

SCIENTIFIC REPORTS



OPEN

Fuzzy Union to Assess Climate Suitability of Annual Ryegrass (*Lolium multiflorum*), Alfalfa (*Medicago sativa*) and Sorghum (*Sorghum bicolor*)

Hyunae Kim¹, Shin Woo Hyun², Gerrit Hoogenboom ^{3,4}, Cheryl H. Porter³ & Kwang Soo Kim^{2,4,5}

The Law of the Minimum is often implemented using t-norm or fuzzy intersection. We propose the use of t-conorm or fuzzy union for climate suitability assessment of a grass species using annual ryegrass (*Lolium multiflorum* Lam.) as an example and evaluate the performance for alfalfa (*Medicago sativa* L.) and sorghum (*Sorghum bicolor* L.). The OR_F and AND_F models, which are fuzzy logic systems based on t-conorm and t-norm between temperature and moisture conditions, respectively, were developed to assess the quality of climate conditions for crops. The parameter values for both models were obtained from existing knowledge, e.g., the EcoCrop database. These models were then compared with the EcoCrop model, which is based on the t-norm. The OR_F model explained greater variation (54%) in the yield of annual ryegrass at 84 site-years than the AND_F model (43%) and the EcoCrop model (5%). The climate suitability index of the OR_F model had the greatest likelihood of occurrence of annual ryegrass compared to the other models. The OR_F model also had similar results for alfalfa and sorghum. We emphasize that the fuzzy logic system for climate suitability assessment can be developed using knowledge rather than presence-only data, which can facilitate more complex approaches such as the incorporation of biotic interaction into species distribution modeling.

Assessment of climatic suitability of a crop could help introduce a new crop into a cropping system as an adaptation option to climate change^{1–4}. A process-based model that quantifies photosynthetic assimilates under given weather conditions could be used to examine climatic suitability of a crop^{5–7}. Still, quantitative prediction of crop growth would require a detailed set of input data for a process-based model. For example, input regarding local management information such as variety, fertilizer application dates and irrigation application rates, is required for the process-based model although such information is often limited to specific sites⁸. Furthermore, high quality observed data would be needed for calibration and evaluation of the model^{9–11}.

When data availability is limited to occurrence sites of a given species, a species distribution model would be useful for the assessment of climate suitability in a region^{12–14}. For example, Estes *et al.*¹⁵ reported that an empirical model based on the maximum entropy had similar or better accuracy compared with a process-based model in predicting productivity of maize. Jeong *et al.*¹⁶ used the random forest algorithm to predict wheat yield at different spatial scales. However, these approaches require a number of sites to find the relationship between bioclimatic variables and climate suitability of a given species.

A fuzzy logic approach could allow for the development of a climate suitability model for a species using human expert knowledge rather than occurrence data. Models based on fuzzy logic have been developed to predict the distribution of forest herbs¹⁷, to identify optimum management of rangeland¹⁸, and to analyze cropland

¹Institute of Convergence Technology, KT, Seoul, 06763, Korea. ²Department of Plant Science, Seoul National University, Seoul, 08826, Korea. ³Department of Agricultural and Biological Engineering, University of Florida, Gainesville, Florida, 32611, USA. ⁴Institute for Sustainable Food Systems, University of Florida, Gainesville, Florida, 32611, USA. ⁵Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul, 08826, Korea. Correspondence and requests for materials should be addressed to K.S.K. (email: luxkwang@snu.ac.kr)

suitability¹⁹, as well as environmental studies^{20–22}. Fuzzy logic models have been developed using an optimization process, e.g., a genetic algorithm, that depends on training data, which would result in a black box model^{23–25}. In contrast, the fuzzy logic model that has the high-level of interpretability can be developed using knowledge based on the relationship between environmental conditions and growth of the species of interest²⁶.

The Law of the Minimum could provide a reasonable framework to assess climate suitability of crops using a fuzzy logic model²⁷. Fuzzy propositions such as “temperature is suitable” and “precipitation is suitable” can be defined to interpret the Law of the Minimum for climate variables. Literally, the Law of the Minimum is the outcome of fuzzy intersection using the minimum t-norm between those propositions. For example, Hackett & Vanclay²⁸ used the minimum value of indices to estimate growth of plant using the law of the Minimum. However, the Law of the Minimum can be represented using fuzzy intersection based on other t-norms such as the product t-norm to represent the minimum quality of given conditions. For example, Hijmans *et al.*²⁹ developed the EcoCrop model that assesses climate suitability of crops using the product between suitability indices of temperature and precipitation. The EcoCrop model has been used for climate change impact assessment for minor crops³⁰.

Austin³¹ suggested that the Law of the Minimum could have limitations in development of climate suitability models. In particular, a climate suitability model can be developed using climate variables that account for a small variation in plant growth. For example, the EcoCrop model is dependent on annual precipitation (Appendix A in Supporting information) although growth of a grass could be affected by the rainfall distribution during a growing season³². The species parameter values for the EcoCrop model can be obtained from the FAO EcoCrop database, which would require no iterative calibration process. Still, Ramirez-Villegas *et al.*³³ found that the EcoCrop model did not have a clear relationship between the climate suitability index and sorghum yield even after parameter optimization for the crop. Piikki *et al.*³⁴ improved the EcoCrop model using local soil properties after additional optimization procedures.

Here we propose an algorithm based on a fuzzy logic system to assess climate suitability of a forage crop. We attempted to examine if climate suitability assessment could be improved using an alternative rule based on fuzzy union or t-conorm to the Law of the Minimum instead of an iterative calibration process. The objectives of this study were to develop and evaluate the fuzzy systems based on fuzzy intersection and fuzzy union. We compared the fuzzy logic system based on t-conorm (fuzzy union) with the other model based on t-norm (fuzzy intersection) for logical operation between suitability indices of temperature and rainfall. Then, we compared these models with the EcoCrop model, which is based on the fuzzy intersection. The parameter values for the fuzzy logic model were determined using the EcoCrop database to have identical parameter values for both models. In the present study, annual ryegrass (*Lolium multiflorum* Lam.) was selected because it has been considered an important forage crop in temperate regions^{35,36} yet limited modeling efforts so far have been conducted. Furthermore, the feasibility of the proposed algorithm for the climate suitability indices was evaluated for alfalfa (*Medicago sativa* L.), a cool season legume crop, and sorghum (*Sorghum bicolor* L.), a warm season grass, respectively.

Materials and Methods

Rules to evaluate climate conditions for survival and growth of a plant species. Climate suitability of a species has been modelled using a combination of rules that characterizes the environmental envelopes of the given species. Models often depend on the t-norm, which generalizes intersection in a logic operation between two environmental conditions such as temperature and precipitation^{33,37,38}. The EcoCrop model evaluates rules that could be described as “temperature is suitable *and* precipitation is suitable.” The use of a t-norm would take into account the minimum quality of environmental conditions as the Law of the Minimum indicates^{27,39}. The t-norm would be useful to determine the boundary conditions favorable for the survival of a given species⁴⁰.

Alternatively, climate conditions for growth as well as survival of a species could be assessed using a t-conorm, which is a logical union operation³⁷. When the quality of environmental conditions is greater than the threshold of survival for a species, the species would have additional growth performance at the degree to which these conditions meet the requirements for optimum growth of the species. For example, a long duration of optimum temperature would have a positive impact on plant growth unless drought conditions are prolonged. Although a temperature condition during a month could be suboptimal, a grass species could benefit from the optimum amount of rainfall in a given month. Thus, a rule based on t-conorm, e.g., “temperature is suitable *or* precipitation is suitable,” could be useful for delineating factors in the assessment of climate suitability of the given species.

Rule statements were designed to represent survival and growth of a grass species in terms of precipitation and temperature (Table 1). Growth of the species was assessed using a statement representing the moisture conditions under favorable temperature conditions. For example, a statement such as “rainfall amount is adequate for the time period during which temperature conditions would be favorable,” would be useful to represent a moisture condition for growth of a species. The rule statement was rephrased as “rainfall is suitable, maximum temperature is favorable, and minimum temperature is favorable.”

Another rule was defined to represent the temperature suitability for growth of a species (Table 1). At first, the rule statement, “temperature is suitable,” was used to determine the degree to which temperature conditions were suitable for growth of a species. Then, a modifier, “for a less dry period” was added to the predicate “suitable” to take into account a negative impact of dry conditions for consecutive months on crop growth.

Additional rule statements were defined to take into account suboptimal and extreme conditions (Table 1). The rules of heat and cold stresses including “maximum temperature is stressful” and “minimum temperature is stressful,” respectively, were used to represent the negative impact of suboptimum temperature conditions on plant growth. Extreme temperature conditions during a season, which would affect survival of a species, were also characterized using rule statements such as “maximum temperature was harmful” and “minimum temperature

Rule	Statement	Eq.*	Variable
1	Precipitation is suitable and Maximum Temperature is favorable and Minimum Temperature is favorable	15	β
2	Temperature is suitable for a less dry period	16	Θ
3	Maximum Temperature is stressful or Minimum Temperature is stressful	19	τ
4	Maximum Temperature is harmful or Minimum Temperature is harmful	20	ξ

Table 1. Rule statements to evaluate the climate suitability for a grass species. *Eq. is the equation number in the text.

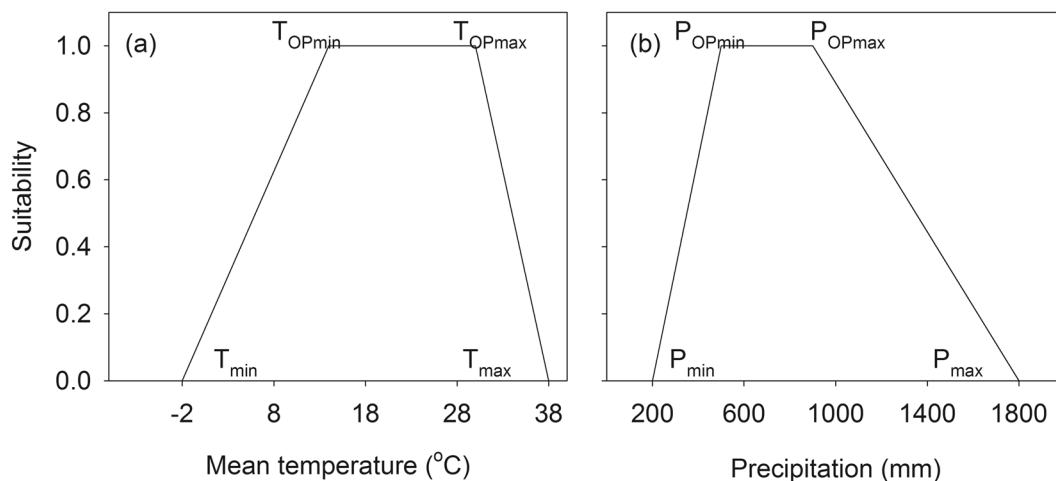


Figure 1. Examples of membership functions for temperature (a) and precipitation (b). These same functions were used to determine the suitability of temperature and precipitation for the EcoCrop model.

was harmful.” In addition, the rule statement “temperature was reasonable for early growth” was used to evaluate whether or not a given species could be established at the beginning of a season.

Evaluation of rule statements. Each rule statement was evaluated using a linguistic term of the predicates, which represents specific conditions for survival or growth of a species. Using a fuzzy set, each linguistic term was quantified for a corresponding climate variable using a crisp value of the given variable. A fuzzy set is a set whose members have a degree of membership between 0 and 1 instead of bivalent membership such as yes or no⁴¹. A membership function assigns a degree of membership to the value of a given variable. For example, the proposition “temperature is suitable” can be evaluated to have the degree of truth, e.g., 0.75, for a mean temperature in a month, e.g., 11 °C, at a site using a membership function illustrated in Fig. 1a. Similarly, suitability of precipitation is also determined using another membership function (Fig. 1b). A detailed description of fuzzy sets can be found in Klir & Yuan³⁷.

Fuzzy sets for each linguistic term were defined using different shapes (Fig. 2; Appendix B). The shape of fuzzy sets was designed to minimize the number of parameters and to maximize the existing knowledge on climate prediction. For example, a trapezoid shape was used to evaluate suitability of precipitation because the ranges of precipitation can be obtained from the EcoCrop database (Fig. 1b). The degree of temperature suitability was determined using a right triangle shape based on the duration of hours during which the optimal temperature conditions were met (Fig. 2a).

The combination of predicates defined in a rule statement was evaluated using fuzzy operations. The connectives in the rule statements, *and* and *or*, were defined using a t-norm and a t-conorm, respectively, as follows:

$$F_x \text{ and } F_y = F_x \cdot F_y \text{ and} \quad (1)$$

$$F_x \text{ or } F_y = F_x + F_y - F_x \cdot F_y, \quad (2)$$

where $F \in [0, 1]$. The rule statements for moisture and temperature suitability were rewritten as follows:

$$\beta(X_m, N_m, P_m) = P_{\text{suitable}}(P_m) \cdot X_{\text{favorable}}(X_m) \cdot N_{\text{favorable}}(N_m), \quad (3)$$

$$\Theta(X_m, N_m, d_m) = d_m^{-1} \cdot T_{\text{suitable}}(X_m, N_m), \quad (4)$$

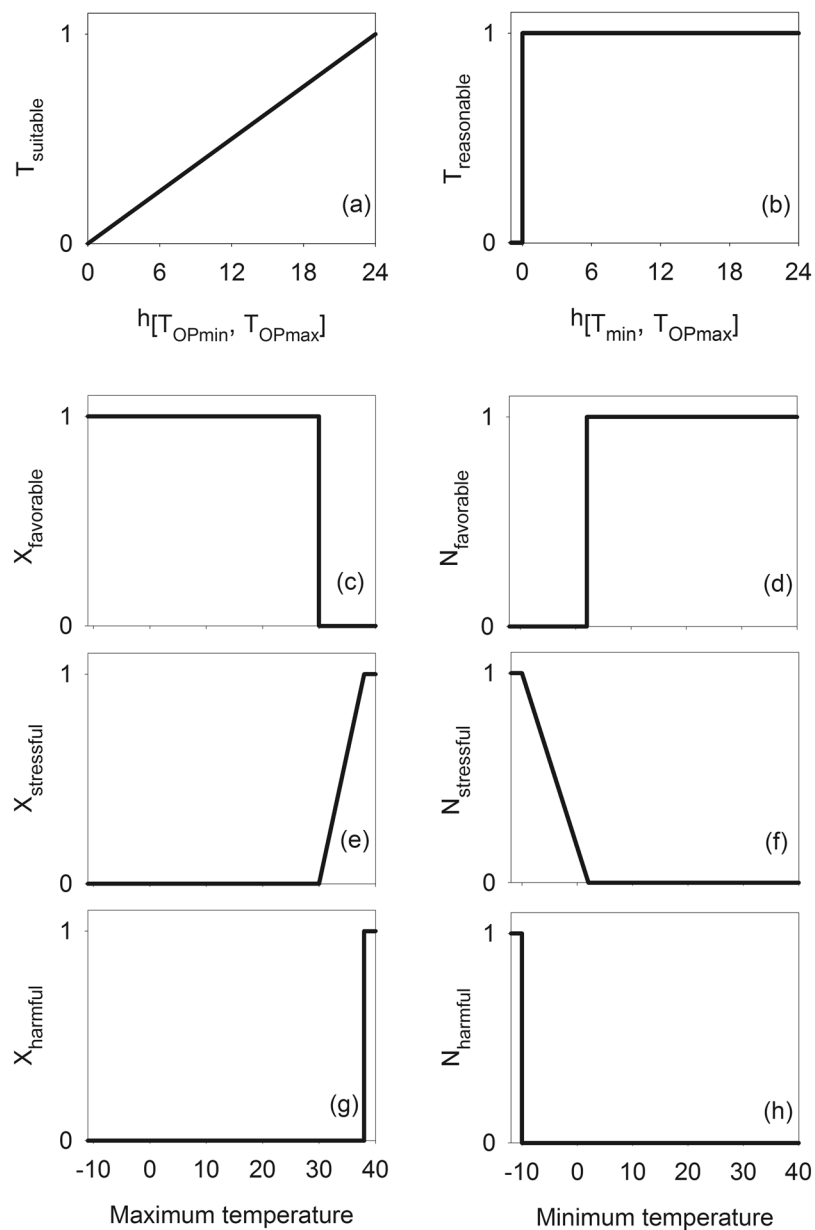


Figure 2. Membership functions of the model based on a fuzzy logic system. X and N represent the membership function that uses maximum and minimum temperature as inputs, respectively. $h_{(a,b)}$ indicates the number of hours during which the temperature ranged between a and b .

where β , Θ , and d represent the outcome of rule evaluation for moisture, temperature suitability, and a weight for a prolonged dry condition respectively. The value of d was determined to be the number of consecutive months during which β value was zero by the time of evaluation. $P_{suitable}$, $X_{favorable}$, $N_{favorable}$, and $T_{suitable}$ are the membership function of fuzzy sets for suitability of precipitation, favorability of maximum temperature, favorability of minimum temperature, and suitability of temperature in a month, respectively (Fig. 2). P_m , X_m , and N_m indicate precipitation, maximum temperature, and minimum temperature in a month m , respectively.

Calculation of climate suitability index. The architecture of the models based on t-norm and t-conorm was designed to evaluate cumulative climate quality during plant growth. For example, the outcome of rules evaluated by each month was accumulated over a growing period of a species. The architecture for the fuzzy models differed by that of the EcoCrop model because the latter determines climate suitability by selecting the minimum climate quality during the growing period. As a result, the architecture used by the fuzzy models would allow for a reasonable evaluation of the conditions for plant growth compared with the EcoCrop model although both models depend on monthly climate conditions. However, it is likely that the fuzzy models would have inferior prediction of plant growth compared to the process-based crop models that rely on more detailed input data and parameters^{8–11}. Nevertheless, the fuzzy model could be useful to assess climate conditions for crop growth when

detailed information on crop management and environmental conditions are rarely available or a detailed crop model does not exist, such as for many vegetable, fruit and ornamental crops.

The outcomes of rule evaluation for moisture and temperature suitability were combined using two types of logical operations, t-norm or t-conorm. To represent the Law of the Minimum, the rule statements were combined using a t-norm, S_{AND} , as follows:

$$S_{AND}(X_m, N_m, P_m, d_m) = \Theta(X_m, N_m, d_m) \cdot \beta(X_m, N_m, P_m) \quad (5)$$

We used the standard intersection t-norm because it provides a low value for suitability indices, which would be useful to represent the Law of the Minimum in practice. Climate suitability assessed using the standard union t-conorm, S_{OR} , was also defined as follows:

$$S_{OR}(X_m, N_m, P_m, d_m) = \Theta(X_m, N_m, d_m) + \beta(X_m, N_m, P_m) - \Theta(X_m, N_m, d) \cdot \beta(X_m, N_m, P_m) \quad (6)$$

Models based on t-conorm and t-norm were denoted by OR_F and AND_F models, respectively.

Climate suitability index in a season was determined by combining the degrees of cumulative suitability and stresses. The monthly stress condition τ was determined as follows:

$$\tau(X_m, N_m) = N_{stressful}(N_m) + X_{stressful}(X_m) - N_{stressful}(N_m) \cdot X_{stressful}(X_m), \quad (7)$$

where $X_{stressful}$ and $N_{stressful}$ indicate the membership function of “stressful” for maximum and minimum temperatures, respectively (Fig. 2e,f). It was assumed that the stress would have a long-term memory on the climate suitability in following months depending on the magnitude of stress conditions. Thus, the degree of cumulative monthly suitability Φ_m was calculated as follows:

$$\Phi_m = \max(\Phi_{m-1} + S_m - \tau_m, 0), \quad (8)$$

where S_m and τ_m denote $S(X_m, N_m, P_m, d_m)$ and $\tau(X_m, N_m)$. If the value of Φ_m was less than 0, it was set to be zero under the assumption that growth would be ceased by climatic stresses. The value of Φ_0 was assumed to be zero.

The climate suitability index was calculated for a growing period. For the growing period, climate suitability would be dependent on occurrence of extreme conditions and favorable conditions for emergence. The suitability index G_p for a growing period p was calculated as follows:

$$G_p = T_{reasonable}(X_0, N_0) \cdot \Phi_p / n_p \cdot \prod_m (1 - \kappa(X_m, N_m)), \quad (9)$$

where Φ_p and n_p indicate the suitability index and the number of month in p , respectively. The extreme conditions κ against survival of a given species in m was determined as follows:

$$\kappa(X_m, N_m) = X_{harmful}(X_m) + N_{harmful}(N_m) - X_{harmful}(X_m) \cdot N_{harmful}(N_m), \quad (10)$$

where $X_{harmful}$ and $N_{harmful}$ represent the membership function of “harmful” for maximum and minimum temperatures, respectively (Fig. 2f,g). The product of degree to which monthly conditions were not harmful, i.e., $1 - \kappa(X_m, N_m)$, was used to take into account only the growing period during which no extreme condition occurred. $T_{reasonable}$ represents the membership function of “reasonable” for establishment in an early season (Fig. 2b). $T_{reasonable}$ was also used to avoid calculating the climate suitability index at sites where climate condition would be unfavorable for establishment of a given species during the first month.

The length of a growing period can vary by region and management. When no harvest date was available, multiple values of climate suitability index were calculated for a set of potential growing periods within a growing season. When a growing season started in March and the length of a growing period for a crop would range from 3–5 months, the values of G_p were determined for a period from March to May, March to June, and March to July, respectively. The median value of G_p values was used to represent a seasonal suitability index G_s for a season s as follows:

$$G_s = A(\mathbf{G}), \quad (11)$$

where $\mathbf{G} = \{G_p | p = 1, 2, \dots, n_p\}$. A represents a function to determine a median value. A was used to determine a central tendency of climate conditions for the grass species in a season when the length of a growing period was unknown.

A growing season would start from a planting date. Still, planting date is often unknown, especially in regions where a grass species is not cultivated. When no planting date was known for the sites of interest, the planting date was set to be the first day of each season. For example, the planting date was assumed to be the first day of the first month for an individual growing season when the climate suitability index was determined using monthly data. The final suitability index G at a site was chosen to be the maximum value of G_s as follows:

$$G = \max\{G_s | s = 1, 2, \dots, n\} \quad (12)$$

where n is the number of potential seasons. When no specific season was defined, the G_s values were determined for every potential season. For example, G_s was calculated for 12 seasons that start each month in a year ($n = 12$). Customized software written in C++ was used to determine the value of G .

The parameter values for annual ryegrass based on existing knowledge. The climate suitability index for annual ryegrass was determined using the OR_F and AND_F models. Klir & Yuan³⁷ suggested that

Name	Description	Annual Ryegrass ^b	Alfalfa ^b	Sorghum ^b
G _{min}	minimum growing period (d)	90	100	90
G _{max}	maximum growing period (d)	270	210	300
T _{kill}	lethal temperature (°C)	−10 ^c	−25	0
T _{min}	minimum absolute temperature (°C)	2	5	8
T _{max}	maximum absolute temperature (°C)	38	45	40
T _{OPmin}	minimum optimal temperature (°C)	14	21	27
T _{OPmax}	maximum optimal temperature (°C)	30	27	35
R _{min} ^a	minimum absolute rainfall (mm)	200	350	300
R _{max} ^a	maximum absolute rainfall (mm)	1800	2700	3000
R _{OPmin} ^a	minimum optimal rainfall (mm)	500	600	500
R _{OPmax} ^a	maximum optimal rainfall (mm)	900	1200	1000

Table 2. The model parameters to assess the climate suitability of forage crops including annual ryegrass, alfalfa, and sorghum obtained from existing knowledge. ^aThe values of precipitation parameters were divided by 12 for the models based on fuzzy sets to represent monthly precipitation. ^bThe EcoCrop database operated by the Food and Agriculture Organization (FAO) was used to determine the parameter values of each crop. ^cThe lethal temperature was obtained from Eagles *et al.*⁴².

an expert's knowledge or existing data would be useful to determine parameters of a membership function. In the present study, the parameters of membership functions for annual ryegrass were determined using the FAO-EcoCrop database (Table 2). Because the EcoCrop database provides the range of annual precipitation, these parameter values were divided by 12 months for the model based on fuzzy logic. It was assumed that even distribution of rainfall across months would represent monthly climate envelopes for growth of a grass species. The lethal temperature for annual ryegrass from the EcoCrop database was -4°C , which is relatively higher than that in previous reports. For example, Eagles *et al.*⁴² reported that a lethal temperature for annual ryegrass was -10°C after a week of hardening. Because climate sensibility index would be calculated for prolonged winter chill conditions rather than short-lived winter spells, it was assumed that annual ryegrass would go through hardening for low temperature. Thus, the lethal temperature was set to be -10°C . These parameter values were used for models based on fuzzy logic systems and the EcoCrop model to determine climate suitability index for annual ryegrass.

Evaluation of climate suitability index using yield of annual ryegrass at site-years. The determination coefficient between climate suitability index and yield of annual ryegrass was calculated to examine if the model based on the fuzzy logic system would have a reasonable relationship between the climate suitability index and yield. Yield of annual ryegrass was obtained for 18 sites in three countries, including the USA, Belgium, and Australia (Appendix C).

Daily weather data were compiled to prepare inputs to the climate suitability models (Table 3). Average of temperature and sum of precipitation for a 30-day period were used as inputs for the models. These summary data were obtained for the duration of growing period (Appendix D). For the EcoCrop model, the sum of precipitation for the 12 consecutive 30-day periods was obtained because the model requires annual precipitation as an input. Instead of the first day of a calendar month, a planting date at a site-year was set to be the first day of the given season. Such data would minimize the impact of the uncertainty associated with climate inputs to the models during the season.

Spatial analysis of climate suitability index for annual ryegrass. The spatial distribution of the climate suitability index was also determined. Monthly climate surfaces were used as inputs to the models to generate maps of climate suitability index. The climatic data reported by Fick and Hijmans⁴³ were obtained from the worldclim website (<http://www.worldclim.org>). Climate surfaces at a 1 km spatial resolution were used as inputs to the climate suitability models. Sowing dates for annual ryegrass was unknown for most of the grid cells. Thus, the climate suitability index for each grid cell was determined using Eq. 12.

A likelihood analysis was performed to compare reliability of climate suitability index obtained from each model using the occurrence sites where annual ryegrass has been reported. The occurrence data of annual ryegrass were obtained from the Global Biodiversity Information Facility (GBIF) database (<http://www.gbif.org>). In total, 27,000 sites were compiled after eliminating sites without geographic coordinates. The occurrence site was denoted by $y = 1$. The probability that annual ryegrass would occur at a site was denoted by $\Psi(y = 1)$. The probability of occurrence for a given climate suitability index $\Psi(y = 1|G)$ can be computed using Bayes rule as follows⁴⁴:

$$\psi(y = 1|G) = \pi(G|y = 1) \cdot \psi(y = 1)/\pi(G), \quad (13)$$

where $\pi(G|y = 1)$ indicates the conditional probability distribution of G at occurrence sites. It was assumed that $\pi(G|y = 1)$ would be similar to the probability distribution of G at occurrence sites obtained from the GBIF database. $\pi(G)$ is the probability distribution function of G over a region.

The likelihood of occurrence for a given climate suitability index G_i would vary. The joint likelihood function A can be defined as follows:

Name	Network ^b	ID ^c	Lat ^d	Long ^d
Mutdapilly	BOM	040004	-27.63	152.70
Gatton	BOM	040436	-27.55	152.33
Munte	GSOD	64280	50.93	3.67
Lexington ^a	GHCN:AWOS	USW00093820	38.04	-84.61
Franklinton	GHCN:AWOS	USW00093919	31.18	-90.47
Jeanerette ^a	GHCN:COOP	USC00164674	29.96	-91.71
Rosepine	GHCN:COOP	USC00162367	30.84	-93.29
Winnsboro ^a	GHCN:COOP	USC00169806	32.10	-91.70
Newton	GHCN:COOP	USC00226308	32.34	-89.08
Starkville	GHCN:COOP	USC00228374	33.47	-88.78
Holly Springs	GHCN:COOP	USC00224173	34.82	-89.43
Poplarville	GHCN:COOP	USC00227128	30.84	-89.55
Raymond	GHCN:COOP	USC00226476	32.21	-90.51
Ardmore ^a	GHCN:COOP	USC00340292	34.18	-97.18
Overton	GHCN:AWOS	USW00003901	32.38	-94.71
College Station	GHCN:AWOS	USW00003904	30.59	-96.36
Beaumont	GHCN:COOP	USC00410611	30.10	-94.10
Blackstone	GHCN:COOP	USC00441322	37.04	-77.94

Table 3. List of weather stations for which weather data were obtained. ^aMissing data were replaced by weather observations at alternative weather stations for Lexington (GSOD; 724220), Ardmore (GHCN:COOP; USC00340312), Jeanerette (GHCN:AWOS; USW00053915), and Winnsboro (GHCN:COOP; USC00169804). ^bBOM, GSOD, GHCN, AWOS, COOP indicate the Bureau of Meteorology, Australia, Global Surface Summary Of the Day, Global Historical Climatology Network, Automated Weather Observing System, and Cooperative Observer Network, respectively. ^cThe identification code for the corresponding weather station network. ^dLat and Long represent latitude and longitude, respectively.

$$\Lambda = \psi(y = 1) \cdot L, \quad (14)$$

where L is $\prod \pi(G_i|y_i = 1)/\pi(G_i)$. The value of Λ for a given model differs by L because $\Psi(y = 1)$ would be constant for a given species. The smaller value of the negative logarithm of L results in the greater value of Λ . Thus, the model that has the smallest value of $-\log(L)$ would have the greatest reliability, e.g., the greatest likelihood of occurrence for a given index, compared with other models. The values of L for the OR_B, AND_B, and EcoCrop models were denoted by L_{OR} , L_{AND} , and L_{ECO} , respectively.

The empirical cumulative distribution function (ECDF) of climate suitability index was obtained for each model. The ECDF of climate suitability index in grid cells corresponding to the occurrence sites E_O was determined to calculate $\pi(G_i|y_i = 1)$. Another ECDF of the index for all grid cells E_R were also determined to calculate $\pi(G_i)$. R, which is the open source statistical package⁴⁵, was used to obtain these ECDFs. The value of L was estimated as follows:

$$L = E_O(G_0)/E_R(G_0) \cdot \prod_{i=1}^{100} [E_O(G_i) - E_O(G_{i-1})]/[E_R(G_i) - E_R(G_{i-1})], \quad (15)$$

where $G_i = 0.01 \cdot i$.

Spatial analysis of the climate suitability index for alfalfa and sorghum. The fuzzy logic system was applied to alfalfa and sorghum to examine if the fuzzy logic system would be applicable to other crops. The parameters of the fuzzy logic systems for both crops were obtained from the EcoCrop database (Table 2). The parameter values of the EcoCrop model for alfalfa were obtained from the EcoCrop database, whereas those for sorghum were determined using Ramirez-Villegas *et al.*³³. A combined suitability index of the EcoCrop model was calculated for sorghum as Ramirez-Villegas *et al.*³³ suggested. The suitability index for alfalfa and sorghum was calculated at a global scale using the same gridded climate data. In total, 31,000 occurrence sites for alfalfa and 8,800 for sorghum were obtained from the GBIF database. The ECDF of climate suitability index at the occurrence sites was determined using maps of the climate suitability index at a 1 km spatial resolution. The ECDF was used to compare the likelihood of occurrence for climate suitability index of the EcoCrop model and the model based on the fuzzy logic system. Results datasets are accessible from the following link: https://figshare.com/articles/climate_suitability_of_annual_ryegrass_alfalfa_and_sorghum/5500876, (<https://doi.org/10.6084/m9.figshare.5500876>)

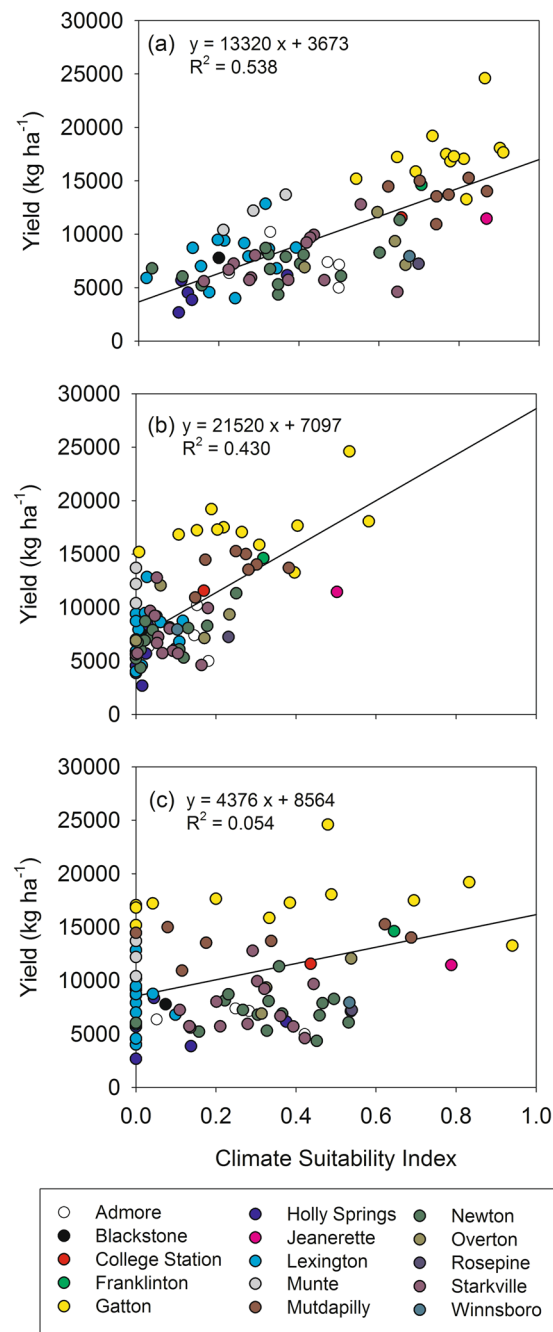


Figure 3. Distribution of observed yield at site-year for climate suitability index values of the OR_F model (a), the AND_F model (b), and the EcoCrop model (c), respectively. A line in each plot represents a regression line between yield and the climate suitability index of the corresponding model.

Results

Comparison between climate suitability index and yield of annual ryegrass. The climate suitability index of the OR_F model had the largest coefficient of determination for yield among the three models (Fig. 3). For example, the OR_F model accounted for 54% of the explained variation in yield for the 84 site-years except for sites where disease occurrence has been reported (Appendix C). The values of the climate suitability index were relatively low for site-years where low yield was observed. In contrast, the greater values of the climate suitability index did not necessarily coincide with high yield, which created a yield distribution similar to a triangular shape.

The AND_F and EcoCrop models, which are based on t-norm, tended to have zero climate suitability index for site-years where a relatively large yield was observed. The climate suitability index of the AND_F model was zero at Munte where a high yield above 10,000 kg/ha was observed for all of the study periods. The EcoCrop model had zero values of climate suitability index for more than 50% of the site-years although the yield ranged from 2,672 to 17,055 kg ha^{-1} for the corresponding periods. As a result, the coefficient of determination between the climate

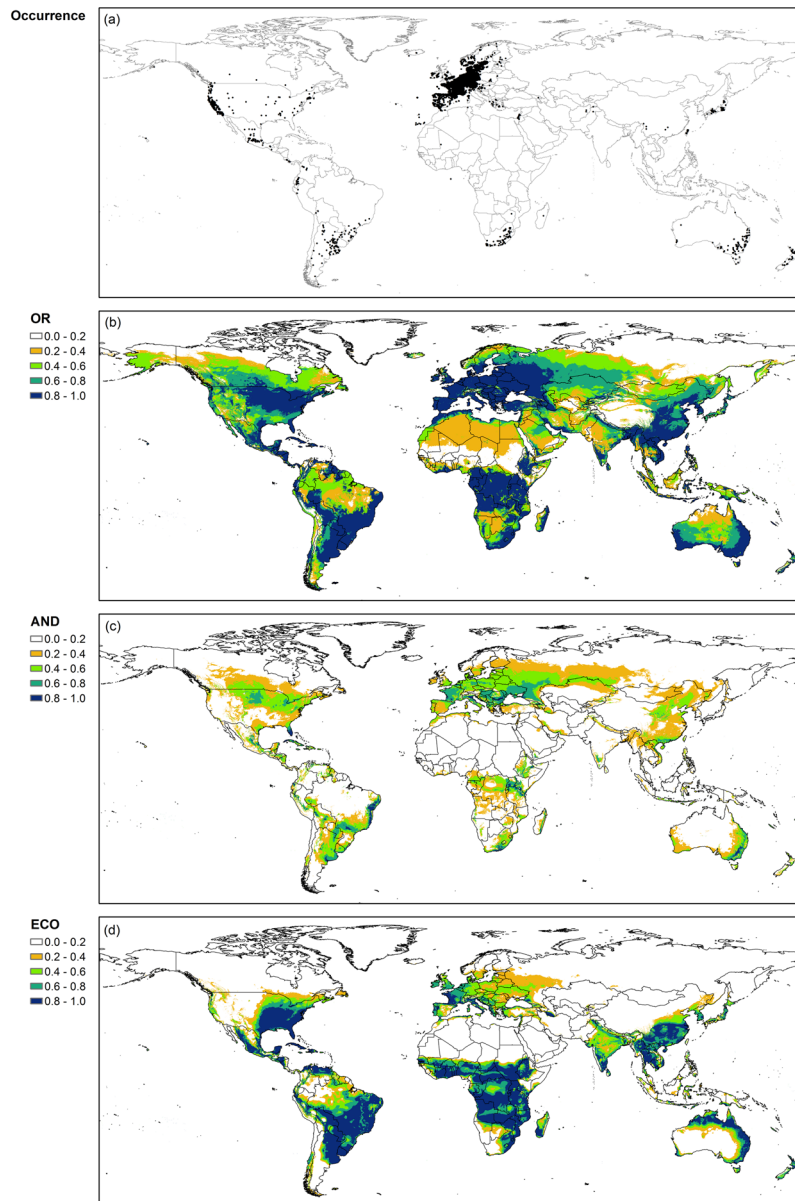


Figure 4. Map of global occurrence (a) and climate suitability index for annual ryegrass (b–d) obtained from the OR_F model (OR; b), the AND_F model (AND; c), and the EcoCrop model (ECO; d), respectively. ArcMap (version 10.0; <http://desktop.arcgis.com/en/arcmap/>) was used to create the maps.

suitability index and yield at the site-years was considerably low (0.05) for the EcoCrop model compared to the other models.

Spatial distribution of the climate suitability index. Both OR_F and AND_F models tended to have a greater suitability index for mid and high latitude areas where annual ryegrass grows readily (Figs 4 and 5a). For example, the average climate suitability index of the OR_F model was >0.6 between 35° and 55°. Spatial distribution of climate suitability index of the AND_F model was similar to that of the OR_F model although the magnitude of the index for the AND_F model was smaller than that for OR_F model. In contrast, the EcoCrop model tended to have a greater climate suitability index for low latitude areas, e.g., between –20° and 20°, than for mid-high latitude areas.

The climate suitability index of the OR_F model for alfalfa was relatively high for mid latitude areas where the most occurrence sites were located (Figs 5b and 6). Areas with a high suitability index, e.g., >0.90, were large in Europe and the northern part of North America. The values of climate suitability index >0.90 were obtained only for small areas in Brazil and Africa. Similarly, the high values of the index for the AND_F model tended to coincide with sites where alfalfa has been grown. On the other hand, the EcoCrop model had considerably high values for the suitability index for the southern hemisphere where only a small number of occurrence sites have been recorded.

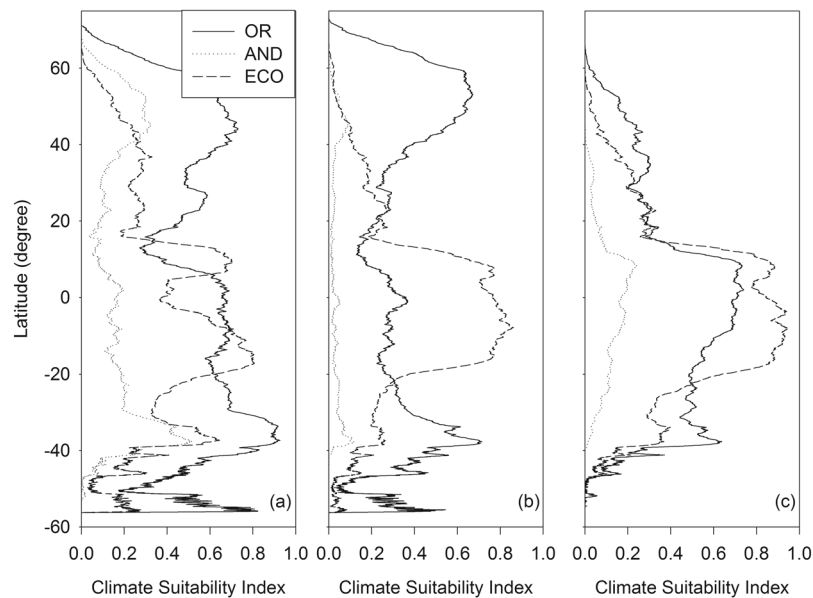


Figure 5. Average climate suitability index of annual ryegrass (a), alfalfa (b), and sorghum (c) by latitude. The values of climate suitability index were obtained from the OR_F model (OR), the AND_F model (AND), and the EcoCrop model (ECO), respectively.

The spatial distribution of climate suitability indices for both OR_F and AND_F models differed between alfalfa and sorghum, whereas that for the EcoCrop model was relatively similar for both crops. The OR_F model had relatively greater values of climate suitability index for sorghum in low latitude areas where the most occurrence sites located (Figs 5–7). The AND_F model had a similar distribution of climate suitability index over latitude with the OR_F model although the magnitude of climate suitability index was considerably smaller for the AND_F model than the OR_F model. The EcoCrop model also tended to have a relatively similar distribution of climate suitability index for sorghum with the OR_F model and AND_F model, although the climate suitability index was relatively low for the northern Europe.

The OR_F model had a climate suitability index >0 at occurrence sites for alfalfa and sorghum (Figs 6 and 7). In contrast, the AND_F model and the EcoCrop models had zero values of climate suitability index at a considerably large number of occurrence sites. For example, the climate suitability index of the AND_F model was zero at about 51% of occurrence sites for alfalfa and 28% for sorghum. In addition, the climate suitability indices for both crops were relatively low mostly for areas with dry climate conditions where irrigation is usually performed, such as the northwestern US for alfalfa and northern India for sorghum.

The climate suitability index of the OR_F model had the greatest likelihood of occurrence for annual ryegrass, alfalfa, and sorghum of the three models (Fig. 8). For example, the value of $-\log(L_{OR})$ for annual ryegrass was -196 whereas those for

$-\log(L_{AND})$ and $-\log(L_{ECO})$ were -5 and 2 , respectively. The negative logarithm of L tended to be smaller for annual ryegrass and alfalfa than that for sorghum for all models. The value of $-\log(L_{ECO})$ was smaller than that of $-\log(L_{AND})$ only for alfalfa.

Discussion

This study demonstrated that the use of t-conorm such as the OR_F model was beneficial in classifying sites where climate quality would be suitable for growth of a crop. The OR_F model had the greatest likelihood of occurrence for the climate suitability index of the three models used in this study. The spatial distribution of climate suitability index for the OR_F model was similar to that of occurrence sites. The climate suitability index of the OR_F model was also highly correlated with yield of a crop, e.g., annual ryegrass, when weather data for the site were used. These findings suggest that the OR_F model based on t-conorm provides a more reliable assessment of climate suitability for the crops of interest compared to the other models based on the t-norm. The application of the t-conorm allowed for a reasonable representation of plant response to temperature and precipitation conditions in the assessment of climate suitability. The same rule statements and parameter values were applied to both OR_F and AND_F models, except for the logical operation between suitability of temperature and moisture. Still, the OR_F model, which is based on the t-conorm, had the most reliable assessment of climate suitability of three crops compared with the AND_F models based on the other logical operation, i.e., the t-norm. Thus, a choice of logical operation, e.g., t-conorm over t-norm, in assessment of climate quality for survival and growth of a plant would have a considerable impact on the reliability of a climate suitability index.

Our results indicated that incorporation of rules to represent climate conditions for growth of a species allowed reasonable assessment of climate suitability. Temperature and precipitation suitability of the fuzzy logic system was determined to take into account the response of a grass species to given climate conditions in terms of survival and growth. As a result, the climate suitability index based on the AND_F model had a greater correlation

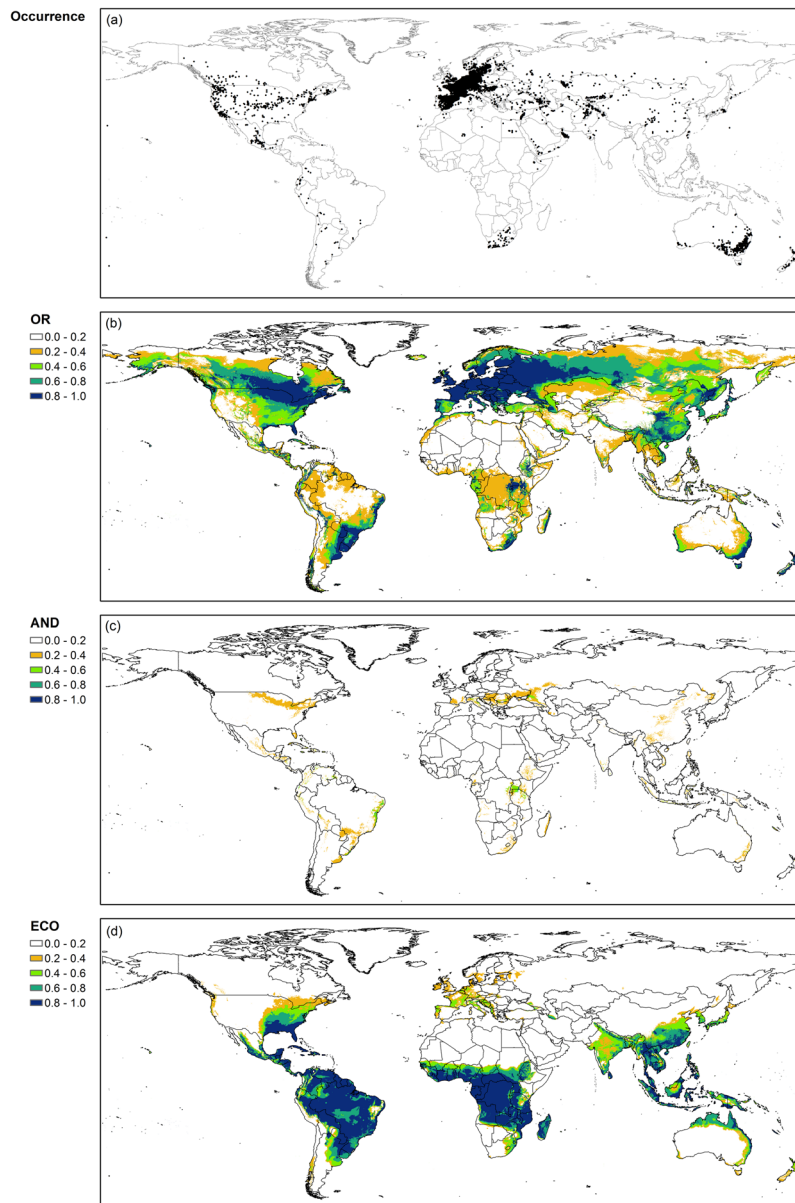


Figure 6. Map of global occurrence (a) and climate suitability index for alfalfa (b–d) obtained from the OR_F model (OR; b), the AND_F model (AND; c), and the EcoCrop model (ECO; d), respectively. ArcMap (version 10.0; <http://desktop.arcgis.com/en/arcmap/>) was used to create the maps.

with yields of annual ryegrass than the one based on the EcoCrop model although the parameter values of both models were identical. Such observations would result from, in part, the fact that the EcoCrop model depended on climate variables that could explain a relatively small variation in growth of a species for the Law of the Minimum.

Although it is possible for a species to survive in an area where climate conditions are marginal, those marginal conditions would be the subset of conditions suitable for the species. Because a t-norm is used to determine an intersection of those conditions, it would be challenging to take into account climate conditions marginally suitable for growth of a species. As a result, a climate suitability index of zero was obtained even for sites where crop yield was considerably high when the t-norm was used to determine climate suitability index. Estes *et al.*¹⁵ reported that a species distribution model explained more variation in yield when the model was calibrated using sites with a high yield rather than all the occurrence sites. This previous finding could be related to those problems associated with a model based on t-norm.

Approaches based on a t-norm such as the AND_F model and the EcoCrop model, would suffer from collinearity between temperature and precipitation in relation to plant growth in assessing climate suitability of a species. Chakraborty *et al.*⁴⁶ reported that temperature variables had a greater impact on growth performance than precipitation-related variables. Under high-temperature conditions, for example, the growth of annual ryegrass

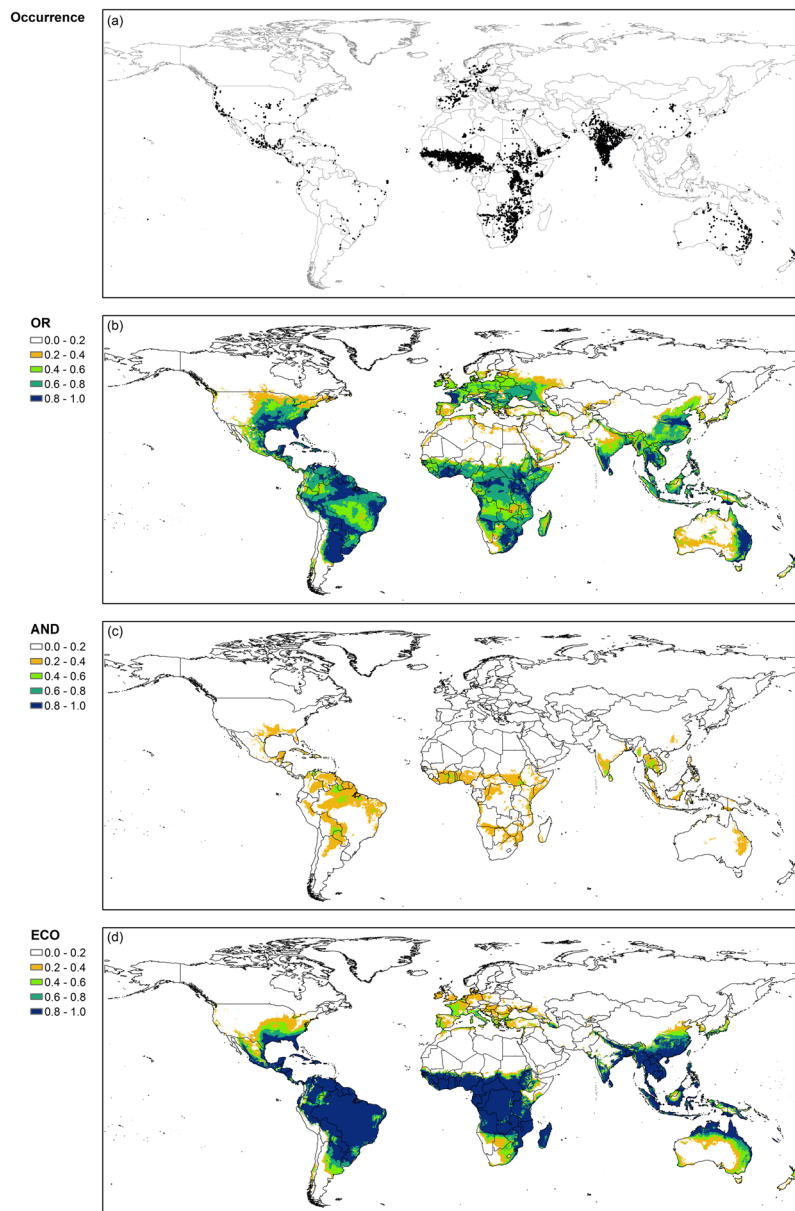


Figure 7. Map of global occurrence (a) and climate suitability index for sorghum (b–d) obtained from the OR_F model (OR; b), the AND_F model (AND; c), and the EcoCrop model (ECO; d), respectively. ArcMap (version 10.0; <http://desktop.arcgis.com/en/arcmap/>) was used to create the maps.

could benefit from ample soil moisture, which could result from optimum rainfall. Such an interaction between temperature and precipitation could not be realized using a t-norm.

The OR_F model would have limitation in assessing climate suitability of a crop due to its assumption such that a grower would manage a crop to maximize its yield under a rain-fed condition. Accurate assessment of climate suitability for a crop would require managements information such as planting date or irrigation options. However, such information is usually available only for major crops and not for other crops such as forages⁴⁷. This would result in overprediction for the OR_F model because growers may be unable to select the best management option due to a lower priority for the forage crops or limited knowledge on climate. On the other hand, a crop can be grown in an irrigated area with low climate suitability, e.g., under a semi-arid climate condition. In such an area, it is likely that the OR_F model would underpredict the climate suitability index for a crop.

The OR_F model can be extended to improve assessment of climate suitability using other abiotic variables associated with crop productivity. For example, solar radiation has been used a required input variable for a process-based crop model in order to simulate photosynthesis and growth accurately⁹. Soil conditions can also play an important role in assessing crop suitability³⁴. These variables can be incorporated into a climate suitability model when a large field data become available through an optimization process for the model. Benitez and Casillas²³ and Fernandez *et al.*²⁴ used an optimization approach to develop a hierarchical fuzzy model that has a

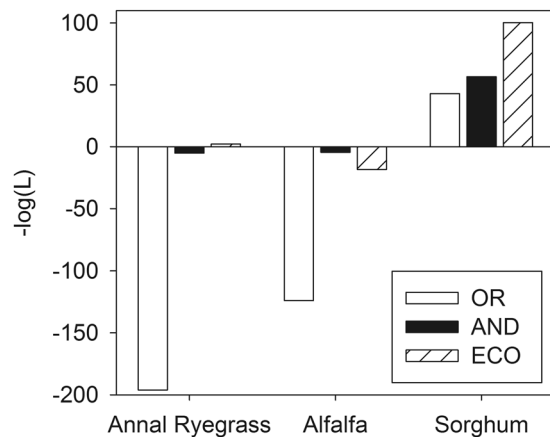


Figure 8. Comparison of the negative logarithm of L to maximize the joint likelihood of occurrence for a given climate suitability index. The climate suitability index for the models based on t-conorm and t-norm, and the EcoCrop model are denoted by OR , AND , and ECO , respectively.

similar architecture to the OR_F model. A hybrid approach including combination of fuzzy logic, clustering, and classification as well as soft computing can also be used to improve the OR_F model^{48,49}. In addition, distribution of multiple species can be assessed using the fuzzy logic system that is derived from knowledge on interaction between multiple species as well as observation data⁵⁰. A sensitivity analysis would be needed for further application of the OR_F model. The OR_F model can be used to support decision making for the design of a new cropping system in a region, which could be used for the development of climate change adaptation options⁴. For example, Rippeke *et al.*³⁰ showed that assessment of climate suitability for a crop would help with the development of climate change adaptation policies. Still, the uncertainty of the OR_F model needs to be analyzed before it is used in practice⁵¹. Janssen *et al.*⁵² suggested that the uncertainty of a fuzzy logic system can be assessed in terms of context, model structure, parameter, and input data. For example, a temporal resolution of climate input data could have a considerable impact on the uncertainty of the OR_F model (Appendix E). Thus, the framework for uncertainty analysis such as suggested by Walker *et al.*⁵³ can be used for the OR_F model.

It would be preferable to examine the circumstances under which different types of logical operation would be better suited to represent the impact of temperature and rainfall conditions on climate suitability. For example, Genesis and Jonas⁵⁴ used the Fuzzy Gamma overlay operations to predict potential yields of food and income crops. Zabel *et al.*³⁸ assessed land suitability for the cultivation of various crops using a t-norm. Combination of different logical operations, e.g., t-conorm and gamma overlay function, could improve the reliability of the climate suitability assessment of a species. For example, an ensemble approach could be used for the assessment of climate suitability because a diverse set of methods has been developed for quantifying spatial distribution of a species⁵⁵. Still, the OR_F model had advantages over other approaches because its parameters could be obtained from an existing knowledge recorded in the Ecocrop database and published literature.

The results from this study showed that a model based on t-conorm resulted in more reliable assessment of climate suitability for annual ryegrass, alfalfa, and sorghum compared with a model based on t-norm. This suggested that climate suitability assessment can be improved by taking into account the quality of environmental conditions for growth as well as survival of a species. In particular, the proposed algorithm in this study can provide an alternative approach to the Law of the Minimum that has been used to determine spatial distribution of a species. Our results also suggested that climate suitability of a species can be assessed without an iterative calibration process using existing knowledge, which could facilitate the impact assessment of biotic interaction on species distribution.

References

- Howden, S. M. *et al.* Adapting agriculture to climate change. *Proc. Natl. Acad. Sci.* **104**, 19691–19696 (2007).
- Lemaire, G. *et al.* Integrated crop–livestock systems: Strategies to achieve synergy between agricultural production and environmental quality. *Agric. Ecosyst. Environ.* **190**, 4–8 (2014).
- Lobell, D. B. *et al.* Prioritizing climate change adaptation needs for food security in 2030. *Science* **319**, 607–610 (2008).
- Olesen, J. E. *et al.* Impacts and adaptation of European crop production systems to climate change. *Eur. J. Agron.* **34**, 96–112 (2011).
- Cullen, B. R. *et al.* Climate change effects on pasture systems in south-eastern Australia. *Crop Pasture Sci.* **60**, 933–942 (2009).
- Kiniry, J. R. *et al.* EPIC model parameters for cereal, oilseed, and forage crops in the northern Great Plains region. *Can. J. Plant Sci.* **75**, 679–688 (1995).
- Ojeda, J. J. *et al.* Evaluation of the agricultural production systems simulator simulating Lucerne and annual ryegrass dry matter yield in the Argentine Pampas and south-eastern Australia. *Agr. Syst.* **143**, 61–75 (2016).
- Abraha, M. G. & Savage, M. J. The soil water balance of rainfed and irrigated oats, Italian rye grass and rye using the CropSyst model. *Irrigation Sci.* **26**, 203–212 (2008).
- Hoogenboom, G., Jones, J. W., Traore, P. C. S. & Boote, K. J. Experiments and data for model evaluation and application in *Improving Soil Fertility Recommendations in Africa using the Decision Support Systems for Agrotechnology Transfers (DSSAT)* (ed. Kihara, J. *et al.*) 9–18 (Springer, 2012).
- Hunt, L. A., White, J. W. & Hoogenboom, G. Agronomic data: advances in documentation and protocols for exchange and use. *Agr. Syst.* **70**, 477–492 (2001).

11. White, J. W. *et al.* Integrated description of agricultural field experiments and production: The ICASA Version 2.0 data standards. *Comput. Electron. Agr.* **96**, 1–12 (2013).
12. Jarvis, A., Lane, A. & Hijmans, R. J. The effect of climate change on crop wild relatives. *Agric. Ecosyst. Environ.* **126**, 13–23 (2008).
13. Phillips, S. J. *et al.* Opening the black box: an open-source release of Maxent. *Ecography* **40**, 887–893 (2017).
14. Ward, G. *et al.* Presence-only data and the EM algorithm. *Biometrics* **65**, 554–563 (2009).
15. Estes, L. D. *et al.* Comparing mechanistic and empirical model projections of crop suitability and productivity: implications for ecological forecasting: Crop model intercomparison. *Global Ecol. Biogeogr.* **22**, 1007–1018 (2013).
16. Jeong, J. H. *et al.* Random forests for global and regional crop yield predictions. *PLoS One* **11**, e0156571 (2016).
17. Skov, F. & Svenning, J. Potential impact of climatic change on the distribution of forest herbs in Europe. *Ecography* **27**, 366–380 (2004).
18. Azadi, H., Shahvali, M., van den Berg, J. & Faghih, N. Sustainable rangeland management using a multi-fuzzy model: How to deal with heterogeneous experts' knowledge. *J. Environ. Manage.* **83**, 236–249 (2007).
19. Nisar Ahamed, T. R., Gopal Rao, K. & Murthy, J. S. R. GIS-based fuzzy membership model for crop-land suitability analysis. *Agr. Syst.* **63**, 75–95 (2000).
20. Jun, K.-S., Chung, E.-S., Kim, Y.-G. & Kim, Y. A fuzzy multi-criteria approach to flood risk vulnerability in South Korea by considering climate change impacts. *Expert Syst. Appl.* **40**, 1003–1013 (2013).
21. Kazakis, N., Spiliotis, M., Voudouris, K., Pliakas, F.-K. & Papadopoulos, B. A fuzzy multicriteria categorization of the GALDIT method to assess seawater intrusion vulnerability of coastal aquifers. *Sci. Total Environ.* **621**, 524–534 (2018).
22. Tsakiris, G. & Spiliotis, M. Planning Against Long Term Water Scarcity: A Fuzzy Multicriteria Approach. *Water Resour. Manag.* **25**, 1103–1129 (2011).
23. Benítez, A. D. & Casillas, J. Multi-objective genetic learning of serial hierarchical fuzzy systems for large-scale problems. *Soft Comput.* **17**, 165–194 (2013).
24. Fernández, A., del Jesus, M. J. & Herrera, F. Hierarchical fuzzy rule based classification systems with genetic rule selection for imbalanced data-sets. *Int. J. Approx. Reason.* **50**, 561–577 (2009).
25. Zhou, S.-M. & Gan, J. Q. Low-level interpretability and high-level interpretability: a unified view of data-driven interpretable fuzzy system modelling. *Fuzzy Set. Syst.* **159**, 3091–3131 (2008).
26. Cordon, O. A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems. *Int. J. Approx. Reason.* **52**, 894–913 (2011).
27. Hijmans, R. J. & Graham, C. H. The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biol.* **12**, 2272–2281 (2006).
28. Hackett, C. & Vanclay, J. K. Mobilizing expert knowledge of tree growth with the PLANTGRO and INFER systems. *Ecol. Model.* **106**, 233–246 (1998).
29. Hijmans, R. J., Guarino, L., Cruz, M. & Rojas, E. Computer tools for spatial analysis of plant genetic resources data. 1. DIVA-GIS. *Plant Genet. Resour. Newsl.* **127**, 15–19 (2001).
30. Rippke, U. *et al.* Timescales of transformational climate change adaptation in sub-Saharan African agriculture. *Nat. Clim. Change* **6**, 605–609 (2016).
31. Austin, M. Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecol. Model.* **200**, 1–19 (2007).
32. Cable, D. R. Influence of precipitation on perennial grass production in the semi-desert Southwest. *Ecology* **56**, 981–986 (1975).
33. Ramirez-Villegas, J., Jarvis, A. & Läderach, P. Empirical approaches for assessing impacts of climate change on agriculture: The EcoCrop model and a case study with grain sorghum. *Agr. Forest Meteorol.* **170**, 67–78 (2013).
34. Piikki, K. *et al.* Improvement of spatial modelling of crop suitability using a new digital soil map of Tanzania. *South African J. Plant Soil* **34**, 243–254 (2017).
35. Kallenbach, R. L. *et al.* Stockpiled annual ryegrass for winter forage in the lower Midwestern USA. *Crop Sci.* **43**, 1414–1417 (2003).
36. Ward, J. D., Readfern, D. D., McCormick, M. E. & Cuomo, G. J. Chemical composition, ensiling characteristics, and apparent digestibility of summer annual forages in a subtropical double-cropping system with annual ryegrass. *J. Dairy Sci.* **84**, 177–182 (2001).
37. Klir, G. J. & Yuan, B. Fuzzy sets and fuzzy logic (Prentice-Hall 1995).
38. Zabel, F., Putzenlechner, B. & Mauser, W. Global agricultural land resources—a high resolution suitability evaluation and its perspectives until 2100 under climate change conditions. *PLoS One* **9**, e107522 (2014).
39. van der Ploeg, R. R., Bo ¨hm, W. & Kirkham, M. B. On the origin of the theory of mineral nutrition of plants and the law of the minimum. *Soil Sci. Soc. Am. J.* **63**, 1055–1062 (1999).
40. Fukuda, S. *et al.* Effect of model formulation on the optimization of a genetic Takagi–Sugeno fuzzy system for fish habitat suitability evaluation. *Ecol. Model.* **222**, 1401–1413 (2011).
41. Zadeh, L. A. Fuzzy sets. *Inform. Control* **8**, 338–353 (1965).
42. Eagles, C. F., Williams, J. & Louis, D. V. Recovery after freezing in *Avena sativa* L., *Lolium perenne* L. and *L. multiflorum* Lam. *New Phytol.* **123**, 477–483 (1993).
43. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas: New climate surfaces for global land areas. *Int. J. Climatol.* <https://doi.org/10.1002/joc.5086> (2017).
44. Royle, J. A., Chandler, R. B., Yackulic, C. & Nichols, J. D. Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. *Methods Ecol. Evol.* **3**, 545–554 (2012).
45. Culpepper, S. A. & Aguinis, H. R is for revolution: A cutting-edge, free, open source statistical package. *Organ. Res. Meth.* **14**, 735–740 (2011).
46. Chakraborty, D. *et al.* Selecting populations for non-analogous climate conditions using universal response functions: the case of douglas-fir in central Europe. *PLoS One* **10**, e0136357 (2015).
47. Sacks, W. J., Deryng, D., Foley, J. A. & Ramankutty, N. Crop planting dates: an analysis of global patterns: Global crop planting dates. *Global Ecol. Biogeogr.* **19**, 607–620 (2010).
48. Anezakis, V.-D., Demertzis, K., Iliadis, L. & Spartalis, S. Hybrid intelligent modeling of wild fires risk. *Evolving Systems.* <https://doi.org/10.1007/s12530-017-9196-6> (2017).
49. Bougoudis, I., Demertzis, K., Iliadis, L., Anezakis, V.-D. & Papaleonidas, A. FuSSFFra, a fuzzy semi-supervised forecasting framework: the case of the air pollution in Athens. *Neural Comput. Appl.* **29**, 375–388 (2018).
50. Kissling, W. D. *et al.* Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents: Modelling multispecies interactions. *J. Biogeogr.* **39**, 2163–2178 (2012).
51. Chen, Y., Yu, J. & Khan, S. Spatial sensitivity analysis of multi-criteria weights in GIS-based land suitability evaluation. *Environ. Modell. Softw.* **25**, 1582–1591 (2010).
52. Janssen, J. A. E. B. *et al.* Assessment of uncertainties in expert knowledge, illustrated in fuzzy rule-based models. *Ecol. Model.* **221**, 1245–1251 (2010).
53. Walker, W. E. *et al.* Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integrated Assess.* **4**, 5–17 (2003).
54. Genesis, T. Y. & Jonas, A. Crop yield gaps in Cameroon. *Ambio* **43**, 175–190 (2014).
55. Yalcin, S. & Leroux, S. J. Diversity and suitability of existing methods and metrics for quantifying species range shifts. *Global Ecol. Biogeogr.* **26**, 609–624 (2017).

Acknowledgements

This study was supported by the Korea Ministry of Environment (MOE) as “Climate Change Correspondence Program” (Project Number 2014001310008). We also thank Ji-Yoon Park, Julian Ramirez Villegas, and Patricia Moreno Cadena for their comments to improve our manuscript.

Author Contributions

H.K. and K.S.K. conceived and designed algorithm. S.W.H. created and analyzed global maps. G.H. and C.H.P. provided data and helped with the manuscript. K.S.K. wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-28291-3>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018