DATA CURATION FOR MODELING TALL FESCUE BIOMASS DYNAMICS WITH DSSAT-CSM

M. B. Hanson
Auburn University

P. D. Alderman
Oklahoma State University

T. J. Butler
Oklahoma State University

A. Caldeira Rocateli
Oklahoma State University

The XXV International Grassland Congress (IGC) "Grasslands for Animal, Soil and Human Health" took place in Covington, Kentucky from May 14 through May 19, 2023.
Proceedings edited by the American Forage and Grassland Council and Published by Curran Associates, Inc ©2023

DATA CURATION FOR MODELING TALL FESCUE BIOMASS DYNAMICS WITH DSSAT-CSM

Henson, M.B.[1], Alderman, P.D[2]., Butler, T.J. [3], Caldeira Rocateli, A. [4]
Auburn University[1]; Okalamoma State University[2,3,4]

Key words: Tall Fescue; Data; Model; Biomass

Abstract

While models for predicting forage production are available to aid management decisions for some forage crops, there is limited research for a yield model designed specifically for tall fescue (*Schedonorus arundinaceus*). Therefore, our objective was to adapt an existing perennial forage model, the Decision Support System for Agrotechnology Transfer Cropping Systems Model (DSSAT-CSM) for predicting forage biomass of tall fescue in the southern Great Plains. To evaluate model performance, there must first be a high level of data manipulation and cleaning. In this project, a cohesive dataset combining biomass, weather, soil, and management data were structured into DSSAT standard file format to be used in future tall fescue crop modeling analysis.

Introduction

In the southern Great Plains, cool-season perennial forages serve an essential role as early season complements to winter annual forages. To achieve optimal forage yield and nutritive value of a pasture in a grazing system, it would be beneficial to have a model that would predict harvestable biomass. However, there are limited long-term cool-season perennial forage trials available, making it difficult to obtain sufficient input data for a dynamic crop model. To adequately evaluate a model's qualitative and quantitative accuracy, there must first be a high level of data manipulation and cleaning to compile a curated dataset. The objective is to produce a comprehensive documentation of a dataset developed for modeling tall fescue harvestable biomass in the southern Great Plains. Therefore, we have created a diverse dataset including biomass data obtained from the Noble Research Institute paired with corresponding weather data from the Oklahoma Mesonet (Brock et al., 1995; McPherson et al., 2007) and the University of Georgia Weather Network (Knox et al., 2020), as well as soil data from the Natural Resource Conservation Service Soil Survery Geographical database (NRCS SSURGO) (Soil Survey Sta, 2020). Through this process, a high quality dataset of biomass, weather, soil, and management data were then structured into a single condensed comprehensive dataset that was converted into DSSAT standard file format which has the potential to be used for crop simulation modeling of tall fescue.

Methods

Noble Research Institute Tall Fescue Experiments

Table 1.1: Summary data of sites in the dataset: location, latitude (Lat, decimal degrees), longitude (Long, decimal degrees), elevation (Elev, m), maximum temperature (TMAX) in ºC, minimum temperature (TMIN) in ºC, and seasonal cumulative rainfall (Rain, mm).

Site	Lat	Long	Elev	TMAX	TMIN	Rain
Ardmore	34.19	-97	266	23.8	11.7	879
Tifton	31.49	-83.5	118	26.8	15.6	785
Woodward	36.42	-99.4	625	23.7	9.8	410

Weather data is described in the summary table by location (Table 1.1). Weather data sources were the Oklahoma Mesonet (Brock et al., 1995; McPherson et al., 2007) and the University of Georgia Weather Network (Knox et al., 2020). The soil data for each site was obtained from the Soil Survey Geographical Database [SSURGO; Soil Survey Sta (2020)]. Management practices vary across sites when evaluating planting technique, fertilization, and utilization. The breeder sites are more similar to one another in management strategies compared to the agronomic site. Tall fescue harvestable biomass samples were collected at four study sites across three locations. The Ardmore location had two studies, an agronomic (AGR) study and a breeder (BRD) trial. The breeder studies consisted of Ardmore, OK (BRD) from 2012-2014, Tifton, GA from 2012-2013, and Woodward, OK from 2012-2014. The agronomic study, Ardmore, OK (AGR), conducted from 2015-2020, had more complex harvest frequencies and nitrogen treatments.

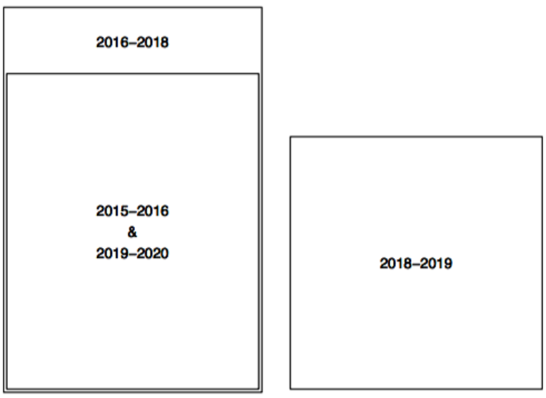
Agromonic Site

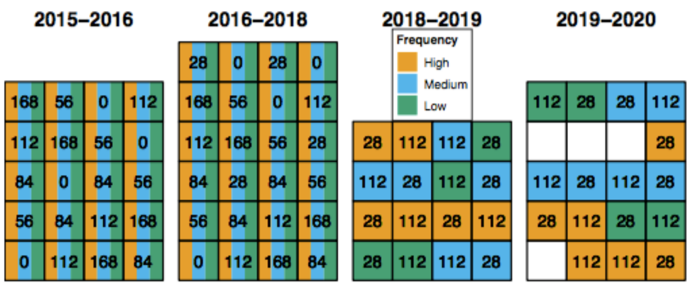
The Ardmore, OK (AGR) site was located on a Heiden Clay and was planted to a novel endophyte variety of tall fescue, 'Flecha'. Harvest frequency varied across AGR with high to low frequency as well as across seasons. At harvest, each plot was mowed to 2.5-cm stubble height and vegetation was removed. To measure the average canopy height (cm) the harvestable section was randomly measured with a meter stick. All forage yield data were recorded on a dry matter (DM) basis. Forage yield was measured as oven-dried clipped forage mass in kg ha^{-1} based on the harvested area.

The study was a nitrogen response trial starting with five levels of nitrogen in the form of ammonium nitrate (34-0-0, NPK) at the start of the experiment: 0, 56, 112, 168, and 224 kg N ha^{-1} each applied as a split rate. An additional N level at 28 kg N ha^{-1} was added starting in the fall of 2016 using the same management. These six levels of nitrogen, 0, 28, 56, 112, 168, and 224 kg N ha^{-1}, with a split application, were maintained until the fall of 2018. From the fall of 2018 through 2020, full rates of N were applied at two different levels: 28, and 112 kg N ha^{-1}. Figure 1.1 and Figure 1.2 illustrate the layout of the experiments across seasons. The experiments were conducted in two main phases: 2015 to 2018 and 2018 to 2020. The first phase was initiated in the fall of 2015 with twenty plots arranged into five rows and four columns (Fig. 1.2). Each column was treated as a block, within which, the five N levels were randomly assigned. Within each block, subplots were created by assigning the three different harvest frequencies in decreasing frequency from high to low (Fig. 1.2). For the latter two seasons of phase one (2016-2018), a set of four plots were added directly adjacent to the existing set of plots. This addition resulted in a total of 24 plots with six N levels for the seasons 2016-2018. In the 2018-2019 season, the second phase of experiments was established in an area adjacent to but separate from the existing plots from phase one (Fig. 1.1). The trial established in 2018-2019 had sixteen plots with four rows and four columns. In this second phase, two N levels were combined with three harvest frequencies to create six unique treatment levels. These unique treatment combinations were then randomly assigned to the sixteen plots (Fig. 1.2). In 2019-2020, the trial was moved back to the original area from the 2015-2016 season and the six treatment combinations were randomly assigned to 16 of the original 20 plots. Four of the original 20 plots were excluded from the study due to issues with pasture persistence and weedy encroachment.

Figure 1.1: (Left) Area diagram of the agronomic plots at Ardmore, OK (AGR) and which season each field was host of the study.

Figure 1.2: (Right) Plot layout of each field organized by season, harvest frequency, and nitrogen level.





Breeder Site Management

The breeder sites, Ardmore, OK (BRD), Tifton, GA, and Woodward, OK had similar management practices. The experiments were set up in 2011 using a randomized complete block design (Trammell et al., 2018). With a small plot cone-drill, 17 kg ha^{-1} of seed were put into 7 rows of clean, tilled seedbeds. The Ardmore (BRD) and Tifton locations were set up in a 1.5 m by 6.1 m plot with four replications. Unlike the other breeder sites, Woodward had five replications in a 1.5 by 7.6 m plot. For each of the breeder sites, there was a single application of ammonium nitrate at the rate of 46 kg N ha^{-1} at the time of sowing each season. Samples were collected leaving a residual of approximately 7-cm (Trammell et al., 2018).

Data Aquisition and Quality Control

Data for each location were collected and captured using manual data entry into a spreadsheet format, yet text and numerical modifications were performed using R code, to avoid manual typographical errors and to ensure reproducibility of the workflow. As each variable of the database was brought in, a screening process was performed using R version 4.1.2 (R Core Team, 2021) to clean and curate the data; any inconsistencies were standardized through a screening process correcting duplication and spelling to produce a harmonized dataset.

Weather Data

Weather data were obtained from the Oklahoma Mesonet (Brock et al., 1995; McPherson et al., 2007) and the University of Georgia Weather Network (Knox et al., 2020). Mesonet time series files were downloaded from the Oklahoma Mesonet website, imported into R and summarized to produce daily values for solar radiation (MJ $m^{-2}d^{-1}$), rainfall (mm d^{-1}), average relative humidity (percent), wind speed (km d^{-1}), and maximum and minimum temperature (ºC). For the Oklahoma sites, weather data were obtained from the following mesonet stations: Ardmore (ARD2) and Woodward (WOOD). The Georgia weather data were received from the Tifton (TIFT) site. Occasional gaps in daily weather records were filled using linear interpolation across days for all variables except rainfall. Missing rainfall data were assumed to be zero. The daily weather data and weather station metadata for each site were combined to a DSSAT standard format weather file, and were written using the function write_wth from the DSSAT package (Alderman, 2020; Alderman, 2021).

Soil Data

Soil names were provided in the tall fescue datasets corresponding with each location. The soil names from each location were used to pull the soil profiles from the SSURGO (Soil Survey Sta, 2020) for each location through a custom utility function written for this purpose. Soil input data were written to the DSSAT standard soil file format using the write_sol() function from the DSSAT version 0.0.4 R package (Alderman, 2020; Alderman, 2021).

Results and Discussion

Data File Description

This paper describes the process of compiling one comprehensive dataset in the form of DSSAT standard format for observed biomass (FileT), management data (FileX), weather data (.WTH), and soil data (.SOL) from a range of environments that will permit further model development. Variable descriptions for weather files are provided in the dataset. Similarly, soil entries are stored in a DSSAT-formatted soil file, using the extension SOL, which stores whole-profile and layer-specific soil variables. File names are specified as an eight-digit code that is unique to each location, year and management combination followed by a three-digit file extension of either U2X (FileX) or U2T (FileT). The FileT contains columns for treatment number (TRNO), date of collection (DATE), and harvestable biomass kg ha^{-1} (FHWAH). The FHWAH column contains data of harvested biomass which excludes a 7-cm residual stubble for breeder sites and a 2.5-cm residual stubble for the agronomic site. Within the FileX, the treatment number, within TREATMENTS, corresponds to TRNO of the FileT; the ID_SOIL column, in FIELDS, links the soil type in the soil data file, and WSTA, in FIELDS, links to the corresponding weather file. The PDATE and HDATE columns provide the planting and harvest dates, respectively, for the specific location, year and management combination.

Conclusions

There has been limited research in the area of tall fescue modeling. Because there are few long-term tall fescue trials that provide adequate characterization needed for modeling, we developed a comprehensive dataset that can be used for future modeling of tall fescue. For this project, biomass data were provided by the Noble Research Institute and included four experiments across three dierent locations, Ardmore, OK, Tifton, GA, and Woodward, OK. This paper provides an exhaustive description of each site, and the management and sampling practices conducted. The paper also serves to explain the process by which weather and soil data were obtained and manipulated. Soil names from each location were used to pull the soil data from the NRCS-SSURGO (Soil Survey Sta, 2020). Daily weather data and weather station metadata were received from the Oklahoma Mesonet and the University of Georgia Weather Network. These inputs were combined for three locations and multiple growing seasons to compile a curated dataset. The dataset documented here provides DSSAT standard format for observed biomass (FileT), management data (FileX), weather data (.WTH), and soil data (.SOL) from a range of environments that will permit further model development, parameterization and evaluation.

Acknowledgements

I would like to thank my advisor Dr. Phillip Alderman.

I would also like to extend my sincere appreciation to Twain Butler, Sindy Interrante, Mike Trammell, Erin Hubbard, and Jimmy Stein for providing data, and the Noble Research Institute, LLC, Ardmore, OK for funding my research.

References

Alderman, Phillip D. (2020). "A comprehensive R interface for the DSSAT Cropping Systems Model". In: *Computers and Electronics in Agriculture* 172, p. 105325. issn: 0168-1699. url:http://www.sciencedirect.com/science/article/pii/S0168169919323075.

— (2021). *DSSAT: A Comprehensive R Interface for the DSSAT Cropping Systems Model*. R package version 0.0.4. url: https://CRAN.R-project.org/package= DSSAT.

Brock, Fred V et al. (1995). "The Oklahoma Mesonet: a technical overview". In: *Journal of Atmospheric and Oceanic Technology* 12.1, pp. 5–19.

Knox, Pamela et al. (2020). "The University of Georgia Weather Network: Providing 30 Years of Data Products and Applications to Southeastern Climate Data Users". In: *100th American Meteorological Society Annual Meeting*.

McPherson, Renee A et al. (2007). "Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet". In: *Journal of Atmospheric and Oceanic Technology*24.3,pp.301–321.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. url: https://www.R- project.org/.

Soil Survey Sta, Natural Resources Conservation Service, United States Department of Agriculture (2020). *Soil Survey Geographic (SSURGO) Database for Oklahoma. Available Online.*Accessed:2020-08-13. url:https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/?cid=nrcs142p2_053627.

Trammell, Michael A et al. (2018). "Registration of 'Chisholm'Summer-Dormant Tall Fescue". In: *Journal of Plant Registrations* 12.3, pp. 293–299.