

MODEL DESCRIPTION

Overall Model Structure

SAVANNA is a spatially explicit, process-oriented model of grassland, shrubland, savanna, and forested ecosystems. Versions 1 and 2 were developed to study a nomadic pastoral ecosystem in Kenya (Coughenour 1992). Version 3 was developed for Elk Island National Park in Alberta (Buckley et al. 1995), and Version 4 was developed and used for the Pryor Mountain Wild Horse Range (Coughenour 2000), northern Colorado (Weisberg et al. 2002), northern Australia (Ludwig et al. 2002, Leidloff et al. 2001), South Africa (Kiker 1998), Tanzania (Boone et al. 2001), and Inner Mongolia (Christensen 2001). A more developed Version 4 (f) was used for Rocky Mountain National Park (Coughenour 2000, Pienetti 1999, Weisberg and Coughenour 2003). Here, version 5a is used. SAVANNA is normally used to simulate ecosystems at landscape through regional spatial scales over 5-100 year periods. The model is comprised of site water balance, plant biomass production, plant population dynamics, litter decomposition, and nitrogen cycling, ungulate herbivory, ungulate spatial distribution, ungulate energy balance, and ungulate population dynamics, predation, and predator population dynamics submodels (Figure 4A). The time-step of the model is one week.

SAVANNA has a hierarchical spatial structure (Figure 4B). It is spatially explicit (i.e., it is sensitive to spatial position) at the landscape scale and it is spatially inexplicit at patch scales. The mosaic of grid-cells covers landscapes or regional-scale ecosystems. Grid-cell size is scaled to the spatial extent of the simulated spatial domain, so as to maintain a sufficient degree of realism in capturing landscape-scale heterogeneity, while also maintaining a computationally feasible total number of grid-cells.

Within each grid-cell the model simulates three vegetation patch types or “facets.” These are defined by the fractional covers of herbaceous plants, shrubs, and trees. Since shrub and tree covers are simulated variables, facet cover is a dynamic outcome of woody plant growth and mortality. Facet locations are not modeled, only the fractions of grid-cells that are covered by the facets. Within each facet, the model simulates plant growth and soil water budgets. Variable results are scaled-up to the grid-cell level by multiplying the variables by the fractions of the grid-cell area covered by each facet, and summing the products over all facets.

The area of land within each grid-cell that is covered by trees or shrubs varies in response to changes in tree and shrub numbers and sizes. As a consequence, direct competitive or facilitative interactions among established plants of these three life-forms are limited to patches where established plants are rooted. Trees and herbs, for example, interact directly on tree-dominated patches but do not interact on herbaceous-dominated patches except during the establishment phase. Tree and shrub vegetation covers are defined in terms of rooted area, i.e., the area of ground which corresponds to the exploited soil volume. Cover is defined in terms of rooted area mainly because soil water and nutrient budgets are computed on a soil volume basis.

The vertical spatial structure of the model is defined by soil and plant canopy layers. The soil is divided into three layers. The top soil layer is a zone of potential bare soil evaporation as well as root uptake. The second layer is generally the deepest layer that is exploited by herbaceous roots. The bottom layer is generally occupied only by shrub and tree roots. Plant canopies are organized into herb, shrub and tree strata. In turn, each of these are divided into three substrata to compute light intensity.

Weather

SAVANNA is driven by monthly weather data. Each month, linear regressions are performed between precipitation and elevation using data from all weather stations in each spatial precipitation domain. A study area may contain more than one precipitation domain if it is likely that precipitation responds differently to elevation in each domain. If the correlation coefficient (r^2) is >0.2 , the slope of the line (i.e., the rate of increase in precipitation with elevation) and the elevation difference is used to estimate precipitation for the weather station as if it were at the same elevation as the point in question. The elevation-corrected weather station data are then spatially interpolated using inverse distance weighting on the six nearest stations with data. This procedure is followed for each grid-cell. Temperature for each grid-cell is similarly estimated from the temperatures at the nearest weather stations by applying adiabatic lapse rates to the station temperatures to estimate temperature at the same elevation as the point in question. The corrected temperature data are spatially interpolated using inverse distance weighting.

Stochastic weather data are created by randomly sampling from the weather files. Each simulated year, a random year of data is drawn from the files, and an additional amount of random variation is added to the data. The sampling and addition of variance are done in such a way that data from all weather stations are affected together, thus preserving the spatial pattern in the original data.

Atmospheric water vapor content is calculated from relative humidity and temperature of the main base station. Relative humidity is calculated from the water vapor concentration, and the temperature at each landscape location. Solar radiation is calculated from monthly cloud cover, latitude, and day of year, correcting for slope and aspect using the methods described by Nikolov and Zeller (1992). Cloud cover is estimated from monthly precipitation and water vapor content.

Water Budget

The water budget at each site is calculated from precipitation, interception, runoff, runoff, infiltration, deep drainage, bare soil evaporation, and transpiration. Incoming precipitation is partitioned between snow and rain depending upon air temperature. Rain interception amounts (mm) are calculated based upon litter biomass, herbaceous biomass of each herbaceous plant type, and leaf area indices (LAI) of each woody plant type. Fraction of snow intercepted by woody plants increases in direct proportion to LAI. The rate of snow melt depends upon temperature and solar radiation. Snowmelt water is treated the same as liquid precipitation with

regards to infiltration and runoff. Bare soil evaporation is simulated using the Ritchie (1972) model (Wight and Skiles 1987). Evaporation also increases under low plant and litter cover. The curve number method of the USDA Soil Conservation Service (now the Natural Resource Conservation Service) is used to calculate runoff (Wight and Skiles 1987). Runoff depends on daily rainfall, the quantity and distribution of water in the soil relative to water holding capacity, and the condition curve number for the soil. Runoff curve numbers are based on range site classifications and range condition such that runoff increases under low plant and litter cover. A higher runoff curve number is specified for fair range condition.

Infiltrated water is routed to three soil layers using a simple “tipping bucket” approach that drains water in excess of field capacity to subsequently deeper layers. The amount draining out of the lowest layer is considered to be deep drainage. Baseflow is derived from deep drainage (the amount of water draining from the bottom soil layer when it is in excess of field capacity). The baseflow water flows into the stream at a rate calibrated to match general streamflow patterns. Surface runoff, computed as described above, also flows into the stream. If the landscape redistribution function is being used, only runoff from the lowest elevation band flows into the stream. The output is total streamflow volume for that watershed over time. A physically-based approach to modeling ground water distribution and water tables is beyond the scope of this model. Instead, water table depth is empirically specified by month for each soil type or vegetation/habitat type where the ground water is shallow enough for plant water utilization.

Microclimate and Light

The microclimate submodel simulates wind speed profiles and aerodynamic resistances to water vapor diffusion. These depend upon canopy height, and reference height wind speed. Potential evapotranspiration (PET) and transpiration are calculated using the Penman-Monteith formula (Monteith 1973). PET is derived using Penman-Monteith with zero stomatal resistance. Transpiration is calculated from canopy-scale stomatal conductance. PET and transpiration are affected by radiation, temperature, humidity, and wind speed.

The light submodel simulates shading within and among plant canopies. On tree-covered facets, incident radiation first passes through the tree canopy, then the shrub understory and finally the herbaceous understory. Light extinction follows an exponential decay function (Beer’s Law) which depends upon leaf area index and a light extinction coefficient. The model tracks the relative heights of woody plants in different age classes, such that small shrubs and trees are more shaded than large shrubs and trees.

Plant Biomass Production

The net primary production (NPP) submodel simulates plant biomass flows and dynamics. Plant biomass production is affected by light, water, temperature, nitrogen and herbivory. The NPP submodel is explicitly linked to the water budget submodel through transpiration and plant water use.

Potential net photosynthesis is calculated as

$$P_{sl} = PAR / \sqrt{1 + (\alpha \times PAR / P_{smx})^2}$$

$$P_s = P_{sl} \times E_{fwp}(SWC) \times E_{ftp}(T_{day}) \times E_{fnp}(P_{nb})$$

where P_{smx} is maximum net photosynthesis rate, PAR is photosynthetically active radiation, α is the initial slope of the light response curve, and P_{sl} is light-limited photosynthetic rate. The initial slope α is also equal to the light use efficiency, and the light response asymptotes at saturating light intensities. The 0-1 function E_{fnp} describes how P_s decreases as plant nitrogen concentration (P_{nb}) decreases. The 0-1 function E_{fwp} reduces P_s due to water stress. SWC is volumetric soil water content, weighted by root distribution in the three layers. E_{ftp} is a bell-shaped temperature effect, generated from optimum, minimum, and maximum temperatures for photosynthesis.

After potential photosynthesis is calculated, stomatal conductance is calculated based upon the equation of Ball et al. (1997), in which stomates open wider as photosynthesis (P_s) and humidity (H) increase, and as atmospheric CO_2 (C) decreases.

$$g_s = e_0 + e_1 \times (P_s \times H / C)$$

From stomatal conductance, atmospheric water vapor pressure deficit, windspeed, and radiation, the model uses the Penman-Monteith approach to calculate transpiration. From transpiration and net photosynthesis, the model calculates water use efficiency.

Root water uptake to supply transpiration demand is allocated among soil layers in proportion to the products of available water and root biomass among layers. The water demands for competing plants are summed, and if total demand exceeds supply, water is allocated among species in proportion to demands. Photosynthesis is then recalculated from water use using the water use efficiency calculated above. Thus, plants rooted in the same soil layer compete for water, but when water is limiting, it is allocated among species in proportion to their demands. The demand for water in a layer thus depends upon transpiration rate, and the proportions of roots and water in that layer.

Photosynthate is allocated between roots and shoots to preserve a functional balance in herbaceous plants, and allometric relationships in woody plants. Phenology affects shoot:root partitioning, with an increased fraction being diverted to roots in later phenophases. Allocation to roots is influenced by nutrient and water availability, with slightly more allocation to roots when soil resources are limiting. In woody plants, allocation varies with plant size to preserve the proper allometric relationships. Herbaceous biomass is allocated to leaves, stems and roots. Woody plants allocate to leaves, fine branch, coarse branch, fine root, and coarse root. The allocation to various tissues in woody plants allometrically depends upon plant size..

Plants maintain a pool of labile carbon, largely consisting of carbohydrate carbon, but also including the carbon in metabolic proteins, amino acids, and other labile compounds. Early

regrowth in the spring or following grazing is supported by this pool until growth demands can be met by photosynthesis. The labile carbon pool is regulated in relationship to phenology. The pool is replenished towards the end of the growing season to ensure plants have a reserve to draw upon when they break dormancy the following spring. Conversely, during early regrowth, the labile pool can be drawn down to minimal levels.

Respiration is separated into growth and maintenance components. Maintenance respiration is proportional to nitrogen concentration (Ryan 1991), and it increases with temperature with a Q₁₀ of 2 (every 10 degrees it doubles). Maintenance respiration cannot exceed current photosynthate in herbaceous plants, and 50% of current photosynthate in woody plants. Growth respiration is assumed to be 25% of the biomass carbon that is synthesized

Plant tissues senesce at rates affected by water stress, temperature, and phenology. Tissues turn over at a nominal rate that reflects maximum longevity. Shoot tissues (leaves and stems of herbs and shrubs, leaves of trees) are transferred to "standing dead" upon senescence. Then the standing dead falls to "litter" on the soil surface. Dying roots are transferred directly to belowground "litter".

Woody Plant Population Dynamics

The woody plant population model represents six age, size, or age/size classes of plants. The recommended approach, and the one taken here, is to simulate age classes, and allow plant size to vary within each age class. Alternatively, size classes may be simulated where plants are promoted irrespective of age, or plants may be promoted to subsequent age/size classes if they reach a certain age, and they are also of a certain size.

Establishment rate is affected by water, temperature, and woody cover. As cover increases, establishment rate declines to reflect effects of shading and competition for space and soil resources. Woody establishment is also affected by herbaceous biomass, to represent competition between seedlings and herbaceous plants. Because woody establishment only occurs on the herbaceous (non-woody) facet, when a closed forest is simulated, establishment only occurs in gaps or openings resulting from tree death.

Plant mortality occurs at age or size-specific nominal rates related to characteristic plant life spans and survivorship curves. Water stress mortality is added to the nominal mortality. Sapling mortality (age/size class 1) can be caused by cold temperatures. Slow growth, expressed as mm of diameter increment per unit time, is a general indicator of stress that can be used to calculate a mortality rate. Self-thinning can be simulated using slow growth mortality, since as trees become larger they increasingly compete for limited resources (water, nitrogen, light) thus reducing individual tree growth rate.

Browsing and fire both result in losses of stem and branch biomass, thus reducing plant size, or reducing the rate of increase in plant size. Fire or heavy browsing can kill large branches, as well as the entire aboveground structure, while leaving the plant alive belowground. Heavy browsing of plant tissues, particularly stems and current annual growth, may contribute to plant

mortality due to associated stresses of loss of meristems, branch breaking, and debarking. This source of plant mortality is expressed as a function of the fraction of the shoot biomass or current annual growth that is eaten. Separate rates are given for saplings (first age class) and for mature (subsequent age class) plants. Seedlings of woody plants are usually assumed to be more susceptible to browsing than older plants.

Herbaceous Plant Population Dynamics

A dynamic index of basal cover is simulated, measured in units of g m^{-2} of maximum shoot biomass. The logic of this is that as basal cover increases so does the maximum shoot biomass because shoots reach a maximum size, and basal cover is determined by the number of shoots. Maximum shoot biomass is a measurement that is readily understandable. Vegetative (asexual) reproduction is simulated by adding to maximum shoot biomass. Maximum shoot biomass increases when root biomass becomes large relative to potential shoot biomass. The maximum shoot biomass decreases when root biomass is low relative to maximum shoot biomass. Shoot biomass production in a given year cannot exceed the maximum shoot biomass. In this way, stress due to climate or overgrazing may reduce root biomass relative to potential shoot biomass. This causes a decrease in maximum shoot biomass, and thus productive potential. Under good growing conditions roots increase, and the maximum biomass and thus productive potential recovers. This simulates time-lags of vegetation recovery due to constraints on the rate of vegetative reproduction.

An exogenous seed rain is assumed to replenish herbaceous plant populations (basal area). The composition of the seed rain is assumed to be invariant, and is usually assumed to have the same composition as initial conditions. The seed rain basis is expressed in the units of grams of shoot biomass per m^2 . These basis values of g m^{-2} shoot biomass are then multiplied by the exogenous seed-based reproduction rates (grams of new potential shoot biomass produced per gram of basis shoot biomass, $\text{g g}^{-1} \text{yr}^{-1}$) to derive a rate in $\text{g m}^{-2} \text{yr}^{-1}$ of new potential shoot biomass. Similarly, an endogenous seed-based reproduction rate is modeled by multiplying an endogenous reproduction rate ($\text{g g}^{-1} \text{yr}^{-1}$) by the current potential shoot biomass (g m^{-2}) to give $\text{g m}^{-2} \text{yr}^{-1}$ of new potential shoot biomass.

Litter Decomposition, Nitrogen Cycling, and Soil Organic Matter Dynamics

A decomposition and nitrogen cycling submodel simulates the breakdown of dead plant materials and animal feces and urine, microbial respiration, nitrogen mineralization, and the formation and turnover of soil organic matter (SOM). The decomposition submodel is based upon the CENTURY model (Parton et al. 1987, 1993). Dying plant tissues are divided into fast (metabolic) and slowly (structural) decomposing litter components, as well as fine and coarse branches and roots. The fractionation between metabolic and structural litter depends upon the ratio of lignin to nitrogen in the dying tissues. As lignin increases and nitrogen decreases, the fraction partitioned to structural litter increases. The two litter fractions are decomposed at rates that depend upon soil moisture and temperature, with faster decomposition rates under warm, wet conditions (Parton et al. 1987). A fraction of the decomposed carbon is respired by microbes. Nitrogen is released during decomposition to mineral forms that can be taken up by the plants or

incorporated into the active SOM pool. Three SOM pools having different turnover rates are simulated. Decomposing metabolic litter flows to active SOM and respiration, while decomposing structural litter flows to the active and resistant pools and respiration. Decomposing active SOM flows to the intermediate and slow turnover SOM pools. Decomposing intermediate SOM flows to active and slow SOM pools, and respiration. Decomposing slow pool SOM flows to the active pool, and respiration. Inorganic nitrogen (ammonium and nitrate) is released upon decomposition if substrate nitrogen concentration is higher than the N concentration of microbial biomass. Mineral N is taken up from the soil if substrate nitrogen concentration is lower than microbial N concentration. A small fraction of N mineralization is assumed to be denitrified and lost to the atmosphere. Parameters for the decomposition model are taken from Parton et al. (1987, 1993).

Nitrogen consumed by herbivores is recycled as either urine or feces. It is assumed that 50% of N flows to feces and 50 % to urine. The undigestible fraction of carbon intake flows to feces. Fecal C and N are partitioned to structural and metabolic litter. Urine N flows to inorganic N (ammonium and nitrate).

Nitrogen enters the system through atmospheric wet and dry deposition and biotic fixation, and it leaves the system through denitrification and the volatilization of ammonia from urine. Denitrification losses are simulated as a 5% fraction of N gross mineralization, as in Parton et al. (1987, 1993). Nitrogen lost to urine volatilization is assumed to be 20% of urine N. Biotic fixation is modeled as a function of annual precipitation, as in Parton et al. (1987, 1993).

An added amount of N input to riparian and water subsidized communities is necessary to support observed levels of plant growth and organic matter accumulation in such habitats. This N likely is transported into the system as N dissolved in ground or stream water. The N would be assumed to originate from precipitation, and leachate from upstream soils.

Soil carbon and nitrogen are initialized by providing values for %C and %N (weight) for the top 20cm. These are converted to gC/m² and gN m⁻² using soil bulk density. The data may be tied to habitat types on the vegetation input map, or to a map of soil fertility classes, which may or may not be identical to the soils map. Soil C is initially split among active, intermediate and slow pools in a 2:59:39 split. The C:N ratios of the 3 pools are initialized, for each vegetation type, to values that have been predicted in long runs of the CENTURY model with similar vegetation (eg. herbaceous, deciduous shrub, deciduous forest, coniferous forest).

Plant nitrogen uptake is represented using an equation that specifies a maximal uptake rate per gram of root tissue. The uptake rate is reduced at low soil N concentrations using a Michaelis-Menton function. Only mineral N is taken up (i.e., the N released during decomposition of litter and soil organic matter, assumed to be ammonium and nitrate). Competition for soil N is calculated when the total plant demand exceeds supply. As with water, uptake is then allocated among competitors in proportion to demands.

Plant nitrogen is regulated within bounds by curtailing uptake as plant N concentration increases to the maximum value. Nitrogen in green leaves is calculated based on phenology. In

most plants, young leaves have a higher N concentration than old leaves. N in dying plant tissues is transferred to litter based on a specified N concentration of dying tissues. Because the N concentration of dying tissues is usually lower than that of the live plant, N is effectively conserved through retranslocation to roots when leaf tissues die.

Fire Responses

Plants respond to fire by being combusted, by being killed, and by either germination or resprouting of new plants. Fire severity maps are read in for each fire, at the simulated time that they occurred. These maps contain fire severity codes for each grid-cell. For herbaceous plants, a fraction of aboveground tissues lost to combustion, and a fraction of plants dying are specified for each fire severity class. For woody plants, the effects of fire are specified as fractions of pre-fire cover or number that will be present post-fire, for each fire severity class.

To represent the 1988 fires, the burned area map (Despain et al. 1989) was used, with codes for canopy burn, mixed burn, non-forested burn, and unknown burn type outside the park. Higher elevation areas burned in August of 1988 while low elevation areas (the northern range) burned in September.

To represent herbaceous responses it was assumed that 80% of herbaceous roots are in the top 20 cm and 100% of meristems are in the top 10 cm. Consequently, for canopy burn, with a mean char depth of 2.5 cm, 10% of the roots might be charred and 20% of the meristems. Similarly for non-forested burn, if the mean char depth is 0.5 cm, then 2% of root and 4% of meristems would be charred. For mixed burn, the mean char rates of canopy, surface and unburned was applied. For unknown burn type, it was assumed half was canopy burn and half was mixed burn. It was assumed that post fire canopy cover was 5% of prefire cover for trees and *Vaccinium* shrubs, due to germination of new plants. For sagebrush and deciduous it was assumed that post fire plant numbers were 75% of prefire numbers due entirely to germination of new plants from seed. For willows it was assumed postfire cover was 5% of prefire cover, and 10% was due to germination and 90% was due to resprouting.

Carbon and nitrogen in combusted tissues were assumed to be 98% volatilized, with the rest being returned to the soil in ash.

Herbivory Submodel

Forage intake rate responds to forage biomass according to a functional response (Holling 1959, Caughley and Lawton 1975). As forage biomass increases, forage intake rate increases, with intake rate eventually reaching a maximal value (Figure 1A,B, Table 1). A Type-I response is used, which linearly increases from 0 at 2 g/m² forage biomass up to the maximal intake rate at some saturating value of forage biomass. Forage intake rate declines with decreasing forage quality due to increased rumen retention time (Feist 2000). The function was set for both bison and elk, to reduce intake rate to 70% of normal at a mean DMD of 0.3, increasing to 100% of

normal at a mean DMD of 0.6. Except for early in the growing season, forages are most often mixtures of live and dead leaves and stems, so digestibility is rarely maximal.

Forage intake rate is affected by snow (Figure 1C,D, Table 1). Above a certain snow depth, intake rate declines linearly with increasing depth. Fine-grained topographic variation leads to the occurrence of wind-blown areas in the particular landscapes of Yellowstone elk and bison winter ranges. This fine-grained variation is not represented in the model. For this reason, the effect of snow was set more liberally than data would suggest.

Forage intake rate is partitioned among plant types and tissue types as affected by preferences and relative availabilities. Diet composition is determined by using preference weights, similar to Ellis et al. (1976). Preference weights are multiplied by biomass densities and the products are summed and normalized to give a sum of fractions totally 1.0. Preference weights were calibrated to match observed diets (Table 2).

Herbivore Energy Balance

The herbivore energy balance submodel simulates body weight of the mean animal of each herd or subpopulation, based on differences between energy intake and energy expenditure. The energy balance model is partly based upon the models of Coppock et al. (1986) and Hobbs (1989). Energy intake depends on forage biomass intake and forage digestibility. Expenditures depend on body weight and travel patterns. The body weight of the mean animal is used to derive an animal condition index, which affects ungulate population dynamics.

Metabolizable energy intake from forage consumption is the product of kg total forage intake per animal per day, the mean digestibility of the forage, the gross energy content of digestible plant matter, and its metabolizability. Dry matter digestibility values are specified by plant functional group and tissue type. The DMD of live tissues is assumed to be positively correlated to leaf tissue nitrogen (protein) concentration, which declines from a maximum value in early growth stages to a minimum value mid-way through the growing season. Leaf nitrogen and thus DMD are also increased in the regrowth following herbivory,

Energy requirements consist of a "base cost" (not basal metabolic cost) metabolic energy demand per unit body weight per day plus optional added energetic costs of travel, thermoregulation, and gestation. To compute the base cost, a minimum and maximum energy requirement is specified, with the minimum corresponding to a resting metabolic rate. The maximum includes energy for activities, not counting travel costs if those are specified. Energy use varies between the minimum and maximum in relationship to condition index, according to a function "Enuserate". If the condition index is high, animals use more energy on activity. If condition index is low, animals conserve energy by reducing activity.

Gestation costs are calculated based on the equation in Robbins (1983), also used by Hobbs (1989). The percentage increase in energy costs above the base cost is

$$\text{Gest (\%)} = 0.000024 * \text{Pergest}^{3.13}$$

where Pergest is the percentage of gestation period completed. This is applied to the fraction of the herd that is comprised of pregnant females.

Thermoregulatory costs are calculated using the same approach used by Hobbs (1989). Below a specified lower critical temperature, the animal loses heat energy at a rate determined by the temperature difference below critical and the thermal conductance of the animals coat. Different critical temperatures and conductances can be specified for bedded vs. non-bedded activity.

Travel costs may be computed in three different ways - 1) based upon specified distances traveled per day, 2) based upon computed distance to water, or 3) based upon specified times spent walking and feeding per day and mean rates of travel for both activities (Hobbs 1989). The third option was used here. The distance walked per day is computed as the product of a speed of walking or walking during feeding (m/min) times minutes per day spent walking or feeding. The latter are specified by month of year. The cost of walking is then computed from body weight and distance walked, using an equation developed by Parker and Robbins (1984). Travel cost per kg body weight is

$$\text{MJ/km/kg} = 0.012 / (W^{0.34})$$

with W in kg of body weight. Total travel cost per km is then

$$\text{MJ/km} = \text{MJ/km/kg} * W$$

and total energetic cost per day is

$$\text{Etrv (MJ/d)} = \text{MJ/km} * \text{km/d.}$$

The fractional increase in travel costs associated with traveling in snow (Fsnw) is also based upon an equation from Parker et al. (1984),

$$\begin{aligned} \text{Etrv}_{\text{snow}} &= \text{Etrv} * (1.0 + \text{Fsnw}) \\ \text{Fsnw} &= (\text{Ftrv}_{\text{brk}} * \text{Fsnw}_{\text{brk}}) + ((1 - \text{Ftrv}_{\text{brk}}) * \text{Fsnw}_{\text{unbrk}}) \end{aligned}$$

where Etrv is cost of travel on bare ground, $\text{Etrv}_{\text{snow}}$ includes the cost of travel on snow, Ftrv_{brk} is the fraction of travel that is broken vs. unbroken snow, and Fsnw_{brk} and $\text{Fsnw}_{\text{unbrk}}$ are the fractional increases in cost due to travel in broken and unbroken snow.

$$\begin{aligned} \text{Snden} < 0.3 & \quad \text{Fsnw}_{\text{unbrk}} = 0.0071 * \text{Ru} * e^{0.019 * \text{Ru}} \\ & \quad \text{Fsnw}_{\text{brk}} = 0.0071 * \text{Rb} * e^{0.019 * \text{Rb}} \\ \text{Snden} \geq 0.3 & \quad \text{Fsnw}_{\text{unbrk}} = 0.0123 * \text{Ru} * e^{0.0223 * \text{Ru}} \\ & \quad \text{Fsnw}_{\text{unbrk}} = 0.0123 * \text{Rb} * e^{0.0223 * \text{Rb}} \end{aligned}$$

where Snden is snow density (mm/mm) and where Ru and Rb are relative sinking depths in

unbroken and broken snow, defined as

$$\begin{aligned}R_u &= 100 * Sndp_{brk} / Brht \\ R_d &= 100 * Sndp_{brk} / Brht\end{aligned}$$

where $Sndp_{brk}$ and $Sndp_{unbrk}$ are snow depths in broken and unbroken snow and $Brht$ is bricket height. It is assumed that snow depth in broken snow is one half of bricket height.

The net outcome of energy intake and expenditure is a change in stored energy reserves. If energy intake exceeds requirements, then the animals convert the excess energy to fat. If energy intake is less than requirements, the deficit is drawn from fat reserves. The level of stored energy reserves is indicative of body condition. Animal condition index is a number that varies between 0.0 and 1.1, and is calculated as the ratio of kg's below maximum body weight to the difference between maximum and minimum body weights (kg). At condition index 0.0, animals are at the minimum weight while at condition index of 1.0, animals are at maximum weight. Above 100% of maximum body weight, forage intake rate is reduced due to satiation.

Ungulate Spatial Distribution Submodel

The herbivore spatial distribution submodel dynamically simulates animal distributions over the simulated landscape or region. Essentially, it is a dynamic habitat selection model. There are two common approaches for modeling habitat selection. One is to model habitat suitability or habitat effectiveness as a function of key habitat variables. Habitat suitability models have been in use for decades (U.S. Fish and Wildlife Service 1981). The second is what has come to be known as resource selection functions (Manley et al. 2002 , Boyce et al. 2003 , Olexa et al. 2003), which are basically logistic regressions (Keating and Cherry 2004). Here, the habitat suitability approach was taken because it is more intuitive and it has been successfully used in SAVANNA in many applications.

The approach is to compute the habitat suitability index (HSI) of each grid-cell each week. This index is a dimensionless scalar where 0 indicates no use, and 1.0 indicates maximum preference. The HSI's are summed up across all grid cells and normalized so they sum to 1.0. Then animals in the population or herd is distributed according to the normalized HSI. The habitat suitability index of a grid-cell is affected by 0-1 scalar functions of preference-weighted forage biomass, green herbaceous biomass, snow depth, slope, tree cover, range location, and humans. Importantly, what is being modeled here is suitability for foraging, since forage intake rate is multiplied by animal density to determine offtake. Use of forests for bedding, for example, is not represented here. It is also important to realize that what is being modeled is distribution on a weekly time-step, that is, the time-integrated use of an area over a week. This is much different than modeling individual movements and habitat selection at time steps of a day or less.

$$\begin{aligned}HSI &= F(\text{forage}) \times F(\text{green herb}) \times F(\text{snow}) \\ &\times \min[F(\text{slope}), F(\text{elevation}), F(\text{tree cover})] \\ &\times F(\text{range location}) \times F(\text{humans}) \times F(\text{random variable})\end{aligned}$$

The forage function simply takes the value of the sum of the products of dietary preference weights times forage biomass over all dietary items in the grid-cell. If forage biomass was the only variable in the HSI equation and dietary preference weights were not used, the animal distribution would be in direct proportion to forage abundance, which is also known as matching, or an ideal, free distribution (Senft et al. 1987). The inclusion of the dietary preference weight causes animals to be distributed more heavily towards locations with preferred food items. The function of green herbaceous biomass (Figures 2,3) represents selection for a greater abundance of green biomass. It is known, for example, that elk follow a wave of green up as it moves upslope during snowmelt in the Spring (Frank and McNaughton 1992). The function has a small non-zero value even with zero green biomass so that with zero green biomass, distributions will still be affected by the distributions of total forage biomass.

The effect of snow depth for elk was set based upon previous modeling and literature review (Coughenour 1994, Coughenour and Singer 1996b). The effect of snow on bison habitat selection was calibrated to attain realistic distributions. It was found that deep snow would have to be tolerated in order for bison to winter in places such as Hayden and Pelican Valleys.

The minimum of the slope, elevation, shade cover, and thicket cover effects is taken, hypothesizing that any one of these effects can override the others. For example, if the slope is too steep it doesn't matter what the cover is, and if cover is unsuitable it doesn't matter what the slope is. The effect of slope is parameterized to reflect the fact that bison prefer gentle or level topography (Barmore 2003), but do make use of slope up to 25% (Figure 2, Table 1). Elk are less constrained by topography, making use of slopes as steep as 45% (Figure 3, Table 1), based upon an unpublished habitat preference analysis by K. Symonds and F. Singer. The tree cover effect is used to express a preference by bison for open areas (Figure 3). Bison may use woodlands, but much less frequently than open grasslands. No tree cover effect is used for elk since they commonly use woodlands, especially during summer. Note that these functions are calculated based upon tree cover for 1 km or 500 m grid-cells. It is common for these large grid-cells to be only partly covered by trees, so these grid-cells should not be excluded. Also, the forage function will have an overriding effect on distribution, so if there is more forage in grassland grid-cells, the outcome will be an avoidance of trees.

Elevation was used to affect elk distributions during summer (Figure 3, Table 1). This proved to be necessary to simulate observed densities of elk during summer in Hayden Valley (Olenicki and Irby 2003). This area has a high biomass of forage and is otherwise a highly suitable habitat, yet elk densities are relatively low. Similarly, some explanation is necessary for elk avoidance of the lower elevation winter range during summer. Causes could include avoidance of insects, avoidance of waterlogged ground, or avoidance of predators. However, little is known about these factors so a simple elevation surrogate sufficed.

A uniform random variate (0.8-1.0) is introduced that prevents animals from attaining an ideal free (ie. perfect) distribution. There are many reasons why animals may not achieve a perfect distribution such as imperfect knowledge of resource distributions, physical inaccessibility, and time lags for travel.

The range function is a binary 0 or 1 that represents whether the grid-cell is in or out of the population's range. Population ranges may be an outcome of knowledge, behavioral density regulation, primary wintering and rutting areas, physical barriers, or geographical constraints on interchange between subpopulations but none of these factors are represented explicitly. Range maps are prescribed to take effect at different times of the year, and to vary over years. For example winter and summer ranges may be specified, and for bison they may change over time due to range expansion (Chapter 7).

After animals are distributed according to the HSI, the model checks to see if densities are either above a behavioral maximum or below a behavioral minimum. The maximum and minimum values are applied to represent social aspects of animal behavior. Maximum values represent behavioral spacing or "mobile territories". Minimum values represent tendencies for group or herd formation as opposed to solitary habits. Maximum densities for bison and elk were set based upon aerial survey data (Taper et al. 2000, Hess 2002, Coughenour and Singer 1996a). Higher values were used for bison during the rutting season (Table 1). Given both species tend to occur in herds or groups, minimum values were merely set to prevent the occurrence of extremely low densities on a weekly basis (Table 1). Animals in excess of the maximum density or below the minimum density are redistributed according to the HSI of cells with densities below the maximum.

Herbivore Population Dynamics

The herbivore population dynamics submodel is an age-sex class model with one age class for each year, for each sex. Recruitment rates and death rates are affected by animal condition indices. Recruitment rather than birth is simulated, because very early calf/fawn survival rate is not distinguished. As condition index increases, recruitment rates increase and death rates decline. The effect of condition index on recruitment rate is multiplicative. For example, a condition index of 0.2 might reduce recruitment rate to 35% of the maximum. The effect of condition on death rate supercedes the nominal death rate if it is larger. The effects of condition index represent population responses to ecological conditions governing forage availability (e.g., forage production, snow depth) and intra- and interspecific competition. As competing animals can reduce forage supply, forage intake rate, and thus body condition, they can consequently reduce population growth rate of the species in question.

The equation for recruitment is

$$\text{Recruit} = \sum \text{Female}_i \times \text{Birthrt}_i * F(\text{CI}_{\text{gest}})$$

where Recruit is number of surviving newborns in the season, Female_i is the number of females in age class i , Birthrt_i is the maximum birth rate by age class, and CI_{gest} is the condition index averaged over the period of gestation. Recruitment is assumed to occur in a single month each year. For Yellowstone bison and elk, that month is May.

The equation for monthly mortality is

$$\text{Deaths}_i = \text{Animals}_i * \max[\text{Deathrt}_{\text{Nom}}, \text{Deathrt}(\text{CI})]$$

where Deaths_i is the number of animals dying in age class i , Animals_i is number of animals in age class i , $\text{Deathrt}_{\text{Nom}}$ is the nominal death rate, and $\text{Deathrt}(\text{CI})$ is the death rate as a function of the monthly body condition. Different $\text{Deathrt}(\text{CI})$ functions are specified for calves, mature females, and mature males.

The monthly nominal mortality rate is computed from the annual survivorship rate as follows.

$$\text{Deathrt}_{\text{Nom}} = \min (\text{SurvCal}_i * \text{Surv}_i^{(1/12)}, 1.0)$$

Monthly survivorship rate is equal to the annual survivorship to the 1/12 power, since the monthly survivorship to the 12th power should equal the annual survivorship. The parameter SurvCal is a calibration parameter that is found by fitting the model to the observed data. A calibration parameter of >1.0 is expected because the nominal survivorships are based upon data from field conditions that include mortalities due to environmental stress while here, stress related mortality is accounted for in $\text{Deathrt}(\text{CI})$.

Animals can be removed by hunting or culling using a specified “trigger” population size. If the population is at or above that value the population will be reduced to a specified “target” number. This method is normally used to simulated management removals to keep a population within a certain range. Removals can also be prescribed with an observed number of animals per year, or an observed fraction of animals per year. Simulations of historical population dynamics use actual removal numbers. Simulations of hypothetical hunting or culling scenarios usually use removal rates.

An emigration rate can be specified as the fraction of the population per month that leaves subpopulation and moves to another subpopulation with a different geographic range.. The emigration rate is a function of condition index, with higher rates when animals are in poor condition, the logic being that as animals increasingly become food-limited, they will more frequently disperse in order to find forage.