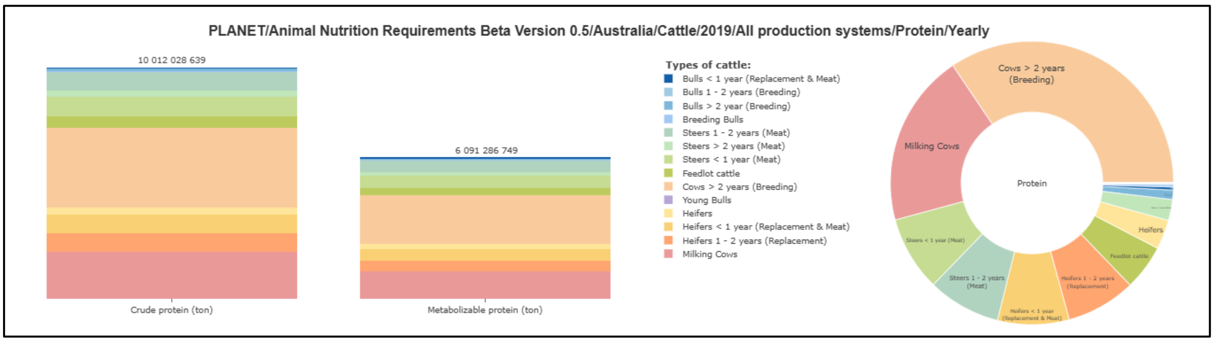



**Figure 2.12: Nutritional requirement displays of different cattle classes: Protein**

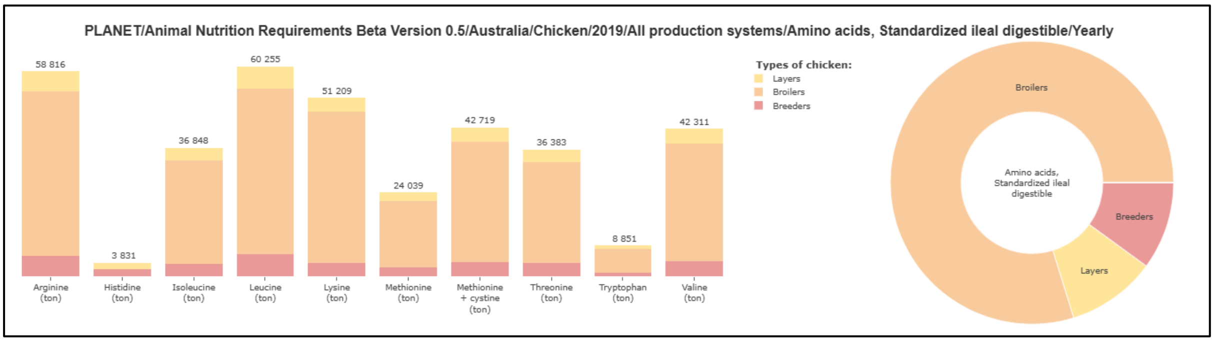
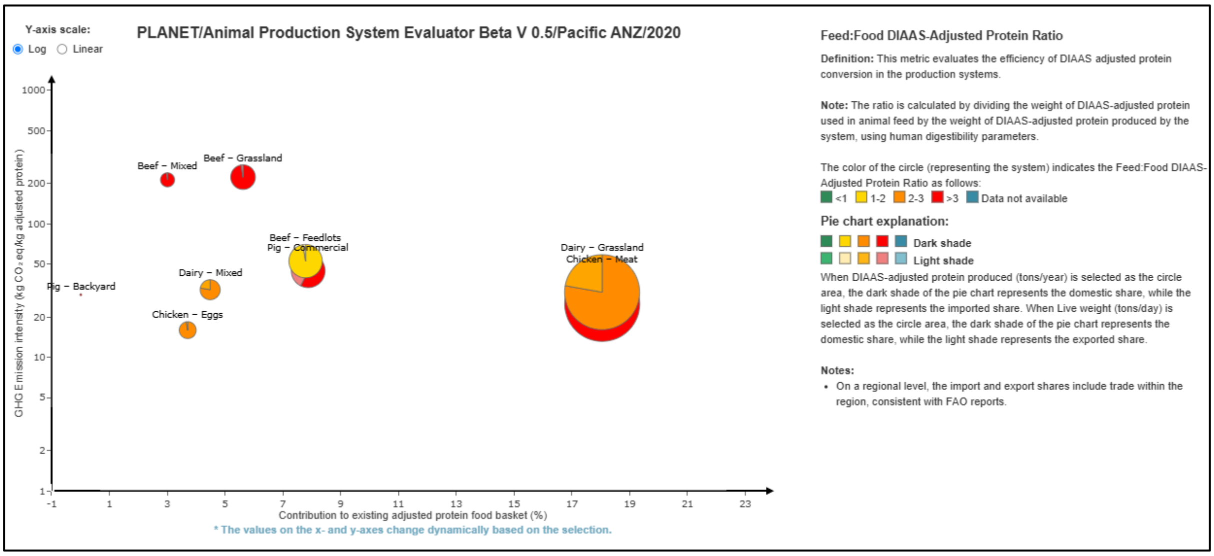


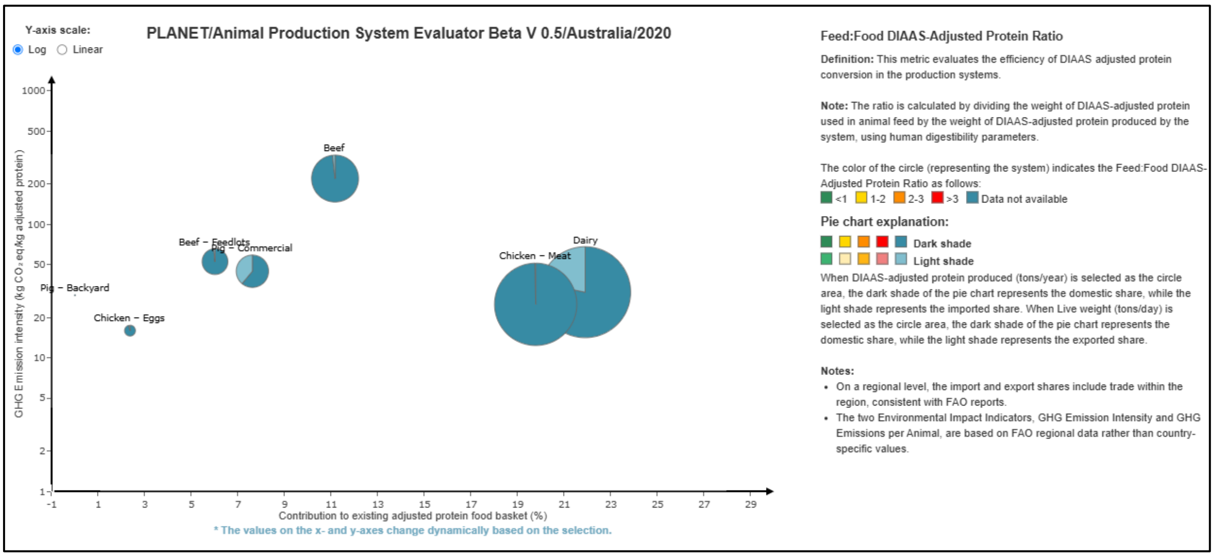

**Figure 2.13: Nutritional requirement displays of different chicken classes: Amino Acids**

## Animal Production System Evaluator (APSE) tool Figures:

Figures 2.14 to 2.15 illustrate all the main features of the tool with a regional and country example.




### Figure 2.14: APSE tool: Pacific ANZ region




### Figure 2.15: APSE tool: Australia

## 7.2 Supplementary File 1

**Table S1: The variable importance of each variable for South Africa predicted livestock**

| Variable, month | All ruminants | Cattle | Goats | Sheep |
|---|---|---|---|---|
| DMP.1 | 68,71556093 | 22,84600291 | 3,633503234 | 25,28979173 |
| DMP.2 | 14,85912445 | 2,893847169 | 0,572403209 | 6,143980795 |
| DMP.3 | 19,72628822 | 4,823594993 | 0,982559764 | 8,32158654 |
| DMP.4 | 46,10566702 | 12,26610209 | 1,493243734 | 11,31117014 |
| DMP.5 | 36,71959899 | 6,514431356 | 1,142662846 | 11,89610336 |
| DMP.6 | 47,34884518 | 11,2331272 | 2,525199897 | 21,40132057 |
| DMP.7 | 17,25545723 | 2,686945887 | 0,5518951 | 6,124147291 |
| DMP.8 | 31,62235253 | 3,219602686 | 0,611366617 | 12,89306961 |
| DMP.9 | 21,25791689 | 3,44765127 | 0,778468786 | 6,270479909 |
| DMP.10 | 37,59739757 | 8,337888666 | 1,409036465 | 9,196675945 |
| DMP.11 | 44,8409678 | 8,723816883 | 1,173104263 | 13,45622325 |
| DMP.12 | 19,08768098 | 3,722982004 | 0,837580072 | 7,783441408 |
| Rain.1 | 88,49195456 | 34,15492544 | 4,962558761 | 45,35671384 |
| Max.Temp.1 | 47,63456735 | 9,847089894 | 2,376373254 | 16,67264417 |
| Min.Temp.1 | 49,68742488 | 6,227054377 | 2,167099551 | 18,9482264 |
| Rain.2 | 109,3115993 | 49,13283146 | 8,787199139 | 51,03775708 |
| Max.Temp.2 | 51,24476844 | 11,61052873 | 2,807996706 | 19,33337291 |
| Min.Temp.2 | 47,64919737 | 5,963537933 | 1,907590577 | 20,4116776 |
| Rain.3 | 124,6587494 | 11,6545997 | 3,386272252 | 68,39650723 |
| Max.Temp.3 | 42,1380077 | 6,229102096 | 1,798337168 | 14,21653186 |
| Min.Temp.3 | 48,34665404 | 6,473819046 | 2,15463068 | 23,55657 |
| Rain.4 | 193,2960084 | 19,27011021 | 9,237934656 | 62,23457968 |
| Max.Temp.4 | 42,0950383 | 6,28200931 | 2,307025012 | 15,06227447 |
| Min.Temp.4 | 58,68993764 | 10,48023661 | 3,348019673 | 29,81825988 |
| Rain.5 | 122,3223067 | 20,25214942 | 9,037946551 | 50,87443077 |
| Max.Temp.5 | 37,52580899 | 7,270713997 | 2,235621451 | 19,54563221 |
| Min.Temp.5 | 50,02315844 | 7,603907475 | 3,373913502 | 31,00163715 |
| Rain.6 | 107,1943024 | 16,31731391 | 6,819547114 | 42,46261634 |
| Max.Temp.6 | 46,98644739 | 7,875570031 | 2,003683044 | 24,96742625 |
| Min.Temp.6 | 54,76629377 | 8,955335997 | 2,875796824 | 40,35616495 |
| Rain.7 | 119,0668993 | 27,77102328 | 8,103538227 | 50,59517032 |
| Max.Temp.7 | 41,09119755 | 9,757573817 | 2,385118833 | 22,65949521 |
| Min.Temp.7 | 51,60464762 | 11,65080734 | 3,312741488 | 31,23308937 |
| Rain.8 | 83,51258487 | 14,47192403 | 5,745810118 | 35,81398848 |
| Max.Temp.8 | 41,65374477 | 16,28544816 | 2,629408681 | 23,23139759 |
| Min.Temp.8 | 59,16070978 | 11,39897316 | 2,38325495 | 29,29598446 |
| Rain.9 | 192,3456932 | 23,9273727 | 11,46397615 | 64,4898421 |
| Max.Temp.9 | 62,31698457 | 18,37862203 | 5,294706179 | 18,38431866 |
| Min.Temp.9 | 63,87749436 | 11,79609883 | 2,383807604 | 20,98640528 |
| Rain.10 | 65,8339693 | 8,815009637 | 3,15845653 | 38,25009238 |
| Max.Temp.10 | 53,46056362 | 7,371698237 | 2,509667401 | 16,99851136 |
| Min.Temp.10 | 53,16537927 | 10,19931653 | 2,132361167 | 18,68215455 |

| | | | | |
|---|---|---|---|---|
| Rain.11 | 92,41326195 | 10,75044648 | 4,8264261 | 72,83554747 |
| Max.Temp.11 | 97,09160287 | 18,0750708 | 5,801638326 | 27,3097031 |
| Min.Temp.11 | 83,78335242 | 15,12076036 | 2,53653365 | 23,51871019 |
| Rain.12 | 94,15459025 | 22,32780007 | 4,780908796 | 59,68859812 |
| Max.Temp.12 | 41,6432751 | 5,642590723 | 1,947649003 | 18,14639601 |
| Min.Temp.12 | 49,72681501 | 7,056191893 | 2,131792116 | 18,50259869 |
| Elevation | 85,02031574 | 16,50381555 | 3,080929872 | 41,62556497 |
| Slope | 19,02091991 | 3,20898192 | 0,978305482 | 6,551465149 |
| Water | 11,0932653 | 3,820590869 | 0,630282936 | 5,353293951 |
| Land.Cover | 35,55594033 | 6,078594477 | 0,56651312 | 8,802568319 |

**Table S2: Land cover classes and their associated grazing utilisation factor**

| | | |
|---|---|---|
| Cropland, rainfed | 0.15 | |
| Cropland, rainfed: Herbaceous cover | 0.15 | |
| Cropland, rainfed: Tree or shrub cover | 0.15 | |
| Cropland, irrigated or post-flooding | 0.15 | |
| Mosaic cropland (>50%) / natural vegetation (tree, shrub, herbaceous cover) (<50%) | 0.2175 | * |
| Mosaic natural vegetation (tree, shrub, herbaceous cover) (>50%) / cropland (<50%) | 0.165 | * |
| Tree cover, broadleaved, evergreen, closed to open (>15%) | 0.32 | * |
| Tree cover, broadleaved, deciduous, closed to open (>15%) | 0.32 | * |
| Tree cover, broadleaved, deciduous, closed (>40%) | 0.195 | * |
| Tree cover, broadleaved, deciduous, open (15 - 40%) | 0.32 | * |
| Tree cover, needleleaved, evergreen, closed to open (>15%) | 0.32 | * |
| Tree cover, needleleaved, evergreen, closed (>40%) | 0.195 | * |
| Tree cover, needleleaved, evergreen, open (15 - 40%) | 0.32 | * |
| Tree cover, needleleaved, deciduous, closed to open (>15%) | 0.32 | * |
| Tree cover, needleleaved, deciduous, closed (>40%) | 0.195 | * |
| Tree cover, needleleaved, deciduous, open (15 - 40%) | 0.32 | * |
| Tree cover, mixed leaf type (broadleaved and needleleaved) | 0 | * |
| Mosaic tree and shrub (>50%) / herbaceous cover (<50%) | 0.3 | * |

| | | |
|---|---|---|
| Mosaic herbaceous cover (>50%) / tree and shrub (<50%) | 0.4 | * |
| Shrubland | 0.35 | * |
| Evergreen shrubland | 0.4 | * |
| Deciduous shrubland | 0.4 | * |
| Grassland | 0.5 | |
| Lichens and mosses | 0 | * |
| Sparse vegetation (tree, shrub, herbaceous cover) (<15%) | 0.2 | * |
| Sparse tree (<15%) | 0 | |
| Sparse shrub (<15%) | 0.15 | * |
| Sparse herbaceous cover (<15%) | 0.15 | * |
| Tree cover, flooded, fresh or brakish water | 0.2 | * |
| Tree cover, flooded, saline water | 0.2 | * |
| Shrub or herbaceous cover, flooded, fresh/saline/brakish water | 0.2 | * |
| Urban areas | 0.15 | * |
| Bare areas | 0.15 | * |
| Consolidated bare areas | 0 | |
| Unconsolidated bare areas | 0 | |
| Water bodies | 0.45 | * |
| Permanent snow and ice | 0 | |
| No data | 0.45 | * |

Note: * indicates where current livestock are 0, the utilisation factor is 0.

**Table S3: Herd structures, assigned intakes (percentages bodyweight (BW) ) of grazing and supplementation**

| **North America: Beef** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Nursing Calves | 0–6 mo | 120 | 20.00% | 24 | 1.00% | 0.24 | 0.00% | 0 |
| Weaned Calves (Stockers) | 6–12 mo | 275 | 10.00% | 27.5 | 2.50% | 0.6875 | 0.00% | 0 |
| Yearlings (Backgrounding) | 12–18 mo | 400 | 10.00% | 40 | 3.00% | 1.2 | 0.00% | 0 |
| Finishing Steers & Heifers | 18–24 mo | 525 | 10.00% | 52.5 | 2.00% | 1.05 | 2.00% | 1.05 |
| Mature Cows (Breeding, 2+ yrs) | ≥2 yrs | 600 | 40.00% | 240 | 3.00% | 7.2 | 0.00% | 0 |
| Mature Bulls (Breeding, 2+ yrs) | ≥2 yrs | 800 | 5.00% | 40 | 3.00% | 1.2 | 0.00% | 0 |
| Cull Cows & Bulls (awaiting slaughter) | ≥2 yrs | 700 | 5.00% | 35 | 3.00% | 1.05 | 0.00% | 0 |
| | | | **100.00%** | **459** | | **12.6275** | | **1.05** |
| **North America: Dairy** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Male Calves Slaughtered Immediately (at birth) | 0 days | 40 | 0.00% | 0 | 0.00% | 0 | 0.00% | 0 |
| Male Calves for Veal (0–6 mo) | 0–6 mo | 120 | 5.00% | 6 | 1.00% | 0.06 | 0.00% | 0 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Male Calves for Beef Finishing (0–24 mo) | 0–24 mo | 320 | 8.00% | 25.6 | 2.50% | 0.64 | 0.50% | 0.128 |
| Female Calves (not yet weaned, ~0–2 mo) | 0–2 mo | 60 | 10.00% | 6 | 0.00% | 0 | 0.00% | 0 |
| Heifer Growers (weaned females, ~2–12 mo) | 2–12 mo | 165 | 10.00% | 16.5 | 1.75% | 0.28875 | 0.00% | 0 |
| Replacement Heifers (12–24 mo) | 12–24 mo | 350 | 15.00% | 52.5 | 3.00% | 1.575 | 0.00% | 0 |
| Lactating Cows (2+ yrs) | ≥2 yrs | 600 | 40.00% | 240 | 2.00% | 4.8 | 2.00% | 4.8 |
| Dry Cows (2+ yrs) | ≥2 yrs | 600 | 5.00% | 30 | 3.00% | 0.9 | 0.00% | 0 |
| Breeding Bulls (2+ yrs, if kept on-farm) | ≥2 yrs | 800 | 3.00% | 24 | 3.00% | 0.72 | 0.00% | 0 |
| Cull Cows & Bulls (awaiting slaughter) | ≥2 yrs | 700 | 4.00% | 28 | 3.00% | 0.84 | 0.00% | 0 |
| | | | **100.00%** | **428.60** | | **9.82** | | **4.93** |
| **North America: Sheep** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Lambs (<1 yr) | 15–45 kg | 30.00 | 30.00% | 9 | 2.00% | 0.18 | 0.50% | 0.045 |
| Yearlings (1–2 yrs) | 45–60 kg | 50.00 | 15.00% | 7.5 | 2.50% | 0.1875 | 0.00% | 0 |
| Mature Ewes (≥2 yrs) | 60–90 kg | 75.00 | 45.00% | 33.75 | 2.50% | 0.84375 | 0.50% | 0.16875 |
| Rams (≥2 yrs) | 80–120 kg | 100.00 | 5.00% | 5 | 2.50% | 0.125 | 0.00% | 0 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Geriatric Ewes (>6 yrs)* | 55–80 kg (varies) | 67.00 | 5.00% | 3.35 | 2.50% | 0.08375 | 0.00% | 0 |
| | | | 100.00% | **58.6** | | **1.42** | | **0.21375** |
| **North America: Goats** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Kids (<1 yr) | 10–25 kg | 18.00 | 30.00% | 5.4 | 2.00% | 0.108 | 0.50% | 0.027 |
| Yearlings (1–2 yrs) | 25–35 kg | 30.00 | 15.00% | 4.5 | 2.50% | 0.1125 | 0.00% | 0 |
| Mature Does (≥2 yrs) | 35–60 kg | 48.00 | 45.00% | 21.6 | 2.50% | 0.54 | 0.50% | 0.108 |
| Bucks (≥2 yrs) | 50–80 kg | 65.00 | 5.00% | 3.25 | 2.50% | 0.08125 | 0.00% | 0 |
| Older Does (>6 yrs)* | 30–50 kg | 40.00 | 5.00% | 2 | 2.50% | 0.05 | 0.00% | 0 |
| | | | 100.00% | **36.75** | | **0.89175** | | **0.135** |
| | | | | | | | | |
| **Europe: Beef** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Nursing Calves | 0–6 mo | 110 | 20.00% | 22 | 1.00% | 0.22 | 0.00% | 0 |
| Weaned Calves (Stockers) | 6–12 mo | 250 | 10.00% | 25 | 2.50% | 0.625 | 0.00% | 0 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Yearlings (Backgrounding) | 12–18 mo | 370 | 10.00% | 37 | 3.00% | 1.11 | 0.00% | 0 |
| Finishing Steers & Heifers | 18–24 mo | 510 | 10.00% | 51 | 2.00% | 1.02 | 2.00% | 1.02 |
| Mature Cows (Breeding, 2+ yrs) | ≥2 yrs | 550 | 40.00% | 220 | 3.00% | 6.6 | 0.00% | 0 |
| Mature Bulls (Breeding, 2+ yrs) | ≥2 yrs | 775 | 5.00% | 38.75 | 3.00% | 1.1625 | 0.00% | 0 |
| Cull Cows & Bulls (awaiting slaughter) | ≥2 yrs | 675 | 5.00% | 33.75 | 3.00% | 1.0125 | 0.00% | 0 |
| | | | **100.00%** | **427.5** | | **11.75** | | **1.02** |
| **Europe: Dairy** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Male Calves Slaughtered Immediately | 0 days | 40 | 0.00% | 0 | 0.00% | 0 | 0.00% | 0 |
| Male Calves for Veal (0–6 mo) | 0–6 mo | 110 | 6.00% | 6.6 | 1.00% | 0.066 | 0.00% | 0 |
| Male Calves for Beef Finishing (0–24 mo) | 0–24 mo | 295 | 8.00% | 23.6 | 2.50% | 0.59 | 0.50% | 0.118 |
| Female Calves (not yet weaned) | 0–2 mo | 53 | 10.00% | 5.3 | 0.00% | 0 | 0.00% | 0 |
| Heifer Growers (2–12 mo) | 2–12 mo | 145 | 10.00% | 14.5 | 1.75% | 0.25375 | 0.00% | 0 |
| Replacement Heifers (12–24 mo) | 12–24 mo | 320 | 15.00% | 48 | 3.00% | 1.44 | 0.00% | 0 |
| Lactating Cows (2+ yrs) | ≥2 yrs | 575 | 40.00% | 230 | 2.00% | 4.6 | 2.00% | 4.6 |
| Dry Cows (2+ yrs) | ≥2 yrs | 575 | 5.00% | 28.75 | 3.00% | 0.8625 | 0.00% | 0 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Breeding Bulls (2+ yrs) | ≥2 yrs | 775 | 3.00% | 23.25 | 3.00% | 0.6975 | 0.00% | 0 |
| Cull Cows & Bulls (awaiting slaughter) | ≥2 yrs | 700 | 3.00% | 21 | 3.00% | 0.63 | 0.00% | 0 |
| | | | **100.00%** | **401.00** | | **9.14** | | **4.72** |
| **Europe: Sheep** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Lambs (<1 yr) | 15–45 kg | 30.00 | 35.00% | 10.5 | 2.00% | 0.21 | 0.50% | 0.0525 |
| Yearlings (1–2 yrs) | 45–60 kg | 50.00 | 15.00% | 7.5 | 2.50% | 0.1875 | 0.00% | 0 |
| Mature Ewes (≥2 yrs) | 60–90 kg | 75.00 | 45.00% | 33.75 | 2.50% | 0.84375 | 0.50% | 0.16875 |
| Rams (≥2 yrs) | 80–120 kg | 100.00 | 5.00% | 5 | 2.50% | 0.125 | 0.00% | 0 |
| | | | 100.00% | **56.75** | | **1.36625** | | **0.22125** |
| **Europe: Goats** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Kids (<1 yr) | 8–25 kg | 17.00 | 35.00% | 5.95 | 2.00% | 0.119 | 0.50% | 0.02975 |
| Yearlings (1–2 yrs) | 25–40 kg | 33.00 | 15.00% | 4.95 | 2.50% | 0.12375 | 0.00% | 0 |
| Mature Does (≥2 yrs) | 40–60 kg | 50.00 | 45.00% | 22.5 | 2.50% | 0.5625 | 0.50% | 0.1125 |
| Bucks (≥2 yrs) | 60–90 kg | 75.00 | 5.00% | 3.75 | 2.50% | 0.09375 | 0.00% | 0 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | 100.00% | **37.15** | | **0.899** | | **0.14225** |
| | | | | | | | | |
| **Asia: Beef** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Nursing Calves | 0–6 mo | 90 | 25.00% | 22.5 | 1.00% | 0.225 | 0.00% | 0 |
| Weaned Calves (Stockers) | 6–12 mo | 200 | 10.00% | 20 | 2.50% | 0.5 | 0.00% | 0 |
| Yearlings (Backgrounding) | 12–18 mo | 325 | 10.00% | 32.5 | 3.00% | 0.975 | 0.00% | 0 |
| Finishing Steers & Heifers | 18–24 mo | 475 | 10.00% | 47.5 | 2.00% | 0.95 | 2.00% | 0.95 |
| Mature Cows (Breeding, 2+ yrs) | ≥2 yrs | 475 | 35.00% | 166.25 | 3.00% | 4.9875 | 0.00% | 0 |
| Mature Bulls (Breeding, 2+ yrs) | ≥2 yrs | 700 | 5.00% | 35 | 3.00% | 1.05 | 0.00% | 0 |
| Cull Cows & Bulls (awaiting slaughter) | ≥2 yrs | 575 | 5.00% | 28.75 | 3.00% | 0.8625 | 0.00% | 0 |
| | | | | **352.5** | | **9.55** | | **0.95** |
| **Asia: Dairy** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Male Calves Slaughtered Immediately | 0 days | 35 | 1.00% | 0.35 | 0.00% | 0 | 0.00% | 0 |
| Male Calves for Veal (0–6 mo) | 0–6 mo | 95 | 5.00% | 4.75 | 1.00% | 0.0475 | 0.00% | 0 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Male Calves for Beef Finishing (0–24 mo) | 0–24 mo | 270 | 8.00% | 21.6 | 2.50% | 0.54 | 0.50% | 0.108 |
| Female Calves (not yet weaned) | 0–2 mo | 48 | 12.00% | 5.76 | 0.00% | 0 | 0.00% | 0 |
| Heifer Growers (2–12 mo) | 2–12 mo | 135 | 10.00% | 13.5 | 1.75% | 0.23625 | 0.00% | 0 |
| Replacement Heifers (12–24 mo) | 12–24 mo | 275 | 15.00% | 41.25 | 3.00% | 1.2375 | 0.00% | 0 |
| Lactating Cows (2+ yrs) | ≥2 yrs | 475 | 35.00% | 166.25 | 2.00% | 3.325 | 2.00% | 3.325 |
| Dry Cows (2+ yrs) | ≥2 yrs | 475 | 5.00% | 23.75 | 3.00% | 0.7125 | 0.00% | 0 |
| Breeding Bulls (2+ yrs) | ≥2 yrs | 675 | 3.00% | 20.25 | 3.00% | 0.6075 | 0.00% | 0 |
| Cull Cows & Bulls (awaiting slaughter) | ≥2 yrs | 600 | 5.00% | 30 | 3.00% | 0.9 | 0.00% | 0 |
| | | | **99.00%** | **327.46** | | **7.61** | | **3.43** |
| **Asia: Sheep** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Lambs (<1 yr) | 10–40 kg* | 25.00 | 40.00% | 10 | 2.00% | 0.2 | 0.50% | 0.05 |
| Yearlings (1–2 yrs) | 40–55 kg | 48.00 | 10.00% | 4.8 | 2.50% | 0.12 | 0.00% | 0 |
| Mature Ewes (≥2 yrs) | 55–80 kg | 68.00 | 45.00% | 30.6 | 2.50% | 0.765 | 0.50% | 0.153 |
| Rams (≥2 yrs) | 70–110 kg | 90.00 | 5.00% | 4.5 | 2.50% | 0.1125 | 0.00% | 0 |
| | | | | **49.9** | | **1.1975** | | **0.203** |

| **Asia: Goats** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Category** | **Avg. Age Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Kids (<1 yr) | 8–20 kg | 14.00 | 40.00% | 5.6 | 2.00% | 0.112 | 0.50% | 0.028 |
| Yearlings (1–2 yrs) | 20–35 kg | 28.00 | 10.00% | 2.8 | 2.50% | 0.07 | 0.00% | 0 |
| Mature Does (≥2 yrs) | 35–55 kg | 45.00 | 45.00% | 20.25 | 2.50% | 0.50625 | 0.50% | 0.10125 |
| Bucks (≥2 yrs) | 50–80 kg | 65.00 | 5.00% | 3.25 | 2.50% | 0.08125 | 0.00% | 0 |
| | | | | **31.9** | | **0.7695** | | **0.12925** |
| | | | | | | | | |
| **Africa: Beef** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Nursing Calves | 0–6 mo | 85 | 20.00% | 17 | 1.00% | 0.17 | 0.00% | 0 |
| Weaned Calves (Stockers) | 6–12 mo | 195 | 10.00% | 19.5 | 2.50% | 0.4875 | 0.00% | 0 |
| Yearlings (Backgrounding) | 12–18 mo | 300 | 10.00% | 30 | 3.00% | 0.9 | 0.00% | 0 |
| Finishing Steers & Heifers | 18–24 mo | 425 | 10.00% | 42.5 | 2.00% | 0.85 | 2.00% | 0.85 |
| Mature Cows (Breeding, 2+ yrs) | ≥2 yrs | 400 | 40.00% | 160 | 3.00% | 4.8 | 0.00% | 0 |
| Mature Bulls (Breeding, 2+ yrs) | ≥2 yrs | 600 | 5.00% | 30 | 3.00% | 0.9 | 0.00% | 0 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Cull Cows & Bulls (awaiting slaughter) | ≥2 yrs | 500 | 5.00% | 25 | 3.00% | 0.75 | 0.00% | 0 |
| | | | **100.00%** | **324** | | **8.8575** | | **0.85** |
| **Africa: Dairy** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Male Calves Slaughtered Immediately | 0 days | 30 | 0.00% | 0 | 0.00% | 0 | 0.00% | 0 |
| Male Calves for Veal / Short-Fed (0–6 mo) | 0–6 mo | 78 | 4.00% | 3.12 | 1.00% | 0.0312 | 0.00% | 0 |
| Male Calves for Beef Finishing (0–24 mo) | 0–24 mo | 243 | 8.00% | 19.44 | 2.50% | 0.486 | 0.50% | 0.0972 |
| Female Calves (not yet weaned) | 0–2 mo | 43 | 12.00% | 5.16 | 0.00% | 0 | 0.00% | 0 |
| Heifer Growers (2–12 mo) | 2–12 mo | 120 | 10.00% | 12 | 1.75% | 0.21 | 0.00% | 0 |
| Replacement Heifers (12–24 mo) | 12–24 mo | 265 | 15.00% | 39.75 | 3.00% | 1.1925 | 0.00% | 0 |
| Lactating Cows (2+ yrs) | ≥2 yrs | 425 | 40.00% | 170 | 2.00% | 3.4 | 2.00% | 3.4 |
| Dry Cows (2+ yrs) | ≥2 yrs | 425 | 5.00% | 21.25 | 3.00% | 0.6375 | 0.00% | 0 |
| Breeding Bulls (2+ yrs) | ≥2 yrs | 600 | 3.00% | 18 | 3.00% | 0.54 | 0.00% | 0 |
| Cull Cows & Bulls (awaiting slaughter) | ≥2 yrs | 525 | 3.00% | 15.75 | 3.00% | 0.4725 | 0.00% | 0 |
| | | | **100.00%** | **304.47** | | **6.97** | | **3.50** |
| **Africa: Sheep** | | | | | | | | |

| Category | Average Range | Average Weight (kg) | % of Herd | Fractioned Weight (kg) | Grazing BW % | Grazing intake (kg) | Cropland BW% | Cropland intake (kg) |
|---|---|---|---|---|---|---|---|---|
| Lambs (<1 yr) | 10–35 kg | 23.00 | 35.00% | 8.05 | 2.00% | 0.161 | 0.50% | 0.04025 |
| Yearlings (1–2 yrs) | 35–50 kg | 43.00 | 15.00% | 6.45 | 2.50% | 0.16125 | 0.00% | 0 |
| Mature Ewes (≥2 yrs) | 50–70 kg | 60.00 | 45.00% | 27 | 2.50% | 0.675 | 0.50% | 0.135 |
| Rams (≥2 yrs) | 60–90 kg | 75.00 | 5.00% | 3.75 | 2.50% | 0.09375 | 0.00% | 0 |
| | | | | **45.25** | | **1.091** | | **0.17525** |
| **Africa: Goats** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Kids (<1 yr) | 8–20 kg | 14.00 | 35.00% | 4.9 | 2.00% | 0.098 | 0.50% | 0.0245 |
| Yearlings (1–2 yrs) | 20–30 kg | 25.00 | 15.00% | 3.75 | 2.50% | 0.09375 | 0.00% | 0 |
| Mature Does (≥2 yrs) | 30–50 kg | 40.00 | 45.00% | 18 | 2.50% | 0.45 | 0.50% | 0.09 |
| Bucks (≥2 yrs) | 45–70 kg | 58.00 | 5.00% | 2.9 | 2.50% | 0.0725 | 0.00% | 0 |
| | | | | **29.55** | | **0.71425** | | **0.1145** |
| | | | | | | | | |
| **Latin America: Beef** | | | | | | | | |

| Category | Average Range | Average Weight (kg) | % of Herd | Fractioned Weight (kg) | Grazing BW % | Grazing intake (kg) | Cropland BW% | Cropland intake (kg) |
|---|---|---|---|---|---|---|---|---|
| Nursing Calves | 0–6 mo | 110 | 20.00% | 22 | 1.00% | 0.22 | 0.00% | 0 |
| Weaned Calves (Stockers) | 6–12 mo | 240 | 10.00% | 24 | 2.50% | 0.6 | 0.00% | 0 |
| Yearlings (Backgrounding) | 12–18 mo | 350 | 10.00% | 35 | 3.00% | 1.05 | 0.00% | 0 |
| Finishing Steers & Heifers | 18–24 mo | 475 | 10.00% | 47.5 | 2.00% | 0.95 | 2.00% | 0.95 |
| Mature Cows (Breeding, 2+ yrs) | ≥2 yrs | 475 | 40.00% | 190 | 3.00% | 5.7 | 0.00% | 0 |
| Mature Bulls (Breeding, 2+ yrs) | ≥2 yrs | 700 | 5.00% | 35 | 3.00% | 1.05 | 0.00% | 0 |
| Cull Cows & Bulls (awaiting slaughter) | ≥2 yrs | 575 | 5.00% | 28.75 | 3.00% | 0.8625 | 0.00% | 0 |
| | | | **100.00%** | **382.25** | | **10.4325** | | **0.95** |
| **Latin America: Dairy** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Male Calves Slaughtered Immediately | 0 days | 40 | 0.00% | 0 | 0.00% | 0 | 0.00% | 0 |
| Male Calves for Veal / Short-Fed (0–6 mo) | 0–6 mo | 95 | 5.00% | 4.75 | 1.00% | 0.0475 | 0.00% | 0 |
| Male Calves for Beef Finishing (0–24 mo) | 0–24 mo | 295 | 8.00% | 23.6 | 2.50% | 0.59 | 0.50% | 0.118 |
| Female Calves (not yet weaned) | 0–2 mo | 58 | 10.00% | 5.8 | 0.00% | 0 | 0.00% | 0 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Heifer Growers (2–12 mo) | 2–12 mo | 150 | 10.00% | 15 | 1.75% | 0.2625 | 0.00% | 0 |
| Replacement Heifers (12–24 mo) | 12–24 mo | 320 | 15.00% | 48 | 3.00% | 1.44 | 0.00% | 0 |
| Lactating Cows (2+ yrs) | ≥2 yrs | 550 | 40.00% | 220 | 2.00% | 4.4 | 2.00% | 4.4 |
| Dry Cows (2+ yrs) | ≥2 yrs | 550 | 5.00% | 27.5 | 3.00% | 0.825 | 0.00% | 0 |
| Breeding Bulls (2+ yrs) | ≥2 yrs | 775 | 3.00% | 23.25 | 3.00% | 0.6975 | 0.00% | 0 |
| Cull Cows & Bulls (awaiting slaughter) | ≥2 yrs | 675 | 4.00% | 27 | 3.00% | 0.81 | 0.00% | 0 |
| | | | **100.00%** | **394.90** | | **9.07** | | **4.52** |
| **Latin America: Sheep** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Lambs (<1 yr) | 15–40 kg | 28.00 | 35.00% | 9.8 | 2.00% | 0.196 | 0.50% | 0.049 |
| Yearlings (1–2 yrs) | 40–55 kg | 48.00 | 15.00% | 7.2 | 2.50% | 0.18 | 0.00% | 0 |
| Mature Ewes (≥2 yrs) | 55–80 kg | 68.00 | 45.00% | 30.6 | 2.50% | 0.765 | 0.50% | 0.153 |
| Rams (≥2 yrs) | 70–100 kg | 85.00 | 5.00% | 4.25 | 2.50% | 0.10625 | 0.00% | 0 |
| | | | | **51.85** | | **1.24725** | | **0.202** |
| **Latin America: Goats** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Kids (<1 yr) | 10–25 kg | 18.00 | 35.00% | 6.3 | 2.00% | 0.126 | 0.50% | 0.0315 |
| Yearlings (1–2 yrs) | 25–40 kg | 33.00 | 15.00% | 4.95 | 2.50% | 0.12375 | 0.00% | 0 |
| Mature Does (≥2 yrs) | 40–60 kg | 50.00 | 45.00% | 22.5 | 2.50% | 0.5625 | 0.50% | 0.1125 |
| Bucks (≥2 yrs) | 50–80 kg | 65.00 | 5.00% | 3.25 | 2.50% | 0.08125 | 0.00% | 0 |
| | | | | **37** | | **0.8935** | | **0.144** |
| | | | | | | | | |
| **Oceania: Beef** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Nursing Calves | 0–6 mo | 120 | 20.00% | 24 | 1.00% | 0.24 | 0.00% | 0 |
| Weaned Calves (Stockers) | 6–12 mo | 275 | 10.00% | 27.5 | 2.50% | 0.6875 | 0.00% | 0 |
| Yearlings (Backgrounding) | 12–18 mo | 400 | 10.00% | 40 | 3.00% | 1.2 | 0.00% | 0 |
| Finishing Steers & Heifers | 18–24 mo | 525 | 10.00% | 52.5 | 2.00% | 1.05 | 2.00% | 1.05 |
| Mature Cows (Breeding, 2+ yrs) | ≥2 yrs | 575 | 40.00% | 230 | 3.00% | 6.9 | 0.00% | 0 |
| Mature Bulls (Breeding, 2+ yrs) | ≥2 yrs | 800 | 5.00% | 40 | 3.00% | 1.2 | 0.00% | 0 |
| Cull Cows & Bulls (awaiting slaughter) | ≥2 yrs | 675 | 5.00% | 33.75 | 3.00% | 1.0125 | 0.00% | 0 |
| | | | **100.00%** | **447.75** | | **12.29** | | **1.05** |

| **Oceania: Dairy** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |
| Male Calves Slaughtered Immediately | 0 days | 40 | 0.00% | 0 | 0.00% | 0 | 0.00% | 0 |
| Male Calves for Veal / Short-Fed (0–6 mo) | 0–6 mo | 110 | 5.00% | 5.5 | 1.00% | 0.055 | 0.00% | 0 |
| Male Calves for Beef Finishing (0–24 mo) | 0–24 mo | 320 | 8.00% | 25.6 | 2.50% | 0.64 | 0.50% | 0.128 |
| Female Calves (not yet weaned) | 0–2 mo | 58 | 10.00% | 5.8 | 0.00% | 0 | 0.00% | 0 |
| Heifer Growers (2–12 mo) | 2–12 mo | 165 | 10.00% | 16.5 | 1.75% | 0.28875 | 0.00% | 0 |
| Replacement Heifers (12–24 mo) | 12–24 mo | 350 | 15.00% | 52.5 | 3.00% | 1.575 | 0.00% | 0 |
| Lactating Cows (2+ yrs) | ≥2 yrs | 575 | 40.00% | 230 | 2.00% | 4.6 | 2.00% | 4.6 |
| Dry Cows (2+ yrs) | ≥2 yrs | 575 | 5.00% | 28.75 | 3.00% | 0.8625 | 0.00% | 0 |
| Breeding Bulls (2+ yrs) | ≥2 yrs | 775 | 3.00% | 23.25 | 3.00% | 0.6975 | 0.00% | 0 |
| Cull Cows & Bulls (awaiting slaughter) | ≥2 yrs | 700 | 4.00% | 28 | 3.00% | 0.84 | 0.00% | 0 |
| | | | **100.00%** | **415.90** | | **9.56** | | **4.73** |
| **Oceania: Sheep** | | | | | | | | |
| **Category** | **Average Range** | **Average Weight (kg)** | **% of Herd** | **Fractioned Weight (kg)** | **Grazing BW %** | **Grazing intake (kg)** | **Cropland BW%** | **Cropland intake (kg)** |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Lambs (<1 yr) | 15–45 kg | 30.00 | 30.00% | 9 | 2.00% | 0.18 | 0.50% | 0.045 |
| Yearlings (1–2 yrs) | 45–60 kg | 53.00 | 15.00% | 7.95 | 2.50% | 0.19875 | 0.00% | 0 |
| Mature Ewes (≥2 yrs) | 60–85 kg | 73.00 | 50.00% | 36.5 | 2.50% | 0.9125 | 0.50% | 0.1825 |
| Rams (≥2 yrs) | 80–120 kg | 100.00 | 5.00% | 5 | 2.50% | 0.125 | 0.00% | 0 |
| | | | | **58.45** | | **1.41625** | | **0.2275** |
| Oceania: Goats | | | | | | | | |
| **Age Group** | **Approx. Weight Range** | **Avg weight** | **% of Herd** | **Weight (kg)** | **Grazing BW %** | **Grazing intake** | **Supplement BW%** | **Suppl. intake** |
| Kids (<1 yr) | 10–25 kg | 19.00 | 30.00% | 5.7 | 2.00% | 0.114 | 0.50% | 0.0285 |
| Yearlings (1–2 yrs) | 25–40 kg | 33.00 | 15.00% | 4.95 | 2.50% | 0.12375 | 0.00% | 0 |
| Mature Does (≥2 yrs) | 40–60 kg | 50.00 | 50.00% | 25 | 2.50% | 0.625 | 0.50% | 0.125 |
| Bucks (≥2 yrs) | 60–90 kg | 75.00 | 5.00% | 3.75 | 2.50% | 0.09375 | 0.00% | 0 |
| | | | | **39.4** | | **0.9565** | | **0.1535** |

## 7.3 Supplementary File 2

Author: James Grove, September 2024

# 1. Background

We define the variables that will be used to predict the livestock numbers as the $\mathrm{X_{i1}}, \mathrm{X_{i2}}, \dots, \mathrm{X_{ip}}$, with $p$ being the number of covariates considered, and it being the $i$'th observation of the dataset., we also define $n$ to be the sample size of the dataset. Other common names for the variables that are used to make predictions are: covariates, predictors, input variables or features.

The response of interest is denoted by $y_1, y_2, \dots, y_n$, and a model's prediction for $y$ given $x = (x_1, x_2, \dots, x_n)$ is given by $\hat{f}(x)$. The model it is referring to would be clear from the context or be indicated by a subscript.

The measure of model performance used is the mean square error (MSE), it is given in equation (1). It gives the average square of the difference between the true value and the predicted value. Because the MSE depends on the scale of the data, it is useful to standardize it. The most common standardized measure is the $R^2$, as is given in (2). Unlike MSE, larger values are better in the case of $R^2$. It is usually interpreted as the proportion of the variance explained by the model. Note that if $MSE(\hat{f}) = MSE(\bar{y})$ (where $\bar{y}$ is the sample mean), it equals zero, since predicting just the mean explains none of the variation in the data, and if the model fits perfectly, and it has an MSE of zero, then the $R^2$ value equals 1, since all the variation is then explained. In addition to its interpretability, the $R^2$ is also invariant to scale transformations.

$$MSE(\hat{f}) = \frac{1}{n}\sum_{i=1}^{n}\left(y - \hat{f}(x_i)\right)^2 \tag{1}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y - \bar{y})^2}{\sum_{i=1}^{n}\left(y - \hat{f}(x_i)\right)^2} \tag{2}$$

Sections 1.1 – X are based on a textbook written by Gareth et al. (2013) and they explain the more technical aspects of the report.

### 1.1 Regression Trees

A decision tree is a non-parametric model, which is usually used in the context of supervised learning. It can be used in the context of regression (where it is called a regression tree) or classification. It was introduced by Breiman et al. (1984).

To explain how regression trees work, we first consider how a fitted tree makes predictions, and then we will explain how they are fit.

Consider Figure 1, in each of the eclipses (nodes labelled with the square with a number inside), the top value is the average of the training data's response at that node, then n is how many of the training points are present in that node, and finally the inequality below the node is the splitting criteria. The data is then split according to the criteria, all those that satisfy it to left, the rest go right.

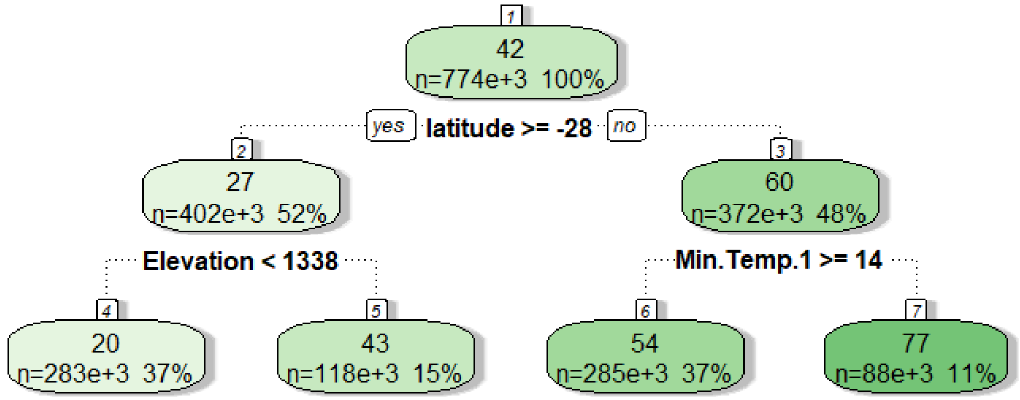



**Figure 1: Regression Tree Example**

For example, suppose we want to find the prediction for $x$ with, $X_{lat} = -26, X_{ele} = 1\,500, X_{Min.T1} = 10$, then the process of finding $\hat{f}(x)$ is as follows. Node 1: $X_{lat} \geq -28$, so go to node 2, here $X_{ele} > 1\,338$, which means that the criteria is not satisfied, so data case is sent to node 5, here the final prediction value is 43.

The nodes that have no splits after them are called leaf nodes, and the depth of a node is the number of splits it takes to get to that node. A when refereeing to a deep tree, it implies that the tree has deep nodes.

The model is trained by making splits on the tree. The split is based on which leaf node, which variable and which value on that variable will minimize the MSE of the data set it is trained on. This is repeated until some stopping criteria is reached, this can be the maximum number of observations in all leaf nodes, the maximum depth of a tree, etc. Then the prediction values are the averages of the data points that are present at each leaf node.

In regular use decision trees are pruned afterwards to reduce its complexity, this is not relevant to this report, because bagging methods exploits the complex versions of decision trees' properties. This paragraph is not required to understand the modelling process, it was just added for completeness.

Deeper trees are more complex and can easily fit to training data very closely, from which it follows that deep trees generally have a very small bias, with the trade-off of having a large variance.

By using ensembles, bagging and random forest exploit this low bias property, along with means decreasing variance to have a model that has low bias and variance, which is then expected to generalize well for unseen data. This is discussed subsequent sections.

## 1.2 Bagging

The main idea behind bagging, is to reduce the variance of model predictions, by taking averages. Because for an independent sample $Z_1, Z_2, \dots, Z_B$, with equal variance, $\sigma^2$, the variance of the mean is $\frac{\sigma^2}{B}$, which decreases as $B$ increases.

To achieve this, a model would have to be trained on many different data sets, with the final prediction being the average output of all the models. Since this is infeasible, what bagging does is take $B$ bootstrap samples from the rows in the original data set, and fit $B$ models one on each bootstrap sample. The final model output is then the average of the $B$ predictions, this is illustrated in equation (3), with $\widehat{f_b}(x)$ being the model fit to the $b$'th bootstrap sample.

$$\widehat{f_{bag}^{*}}(x) = \frac{1}{B}\sum_{b=1}^{n} \widehat{f_b}(x) \tag{3}$$

Bagging is normally applied to deep decision trees, that is, trees with many nodes. Deep trees fit the training data very closely, which leads to them having a very small bias, but high variance. It follows that combining deep trees with bagging leads to a model with little bias and variance reduction, thereby providing a model that gets the best of both worlds, leading to good generalization to unseen data.

The one big problem is that the variance reduction works best when the individual models are independent, random forests are a special case of bagging where this problem is addressed, they are discussed in section 1.3.

## 1.3 Random forests

To overcome the problem of the sub models not being independent, Breiman (2001) introduced random forest, which apply the principle of bagging for deep decision trees, but for each subtree $m$ of the predictors are randomly chosen. When making predictions for a numeric response, $m = p/3$ is usually chosen -- where $p$ is the number of predictors in the data set. It follows that random forest can handle noisy or some uninformative predictors.

## 1.4 Out-of-bag error

Another benefit of bagging is how it estimates unseen prediction error without the need for cross validation, by using out-of-bag error. How this works, is that for each data case in the original data set, the squared error is calculated by taking the average of all the sub models that did not use that data point to train on. This way, the out-of-bag mean squared error, gives a more accurate prediction error estimate, given that the data cases are independent, per observation.

## 1.5 Importance scores

The biggest downside of bagging and random forests, is that the interpretability of the average of many trees, is not as easily interpreted, as an individual decision tree. To solve this problem, importance scores and partial dependence plots, were developed.

Importance scores measure how "important" the final model views each of the predictors. These scores are calculated by viewing each individual tree and summing the amount that all the splits, using the predictor of interest, reduced the model's mean squared error. Taking the average of these values over all the sub models per predicator, gives the importance score of the given predictor. In laymen terms, it is the average amount of error reduction over each sub model.

## 1.6 Partial Dependence Plots

To make a model, with a complex structure, more interpretable, partial dependence plots - as discussed by Molnaret al. (2023) - were also introduced. They show how a model's prediction is affected on average by different values of a given covariate.

More formally, for a covariate $X_s$, at the given value of $x$, we are interested in what output the model will give in general or what we expected the model will predict. It follows that the partial dependence is as given in equation (4), where it is the expected value of the output of the model over all possible values of the other covariates.

$$PD_s(x) = E\left[\hat{f}\left(X_1, X_2, \dots, X_{s-1}, x, X_{s+1}, \dots, X_p\right)\right] \tag{4}$$

Since the true distribution of the covariates are usually not known, the partial dependence must be approximated. This is done by finding the predicted value at each point in the training data, where the covariate of interest is replaced by the given value of $x$. The approximated partial dependence is given in equation (5).

$$\widehat{PD}_s(x) = \frac{1}{n}\sum_{i=1}^{n} \hat{f}\left(x_{1,i}, x_{2,i}, \dots, x_{s-1,i}, x, x_{s+1,i}, \dots, x_{p,i}\right) \tag{5}$$

For example, to see how the "latitude" covariate, $X_{latitude}$, affects how the random forest model predicts the number of livestock on a given 1 by 1 km plot, its partial dependence was calculated at sequential points in the range of latitude -- as given in the training set. Plotting these values yields the partial dependence plot in Figure 2. Note how it captures how the random forest models the non-linear relationship between latitude and the model's prediction of livestock.

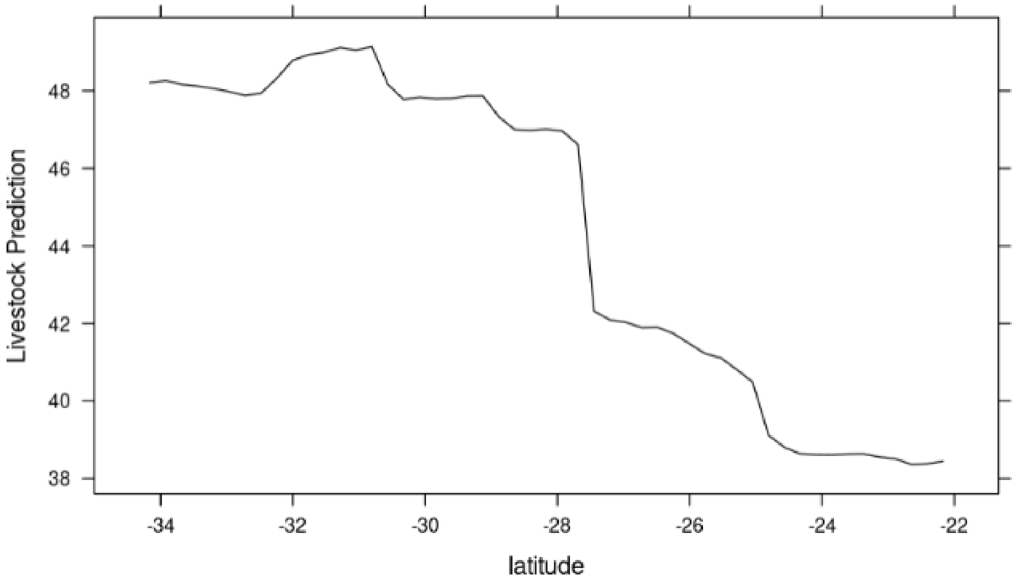



**Figure 2: Example of partial dependence plot**

Notice how the number of predicted livestock only changes by about 10, this due to how partial difference plots only consider the effect of one covariate. To make a prediction, the effect of all covariates gets aggregated to a final prediction. The very steep decrease is due to the nature of pdps, because they are an ensemble of many regression trees.

Partial dependence plots were made using the R package pdp, and its functions, partial and plotPartial. Where the partial function creates a partial object per predictor of interest, which is then plotted using plotPartial.

## 1.7 Spatial K-Fold Cross-Validation

In addition to out-of-bag prediction error, that was discussed in section 1.4, K-fold cross validation is also used to estimate prediction error. Here the data is split into $K$ different groups. This can be done in a variety of different ways, for independent data this can be split uniformly. In this study, the data was split into polygons based on their positions. These polygons were then uniformly sampled. Where all the points in a selected polygon are in that fold. This way, the points in given folds are not as close to points outside the folds, thereby reducing the dependence between the folds that were introduced. This was based on methods introduced by Valavi et al. (2018).

After splitting the data into folds, for each fold the model is fitted on all other points and the mean squared is calculated on the predictions of the values on the folds. The average of the MSE of the $K$ folds is then an approximation of the model's prediction error.

This was the method used to do feature selection and was also used as a performance measure. This was preferred over the out-of-bag prediction error, due to it reducing the dependence introduced by the proximity of the different data points.

Update: this technique was not used any more for the countries after South-Africa and Switzerland, due to it being very computationally expensive. The out-of-bag error was deemed to be sufficient.

# 2. Methodology

## 2.1 Data pre-processing

**Animal Data**

The main response was the sum of cattle, sheep and goats, referred to as the number of livestock, where all the values refer to the head counts of the livestock. For the livestock numbers, pigs and chickens were not considered. This was due to them having mostly very small numbers, except for a few places with very large numbers. This is most likely due to there being some large farms, where they feed the animals' food, which was predominantly bought and not grown on the land itself -- which is not in the scope of this study.

Update: In addition to this, three additional models were fitted for cattle, sheep and goats, individually.

**Dry Matter Productivity**

The dry matter data used was the total dry matter productivity (DMP) for each month of 2021. During exploratory analyses, it was found that having an initial DMP value and then just considering the changes in DMP each month, was the most informative.

For the South African data, the values at the tenth of each month was used as a proxy for the month's DMP.  For Switzerland, there were a lot of missing values, especially in their winter months. The remaining values were accounted for by taking the average DMP at the tenth, twentieth, and last day of the month, that was not present. The few remaining missing values were treated by removing them. This worked well because then when the values were needed, the closets alternatives were considered, since interpolation was used.

Update: For Australia 2022 and 2023's DMP values were also considered.

**Geographical data**

Slope data was used, with the gradients being encoded as is described in Table 1. The Elevation for 1 by 1 km plots were also available in metres. The coordinates for each of these variables were slightly different to the coordinates of the rest of the data set, to account for this interpolation was used. Interpolation is described in section 2.2.

*Table 1: Values corresponding to different gradients*

| Value | Gradient |
|---|---|
| 0 | 0 - 3 % |
| 1 | 3 - 9 % |
| 2 | 9 - 15 % |
| 3 | 15 - 30 % |
| 4 | 30 - 60 % |
| 5 | 60 + % |

**Weather Data**

The weather data that was used, was the daily rainfall in millimetres, and the daily maximum and minimum temperature in degrees Celsius. These values were available for each 5 by 5 km plot in the areas of interest, for each day of the year. In order to use these values alongside the 1 by 1 data, the values had to be interpolated. This was used for 2021.

Since considering all the days of the year and all three variable types would lead to over a thousand variables, these variables were converted into 36 monthly variables.

For the total rain and maximum and minimum temperatures, the average value of each month was calculated and used as covariates. The variable names were chosen as Rain.$m$, Max.$m$, and Min.$m$, where $m$ is the $m$'th month of the year. The water variable  was also included which represents "mm_TAW_in_the_first_meter",

## 2.2 Interpolation

When geospatial data is only available for, say, 5 by 5 km plots, and you need 1 by 1 km plots, interpolation can be applied to transform the geospatial in the correct format. The description of this method is based on the work of Molnar et al. (2023). In this subsection, $x$ , is a vector containing the longitude and latitude, and $z(\boldsymbol{x})$ is the value of the variable of interest at the point $\boldsymbol{x}$.

Since we have a square grid, we consider the four closest points and determine the value at a given point as the weighted average of the four closest points (say $\boldsymbol{x'_1}, \boldsymbol{x'_2}, \boldsymbol{x'_3}, \boldsymbol{x'_4}$), weighted on how far the four points are from the point that is considered, this is given in equation (6). The weights used are given in equation (6), which is inversely proportional to the square of the Euclidean distance between the points.

$$z(x) = \sum_{j=1}^{4} w_j(x) \cdot z(x_j) \quad (6)$$

$$w_j(x) = \frac{1}{d(x, x'_j)^2} \frac{1}{\sum_{k=1}^{4} w_k} \quad (7)$$

This technique is known as inverse distance weighted interpolation. Equation (7) uses four neighbours and parameter $p = 2$, so that Euclidean distance is used. This distance metric is appropriate, because the areas of interest are small enough for the curvature of the earth to have a negligible effect.

The formula has to be adjusted when you are working with extensive variables like total rainfall, where you have to adjust for the change of area by multiplying the final value by $\frac{Area_{new}}{Area_{old}}$. For South African total rainfall, this was $\frac{1}{25}$, and $\frac{1}{100}$ for Switzerland. This is to adjust for the fact that the magnitude of the total rainfall is additive. On the other hand, no adjustment is required for intensive variables, like the maximum or minimum temperature.

Note that extrapolation was not incorporated into the models, this would have only been relevant to coastline plots, which make up a small percentage of the overall dataset, the effect of this is expected to be very small. The process of merging different datasets with just interpolation was already very computationally expensive, so that adding extrapolation seemed unnecessary.

## 2.3 Feature selection

Although random forest are robust against uninformative predictors, the quality was tested using 5-fold cross-validation. Done with and without predictors. This was also used for the linear model. When predictors did not improve the cross-validation error, they were deemed uninformative and excluded from the model for both linear models and random forests.

Cross-validationcwas used for feature selection, because the observations were not independent, making the out-of-bag prediction errors weaker. This was done with a smaller data set, due to the immense computational power required for a full analysis. This methodology led to the removal of NDVI data.

Update: just the out-of-bag error was used for feature selection in countries after South-Africa and Switzerland.

## 2.4 Linear model fitting

The linear model was also applied to the data for comparison. 5-fold cross-validation was also applied here for feature selection. This was done for a comparison of results -- to test the stability of the random forest model as well the performance. Although the linear model is very effective at using the categorical variables, it still is flawed since it must have been trained on all the possible combinations of the variable, if it wants to make predictions that have these labels -- it cannot make predictions if it has never seen a label before. It follows that cross validation could not be done using the categorical variable for the linear model.

## 2.5 Random forest model fitting

The random forests were fitted using the R package ranger, this method is much faster than some of the older packages like randomForest. All the default values were used, except num.trees and importance.

For feature importance to be calculated, the parameter "importance" had to be set equal to the string "permutation", since we are considering the regression case.

**How were the tuning parameters picked**

The number of trees, num.trees, was chosen to be big enough for the cross-validation error to be consistently be small. The only downside for a larger number of trees is that the model takes longer to train and make predictions, since adding more trees will not lead to overfitting -- which is a very useful result for random forests. It follows that the final value of the tuning parameter was rounded up, 1 500 trees appeared to be sufficient.

The possible values for min.node.size and mtry were tuned using the spatial cross validation, but this did not lead to a significant improvement in the error. It follows that the "rule of thumb" methods were used, for consistency among the random forest models of different countries.

## 3. Results for South Africa

Both a linear model and a random forest, was applied to the South African data. The performance and insights drawn from the models are discussed in this section. South Africa has a very diverse geography, which means that it will test the model's ability to train despite heteroscedasticity -- data whose variance changes for different positions.

### 3.1 Linear model performance

Based on the training data, it was found that $R^2 = 0.6045$, this is not very high, especially when you consider that these values are from the data the model was trained. Furthermore, the spatial cross validation $R^2$ is 0.6042 for the linear model, which is very similar to the training error. These small $R^2$ values, are the reason a more complicated model was also fit to the data, since the model clearly picks up that there is a relationship between the response and the features, but it cannot account for the complex relationship between these variables.

Fyi: the insignificant variable is DMP.12 min and max temp 4 is also not significant.

### 3.2 Random forest performance

Fitting a random forest to the South African data yielded out-of-bag $R^2$ values of 0.8850, 0.8903, 0.8915, and 0.9042, for the number of total livestock, of total cattle, of total goats, and of total sheep, respectively. This implies that most of the variance is explained by the model, which is excellent, because the out-of-bag error is calculated on data points that the model was not trained on.

The total model under performs in this case, this can be because the total number of animal heads and does not account for the fact that some animals eat more than others.

For a more robust evaluation of the model, the spatial cross validation error was also tested, taking the dependence of the positions into account. The position of each point and the sub-models do not get to train with latitude and longitude values close to the points it is tested on.

Overall, there was only a marginal improvement when using the coordinate data. Due to the added interpretability of a model without coordinate data, the analysis also included a model without coordinates.

## 3.3 Variable performance

To compare how important different features were to calculate the response, we consider both the linear and random forest's importance, since using two methods with completely different model assumptions, can provide further insights.

**Variable importance scores**

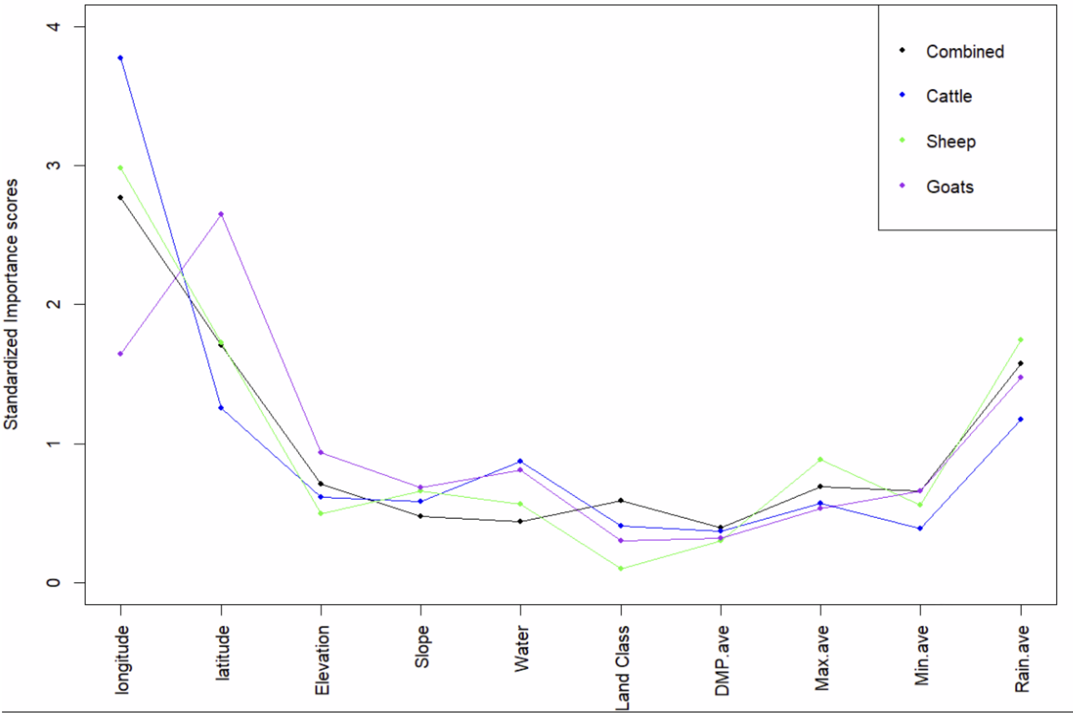



**Figure 2: Variable importance for South Africa**

The variable importance was performed as is described in section \ref{sec:Var Imp}. To make the results more interpretable, the scores were scaled, so that the mean score for all four variants of the models were equal. Since there are many features, all the monthly DMP, rain, maximum temperature, and minimum temperature scores were combined by taking the averaging the monthly mean and the maximum month's importance scores. These values are visualised in figure 2.

In general, all predictors seem to place a similar importance on similar covariates. The importance of elevation on the total number of livestock and cattle is very interesting, it is important to note that cattle are the most prevalent form of livestock, which is why its variable importance is so similar to the combined importance. Notice the emphasis on coordinate data, this was the main motivation behind fitting the models without coordinate data.

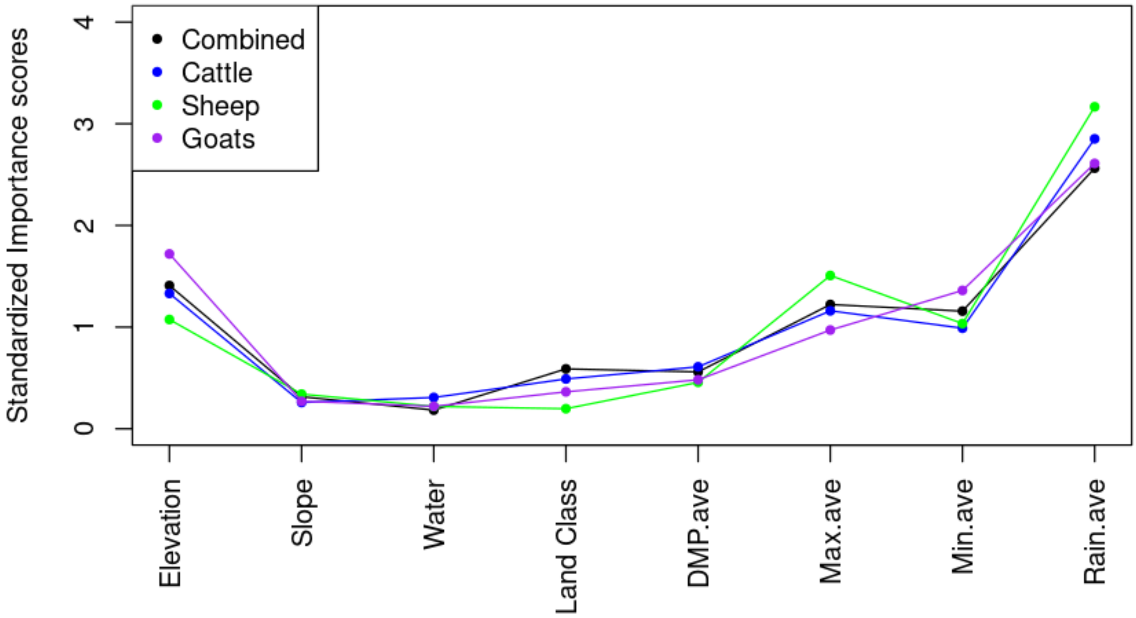



**Figure 3: Variable importance for South Africa no coordinates**

The model without coordinate data's importance scores are summarised in figure3, the same patterns are present. The features seem to affect the different responses very similarly in this plot, with the effects just being highlighted more now. Elevation and weather data is what the model emphasis the most, with rain being the main feature used by the models.

**Linear t-statistics**

To demonstrate the importance, that the linear model placed on each of the variables, a similar graph was drawn. The only difference is that the importance scores were replaced by the absolute value of the t-statistics of each variable.

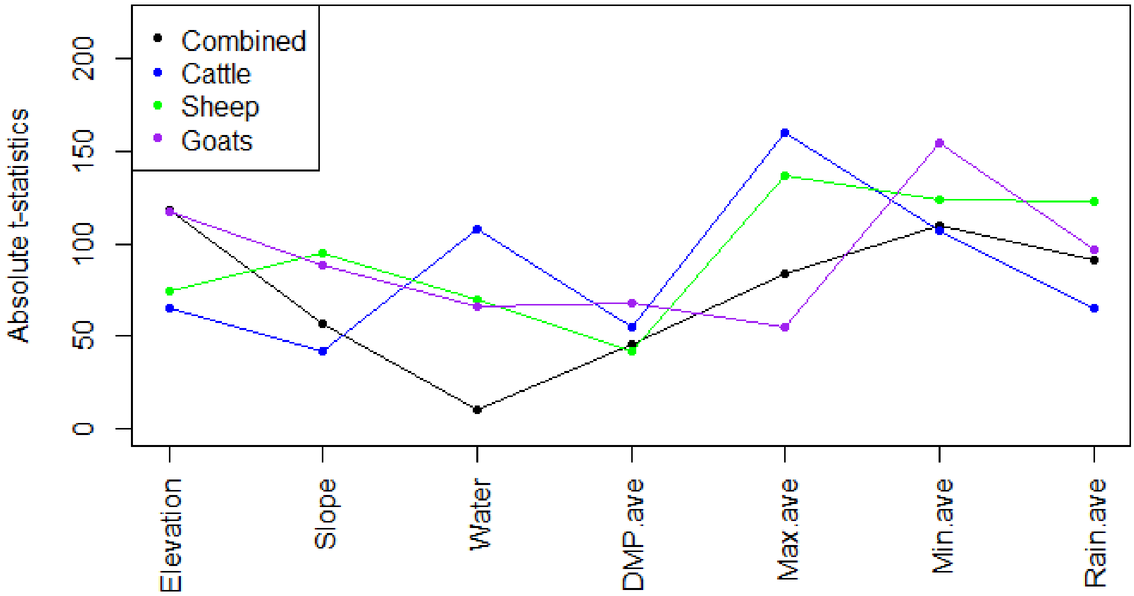



**Figure 4: Variable importance for the South African Linear model with no coordinates**

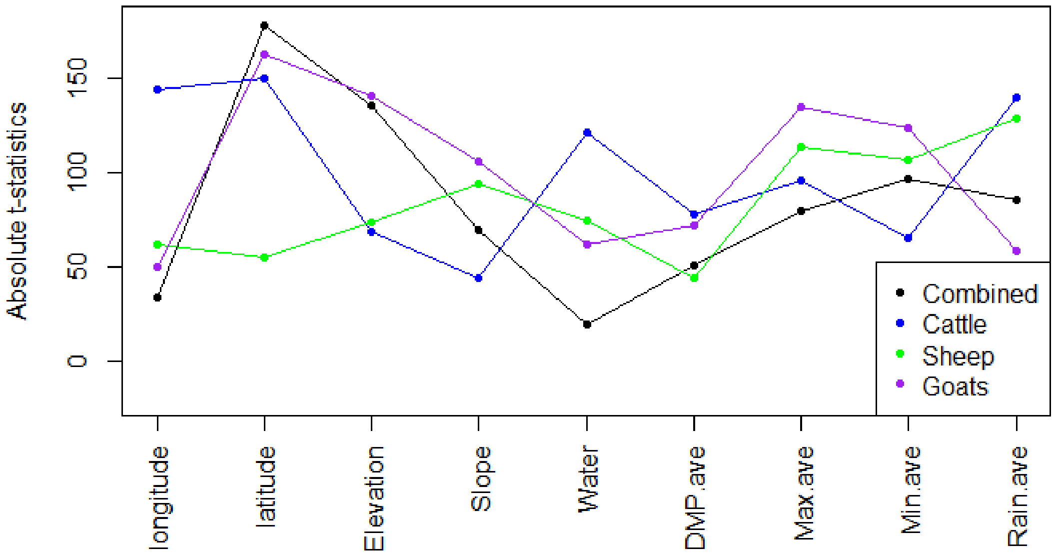



**Figure 5: Variable importance for the South African Linear model with coordinate data**

This is displayed graphically in figure 4. The linear model seems to put a more uniform importance on all variables. It also does not put as much emphasis on the elevation and rain data. For some further insights, the linear model was also fitted with the coordinates, the absolute t-statistics are displayed in figure 5. The model's emphasis seems to be shifted considerably - implying unstable relationships.

# 4. Results for Switzerland

Both a linear model and a random forest, were applied to the Switzerland data. The performance and insights drawn from the models are discussed in this section.

## 4.1 Linear model performance

Based on the training data, it was found that $R^2 = 0.6737$, this is not very high, especially when you consider that these values are from the data the model was trained. This small $R^2$ values, are, again, an indication that a more complicated model is appropriate.

An $R^2$ value of 0.6737, shows that the model has some potential, it can potentially be improved by modelling some interactions or by adding transformations of the given features as new features.

## 4.2 Random forest performance

Fitting a random forest to the Switzerland data yielded out-of-bag $R^2$ values 0.911, 0.907, 0.842, and 0.874, for the number of total livestock, of total cattle, of total goats, and of total sheep, respectively. This implies that the majority of the variance is explained by the model, this is excellent, because the out-of-bag error is calculated on data points that the model was not trained on. It makes sense that the total livestock is predicted the most accurately.

## 4.3 Variable performance

To compare how important different features were to calculate the response, we consider both the linear and random forest's importance, since using two methods with completely different model assumptions, can provide further insights.

### Variable importance scores

The variable importance was performed as is described in earlier sections. To make the results more interpretable, the scores were scaled, so that the mean score for all four variants of the models were equal. Since there are many features, all the monthly DMP, rain, maximum temperature, and minimum temperature scores were combined by taking the averaging the monthly mean and the maximum month's importance scores. These values are visualised in Figure 6.

In general, all predictors seem to place a similar importance on similar covariates. The importance of elevation on the total number of livestock and cattle is very interesting, it is important to note that cattle are the most prevalent form of livestock, which is why its variable importance is so similar to the combined importance.

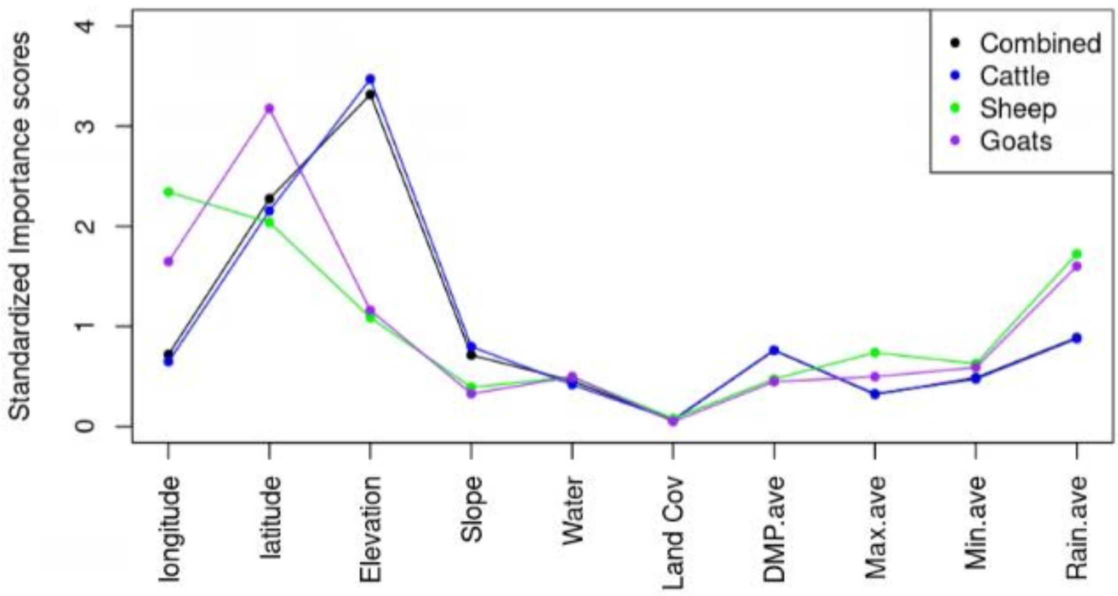



**Figure 6: Variable importance for Switzerland**

**Linear T-statistic**

To demonstrate the importance, the linear model placed on each of the variables, a similar graph was drawn, where the only difference is that the importance scores were replaced by the absolute value of the t-statistics of each variable. The values are given in Figure 7. The linear model put much more of an emphasis on DMP and water, and a lot less on emphasis on the position of the position of the individual plots. Land cover was not added, due to there being so many different possible land classes, and each of them having completely different values for each variable. Incorporating this into the graph was deemed to be above the scope of the illustration.

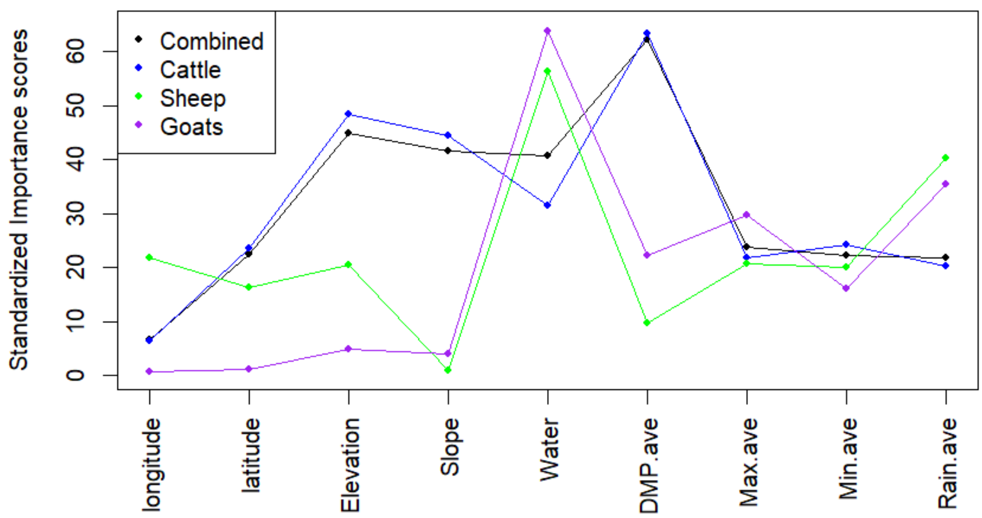



**Figure 7: Variable importance for Switzerland Linear model**

**Partial Dependence Plots**

Partial dependence plots provide a deeper insight into how models make predictions. The first one considered for Switzerland, is of elevation, which was deemed the most important. Its PDP is given in Figure 8, in which it appears that the number of livestock increases with the elevation until 1 000 metres is reached, after which it decreases substantially.

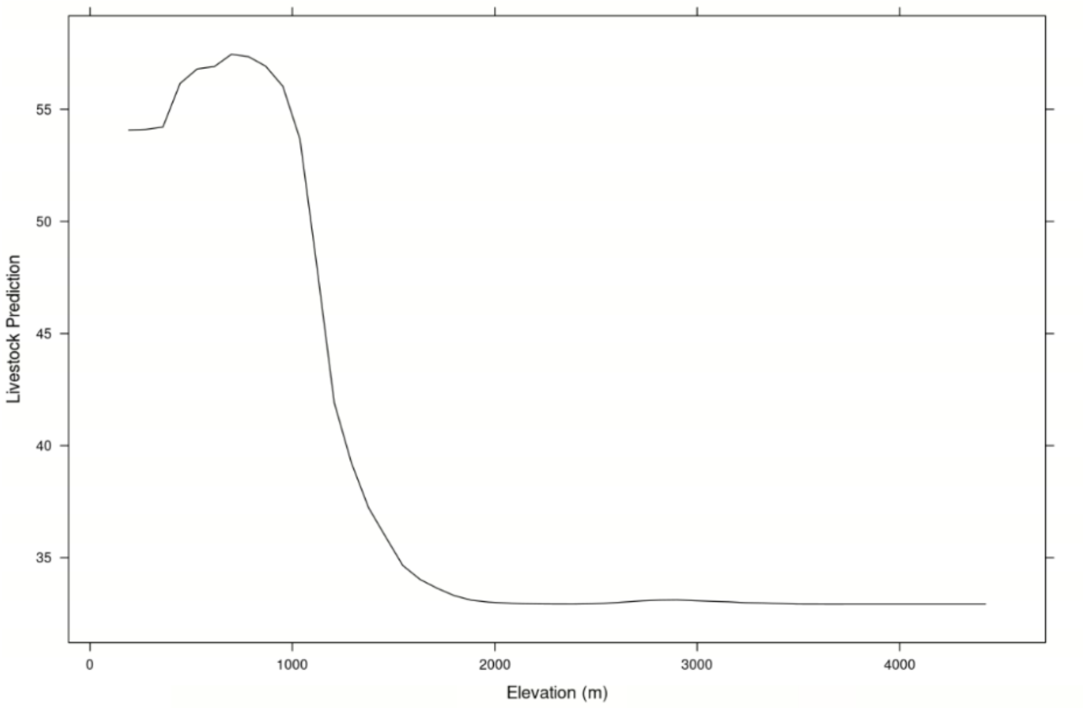



**Figure 8: PDP Elevation Switzerland**

The next interesting covariate, to consider, is the amount of rain in December. Its PDP is given in Figure 9. The relationship appears to be linear, with positive causality, as one would expect.

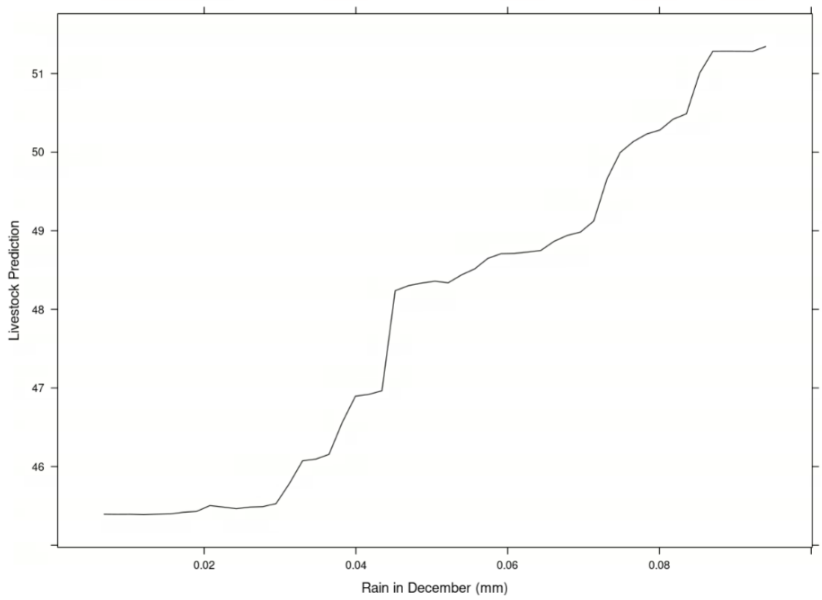



**Figure 9: PDP December Rain Switzerland**

## 5. References

Breiman, L., Friedman, J., Olshen, R.A., & Stone, C.J. (1984). Classification and Regression Trees (1st ed.).

Breiman, L., 2001. Random forests. *Machine learning*, *45*, pp.5-32.

Chapman and Hall/CRC.Hastie, T., 2009. The elements of statistical learning: data mining, inference, and prediction.

Gareth, J., Daniela, W., Trevor, H. & Robert, T. (2013), An introduction to

statistical learning: with applications in R, Spinger.

Molnar, C., Freiesleben, T., König, G., Herbinger, J., Reisinger, T., Casalicchio, G., Wright, M. N. & Bischl, B. (2023), Relating the partial dependence plot and permutation feature importance to the data generating process, in 'World Conference on Explainable Artificial Intelligence', Springer, pp. 456–479.

Valavi, R., Elith, J., Lahoz-Monfort, J. J. & Guillera-Arroita, G. (2018), 'blockcv: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models', Biorxivp. 357798.