*Global Change Biology*

SUPPORTING INFORMATION

Projecting future expansion of invasive species:

Comparing and improving methodologies for species distribution modeling

Kumar P. Mainali, Dan L. Warren, Kunjithapatham Dhileepan, Andrew McConnachie, Lorraine   
 Strathie, Gul Hassan, Debendra Karki, Bharat B. Shrestha, Camille Parmesan

Appendix S1: Additional tables (Tables S1-S5)

 Table S1. Sources of occurrence points.

 Table S2. Environmental predictors used in species distribution modeling.

 Table S3. Sources of and explanation for using non-climatic variables.

 Table S4. Four way analysis of variance for effect of various factors on AUC score.

 Table S5. AUC scores for various regions computed with models using three point sources.

Table S1 Sources of occurrence points.

Sources Number

of records

*Freely available databases*

Arne Witt Commonwealth Agricultural Bureau International (CABI) Africa 2

Australian National Herbarium 163

Australia's Virtual Herbarium (AVH) 168

Global Biodiversity Information Facility (GBIF) 1097

Integrated Pest Management Collaborative Research Support Program (IPM CRSP) Ethiopia 266

Integrated Pest Management Collaborative Research Support Program (IPM CRSP) South Africa 45

The National Commission for Knowledge and Use of Biodiversity, Mexico (CONABIO) 382

South African Plant Invaders Atlas 53

Swaziland's Alien Plant Database 20

*Personal collection*

Bharat Shrestha 67

Debendra Karki 859

Iqbal Zuberi 15

Jim Findlay (private individual) 3

K Dhileepan 332

Maan Rokaya 3

Mark Hyde (private individual) 8

Sue van Rensburg (Grumeti Trust) 1

*Published studies*

Dhileepan K and KAD Wilmot Senaratn. 2009. How widespread is *Parthenium hysterophorus* and 423 its biological control agent *Zygogramma bicolorata* in South Asia? *Weed Research* 49: 557-562

Hassan G and International Linkages Project. 2011. Further update of the status of parthenium 64 weed in Pakistan. *International Parthenium News* 3:9-10

Tang SQ, F Wei, LY Zeng, XK Li, SC Tang, Y Zhong, and Y-P GENG. 2009. Multiple introductions 18 are responsible for the disjunct distributions of invasive *Parthenium hysterophorus* in China:   
evidence from nuclear and chloroplast DNA. *Weed Research* 49: 373-380

TOTAL 3989

Mainali et al. Improving gobal projections for plant invasion - APPENDIX S1 pg. 1

Table S2 Environmental predictors used in species distribution modeling.

1. bio1 = annual mean temperature

2. bio2 = mean diurnal range (mean of monthly (max temp - min temp))

3. bio3 = isothermality (bio2/bio7) (\* 100)

4. bio4 = temperature seasonality (standard deviation \*100)

5. bio5 = max temperature of warmest month

6. bio6 = min temperature of coldest month

7. bio7 = temperature annual range (bio5-bio6)

8. bio8 = mean temperature of wettest quarter

9. bio9 = mean temperature of driest quarter

10. bio10 = mean temperature of warmest quarter

11. bio11 = mean temperature of coldest quarter

12. bio12 = annual precipitation

13. bio13 = precipitation of wettest month

14. bio14 = precipitation of driest month

15. bio15 = precipitation seasonality (coefficient of variation)

16. bio16 = precipitation of wettest quarter

17. bio17 = precipitation of driest quarter

18. bio18 = precipitation of warmest quarter

19. bio19 = precipitation of coldest quarter

20. alt = altitude

21. vegetation = percent tree canopy cover

22. soilmois = soil moisture

23. population = human population density

24. roadsqrt = square root of proximity to road

25. road = proximity to road (linear distance)

Mainali et al. Improving gobal projections for plant invasion - APPENDIX S1 pg. 2

Table S3 Sources of and explanation for using non-climatic variables. WorldClim variables can be

supplemented with other climatic and non-climatic variables to improve model predictions (Peterson & Nakazawa, 2008).

Variables Source Reference Note and explanation

Soil NOAA,

moisture [http://www.esrl.noaa.gov/psd/](http://www.esrl.noaa.gov/psd/data/gridded/data.cpcsoil.html)

data/gridded/data.cpcsoil.html

Percent Global Land Cover Facility,

canopy Collection 4, Version 3, accessed in

cover Dec 2011,

<http://www.glcf.umd.edu/>data/vcf/

Human NASA Socioeconomic Data and population Application Center, Gridded

density Population of the World v3,

<http://sedac.ciesin.columbia.edu/>data/collection/gpw-v3

Proximity Digital Chart of the World,

to nearest accessed on Jan 30, 2012,

road [http://www.diva-gis.org/gData](http://www.diva-gis.org/gdata/)

van den Dool Because precipitation is not the best measure

et al., 2003 of moisture availability to plants, we added a

layer representing soil moisture. Monthly   
values of soil moisture for the period of   
1948-2011 were obtained in grid format at

0.5 degrees resolution. The raster data was resampled to match the resolution and extent of WorldClim predictors

Hansen et al., Percent canopy cover per 500 m MODIS

2003 pixel was obtained for the years 2001-2005.

We averaged the six years’ data. The raster data was resampled to match the resolution and extent of WorldClim predictors

Except in Australia, the plant seems to be   
more abundant in localities with higher

population density. It is often seen growing at high density in areas that receive frequent

human disturbance and domestic waste. To use human population density as a potential predictor of the weed, we obtained gridded population density (2.5 arc minute

resolution, adjusted to match UN totals) in 2000 from

A shapefile of roads of the world was used to generate two types of raster layers: one with linear distance to the nearest road and the other with square root of linear distance (see Appendix 2, Fig S1).

References

Hansen MC, DeFries RS, Townshend JRG, Carroll M, Dimiceli C, Sohlberg RA (2003) Global percent tree cover at   
 a spatial resolution of 500 meters: first results of the MODIS vegetation continuous fields algorithm. *Earth   
 Interactions*, 7, 1-15.

Peterson AT, Nakazawa Y (2008) Environmental data sets matter in ecological niche modelling: an example with   
 *Solenopsis invicta* and *Solenopsis richteri*. *Global Ecology and Biogeography*, 17, 135-144.   
van den Dool H, Huang J, Fan Y (2003) Performance and analysis of the constructed analogue method applied to   
 U.S. soil moisture over 1981-2001. *Journal of Geophysical Research*, 108, D16.

Mainali et al. Improving gobal projections for plant invasion - APPENDIX S1 pg. 3

Table S4 Four way analysis of variance for effect of various factors on AUC score.

Factors Df Sum Sq Mean Sq F value Pr(>F) Sig. codes

PointSource 2 216.48 108.24 450000.00 < 2e-16 \*\*\*

Bias 1 0.03 0.03 132.10 < 2e-16 \*\*\*

Model 3 43.94 14.65 60890.00 < 2e-16 \*\*\*

ExplRoad 1 5.13 5.13 21310.00 < 2e-16 \*\*\*

PointSource:Bias 2 0.01 0.00 18.21 0.00 \*\*\*

PointSource:Model 6 6.33 1.06 4388.00 < 2e-16 \*\*\*

Bias:Model 3 0.00 0.00 2.92 0.03

PointSource:ExplRoad 2 2.67 1.33 5546.00 < 2e-16 \*\*\*

Bias:ExplRoad 1 0.00 0.00 0.06 0.80

Model:ExplRoad 3 0.48 0.16 669.00 < 2e-16 \*\*\*

PointSource:Bias:Model 6 0.00 0.00 2.41 0.02

PointSource:Bias:ExplRoad 2 0.00 0.00 0.44 0.64

PointSource:Model:ExplRoad 6 0.44 0.07 306.00 < 2e-16 \*\*\*

Bias:Model:ExplRoad 3 0.00 0.00 1.32 0.27

PointSource:Bias:Model:ExplRoad 6 0.00 0.00 1.46 0.19

Residuals 119952 28.85 0

*Signif. codes: ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1*

Mainali et al. Improving gobal projections for plant invasion - APPENDIX S1 pg. 4

\*

\*

Table S5 AUC scores of four models built with three point sources when tested with presences and

background points from various continents. The background points were obtained using grid cell area bias (AreaBias). The set of explanatory variables does not include proximity to road (NoRoad). Scores are displayed against background of color ramp (green:low to red:high) that spans the range of scores within each continent. (Colors are not comparable across continents.)

Continent PointSource GAM GLM RF BRT

Africa PNBN 0.695 0.693 0.739 0.749

Africa PWBN 0.641 0.591 0.837 0.794

Africa PWBW 0.797 0.782 0.908 0.852

Asia PNBN 0.66 0.654 0.651 0.661

Asia PWBN 0.565 0.566 0.69 0.653

Asia PWBW 0.799 0.799 0.856 0.837

Australia PNBN 0.573 0.586 0.61 0.617

Australia PWBN 0.774 0.712 0.821 0.799

Australia PWBW 0.825 0.819 0.898 0.867

North America PNBN 0.827 0.826 0.859 0.85

North America PWBN 0.804 0.804 0.864 0.834

North America PWBW 0.806 0.793 0.859 0.83

South America PNBN 0.618 0.614 0.596 0.612

South America PWBN 0.582 0.567 0.578 0.569

South America PWBW 0.622 0.621 0.702 0.636

World PNBN 0.701 0.7 0.731 0.717

World PWBN 0.635 0.587 0.703 0.691

World PWBW 0.794 0.787 0.87 0.835

Mainali et al. Improving gobal projections for plant invasion - APPENDIX S1 pg. 5