17

Predictive Distributional Modeling of *Toricellia Tiliifolia* DC.— A Poorly Known Taxa in the Indo-Himalayan Region

Arun Chettri and S. K. Barik

Abstract

Toricellia tiliifolia DC., is the only species of the genus *Toricellia* endemic to Eastern Asiatic region. In India, it consists of only one species and its distribution is restricted to the hilly terrains of Sikkim, Darjeeling and Arunachal Pradesh, habitually with thin population. In China the genus is represented by two species. An ecological niche modeling approach using a genetic algorithm was chosen to predict the potential range of distribution of the species in the Indo-Himalayan region. The niche model shows that the distributional area of *T. tiliifolia* is concentrated in the Eastern Asiatic region. The distributional model agreements of the species were negatively related to digital elevation model, while isothermality and the precipitation of the driest quarter had a positive effect on the niche model agreements of the species. The predictions of the potential distribution of the species should help in developing strategies for proper censuring the species population.

Key words: Ecological niche modeling, niche visualization, potential distributional areas, *Toricellia tiliifolia*.

Centre for Advanced Studies in Botany, School of Life Sciences,
North - Eastern Hill University, Shillong 793 022, India
Email: sentin_chettri@yahoo.co.in

Introduction

Toricellia tiliifolia, a species of one genus usually placed in the family Cornaceae (Polunin & Stainton, 1984) is confined in the deciduous broad-leaved forest. *Toricellia,* consisting of two species, is endemic to Eastern Asiatic region. However, Menris Wiki (2006) in Nepal Biodiversity yearbook has reported that it is represented by a single species, *T. tiliifolia* (Toricelliaceae). Its systematic position, whether closer to Cornaceae than to Araliaceae or vice versa, has been in dispute. Cytologically it seems closer to Araliaceae (Yan-Cheng et al., 1984), because the basic chromosome number of Cornaceae is x =11, 9, 8 whereas that of Araliaceae is 12 (Raven, 1975). Detailed descriptions of the species have been given in (Flora of China, 2005) and in the Flowers of the Himalaya (Polunin & Stainton, 1984).

The present study aims to develop predictive models of the potential distributional range of the species and to understand the macro scale environmental factors affecting its distribution in the Indo-Himalayan region. The results of this study should impact our understanding of this naturally-occurring species and enhance our expectations of where it may occur and which environmental factors are likely to serves as best attribute for its distributions. Surveys for more *T. tiliifolia* might, profitably be focused in the areas indicated in the maps provided herein. The study would further help in delineating conservation areas of the species in the Indo-Himalayan region.

Material and Methods

Preliminary information about the occurrence of the species *T. tiliifolia* in different parts of the country was collected based on the survey of published literature, herbarium data collected from Botanical Survey of India, Eastern Circle, Shillong, Sikkim and Department of Botany, North-Eastern Hill University, Shillong. Nonetheless, we have also recorded occurrence data on a few individuals of the species in a patchy population in the state of Sikkim, Arunachal Pradesh in India and from Bhutan.

Input data

Occurrence data on *T. tiliifolia* were collected by extensive field visits in Sikkim and Arunachal Pradesh and coordinates were collected with the help of GPS (GARMIN model Map 76). GPS readings on 20 unique localities were collected from Arunachal Pradesh and Sikkim in India. The readings were in DMS (degree, minute, second) format which were later converted to decimal degrees for use in the modeling exercise. We employed an artificial intelligence program, GARP which includes several inferential approaches in an iterative, evolutionary computing environment (Stockwell and Peters, 1999). GARP is a genetic algorithm that produces predictive models for species distribution. However, as GARP is non-deterministic, multiple optimum models are produced, and subsequent runs using the same data produces slightly different result.

33 environmental parameters were used for modeling the potential distributional area, of which 19 bioclimatic variables were derived from globally interpolated datasets of monthly temperature and precipitation downloaded from *http://www.worldclim.org*. The bioclimatic variables included annual mean temperature and precipitation, mean diurnal temperature range, temperature annual range, temperature and precipitation seasonality, isothermality, maximum temperature of warmest month, minimum temperature of coldest month, precipitation of driest and wettest month, and mean temperature and precipitation of wettest, driest, warmest and coldest quarter. Since the species is reported to endemic to the eastern Asiatic region, the environmental layers representing the mean monthly precipitation and temperature for the flowering and fruiting months i.e. May, June, July and August were also included in order to minimize the generality of the bioclimatic variables. It is presumed that precipitation and temperature for these four months play an important role in the vegetative and reproductive phenology of the species and as a result would affect its distribution also. Environmental layers of slope, aspect, digital elevation model (DEM) and compound topographic index from the USGS Hydro-

1K dataset were also included (USGS 2001). All analyses were conducted at the native 30 arc seconds (1x1 km) spatial resolution of the environmental data sets.

Ecological niche modeling

The ecological niche of a species can be defined as the set of ecological conditions within which it is able to maintain its populations without immigration (Grinnell, 1917; Holt & Gaines, 1992). Several approaches have been used to approximate species ecological niches (Austin et al., 1990; Nix, 1986; Scott et al., 1993, 1996, 2002; Walker & Cocks, 1991). Of these, one that has undergone rigorous testing under diverse scenarios is the Genetic Algorithm for Rule-set Prediction (GARP). All modeling in this study was carried out on a desktop implementation of GARP now available publicly for download.

We divided the occurrence data as follows: (1) 10 training points (for rule generation) and (2) 10 intrinsic testing points for model optimization and refinement. The predictive GARP model thus produced was validated by intensive visits to the predicted sites. The occurrence of the species in the predicted regions was used as a measure of model predictive ability. Binomial tests (based on the proportional area predicted present and the number of independent test points successfully predicted) were used to compare observed predictive success with that expected under random (null) models of no association between predictions and test points. As model results are in the form of a 'ramp' of model agreement from 0-10, we repeated binomial tests across all the thresholds of model agreement (prediction levels 1to 10). As some places outside India could not be visited due to logistic constraints, we relied on published reports for validation of the model (Qiuyun & Boufford, 2005; Wangda & Ohsawa, 2006). Further, field validation was done by overlaying the predicted distribution map on Google Earth and subsequently visiting the accessible areas falling within the predicted distributional range.

GARP is designed to work based on presence-only data; absence information is included in the modeling via sampling of pseudo-absence points from the set of pixels where the species has not been detected (Stockwell & Peters, 1999). GARP works in an iterative process of rule selection, evaluation, testing, and incorporation or rejection: first, a method is chosen from a set of possibilities (e.g., logistic regression, bioclimatic rules), and then is applied to the training data and a rule developed; rules may evolve by a number of means (e.g., truncation, point changes, crossing over among rules) to maximize predicitivity. Predictive accuracy is then evaluated based on 1,250 points sampled randomly from the study region as a whole to represent pseudo-absences. The change in predictive accuracy rule should be incorporated into the model and the algorithm run either 1,000 iterations or until convergence.

We developed 100 replicate model runs for *T. tiliifolia*, and filtered out suboptimal models based on characteristics in terms of omission (leaving areas of known predictions) and commission (including areas not actually inhabited) error statistics. Following recent recommendations (Anderson et al., 2003) and also to represent a balance between optimizing model selection and practicalities of computing time required for the analysis, we selected best models in Desktop GARP using a 0% extrinsic omission threshold and 50% commission threshold. Experiments with different thresholds indicate that results are quite robust to minor variation in thresholds chosen. Throughout our analysis, we masked analyses to include only the native range of the species i.e. Sikkim, Bhutan and Western part of Arunachal Pradesh including Tawang and West Kameng districts. The final rules generated were then projected into the whole Indo-Himalayan belt including Nepal, some parts of China, Bhutan and the northeastern states of India. All the modeling exercise was performed in a DELL Precision WORKSTATION 690 (X64 Edition).

Extraction and identification of important macro-environmental factors

Environmental data for the native range of *T. tiliifolia* was extracted from the input raster grids used in the modeling exercise.

We combined the input environmental grids with the final ecological niche model to create a new grid with a distinct value for each unique combination of environments. Then we exported the attributes table associated with this grid in database format for further analysis in the statistical software package. To identify the major factors governing the distribution of *T. tiliifolia* in its native range we run a forward stepwise regression keeping all the input environmental variables *viz.* the 19 bioclimatic variables along with altitude, slope, aspect, digital elevation model and topographic index as predictors and the model agreements (1 to 10) as dependent variable. The analysis was run in SYSTAT 10.2, SYSTAT Software Inc., USA, 2002 (Table 1).

Table1: Results of forward multiple regression showing the effect of three environmental parameters (3, 12 & 17) on *T. tiliifolia* distribution in the Eastern Asiatic region: Dependent Variable: Model N: 108 Multiple R: 0.677 Squared multiple R: 0.458; Adjusted squared multiple R: 0.443 Standard error of estimate: 2.072.

Effect	*df*	*F*	*'P'*
In			
Digital Elevation Model	1	24.34	0.00
Bio3	1	4.75	0.03
Bio17	1	61.45	0.00
Out			
Topography	1	3.90	0.05
Slope	1	0.63	0.43
Aspects	1	0.47	0.49
Bio8	1	2.19	0.14
Bio7	1	2.70	0.10
Bio6	1	1.04	0.31
Bio5	1	2.82	0.10

Bio4	1	0.32	0.58
Bio19	1	1.94	0.17
Bio15	1	2.34	0.13
Bio13	1	0.00	0.95
Bio12	1	0.00	0.96
Bio11	1	1.69	0.20
Bio10	1	2.19	0.14
Bio1	1	1.70	0.20
Altitude	1	0.82	0.37

Ecological niche visualization

To permit visualization of patterns of *T. tiliifolia* ecological niche variation, we used the data extracted in database format by the procedure performed in above section. The data was then imported in MS-EXCEL for graphical exploration.

RESULTS

Predicted distributional range

The habitat suitability of *T. tiliifolia*, as predicted by all 10 best-subsets models is depicted in Figure 1.

Predicted distributions of *T. tiliifolia* were agreeable on the reported distributional range of the species and shows that most of the Indo Himalayan belt has high probability of occurrence of the *T. tiliifolia* species. This species has been reported from Nepal, Bhutan (South Chuka, Punakha, Trongsa, Way to Wangdue Phodrang, Trashiyangtse & East Trashigang) Sikkim (personal field visit), some parts of China and particularly in the Himalayan belt (Hooker, 1979, Grierson & Long, 1991; Polunin & Stainton, 1984; Qiuyun & Boufford, 2005;) and Arunachal Pradesh (personal field visit).

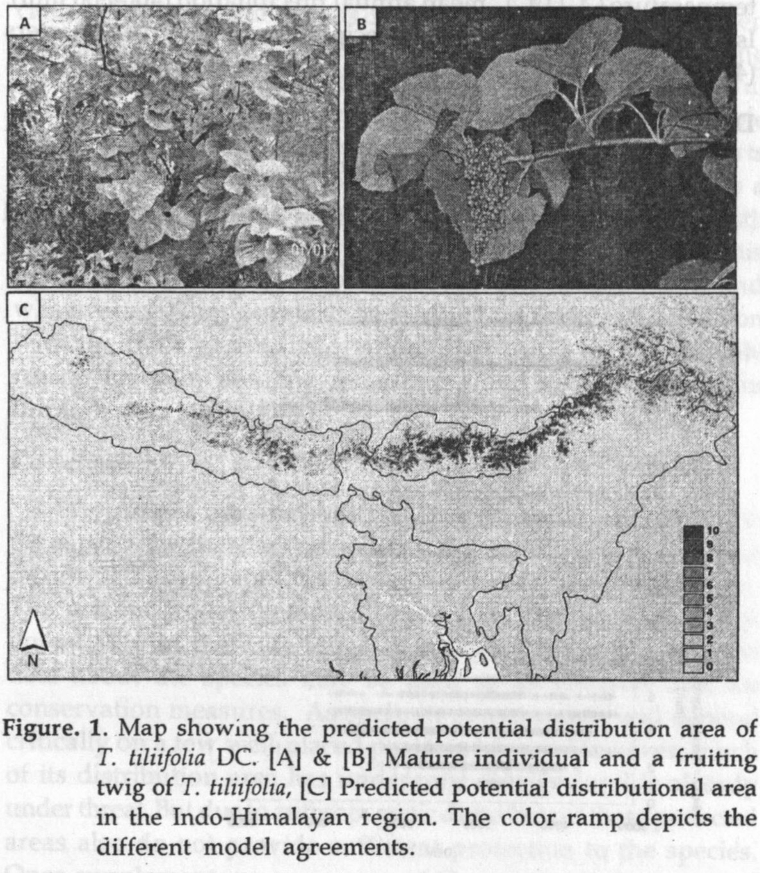

Figure. 1. Map showing the predicted potential distribution area of *T. tiliifolia* DC. [A] & [B] Mature individual and a fruiting twig of *T. tiliifolia,* [C] Predicted potential distributional area in the Indo-Himalayan region. The color ramp depicts the different model agreements.

The environmental visualizations of the models explored the species' distribution in various ecological niches. We compared ecological characteristics of the areas of the species with the environmental range. For the 19 bioclimatic and physiographic variables we regressed the variables and removed highly correlated, redundant variables, and then developed bivariate plots to visualize the species, distribution (Figure 2).

This visualization reveal a very narrow niche for the species - restricted to areas preserving narrow ranges of mean annual temperature (4-11°C), mean annual precipitation (400-860 mm), Isothermality (4.5-4.7 °C) and precipitation of the driest quarters (4-10 mm).

Discussions

3.1 Conservation of *T. tiliifolia*

In all the 20 unique localities from Arunachal Pradesh and

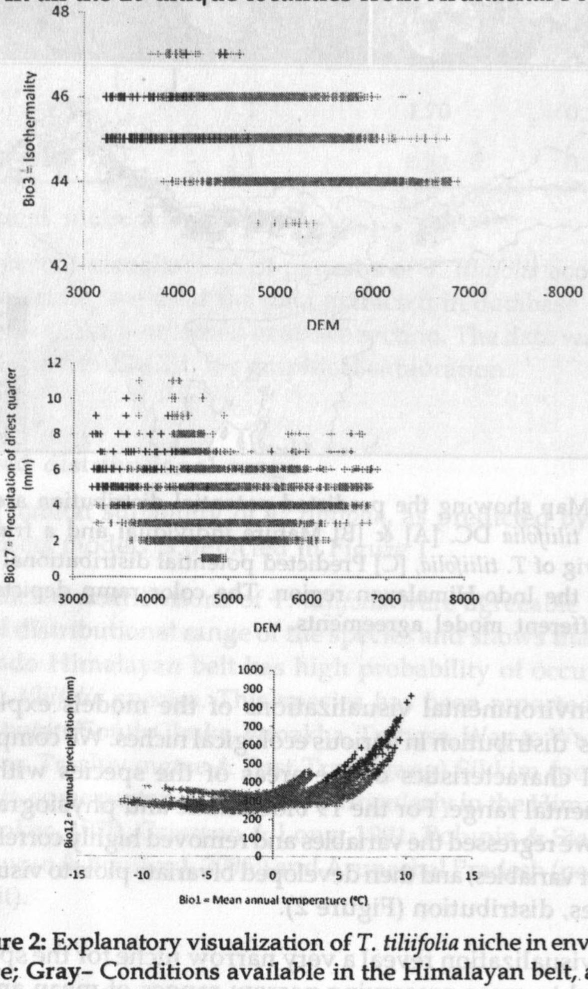

Figure 2: Explanatory visualization of *T. tiliifolia* niche in environmental space; Gray– Conditions available in the Himalayan belt, and Black– modeled potential distribution of species.

Sikkim in India, disturbance in the form of movement by humans, cattle, trekking corridor, roads constructions are common. Further, the extraction of forest trees by humans can reduce flowering trees too far below their naturally occurring densities (Kadavul & Parthasarathy, 1999). Phenology also plays an important role for the perpetuation of the species. In *T. tiliifolia*, fruiting starts during the month of May and the seeds are soon exposed to a long moisture excess condition and high temperature (July-August). This is followed by heavy rainfall (May onwards). During this period (May-August) a substantial amount of seed is lost and whatever seedlings germinate have to undergo intense competition with the thick ground vegetation. This could be the possible reason for poor seedling recruitment and good regeneration through stem sprouting.

Conclusion

We present our analyses and distributional predictions in the expectation that they can assist in assembling a more complete picture of the geographic and ecological distribution of the species. The present study attempted to focus on species distribution potentiality in the Indo-Himalayan region. Thus the baseline data about the species may be used as an effective tool for conservation measures. As such, its conservation will depend critically on a few well-placed protected areas. However, much of its distribution area has undergone changes and is already under threat. But due to anthropogenic disturbances these protected areas also do not provide sufficient protection to the species. Once supplementary occurrence of the species is documented than detailed extend of distribution can be mapped in this regions. The predictions of the potential distribution of the species should help in developing strategies for proper censuring the species population.

Based on the population size, present area of occupancy, the extent of occurrence, habitat under threat, low density of mature trees are some of the parameters that is likely to put the species under threatened category. Our findings indicate that if anthropogenic pressure continues to operate future perpetuation of *T. tiliifolia* in the forest may be threatened. Further, to identify the potential niches of the species using Ecological Niche Modelling (ENM) techniques could be applied for its reestablishment. Hence, proper conservation measures have to be formulated to protect the species.

Acknowledgements

Financial support received from Ministry of Environment and Forests, Government of India under its Biosphere Reserve Programme (Sanction letter No. 8/30/03-CS/BR dated 29.3.2004) is thankfully acknowledged.

REFERENCES

1. Anderson RP, Lew D & Peterson AT. 2003. Evaluating predictive models of species distributions: criteria for selecting optimal models. *Ecological Modeling* 162: 211- 232.
2. Austin MP, Nicholls AO & Marmugules CR. 1990. Measurement of the realized qualitative niche: environmental niches of five *Eucalyptus* species. *Ecological Monograph* 60: 161-177.
3. Yan-Cheng T, Qiu-Yun X & Ya-Ling C. 1984. Cytological studies on some plants of Sichuan and neighbouring regions (I). *Acta Phytotaxonomica Sinica* 22 (5): 343-350.
4. Grinnell J. 1917. Field test of theories concerning distributional control. *American Naturalist* 51: 115-128.
5. Grierson AJC & Long DJ. 1991. *Flora of Bhutan,* volume 2.
6. Holt RD & Gaines MS. 1992. Analysis of adaptation in heterogeneous landscapes: Implication for the evolution of fundamental niches. *Evolutionary Ecology* 6: 433- 447.
7. Hooker JD. 1979. Flora of India. Volume II, pp748.
8. Kadavul K & Parthasarathy N. 1999. Structure and composition of woody species in tropical semievergreen forest of Kalrayan Hills, Eastern Ghats, India. *Tropical Ecology* 40: 247-260.
9. Nix HA. 1986. A biogeographic analysis of Australian elapid snakes. In: Longmore R, ed. Atlas of elapid snakes of Australia. Canberra: *Australian Government publishing Service* pp4-15.

10. Polunin O & Stainton A. 1984. Flowers of the Himalayas. Oxford University Press, New Delhi - 110001.

11. Qiuyun X & Boufford DE. 2005. *Flora of China* 14: 233- 234.

12. Raven PH. 1975. The bases of angiosperm phylogeny, cytology. *Annals of Missouri Botanic Garden* 62(3): 724-764.

13. Scott JM, Davis F, Csuti B, Noss R, Butterfield B, Groves C, Anderson H, Caicco S, D'Erchia F, Edwards Jr TC, Ullman J & Wright RG. 1993. Gap analysis: a geographic approach to protection of biological diversity. *Wildlife Monograph* 123: 1- 41.

14. Scott JM, Heglund PJ, Haufler JB, Morrison M, Raphael MG, Wall WB & Samson F. (eds) 2002. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington DC.

15. Stockwell DRB & Peters D. 1999. The GARP modeling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* 13: 143-158.

16. Walker PA & Cocks KD. 1991. HABITAT: a procedure for modelling a disjoint environmental envelope for a plant or animal species. *Global Ecology and Biogeographic Letter* 1: 108-118.

17. Wangda P & Ohsawa M. 2006. Gradational forest change along the climatically dry valley slopes of Bhutan in the midst of humid eastern Himalaya. *Plant Ecology* 186: 109-128.