Applied Soft Computing 59 (2017) 326–339

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Novel fuzzy linguistic based mathematical model to assess risk of invasive alien plant species

H.O.W. Peiris[a,*], S. Chakraverty[b], S.S.N. Perera[c], S.M.W. Ranwala[d]

[a] Department of Mathematics and Computer Science, Open University, Sri Lanka
[b] Department of Mathematics,National Institue of Technology Rourkela, India
[c] Research & Development Centre for Mathematical Modelling, Department of Mathematics, University of Colombo, Sri Lanka
[d] Department of Plant Sciences, University of Colombo, Sri Lanka

ARTICLE INFO

Article history:
Received 13 April 2016
Received in revised form 11 May 2017
Accepted 4 June 2017
Available online 8 June 2017

Keywords:
Invasive Alien Species
Biological traits
Fuzzy Membership Functions
Fuzzy Averaging Operators
Linguistic variables

ABSTRACT

The purpose of this study is to develop a mathematical model for risk assessment that evaluates the invasion risk of invasive alien plant species based on their biological attributes. Data for most of the attributes were qualitative and in the form of linguistic terms. In order to handle the numerical and linguistic variables, we proposed three models in the fuzzy environment. In the first model, all the selected attributes were considered as equally important to invasiveness and in the second and third models, these are considered to be unequally important. The important weights for the biological attributes in these models were gathered from the group of experts in plant sciences. Proposed Model III incorporates more sophisticated linguistic tool than Model II. Model II gives better predictions in comparison to the first and third models and it is found to be better tracking system for identifying potential invaders as in the case of conventional risk assessment method.

1. Introduction

Invasive Alien Species (IAS) are considered as a serious threat to the existence of native species as they alter physical, chemical and biological components of the environment [1]. Invasive alien plants usually have higher ability of tolerance due to various environmental conditions, strong reproductive potential, efficient dispersal of their seeds and other plant parts which make them established in a great diversity of habitats. Therefore invasive potential of plant species can be recognized by their biological traits or biological attributes [2]. As such, Appendix A provides more discussions on IAS.

To identify invasive alien species from other organisms, scientists have developed different methodologies. Risk assessment is one such approach. Application of risk assessments has become a vital part of comprehensive prevention strategy of IAS [3–5]. It is being used in many countries to identify the potential IAS at the time of introducing any species from one country to another. Risk assessments are in the form of a set of questions with predefined answers for each question. The questions have been developed based on risk factors which are the biological traits that may reflect invasiveness of a species. The outcome of the assessment is the risk value for a particular plant or animal species, which is usually the sum of scores that have been obtained by each question.

Risk assessment is only predictive tool developed using existing information at a particular time and greatly varies among countries, however most are based on qualitative methods. Risk assessments are context dependent and it requires in-depth information about the species to be assessed. When information on biological traits are qualitative in nature, determination of the invasiveness of species is often accompanied with imprecision and uncertainty. In order to minimize the uncertainty and bias towards the risk assessments, these are usually conducted with the opinion of a team of domain experts. In many cases, it has been a manual process that takes a considerable time to complete. We believe that integration of a mathematical approach with risk assessments may assist (at least partly) to overcome this situation and increase the efficacy of the process. For an example; when the risk factors are unquantifiable, it may be stated only in linguistic terms in the conventional method (e.g., when evaluating the "vegetative reproduction strength" of a plant, terms like "low", "medium", "high" are used). In this situation, fuzzy set theory may play a vital role in modeling qualitative information as it is capable to handle fuzziness in qualitative aspects representing them in linguistic variables.

Therefore our aim has been to develop a mathematical model which incorporates quantitative as well as qualitative biological

* Corresponding author.
E-mail address: oshivida@yahoo.com (H.O.W. Peiris).

http://dx.doi.org/10.1016/j.asoc.2017.06.006
1568-4946/

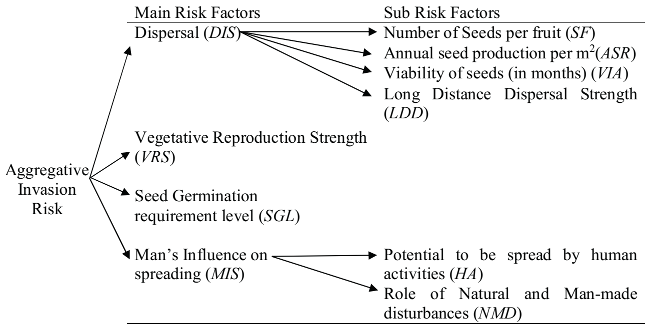

Fig. 1. Structure of the Invasion risk Model used in this study.

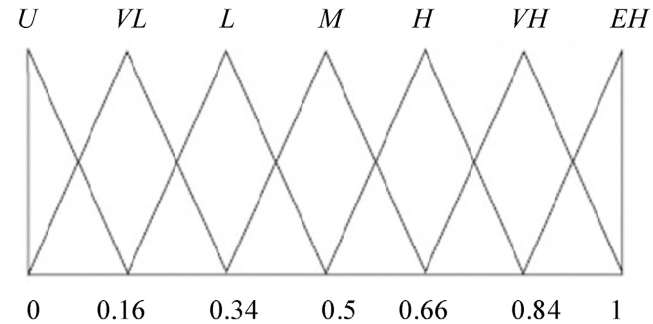

Fig. 2. Membership functions of linguistic term set *S*.

risk factors to evaluate the invasiveness risk of Invasive alien plant species. In this study, eight biological traits of plants that contribute heavily for the invasiveness were chosen as risk factors. The set of risk factors consisted of quantitative as well as qualitative variables. In this work, we developed two different models using fuzzy linguistic aggregation operators such as Linguistic Ordered Weighted Average operator (LOWA) and Linguistic Weighted Average (LWA) operator [8,11] to obtain the relationship among the biological traits. The risk factors were categorized according to their importance, i.e. whether they are directly important or indirectly important for invasiveness when developing Model I and Model II respectively. In recent years, new concepts based on linguistic approach were designed to overcome the shortcomings of the aggregation operators [17,28–31]. Hence, Model III was developed by incorporating weighted Majority guided Linguistic IOWA (MLIOWA) operator to see which aggregation technique is more suitable to provide a reliable model.

The models were validated by testing with a set of known invasive plant species and non invasive species found in Sri Lanka. Data for the National Risk Assessment (NRA) by the National Invasive Specialists Group affiliated to Ministry of Environment and Renewable Resources, Sri Lanka (2012) were used.

2. Structure of the model

To develop the risk assessment mathematical model, eight biological traits related to reproduction and dispersal were selected as the risk factors as per the opinion of botany experts. These represented parameters [5] are Number of seeds per fruit (*SF*), Annual seed production per m^2(*ASR*), Viability of seeds (in months) (*VIA*), Long distance dispersal strength (*LDD*), Vegetative reproduction strength (*VRS*), Seed germination requirement level (*SGL*), Potential to be spread by human activities (*HA*) and Role of natural and man-made disturbances (*NMD*).

To establish the structure of the model, the eight risk factors were divided into two categories, under main and sub-risk level (Fig. 1). The data for sub risk factors in dispersal category were quantitative and other factors were in the form of linguistic terms.

Firstly, a sub model which aggregates the risk of dispersal related risk factors was constructed. The output of this sub model has been converted into a linguistic terms. Therefore, data of each main factor takes the same format. The steps for developing the sub model are described in Section 4.

In order to develop the model, related data of twenty seven known invasive species were used. Appendix B provides general facts about these species. In addition, related data including respective risk scores obtained in NRA process of six known invasive species and four non- invasive species were used for validation of the model.

In the next section we present some theoretical aspects of linguistic approach.

3. The choice of linguistic term set for the risk levels

We already mentioned that the final output of the model was expected to be an invasive risk level of plant species where the parameters (main risk factors) could be represented in linguistic variables. Therefore we need to choose the cardinality of the set of linguistic terms which could discriminate the different countings of uncertainty. Below we present the definition of linguistic variable as defined by Zadeh [12].

Definition 1 *([12])*. A linguistic variable is characterized by a quintuple (L, $H(L)$,U,G,M) in which L is the name of the variable; $H(L)$ (or simply H) denotes the term set of L, i.e., the set of names of linguistic values of L, with each values being a fuzzy variable denoted generically by X and ranging across a universe of discourse U which is associated with the base variable u; G is a syntactic rule (which usually takes the form of a grammar) for generating the names of the values of L; and M is semantic rule for associating its meaning with each L, $M(X)$, which is a fuzzy subset of U.

One may note that the meaning of each linguistic term has been given by means of a fuzzy subset defined in the [0, 1] interval, which are usually described by membership functions [8]. The triangular membership functions have been found to be good enough to capture the vagueness of the linguistic assessments. This is parametrically represented by a 3-tuple(a_i, l_i, r_i). The first parameter indicated the membership value of 1; the second and third parameters indicated the left and right width from a_i to the left and right end point of the domain of the membership function respectively.

A set of seven linguistic terms are used by the plant science experts to express their evaluation as below (see Fig. 2.):

$s_0 = U = Unlikely = (0, 0, 0.16)$
$s_1 = VL = Very\ Low = (0.16, 0.16, 0.18)$
$s_2 = L = Low = (0.34, 0.18, 0.16)$
$s_3 = M = Medium = (0.5, 0.16, 0.16)$
$s_4 = H = High = (0.66, 0.16, 0.18)$
$s_5 = VH = Very\ High = (0.84, 0.18, 0.16)$
$s_6 = EH = Extremely\ High = (1, 0.16, 0)$

These linguistic terms are placed symmetrically around the middle term which represent the assessment of "approximately 0.5".

In this kind of ordered structure, it is often required to see whether the linguistic term set satisfies the following additional characteristics as defined in [8]:

- There is a negation operator, e.g., $Neg(s_i) = s_j$, $j = T - i$ ($T + 1$ is the cardinality).
- Maximization operator: $\mathrm{Max}\left(s_i, s_j\right) = s_i$ if $s_i \geq s_j$.
- Minimization operator: $\mathrm{Min}\left(s_i, s_j\right) = s_i$ if $s_i \leq s_j$.

In the next section we present the sub model which evaluates the risk of dispersal related factors of invasive species.

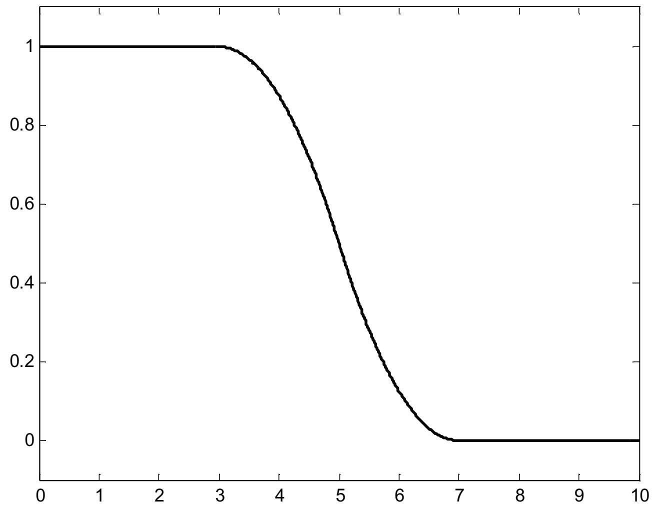

Fig. 3. Z-shaped Membership Function.

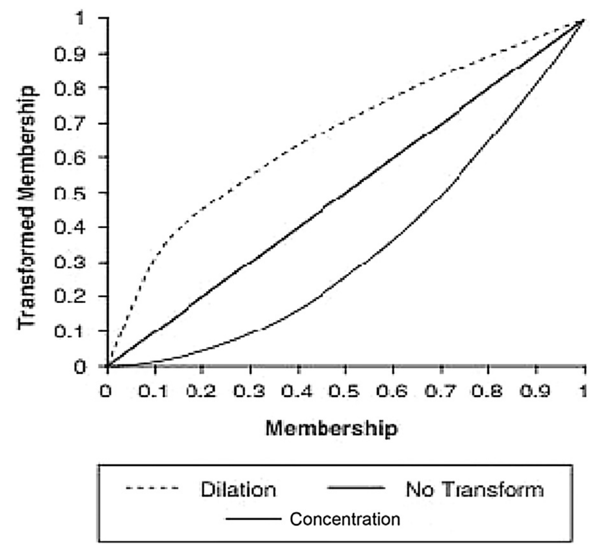

Fig. 4. Concentration and Dilation operators.

4. Sub model

4.1. Fuzzy membership functions (FMF) for dispersal traits

In this subsection we focus on developing the FMF for the four sub risk factors related to dispersal [15]. In order to develop the membership functions for the four risk factors, the lower and upper boundary points have been determined by considering their behavior with regard to the invasive potential. Experts' suggestions were considered when determining the boundary points using the following assumptions.

- lower boundary point is the lowest possible value which has the minimum effect to the invasive potential of a plant species.
- upper boundary point is the extreme value which has the maximum effect to the invasive potential of a plant species. The upper boundary chooses as an unrealistic extreme value in order to compatible to any invasive plant other than in the database.
- the invasive potential of plant species increase when the values of risk factors increase from lower to upper boundary points.

The functions between lower and upper boundary points are predicted based on the compatibility with the actual impacts of biological traits as in Eqs. (2)–(5). Initially, *Z*- shaped membership functions are defined for each parameter. The *Z*-shaped function represents an asymmetrical polynomial curve open to the left as defined in Eq. (1) [7].

$$\mu_Z(x,a,b,c) = \begin{cases} 1, & \text{for } x \le a \\ 1-2\left[(x-a)/(c-a)\right]^2, & \text{for } a \le x \le b \\ 2\left[(x-c)/(c-a)\right]^2, & \text{for } b \le x \le c \\ 0, & \text{for } x \ge c. \end{cases} \quad (1)$$

Graphical representation of a *Z*-shaped function has been depicted in Fig. 3. The process of fitting the actual impacts to the functions is described in Section 4.2. Various FMFs are now introduced as below:

4.1.1. FMF $U_A(x)$ for number of seeds per fruit

$$U_A(x) = \begin{cases} 1 & \text{for } x < 1 \\ 1-2\left[(x-1)/1000\right]^2 & \text{for } 1 \le x \le 501 \\ 2\left[(x-1001)/1000\right]^2 & \text{for } 501 < x \le 1001 \\ 0 & \text{for } x > 1001 \end{cases} \quad (2)$$

4.1.2. FMF $U_B(x)$ for annual seed production per m^2

$$U_B(x) = \begin{cases} 2\dfrac{(10000-x)^2}{8\times10^8} + 0.75 & \text{for } 0 \le x < 10000 \\ 2\dfrac{(100000-x)^2}{5.4\times10^{10}} + 0.45 & \text{for } 10000 \le x < 100000 \\ 2\dfrac{(10\times10^6-x)^2}{4.356\times10^{14}} & \text{for } 100000 \le x \le 10\times10^6 \\ 0 & \text{for } x > 10\times10^6 \end{cases} \quad (3)$$

4.1.3. FMF $U_C(x)$ for viability of seeds in months

$$U_C(x) = \begin{cases} 1 & \text{for } x < 3 \\ 1-2\left[(x-3)^2/2376060\right] & \text{for } 3 \le x < 602 \\ 2\left[(1200-x)^2/1028572\right] & \text{for } 602 \le x \le 1200 \\ 0 & \text{for } x > 1200 \end{cases} \quad (4)$$

4.1.4. FMF $U_D(x)$ for long distance dispersal strength

$$U_D(x) = \begin{cases} 1-2\left[x^2/160\right] & \text{for } 0 \le x < 2 \\ 0.95-2\left[(x-2)^2/60\right] & \text{for } 2 \le x < 5 \\ 2\left[(10-x)^2/77\right] & \text{for } 5 \le x \le 10 \end{cases} \quad (5)$$

Data for long distance dispersal strength of plant species are represented in ten point scale where point 0 denotes the lowest dispersal strength and point 10 the highest dispersal strength.

Fuzzy set for each parameter has been defined using their FMF's as below:

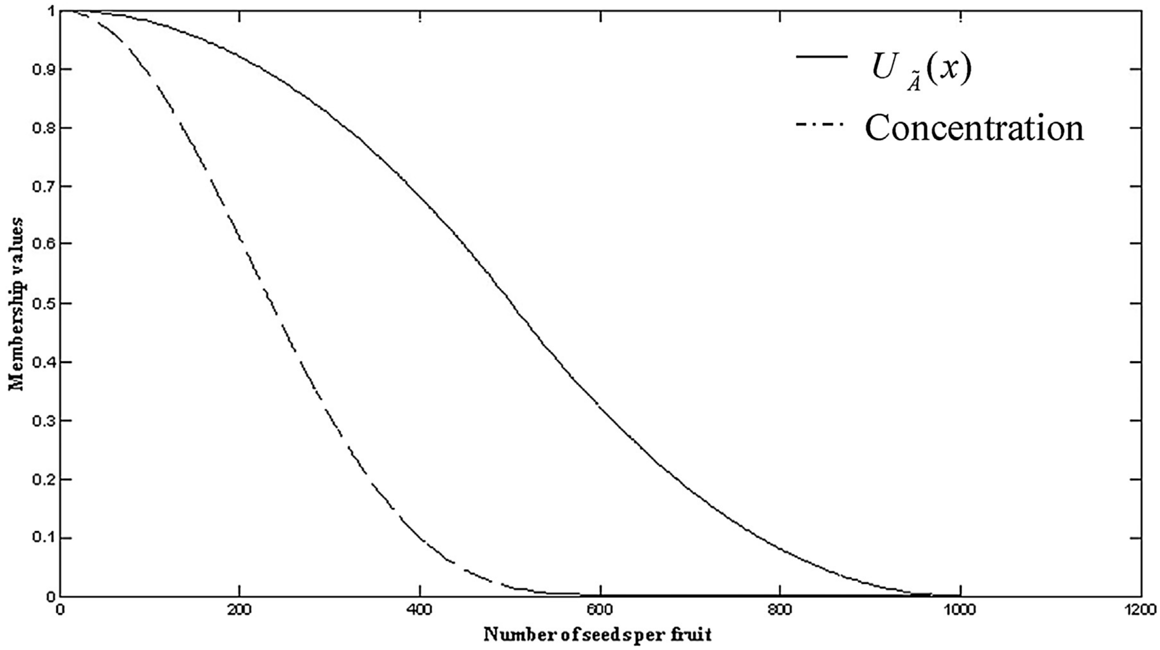

Fig. 5. Concentration of the Membership function for Seeds per fruit (*SF*).

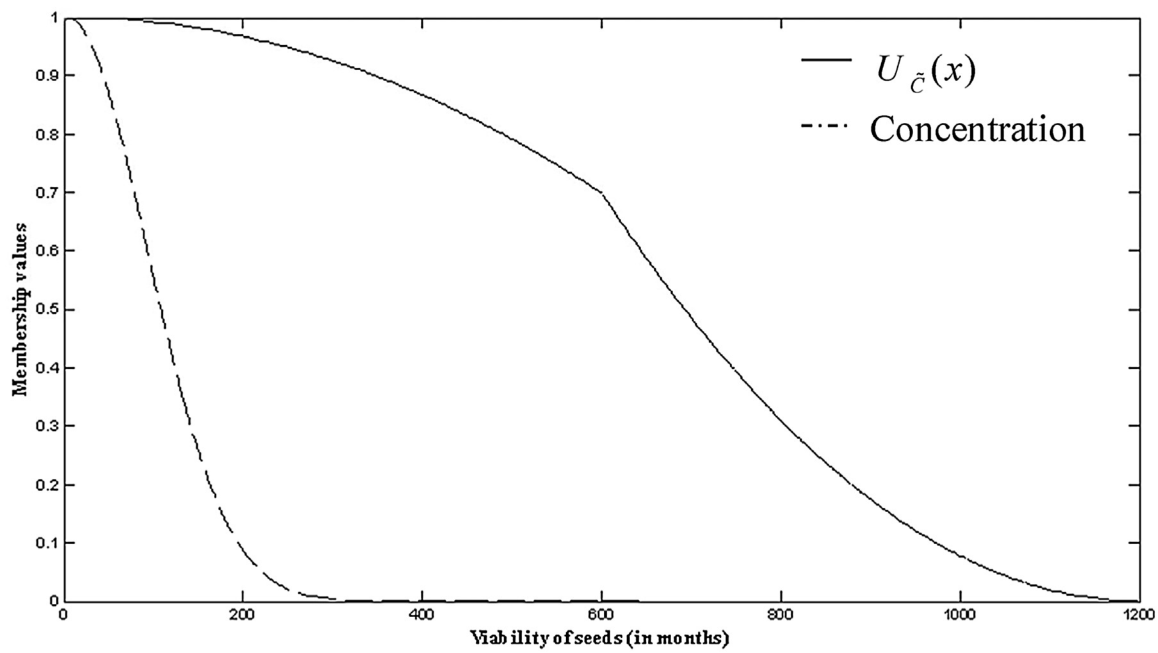

Fig. 6. Concentration of the Membership function for Viability of seeds (*VIA*).

Fuzzy set $\tilde{A}$ for number of seeds per fruit:

$$\tilde{A} = \left\{ (x, U_A(x)) \,|\, x \in R, U_A(x) \in [0, 1] \right\}. \quad (6)$$

Fuzzy set $\tilde{B}$ for annual seed rain per m^2:

$$\tilde{B} = \left\{ (x, U_B(x)) \,|\, x \in R, U_B(x) \in [0, 1] \right\}. \quad (7)$$

Fuzzy set $\tilde{C}$ for viability of seeds in months:

$$\tilde{C} = \left\{ (x, U_C(x)) \,|\, x \in R, U_C(x) \in [0, 1] \right\}. \quad (8)$$

Fuzzy set $\tilde{D}$ for long distance dispersal strength:

$$\tilde{D} = \left\{ (x, U_D(x)) \,|\, x \in R, U_D(x) \in [0, 1] \right\}. \quad (9)$$

4.2. Setting up the model

Dispersal risk of plant species has been generated by the sub model developed by aggregating the fuzzified dispersal related parameters. For this purpose, several parametric fuzzy set theoretic operators such as intersection and union of fuzzy sets [6,16] have been considered. In this study, the parametric operators; Hamacher, Yager and Dombi which allow for cumulative effects, interactions, and compensations between criteria are considered. The reason for introducing compensations, interactions in the aggregation process can be explained as below:

It is not realistic to assume that a species with extremely high viability of seeds has potential to become more invasive. Obviously,

Table 1
Comparison of Dispersal risk levels.

Invasive species	Dispersal Risk Level (Sub Model)	Dispersal Risk Level (NRA)
Alternanthera philoxeroides	*Low*	*Low*
Clidemia hirta	*High*	*High*
Miconia calvescens	*Extremely High*	*Very High*
Alstonia macrophylla	*Very High*	*Medium*
Annona glabra	*Very Low*	*Medium*
Clusia rosea	*Medium*	*Medium*
Dillenia suffructicosa	*Very Low*	*Medium*
Ageratina riparia	*High*	*Medium*
Mimosa invisa	*High*	*Very High*
Myroxylon balsamum	*Medium*	*Medium*
Tithonia diversiflora	*Low*	*Medium*
Mikania micrantha	*High*	*Medium*
Prosopis juliflora	*High*	*High*
Ulex europaeus	*Medium*	*High*
Mimosa pigra	*High*	*Very High*
Chromolaena odorata	*High*	*High*
Parthenium hysterophorus	*High*	*Medium*
Lantana camara	*Medium*	*Medium*
Imperata cylindrica	*High*	*High*
Opuntia stricta	*High*	*Very High*
Colubrina asiatica	*Low*	*Medium*
Pennisetum polystachion	*High*	*Medium*
Sphagneticola trilobata	*Low*	*Medium*
Zizphus mauritiana	*Very Low*	*Low*
Eichhornia crassipes	*High*	*Very High*
Pistia stratiotes	*High*	*Medium*
Leucaena leucocephala	*Medium*	*Medium*

the effect of high viability of seeds can be amplified by the presence of increase of annual seed rain and long distance dispersal strength with above moderate level or the effect can be compensated by low annual seed rain and moderate level of long distance dispersal strength.

The model using Hamacher *t*-norm operator showed significant results than that of the other operators which has been presented in the latter part of this section.

Hamacher *t*-norm operator defines the intersection of two fuzzy sets *A* and *B* by

$$U_H^p(x) = \frac{U_A(x) \cdot U_B(x)}{p + (1-p)\left[U_A(x) + U_B(x) - U_A(x)U_B(x)\right]} \quad 0 \le p \le 1 \tag{10}$$

- The dispersal related risk has been generated using the above Hamacher *t*-norm operator by intersecting four fuzzy sets (Eqs. (6)–(9)). In order to develop the model, data of 27 invasive plant species have been used as test data. These values were compared with NRA scores related to the relevant factors. In the comparison process, the membership functions have been modified using concentration and dilation operators until the scores become more likely to NRA scores.

One may note that these Concentration and Dilation operators can simply upgrade or downgrade the importance of the fuzzy sets.

The operator concentration reduces the grade of membership of all the elements of a fuzzy set that are only partly in the set. This is done in such a way that the less an element is in the set, the more its grade of membership is reduced (see Fig. 4.) [7,19,20]. The concentration of a fuzzy set *A* has been denoted by CON (*A*) and may be expressed as

$$U_{\mathrm{CON}(A)}(x) = U_A^{\alpha}(x) \quad \alpha < 1. \tag{11}$$

Dilation is the opposite of concentration. A fuzzy set is dilated by increasing the grade of membership of all elements that are partly in the set. The dilation of a fuzzy set *A* has been denoted by $\mathrm{DIL}(\tilde{A})$ and may be expressed as

$$U_{\mathrm{DIL}(A)}(x) = U_A^{\alpha}(x) \quad \alpha < 1. \tag{12}$$

In this kind of real world problems, it is not realistic to assume that each parameter has the same importance. As such, let us suppose that the annual seed rain is the better predictor of dispersal risk, while the importance of viability of seeds to be downgraded. Therefore, it is worth to introduce concentration and dilation operators in the aggregation process.

- The modified membership functions with concentrations and dilations are defined in Eqs. (2)–(5). Our aim is to investigate whether these modified membership functions fit well with plant categories having different levels of dispersal status. Therefore the model consists the following categories:

Category I: A plant whose *SF*≤200.
The membership function *SF* concentrates as $U_A^6(x)$ and $U_A(x)$ is replaced by $U_A^6(x)$

Fig. 5 depicts the membership function of *SF* and its concentration. The important part of the domain of concentrated membership function is between 0 and 200 (in months).

Category II: A plant whose *ASR* ≤20000 and *VIA*≤10yrs and *SF*≤200.

The membership function of *SF* dilates as $U_A^6(x)$ and membership function of *VIA* concentrate as $U_C^{73.5}(x)$ (see Fig. 6.) and $U_A(x)$ and $U_C(x)$ are replaced by $U_A^6(x)$, $U_C^{73.5}(x)$ respectively.

Membership function of *VIA* with concentration is depicted in Fig. 6. One may note that the domain of concentrated membership function is between 0 and 120 (in months).

Category III: A plant whose *ASR* ≥100000 and *SF*≥100.

The membership function of *SF* dilates as $U_A^{0.5}(x)$ and $U_A(x)$ is replaced by $U_A^{0.5}(x)$.

Category IV: A plant does not belong to any of the above category.

Use the original fuzzy sets Eqs. (6)–(9).

4.3. Transforming numerical scores into risk levels

It is mentioned earlier that the output of the sub model need to be in the form of linguistic terms in order to convert the "Dispersal" factor into a qualitative/linguistic parameter. Therefore the numerical values of the output are to be transformed into linguistic terms as follows:

One may see that the final value decreases from one to zero when we consider the intersection of fuzzy sets in the sub model. Therefore, in order to have the compatibility with linguistic terms, the compliment of the final value of the sub model $U_{CH}^p(x)$ has been considered as the final output.

Note that the value of $U_{CH}^p(x)$ is in between zero and one. Now the aim is to define a transformation function which maps the numerical value to one of the linguistic labels in the set which has been defined in Fig. 4. The Numerical-Linguistic transformation function for a given numerical value can be explained as follows [9]:

Definition 2 *([9])*. Let $q \in [0, 1]$ be a numerical value and l_i be a label verifying that $\mathrm{h(q, l_i) = min\left\{h(q, l_t)\,|\forall l_t \in L\right\}}$, with

$$h(q, l_t) = \begin{cases} c & \text{if} \quad q \notin \text{Supp} \quad (l_t), \\ \sum_{j=1}^{c}\left(q - G_j(l_t)\right)^2 & \text{if} \quad q \in \text{Supp} \quad (l_t). \end{cases} \tag{13}$$

where *c* is the cardinality of the characteristic function set G_j and Supp () refer the support of a given membership function. The characteristic function set G_j consists of three functions which generate

Table 2
Validation results of sub model.

Category of species	Invasive species	Dispersal Risk Level (Sub Model)	Dispersal Risk Level (NRA)
Invasive	*Austroeupatorium inulifolium*	*Medium*	*Medium*
	Panicum maximum	*Low*	*Medium*
	Cuscuta campestris	*Low*	*Medium*
	Pueraria montana	*Low*	*Low*
	Acacia mearnsii	*High*	*High*
	Myrica faya	*Medium*	*Low*
Non Invasive	*Cassia fistula*	*Low*	*Low*
	Cissus rotundifolia	*Low*	*Low*
	Hedychium gardnerianum	*Unlikely*	*Very Low*
	Magnefera indica	*Very Low*	*Low*

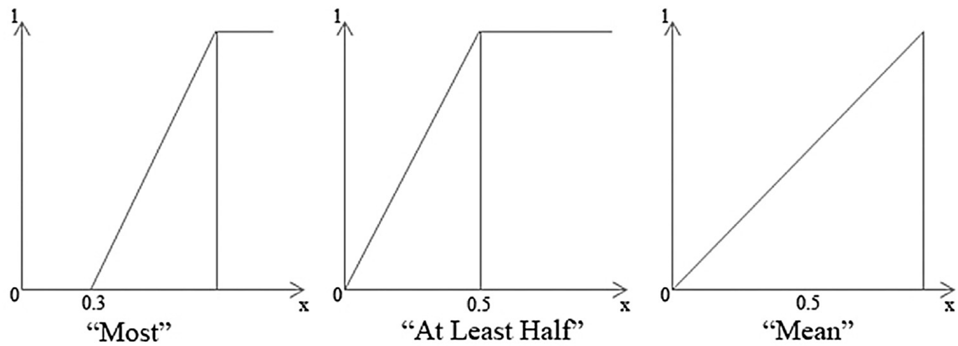

Fig. 7. Fuzzy linguistic quantifiers.

crisp values summarizing the information of a given fuzzy number [10].

In this task, three characteristic functions have been used by setting the cardinality, $c=3$ to evaluate the label l_i. The three characteristics functions G_js are mentioned below [9]:

1. $G_1(l_i)$- It is the method of Center of Gravity. This method summarizes the meaning of a label l_i into a numerical value as:

$$G_1(l_i) = \frac{\int_\nu \nu \mu_{y_{l_i}}(\nu)\,d\nu}{\int_\nu \mu_{y_{l_i}}(\nu)\,d\nu} \tag{14}$$

For triangular fuzzy number, the function $G_1(l_i)$ may be expressed as

$$G_1(l_i) = \begin{cases} x_1, & \text{if } x_1 = x_2 = x_3 \\ \dfrac{{x_3}^2 - {x_1}^2 + x_3x_2 - x_1x_2}{3(x_3 - x_1)} & \text{otherwise} \end{cases}. \tag{15}$$

2. $G_2(l_i)$- This is the method of value of a fuzzy number

$$G_2(l_i) = \int_0^1 s(r)\left(L_{y_{s_i}}(r) + R_{y_{s_i}}(r)\right) dr \tag{16}$$

where $L_y(r)$, $R_y(r)$ are the r-cut representations of y_{s_i} and $s(r)$ is a reducing function.

The simplified form may be expressed by using triangular fuzzy membership function and taking $s(r)=r$ as

$$G_2(l_i) = x_2 + [(x_3 - x_2) - (x_2 - x_1)]/6 \tag{17}$$

3.$G_3(l_i)$- which is the method of *Maximum value*

Let us consider the given label as l_i, with a membership function, $\mu_{y_{s_i}} = (\nu)$, $\nu \in V = [0, 1]$. The height is defined as

$\text{height}(l_i) = Sup\left\{\mu_{y_{s_i}}(\nu), \forall\nu\right\}$.

Using height(l_i), the function may be expressed as

$$G_3(l_i) = \max\left\{\nu | \mu_{y_{s_i}}(\nu) = \text{height}(l_i)\right\} \tag{18}$$

Using Definition 2, the dispersal risk values and NRA score related to dispersal related factors have been converted into linguistic labels.

For example, if the dispersal risk value is $q=0.62$, then the representative linguistic label is "*High*" where

$$min\left\{h(0.62, U), h(0.62, VL), h(0.62, L), h(0.62, M), h(0.62, H), h(0.62, VH), h(0.62, EH)\right\}$$

$$= \min\left\{1.0567, 0.6257, 0.2409, 0.0432, 0.0057, 0.1409, 0.376\right\}$$
$$= 0.0057$$

$$= h(0.62, H).$$

4.4. Test results and validation

Table 1 is a comparison of the dispersal risk level derived from this study with that of NRA for the 27 invasive plant species which have been used to develop the sub model. The model has been validated using data of known six invasive and four non invasive species in Sri Lanka. These results are presented in Table 2.

According to the biological explanations, depth of contribution of dispersal risk may be different from species to species among plants. Among the test species considered, there are species such as *Miconia calvescens*, *Clidemia hirta*, *Alstonia macrophylla*, *Mikania micrantha*, *Prosopis juliflora*, *Parthenium hysterophorus*, which depend highly on dispersal related risk factors. On the other hand, species such as *Annona glabra, Sphagneticola trilobata*, *Colubrina asiatica*, *Dillenia suffructicosa, Zizipus marutinum*, contribute relatively low on dispersal related risk factors. It can be seen that species *Parthenium hysterophorus* and *Sphagneticola trilobata*ba obtain same risk level as in NRA but their contribution toward the

Table 3
Test results of Model I.

Invasive species	Model I			NRA
	Mean	Most	At least half	
Alternanthera philoxeroides	*M*	*M*	*VH*	*H*
Clidemia hirta	*H*	*M*	*H*	*H*
Miconia calvescens	*VH*	*H*	*EH*	*H*
Alstonia macrophylla	*H*	*M*	*VH*	*M*
Annona glabra	*L*	*L*	*M*	*M*
Clusia rosea	*M*	*L*	*M*	*M*
Dillenia suffructicosa	*L*	*VL*	*H*	*M*
Ageratina riparia	*H*	*M*	*H*	*M*
Mimosa invisa	*H*	*M*	*H*	*H*
Myroxylon balsamum	*M*	*L*	*H*	*M*
Tithonia diversiflora	*M*	*L*	*H*	*M*
Mikania micrantha	*H*	*H*	*H*	*H*
Prosopis juliflora	*H*	*M*	*H*	*H*
Ulex europaeus	*M*	*L*	*M*	*M*
Mimosa pigra	*H*	*M*	*H*	*H*
Chromolaena odorata	*H*	*M*	*H*	*H*
Parthenium hysterophorus	*H*	*M*	*H*	*H*
Lantana camara	*M*	*M*	*M*	*M*
Imperata cylindrica	*H*	*H*	*VH*	*VH*
Opuntia stricta	*H*	*M*	*H*	*H*
Colubrina asiatica	*L*	*L*	*M*	*M*
Pennisetum polystachion	*H*	*M*	*H*	*M*
Sphagneticola trilobata	*M*	*L*	*H*	*M*
Zizphus mauritiana	*L*	*VL*	*M*	*L*
Eichhornia crassipes	*H*	*M*	*VH*	*H*
Pistia stratiotes	*H*	*M*	*VH*	*H*
Leucaena leucocephala	*M*	*L*	*H*	*M*

dispersal has been different. Our sub model gives clear discrimination on the risk level of those species.

The species mentioned in the validation results Table 2, such as *Cuscuta campestris*, *Pueraria montana*, and *Panicum maximum* give low contribution on dispersal and our model predict their risk level effectively, better than in NRA. The non-invasive species are the species that show low contribution toward the invasive attributes or invasiveness than the invasive species. It can be clearly seen that the predictions from the present model is good as our expected outcome for the non invasive species.

In the next section we present our main model which aggregates all the linguistic risk factors.

5. Model I

In the previous section, the main factor 'dispersal' has been converted to a linguistic variable. Therefore all the main risk factors including dispersal are linguistic variables. The model has been designed to aggregate the factors considering their contribution as equally important towards invasiveness.

5.1. Selecting aggregation operator

The Linguistic Ordered Weighted Averaging (LOWA) operator has been chosen among the aggregation operators of linguistic non-weighted information [8]. The LOWA operator is a symbolic operator with properties such as increasingly monotonous, commutative and "or-and."

Table 4
Validation results of Model I.

Category of species	Invasive species	Model I			NRA
		Mean	Most	At least half	
Invasive	*Austroeupatorium inulifolium*	*M*	*M*	*H*	*M*
	Panicum maximum	*M*	*L*	*VH*	*H*
	Cuscuta campestris	*M*	*M*	*VH*	*H*
	Pueraria montana	*M*	*M*	*VH*	*H*
	Acacia mearnsii	*H*	*M*	*H*	*M*
	Myrica faya	*M*	*L*	*H*	*M*
Non invasive	*Cassia fistula*	*M*	*L*	*H*	*M*
	Cissus rotundifolia	*L*	*L*	*M*	*L*
	Hedychium gardnerianum	*L*	*VL*	*M*	*L*
	Magnefera indica	*L*	*L*	*H*	*M*

Below we reproduce the Definition 3 as given in the reference [8].

Definition 3 *([8])*. Let $A = \{a_1, \ldots, a_m\}$ be the set of labels to be aggregated, then the LOWA operator, ϕ, is defined as

$$\begin{aligned}\phi(a_1, \ldots, a_m) &= W \cdot B^T = \xi^m \left\{w_k, b_k, k = 1, \ldots, m\right\} \\ &= w_1 \odot b_1 \oplus (1 - w_1) \odot \xi^{m-1} \left\{\beta_h, b_h, h = 2, \ldots, m\right\},\end{aligned}$$

where $W = [w_1, \ldots, w_m]$, is a weighting vector, such that, (i) $w_i \in [0, 1]$ and, (ii) $\sum_i w_i = 1$, $\beta_h = w_h / \sum_2^m w_k$, $h = 2, \ldots, m$, and $B = \left\{b_1, \ldots, b_m\right\}$ is a vector associated to A, such that,

$B = \sigma(A) = \left\{a_{\sigma(1)}, \ldots, a_{\sigma(n)}\right\}$ in which, $a_{\sigma(j)} \le a_{\sigma(i)} \forall i \le j$, with σ being a permutation over the set of labels A. ξ^m is the convex combination operator of m labels and if $m = 2$, then it is defined as

$$\xi^2 \left\{w_i, b_i, \mathrm{i} = 1, 2\right\} = w_1 \odot s_j \oplus (1 - w_1) \odot s_i = s_k, s_i, s_j \in S\,(j \ge i)$$

such that, $k = \min\left\{T, i + round\,(w_1 \cdot (j - i))\right\}$, where "round" is the usual round operation, and $b_1 = s_j, b_2 = s_i$.

As mentioned in [8], the weights W represents the concept of *fuzzy majority* [14] in the aggregation of LOWA operator using *fuzzy linguistic quantifier* [13]. Zadeh suggested that the semantic of a linguistic quantifier can be captured by using fuzzy subsets for its representation [12]. To calculate the weights by means of fuzzy linguistic quantifier, the method proposed by Yager [11] has been used as in the case of a non-decreasing proportional fuzzy linguistic quantifier Q which may be given as below:

$$w_i = Q\left(i/n\right) - Q\left((i-1)/n\right), i = 1, \ldots, n,$$

where the membership function of Q is:

$$Q\left(i/n\right) = \begin{cases} 0 & \text{if } r < a, \\ (r-a)/(b-a) & \text{if } a \le r \le b, \\ 1 & \text{if } r > b. \end{cases} \quad (19)$$

with a, b, r $\in$ [0, 1]. Some examples of quantifiers which have been used in this work are shown in Fig. 7, where the parameters, (a, b) are $(0.3, 0.8)$, $(0, 0.5)$ and $(0.5, 1)$, respectively.

Table 5
Important weights of Main Risk Factors/Sub Risk Factors.

Important Weight of Main Risk Factor	Main Risk Factor	Sub Risk factor	Weights of Sub Risk Factor
Very High	Dispersal (*DIS*)		
Very High	Vegetative Reproduction strength (*VRS*)		
Medium	Seed germination requirement level (*SGR*)		
High	Man's influence on spreading (*MIS*)	Natural & Man made disturbances(*NMD*)	*Medium*
		High potential to spread by human activities(*HA*)	*High*

The membership function of Q for "*Mean*" quantifier is different from other quantifiers as it is a pure averaging quantifier [11]

$$w_i = \frac{1}{n}, \text{ for all } i = 1, \ldots, n,$$

and then we have

$$Q_{mean}(K) = \frac{K}{n}, K = 1, \ldots, n \quad (20)$$

5.2. Setting up model I

First of all the linguistic information for the main risk factor "Man's influence on spreading" has been obtained by aggregating its two sub risk factors using LOWA operator. For this purpose "*Mean*" fuzzy quantifier guider is used. The invasion risk level of plant species due to the four main risk factors has been obtained by aggregating the linguistic information of each factor using LOWA operator. The risk levels of same set of invasive plant species used in Section 4.4 under different types of fuzzy linguistic quantifiers are obtained. The set of seven linguistic terms set depicted in Fig. 2. has been considered as the risk labels of the final outcome of the model.

For example, consider the weighting vector obtained for the aggregation using "*Most*" quantifier guider with the pair (0.3, 0.8) is,

w = [0, 0.4, 0.5, 0.1] where

$w_1 = Q(1/4) - Q(0/4) = Q(0.25) = 0,$

$$w_2 = Q(2/4) - Q(1/4) = Q(0.5) - Q(0.25) = \frac{0.5 - 0.3}{0.8 - 0.3} - 0 = 0.4,$$

$$w_3 = Q(3/4) - Q(2/4) = \frac{0.75 - 0.3}{0.8 - 0.3} - Q(0.5) = 0.9 - 0.4 = 0.5,$$

and $w_4 = Q(1) - Q(3/4) = 1 - Q(0.5) = 1 - 0.9 = 0.1$.

Let us consider the invasive species *Annona glabra* with performance values of (M, VL, M, M) for the main risk factors *DIS*, *VRS*, *SGR*, *MIS* respectively. The aggregate invasion risk of *Annona glabra* has been evaluated using the LOWA with above weighting vector as follows:

By preparing the performance values in descending order we have

(M, M, M, VL)

First consider the pair *M* and *VL*. Appling LOWA

$$k_2 = \min\left\{6, 1 + r\left(\frac{0.5}{0.6} \times 2\right)\right\} = M \qquad \text{(Here j = 3 and i = 1)}$$

For the pair *M* and *M*

$$k_3 = \min\left\{6, 3 + r\left(\frac{0.4}{1} \times 0\right)\right\} = M \qquad \text{(Here j = 3 and i = 3)}$$

For the last pair

$$k_4 = \min\left\{6, 3 + r(0 \times 0)\right\} = M \qquad \text{(Here j = 3 and i = 3)}$$

Then the aggregate risk of the *Annona glabra* under "*Most*" quantifier guider is the level- *Medium.*

Note that, the performance value of the main risk factor 'Man's influence on spreading' has been evaluated using LOWA operator with '*Mean*' quantifier guider prior to the above calculation.

Likewise the aggregate risks of species are obtained under different quantifier guiders. These results are compared with the NRA risk level. The NRA risk level has been obtained by converting the NRA score related to eight risk factors using Definition 2. These results are shown in Table 3. The model was also validated using known invasive and non invasive species in Sri Lanka and are presented in Table 4.

5.3. Discussion – results of Model I

According to Table 3, it can be seen that the risk levels of most of the invasive species obtained from the model are compatible with the NRA risk level if the quantifier guider is "*Mean*". In that case, species predicted risk levels of *Alternanthera philoxeroides*, *Annona glabra*, and *Dillenia suffructicosa*, are one level behind the NRA risk level. From results of Table 3 we can see that most suitable quantifier guider is "*Mean*" and therefore in the validation process, this aspect needs to be considered.

The species in Table 4 such as *Panicum maximum*, *Cuscuta campestris* and *Pueraria montana* show the similar behavior in "*Mean*" quantifier guider. One may note that the non invasive species *Cassia fistula* has taken the risk level "*Medium*" which is same as NRA. Practically, we expected a lower risk level to these species. On the other hand, for the species *Magnefera indica*, model gives a risk level of "*Low*" compared to risk level of "*Medium*" as in NRA. However, our model gives better predictions on non invasive species with "*Most*" quantifier guider but majority of species in Tables 3 and 4 comply with the "*Mean*" quantifier guider. There may be several facts which effect the incompatibility of the outcomes of two different methods (model and NRA). One reason was that our model attempted to mimic the uncertainty of the parameters while NRA did not follow such a mechanism. The other reason may be the fact that, risk factors have been considered as equally important to invasion risk. But in real situations these factors of plants have different important weights to the invasion risk.

In next section, we present Model II considering the parameters with different important weights.

6. Model II

Here we model the invasion risk by aggregating the linguistic information of the main risk factors and sub factors which are not equally important. As to Fig. 1, the main risk factors; dispersal, vegetative reproduction strength, seed germination requirement level and man's influence on spreading were the linguistic parameters of the model. The risk factors under the dispersal are not considered for assigning weights because their aggregation are incorporated through the sub model as mentioned in section 4.

6.1. Important weights of risk factors

The weights for the four main risk factors and two sub factors of the risk factor on 'man's influence on spreading' are given by a group of three plant science experts. For this evaluation the seven labels presented in Fig. 2 are considered as,

$S = \left\{s_0 = U, s_1 = VL, s_2 = L, s_3 = M, s_4 = H, s_5 = VH, s_6 = EH\right\}$.

Table 5 presents the weights obtained for each linguistic parameter.

6.2. Aggregation of weighted factors

In this section, we present the development of the model by using proper aggregation operator which is capable of dealing with weighted parameters. It may be noted that as in Table 5, the important weights of model parameters are linguistic values. When aggregating weighted information, one may need to consider two aggregations as follows [8]:

- the aggregation of important weights of model parameters
- the aggregation of weighted information

In literature we can find several aggregation operators that satisfy the above aspects [11,13]. Among them, *Linguistic Weighted*

Averaging operator (*LWA*) has been chosen and is reproduced here from [11].

Definition 4. [11]: the aggregation of the set of weighted individual opinions,

$\{(w_1, a_1), \ldots, (w_n, a_n)\}$, according to the *Linguistic Weighted Averaging* operator (*LWA*) operator is defined as

$$(w_E, a_E) = LWA\left[(w_1, a_1), \ldots, (w_n, a_n)\right],$$

where the important weights of the group opinion, w_E, is obtained as

$$w_E = \phi_Q(w_1, \ldots, w_n).$$

and, the opinion of the group, a_E, is obtained as

$$a_E = f\left[g(w_1, a_1), \ldots, g(w_n, a_n)\right],$$

where $f = \phi_Q$ and $g \in \{LC_1^{\rightarrow}, LC_2^{\rightarrow}, LC_3^{\rightarrow}\}$.

Here $LC^{\rightarrow}$ is a linguistic conjunction function which is a monotonically non decreasing in the weights and satisfy the properties required for any transformation function *g*.

In this work *classical MIN* operator has been selected as the most appropriate transformation function. The *classical MIN* operator can be stated as follows [11]:

$$LC_1^{\rightarrow}(w, a) = \mathrm{MIN}(w, a)$$

6.3. Setting up Model II

First of all, the linguistic information for the main risk factor "Man's influence on spreading" has been obtained by aggregating two weighted sub risk factors related to this factor using *LWA* operator as defined in the previous sub section 6.2. Here "*Mean*" fuzzy quantifier guider is considered. The invasive risk level of plant species due to the four weighted main risk factors is obtained by aggregating the linguistic information of each attribute using *LWA* operator. The risk levels of same set of invasive plant species which has been used in Model I are obtained under different types of fuzzy linguistic quantifiers.

For example, consider the weighting vector obtained for the aggregation using "*At least half*" quantifier guider with the pair(0, 0.5),

$w = [0.5, 0.5, 0, 0]$, where

$$w_1 = Q(1/4) - Q(0/4) = Q(0.25) = 0.5, \quad \text{since } 0 \le 0.25 \le 0.5.,$$

$$w_2 = Q(2/4) - Q(1/4) = Q(0.5) - Q(0.25) = 1 - 0.5 = 0.5,$$

$$w_3 = w_4 = 0 \qquad .$$

Let us consider the invasive species *Clidemia hirta* with performance values of (H, VL, M, M) for the attributes *DIS*, *VRS*, *SGR*, and *MIS* with their important weights (VH, VH, M, H) as given in Table 5. The performance values of the main risk factors *DIS*, *VRS*, *SGR* and *MIS* with respect to the important weight may be represented as follows:

$$((VH, H), (VH, VL), (M, M), (H, M)).$$

The aggregate invasion risk of *Clidemia hirta* is therefore evaluated using the LWA with above weighting vector as follows:

$$\phi \quad (\min(VH, H), \min(VH, VL), \min(M, M), \min(H, M))$$

(Here the classical min operator has been used for the transformation function in the LWA operator.)

$$= \phi(H, VL, M, M)$$

By preparing the performance values (after the weighting) in descending order we have

$$(H, M, M, VL)$$

First consider the pair (M, VL).

$$k_2 = \min\left\{6, 1 + r\left(\frac{0}{1} \times 2\right)\right\} = VL \qquad (\text{Here j} = 3 \text{ and i} = 1)$$

For the second pair (M, VL)

$$k_3 = \min\left\{6, 1 + r\left(\frac{0.5}{0.5} \times 2\right)\right\} = M \qquad (\text{Here j} = 3 \text{ and i} = 1)$$

For the last pair (H, M)

$$k_4 = \min\left\{6, 3 + r\left(\frac{0.5}{1} \times 1\right)\right\} = H \qquad (\text{Here j} = 4 \text{ and i} = 3)$$

Then the aggregate risk of the *Clidemia hirta* under "*At least half*" quantifier guider is *High.*

We may note that the performance value of the main risk factor 'Man's influence on spreading' has been evaluated using LWA operator with '*Mean*' quantifier guider prior to the above calculation. For example, consider the same species *Clidemia hirta* with performance values of (H, M) for the risk factors *HA, NMD* with their important weights (H, M).

The aggregate risk of *MIS* on *Clidemia hirta* is therefore evaluated using the LWA with above weighting vector as follows:

$$\phi \quad (\min(H, H), \min(M, M))$$

(Here classical Min operator has been used for the transformation function in the LWA operator.)

$$= \phi(H, M)$$

By preparing the performance values (after the weighting) in descending order we have

$$(H, M)$$

For the pair (H, M),

$k_2 = \min\left\{6, 3 + r\left(\frac{0.5}{1} \times 1\right)\right\} = H$ (Here $j = 4$, $i = 3$ and $w_1 = w_2 = 0.5$)

Likewise the aggregate risks of species are obtained under different quantifier guiders.

After analyzing the results, the model is divided into three cases as below:

Case I. If $g(DIS) \ge Neg(g(VRP))$ then aggregate the weighted parameters using LWA with "*Mean*" quantifier guider otherwise use "*At least half*" as the quantifier guider. Note that in either case at least one linguistic value of g(DIS) / g(VRP)should be greater than or equal to the label "*Medium*".

Case II. If g(DIS) = High & (g(VRP) = Medium or Low) & g(SG) = Medium then aggregate the weighted parameters using LWA with "*At least half*" quantifier guider.

Case III. If the linguistic value g(DIS) & g(VRP) less than or equal to the label "*Low*" then aggregate the weighted parameters using LWA with "*Most*" quantifier guider.

Considering these cases, the test results and validation results are presented in Tables 6 and 7 respectively.

In the process of developing this model, we have tried to maintain the risk levels of species used as test data in order to be compatible with NRA risk levels. The reason was that the NRA risk levels for plant species have been obtained through consultation with experienced plant scientists.

Table 6
Test results of Model II.

Invasive species	Model II	NRA
Alternanthera philoxeroides	*High*	*High*
Clidemia hirta	*High*	*High*
Miconia calvescens	*High*	*High*
Alstonia macrophylla	*Medium*	*Medium*
Annona glabra	*Medium*	*Medium*
Clusia rosea	*Medium*	*Medium*
Dillenia suffructicosa	*Medium*	*Medium*
Ageratina riparia	*Medium*	*Medium*
Mimosa invisa	*High*	*High*
Myroxylon balsamum	*Medium*	*Medium*
Tithonia diversiflora	*Low*	*Medium*
Mikania micrantha	*High*	*High*
Prosopis juliflora	*High*	*High*
Ulex europaeus	*Medium*	*Medium*
Mimosa pigra	*High*	*High*
Chromolaena odorata	*High*	*High*
Parthenium hysterophorus	*High*	*High*
Lantana camara	*Medium*	*Medium*
Imperata cylindrica	*High*	*Very High*
Opuntia stricta	*High*	*High*
Colubrina asiatica	*Medium*	*Medium*
Pennisetum polystachion	*Medium*	*Medium*
Sphagneticola trilobata	*Medium*	*Medium*
Zizphus mauritiana	*Low*	*Low*
Eichhornia crassipes	*High*	*High*
Pistia stratiotes	*High*	*High*
Leucaena leucocephala	*Medium*	*Medium*

Table 7
Validation results of Model II.

Category of species	Invasive species	Model II	NRA
Invasive	*Austroeupatorium inulifolium*	*Medium*	*Medium*
	Panicum maximum	*High*	*High*
	Cuscuta campestris	*High*	*High*
	Pueraria montana	*High*	*High*
	Acacia mearnsii	*Medium*	*Medium*
	Myrica faya	*Medium*	*Medium*
Non invasive	*Cassia fistula*	*Low*	*Medium*
	Cissus rotundifolia	*Low*	*Low*
	Hedychium gardnerianum	*Low*	*Low*
	Magnefera indica	*Low*	*Medium*

6.4. Discussion – results of Model II

According to Table 6, it may clearly be seen that, the risk level of most of the plant species obtained from the model were compatible with NRA risk level except for the species *Tithonia diversiflora* and *Imperata cylindrica*.

One may also see that the discrimination occurred in the Model I have been solved in Model II. For example, species *Alternanthera philoxeroides*, and *Dillenia suffructicosa*, predicted risk levels in Model I are different to that of NRA risk level under each linguistic quantifier. But in Model II these levels have become the same as in NRA. Similarly, for the species *Panicum maximum*, *Cuscuta campestris* and *Pueraria montana* in Table 5, the risk levels of these species can be better compared to Model I as given in Table 7. Moreover, Model II maintains the risk level of non invasive species at "*Low*". This fact gives a clear idea of the efficiency of Model II compared to the risk assessment methods used for assessing risk of invasive plant species.

7. Setting up Model III

As mentioned in literature [17,18], the IOWA operator was developed to overcome the problems of the majority guided OWA operator. Since the Model II accompanied with LWA operator, the weighted majority guided linguistic IOWA operator has been used to model the invasion risk. This is an attempt to see whether the Model II works well even with the shortcomings of LWA operator.

Table 8
Order inducing values of important weights of Main Risk Factors.

Main Risk Factor	Important weights (I_i)	$\sup_i$	u_i
DIS	*Very High*	3	4
VRS	*Very High*	3	4
SGR	*Medium*	2	2.5
MIS	*High*	4	4

Below we reproduce the definition of the weighted majority guided linguistic IOWA operator as in [17].

Definition 5. [17]: A weighted MLIOWA operator of dimension n is a function $\Phi_Q^I : (S \times S)^n \to S$,

defined according to the expression

$$\Phi_Q^I((I_1, p_1), \ldots, (I_n, p_n)) = \Phi_Q((u_1, p_1), \ldots, (u_n, p_n))$$

in which

1. The order inducing values are obtained from the linguistic importance degrees associated with the values to be aggregated as

$$u_i = \frac{\sup_i + ind(I_i)}{2} \quad \text{with}$$

$$\sup_i = \sum_{j=1}^{n} \sup_{ij} | \sup_{ij} = \begin{cases} 1 & \text{if } |ind(I_i) - ind(I_j)| < \alpha \\ 0 & \text{otherwise} \end{cases}$$

where I_i the linguistic importance degree of the value p_i to be aggregated.

2. The weighting vector is therefore obtained as

$$w_i = Q\left(\frac{u_{\sigma(i)}}{n}\right) / \sum_{j=1}^{n} Q\left(\frac{u_{\sigma(j)}}{n}\right).$$

Considering the important weights associated with main risk factors/sub risk factors given in Table 5 and the linguistic quantifiers "*Most*" and "*At least half*" defined by the parameters (0.3,0.8), (0,0.5) respectively, the risk levels of same set of invasive species as in Table 6 are obtained using weighted MLIOWA operator with $\alpha = 1$. The results of order inducing values for main risk factors are shown in Table 8.

The weighting vector obtained for the aggregation using most quantifier guider is

$w = \left(\frac{0.65}{3.65}, \frac{1}{3.65}, \frac{1}{3.65}, \frac{1}{3.65}\right)$, where for example,

$w_1 = Q\left(\frac{2.5}{4}\right) / \sum_{j=1}^{n} Q\left(\frac{u_{\sigma(j)}}{n}\right) = \frac{0.65}{3.65}$ and

$$w_3 = Q\left(\frac{4}{4}\right) / \sum_{j=1}^{n} Q\left(\frac{u_{\sigma(j)}}{n}\right) = \frac{1}{3.65}.$$

Let us consider the invasive species *Clidemia hirta* with performance values of (H, VL, M, H) for the main risk factors *DIS*, *VRS*, *SGR*, *MIS* respectively. The aggregate invasion risk of *Clidemia hirta* has been evaluated using the weighted MLIOWA with above weighting vector as follows:

$$\Phi_{Q_2}^I((VH, H), (VH, VL), (M, M), (H, H))$$

$$= round\left(3 * \frac{0.65}{3.65} + 1 * \frac{1}{3.65} + 4 * \frac{1}{3.65} + 4 * \frac{1}{3.65}\right)$$
$$= round(3) = 3 = M = Medium.$$

Table 9
Test results of Model III.

Invasive species	Model III		NRA
	Most	At Least Half	
Alternanthera philoxeroides	*High*	*High*	*High*
Clidemia hirta	*Medium*	*Medium*	*High*
Miconia calvescens	*High*	*High*	*High*
Alstonia macrophylla	*Medium*	*Medium*	*Medium*
Annona glabra	*Low*	*Medium*	*Medium*
Clusia rosea	*Low*	*Medium*	*Medium*
Dillenia suffructicosa	*Medium*	*Medium*	*Medium*
Ageratina riparia	*High*	*High*	*Medium*
Mimosa invisa	*Medium*	*Medium*	*High*
Myroxylon balsamum	*Medium*	*Medium*	*Medium*
Tithonia diversiflora	*Medium*	*Medium*	*Medium*
Mikania micrantha	*High*	*High*	*High*
Prosopis juliflora	*Medium*	*High*	*High*
Ulex europaeus	*Low*	*Medium*	*Medium*
Mimosa pigra	*Medium*	*Medium*	*High*
Chromolaena odorata	*Medium*	*High*	*High*
Parthenium hysterophorus	*High*	*High*	*High*
Lantana camara	*Medium*	*Medium*	*Medium*
Imperata cylindrica	*High*	*High*	*Very High*
Opuntia stricta	*Medium*	*Medium*	*High*
Colubrina asiatica	*Medium*	*Medium*	*Medium*
Pennisetum polystachion	*High*	*High*	*Medium*
Sphagneticola trilobata	*Medium*	*Medium*	*Medium*
Zizphus mauritiana	*Low*	*Low*	*Low*
Eichhornia crassipes	*High*	*High*	*High*
Pistia stratiotes	*High*	*High*	*High*
Leucaena leucocephala	*Medium*	*Medium*	*Medium*

Table 10
Validation results of Model III.

Category of species	Invasive species	Model III		NRA
		Most	At Least Half	
Invasive	*Austroeupatorium inulifolium*	*Medium*	*Medium*	*Medium*
	Panicum maximum	*Medium*	*Medium*	*High*
	Cuscuta campestris	*Medium*	*High*	*High*
	Pueraria montana	*High*	*High*	*High*
	Acacia mearnsii	*High*	*Medium*	*Medium*
	Myrica faya	*Medium*	*Medium*	*Medium*
Non invasive	*Cassia fistula*	*Medium*	*Medium*	*Medium*
	Cissus rotundifolia	*Medium*	*Low*	*Low*
	Hedychium gardnerianum	*Low*	*Low*	*Low*
	Magnefera indica	*Low*	*Medium*	*Medium*

Likewise the aggregate risk for the same set of invasive alien plant species (Table 6) are obtained using weighted MLIOWA for two different linguistic quantifiers such as "*Most*" and "*At least half*" separately. Here we considered only two distinct linguistic quantifier guiders since the weighting vector obtained with "*Mean*" quantifier guider is same as the weighting vector obtained from "*At least half*" guider. The results are shown in Table 9. Furthermore, the model has been validated using same set of species in Table 7 and the results are displayed in Table 10.

Note that the order inducing values for the important weights of sub factors (*HA* and *NMD*) of main factor *MIS* have been calculated using "*Most*" linguistic quantifier and evaluated the performance values of each species on *MIS* using the weighted MLIOWA operator prior to the main aggregation.

7.1. Discussion – results of Model III

According to Table 9, it can be clearly seen that the approximate risk levels of invasive species from Model III with at least half linguistic quantifier guider are same as the NRA risk levels for the species *Clidemia hirta*, *Ageratina riparia*, *Mimosa invisa*, *Mimosa pigra*, *Imperata cylindrica*, *Opuntia stricta*, and *Pennisetum polystachion*.

Unlike in the Model II, there was no specific pattern to divide the model logically into cases considering the performance values of attributes and linguistic quantifiers, in order to present the results of the model more informative. One may conclude that *at least half* linguistic quantifier guider was the most appropriate one for the Model III. The validation results in Table 10 justify the above conclusion. But this model does not provide significant results for non-invasive species. In reality we expect low risk levels for this species category and the NRA tool also has not been able to tackle this matter. For example, non-invasive species such as *Cassia fistula* and *Magnefera indica* show *Medium* risk in Model III (at least half quantifier) as well as in NRA. One may see that the validation of Model II clearly distinguish the risk of invasive and non-invasive species than that of Model I and III.

8. Advantages and drawbacks

In this study, three models have been developed to evaluate the invasion risk of plant species. Among them, we have clearly justified that the Model II is a better tracking system to identify the potential invaders by the conventional risk assessment methods. In conventional risk assessment method, a group of expertise is always required to make judgments and it is often manually conducted. The main advantage of this model is that the single user can make prediction on a particular species if data are available for the parameters.

The main drawback of the method is that it strongly depends on the availability of precise data which could be derived through research and most of the data for the model parameters are approximations or linguistic recommendations. The other reason was the reluctance of some experts to integrate mathematical concepts into their judgments even though the linguistic recommendations may provide a precise way of data conversion.

9. Conclusion

In this paper, three models with linguistic inputs and outputs have been constructed to assess the risk of IAS. We have used fuzzy linguistic aggregation operators to build up the models. Here the risk factors have been considered as equally important to invasiveness in Model I and unequally important in Model II. The experts' group decision has been taken to weight the risk factors in Model II. The Model III has been developed using more sophisticated linguistic aggregation operator than that in Model II to observe the well-fit model.

The proposed models give better prediction of risks of invasive alien plant species when the invasion is dominated by invasive attributes. It is also worth mentioning that Model II has produced significant improved results in comparison to Model I and Model III. Also it gives better prediction than the manually employed risk assessment method. The models need to be further modified by incorporating more parameters other than invasive attributes such as ecology, establishment, and management aspects to evaluate overall invasion risk of invaders. But the limited amount of available data on those factors set serious constraints to the evaluation of overall risk of IAS.

On the other hand, the methods that have been used in these models are based on type-1 fuzzy sets. In future we hope to extend this study using type-2 fuzzy sets since it is a sophisticated tool to capture the uncertainty in linguistic recommendations than type-1 fuzzy sets.

Acknowledgment

Authors would like to acknowledge the University of Colombo Research Grant (AP/3/2012/CG/26) for providing necessary financial support for visiting National Institute of Technology Rourkela, Odisha, India to undertake this collaborative work.

Appendix A.

INVASIVE ALIEN SPECIES

Owing to the increasing travel, trade, and tourism associated with globalization and expansion of the human population have facilitated intentional and unintentional movement of species beyond natural biogeographically barriers. As such, many of these species have become invasive and continue to invade new regions at an unprecedented rate, exerting strong impacts on biodiversity and human welfare [23].

Populations of Invasive alien species can self-sustain and spread over large areas through production of offspring at considerable distances from the parent plants, as they have efficient reproductive and dispersal mechanisms [1].

IAS can change the community structure and species composition of native ecosystems directly by out-competing indigenous species for resources. It may also have important indirect effects through changes in nutrient cycling, ecosystem function and ecological relationships between native species. Such changes alter the flow, availability and quality of nutrient resources in biochemical cycles, which in turn modify food webs [24]. In aquatic ecosystems, dense mats of free floating invaders such as water hyacinth are often quoted as resulting in the deterioration of aquatic biodiversity and changes to water chemistry and oxygen levels. Furthermore aggregate effects of multiple invasive species can have large and complex impacts in an ecosystem such as altering the evolutionary pathway of native species by competitive exclusion, niche displacement, hybridization predation, and ultimately extinction [1].

IAS can be invaded in almost every ecosystem type on Earth, even in agricultural areas. In terms of economic, the costs of invasive alien species are significant. Total annual costs, including losses to crops, pastures and forests, as well as environmental damages and control costs, have been conservatively estimated to be a large amount [1].

In this regard it is very important to understand how species become invasive. Predicting invasive behavior is a very difficult task. Nevertheless, one way is to estimate their invasive potential of inherited characteristics such as rapid growth and maturity; prolific seed production; highly successful seed dispersal, germination and colonization; aggressive vegetative spread; ability to out-compete native species; high cost to remove or control etc [26]. However different types of characteristics may dominate under different climatic conditions. Hence lists of invasive alien species from one country may be different from another.

IAS are particularly important to Sri Lanka as the country is a geographically separated island with greater diversity of habitats, high levels of species endemism and for the highly threatened status of some endemic species and their habitats. On the other hand, open economic policies facilitating international trade, travel and transportation encourage many species introductions [25] which might be a threat if they will not be introduced without checking of their potential invasive status. Government of Sri Lanka has been highly concerned on managing the already established alien invaders and controlling the introduction of potential invasive species to the country over the last two decades. All biological material brought into the country is released only when the specific quarantine period is over for diseases and pest checks [27]. However, steps are being taken to identify potential invaders in addition to the quarantine screening procedure conducted at the time of introduction of species introductions in line with the international standards.

Appendix B.

Here we provide some important facts on twenty seven invasive species which have been used to develop the three models (Tables 11–13) [21,22].

Table 11
Some important facts on invasive species used as test data (a).

Species	Distribution in Sri Lanka	Habit	Means of spread	Annual seed rain per m^2	Method of dispersal	Impact on ecosystems
Ageratina riparia	montane zone	herb	by seeds	100000	wind, water, animals, vehicles and machinery	forms dense stands, produce toxins and affect establishment of other plants
Alternanthera philoxeroides.	up- low-country wet zones	herb	propagate through stem, root buds, seeds	2000	water	*form dense mats with a massive underground root system.* Competes with pasture species.
Annona glabra	lowland wet zone marshlands	tree	by suckers and seeds	1000	water, animals (e.g. large birds and feral pigs)	form thickets and provide excessive shade. Reduce capacity of water storage in marshlands.
Chromolaena odorata	island-wide	herb	by seeds and when contamination of stem with soil	260,000	water, animals, garden wastes, machinery	forms dense thickets, exude chemicals toxic to other plants.
Clidemia hirta	lowland wet zone	shrub	by seeds	1000000	animals	exude toxins that inhibit growth of others species, forms dense stands.
Clusia rosea	sub-montane zone	tree	by seeds, suckers	20000	birds	form dense stands and shade out.
Colubrina asiatiica	island-wide	spreading shrub	by seeds and rooting branches and resprouting stems	5000	birds	produce a thick shade inhibiting germination of other seeds.
Dillenia suffructicosa	low-country wet zone	shrub	by seeds, stem cutting	1000	birds	form dense thickets, create shade and suppress growth and germination of other plants.
Eichhornia crassipes	island-wide	free floating herb	by seeds, offshoots	100	water	form dense floating mats on water inhibiting penetration of light.
Imperata cylindrica	intermediate and dry zones	herb	by seeds and rootstock	200000	wind	exude toxic substances that inhibit the growth of other plant species, forms a ground cover.

Table 12
Some important facts on invasive species used as test data (b).

Species	Distribution in Sri Lanka	Habit	Means of spread	Annual seed rain per m^2	Method of dispersal	Impact on ecosystems
Leucaena leucocephala	wet, dry and intermediate zones	tree	by seeds and vegetative organs	10000	seeds, small animals (rodents and birds) that consume fruits	forms dense thickets and shade out.
Miconia calvescens.	sub-montane zone	small tree	by seeds, layering, resprouting	15000	birds, wind, water, vehicles, and other animals (i.e. small mammals)	create deep shade and suppress growth of other plants.
Mikania micrantha	Island-wide	*climber*	by seeds and rooting at nodes	10000	wind, animals, flood or machinery	smother native vegetation and suppress growth, produces toxins that inhibit plant growth.
Mimosa invisa	lowland wet and dry zones	spreading spiny shrub	by seeds	12000	water, machinery	climb over and smother other plants shading them.
Mimosa pigra	dry and intermediate zones	spiny shrub	by seeds, suckers	100000	animals, water, clothing	form dense stands, interfere with movements of animals.
Myroxylon balsamum	wet and intermediate zones	tall tree	by seeds	20000	wind	produce dense stands of young trees and shade out.
Opuntia stricta	arid zone	succulent spiny shrub	by seeds, succulent	1000	birds	makes the place inaccessible due to spines and interfere with animal movements.
Parthenium hysterophorus	dry and intermediate zones	herb	by seeds, root buds	100000	wind, water, animals, machinery	forms dense stands, produce toxins and affect establishment of other plants.

Table 13
Some important facts on invasive species used as test data (c).

Species	Distribution in Sri Lanka	Habit	Means of spread	Annual seed rain per m^2	Method of dispersal	Impact on ecosystems
Pennisetum polystachyon	dry and intermediate zones	herb	by seeds, creeping stems that touch soil, rootstock	2000	wind, water, animals	increase the risk of damaging fires, forms a thick ground cover.
Pistia stratiotes.	island-wide	free floating herb	by seeds, offshoots	100	water	form dense floating mats on water inhibiting penetration of light.
Prosopis juliflora	arid zone	spreading tree	by seeds	100000	water, wild animals such as birds, bats, monkeys	form dense stands, exude toxins and suppress growth of other plant species.
Sphagneticola trilobata	wet and intermediate zones	herb	by seeds, and stems	10000	garden waste	form dense ground cover and inhibit germination of native seeds.
Tithonia diversiflora	wet, sub montane and intermediate zones	shrub	by seeds	12000	wind	form dense cover sheltering neighbor plants.
Ulex europaeus	montane zone	spiny shrub	by seeds	20000	ants	forms impenetrable thickets, interfere with movement of animals.
Z Zizphus mauritiana	dry zone	spreading thorny shrub or small tree	by seeds, root suckers	500	water, animals	forms dense thickets, interfere with movement of animals.

References

[1] Convention on Biological Diversity, Alien Species That Threaten Ecosystems, Habitats and Species, Article 8[h], Secretariat of the Convention on Biological Diversity, United Nations, 2008.
[2] M. Rejmanek, M.D. Richardson, What attributes make some plants species more invasive, J. Ecol 77 (1996) 1655–1661.
[3] E. Weber, G. Daniel, Assessing the risk of potentially invasive plant species in Central Europe, J. Nat. Conserv. 12 (2004) 177–179.
[4] J. Kelly, C. O'Flynn, C. Maguire, Risk Analysis and Prioritization for Invasive and Non- Native Species in Ireland and Northern Ireland, Northern Ireland Environment Agency, 2013.
[5] S.M.W. Ranwala, Risk assessment for invasive alien species, in: B. Marambe, P. Silva, S. Wijesundera, N. Attapattu (Eds.), Invasive Alien Species-Strengthening Capacity to Control Introduction and Spread in Sri Lanka Biodiversity Secretariat, Ministry of Environment and Natural Resources, Sri Lanka, 2010.
[6] J. Lemaire, Fuzzy insurance, ASTIN 20 (1990) 33–56.
[7] H.J. Zimmermann, Fuzzy Set Theory and Its Applications, fourth ed., Springer (India) Private Limited, New Delhi, 2001.
[8] F. Herrera, E. Herrera-Viedma, Linguistic decision analysis: steps for solving decision problems under linguistic information, J. Fuzzy Set Syst. 115 (2000) 67–82.
[9] M. Delgado, F. Herrera, E. Herrera-Viedma, L. Martinez, Combining numerical and linguistic information in group decision making, J. Inform. Sci. 107 (1998) 177–194.
[10] M. Delgado, A.M. Vila, W. Voxman, On a canonical representation of fuzzy numbers, J. Fuzzy Set Syst. 93 (1998) 125–135.
[11] F. Herrera, E. Herrera-Viedma, Aggregation operators for linguistic weighted information, IEEE T Syst. Man Cyb. 27 (1997) 646–656.
[12] L.A. Zadeh, The concept of a linguistic variable and its applications to approximate reasoning, J. Inform. Sci. 8 (1975) 199–249.
[13] R.R. Yager, On ordered weighted averaging aggregation operators in multicriteria decision making, IEEE T Syst. Man Cyb. 18 (1980) 183–190.
[14] P. Gabriella, R.R. Yager, Modeling the concept of majority opinion in group decision making, J. Inform. Sci. 176 (2006) 390–414.
[15] J. Zhao, K.B. Bose, Evaluation of Membership Functions for Fuzzy Logic controlled induction motor drive, IEEE 2002 28th annual Conference of the Industrial Electronics Society 1 (2002) 229–234.
[16] M. Detyniecki, Fundamentals on Aggregation Operators, Berkeley, 2001.
[17] E. Herrera-Viedma, G. Pasi, A.G. Lopez-Herrera, C. Porcel, Evaluating the information quality of web sites: a methodology based on fuzzy computing with words, J. Am. Soc. Inf. Sci. Technol. 57 (2006) 538–549.
[18] R.R. Yager, The power average operator, IEEE T Syst. Man Cya. 31 (2001) 724–731.
[19] V.N. Huynh, T.B. Ho, Y. Nakamori, A parametric representation of linguistic hedges in Zadeh's fuzzy logic, Int. J. Approx Reason. 30 (2002) 203–223.
[20] L.A. Zadeh, A fuzzy-Set-Theoretic interpretation of linguistic hedges, J. Cybern. 2 (1972) 4–34.
[21] IUCN, SSC, ISSG. 2005. Global Invasive Species Database.Available from: http://www.issg.org/database.
[22] CABI. Invasive Species Compendium. Online at http://www.cabi.org/isc/.
[23] M. Chenje, J. Mohamed-Katerere, Invasive alien species, africa environment outlook 2, United Nations Environment Program. 33 (2006) 331–347.
[24] P.M. Vitousek, Biological invasions and ecosystem processes: towards an integration of population biology and ecosystem studies, Oikos 57 (1990) 7–13.
[25] N.D.R. Weerawardane, J. Dissanayake, Status of forest invasive species in Sri Lanka. In: P. McKenzie, C., brown, S. jianghua and W. jian, (Eds.), the unwelcome guests, Proceedings of the Asia Pacific Forest Invasive Species Conference. (2005) 114–120.

[26] Natural Heritage, Virginia Department of Conservation & Recreation. Available from: http://www.dcr.virginia.gov.
[27] M.C.M. Iqbal, D.S.A. Wijesundera, S.M.W. Ranwala, Climate change, invasive alien flora and concerns for their management in Sri Lanka, Ceylon J. Sci. (Bio. Sci) 43 (2) (2014) 1–15.
[28] Y. Zhang, Z. Xu, H. Wang, H. Liao, Consistency-based risk assessment with probabilistic linguistic preference relation, Appl. Soft Comput. 49 (2016) 817–833.
[29] Y. Dong, E. Herrera-Viedma, Consistency-driven automatic methodology to set interval numerical scales of 2-tuple linguistic term sets and its use in the linguistic GDM with preference relation, IEEE Trans. Cybern. 45 (4) (2015) 780–792.
[30] J. A. Morente-Molinera, J., Mezei, C., Carlsson, E. Herrera-Viedma, Improving supervised learning classification methods using multi-granular linguistic modelling and fuzzy entropy. IEEE Trans. Fuzzy Syst, in press 10.1109/TFUZZ.2016.2594275.
[31] Z. Tao, H. Chen, X. Song, L. Zhou, J. Liu, Uncertain linguistic fuzzy soft sets and their applications in group decision making, Appl. Soft Comput. 34 (2015) 587–605.

H.O.W. Peiris holds a B.Sc. (Honours) Degree in Mathematics special with First class from the University of Sri Jayewardenepura, Sri Lanka. Currently, she is reading for a Doctorate in Mathematics in Field of Mathematical Modeling in Research & Development Centre for Mathematical Modeling at University of Colombo and holds six publications. She started her academic career as an instructor at Department of Mathematics, University of Sri Jayewardenepura. Currently, she is working as Lecturer (Probationary) at Department of Mathematics and Computer Science, The Open University of Sri Lanka. Her research area includes Mathematical modeling, Fuzzy set theory.

Dr. S. Chakraverty, a Mathematics professor in National Institute of Technology Rourkela, India, Ph. D. from IIT Roorkee and post-doctoral research from ISVR, University of Southampton, U.K. and Concordia University, Canada. He was visiting professor at Concordia and McGill universities, Canada and University of Johannesburg, South Africa. He published ten books, 259 research papers, reviewer of many international journals, recipient of CSIR Young Scientist, BOYSCAST, UCOST, Golden Jubilee CBRI, INSA International Bilateral Exchange, ISCA Platinum Jubilee Lecture and Roorkee University gold medal awards. His research area includes mathematical Modeling, Soft computing, Machine Intelligence, Vibration, differential equation and Numerical analysis.

Dr. S.S.N. Perera holds a B.Sc. (Honours) Degree in Mathematics special with a First Class from the University of Colombo, a Postgraduate Diploma in Modeling in Complex Realities from ICTP, Trieste, Italy, a Master Degree from ICTP and SISSA, Italy, a PhD from Kaiserslautern University, Germany. Currently, he holds the position of Senior Lecturer, at the Department of Mathematics, University of Colombo. His expertise areas are Financial Modeling, Actuarial Mathematics and Mathematical Modeling in Biology. He established, Research & Development Centre for Mathematical Modeling and Mathematical Modeling Research Group at University of Colombo. He was a junior associate at ICTP, Italy. He was awarded NRC merit award for scientific publication and SLASS Postgraduate research merit award.

Dr. Sudheera Ranwala B. Sc, (Colombo), PhD (Aberdeen) is a Senior Lecturer in Ecology and Biodiversity Conservation at Department of Plant Sciences, University of Colombo. She is a member of the national Invasive Species Specialist Group of Sri Lanka and National Environment Council. Risk assessments for Invasive alien species has been one of her key research areas which has significantly contributed to prioritize IAS in Sri Lanka and establish a national warning system to prevent entry of invasive alien species. Apart from research, Sudheera has authored few books including the first local book publication on invasive alien species of Sri Lanka.