J. Range Manage.
56: 206-212 May 2003

# Distribution of Russian knapweed in Colorado: Climate and environmental factors

**SARAH C. GOSLEE, K. GEORGE BECK AND DEBRA P.C. PETERS**

*Authors are Ecologist, USDA-ARS, Jornada Experimental Range, Las Cruces, N.M. 88003; Professor, Department of Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, Colo. 80523; and Landscape Ecologist, USDA-ARS, Jornada Experimental Range, Las Cruces, N.M. 88003.*

## Abstract

**Russian knapweed (*Acroptilon repens* (L.) DC.) was introduced to the western United States during the early 1900s. This invasive perennial was a contaminant of alfalfa seed, and was distributed widely across Colorado. Thus, current distributions reflect the climate and soils tolerances of Russian knapweed, and management history, rather than dispersal processes. We surveyed extension and weed agents across Colorado, and were able to locate 528 current or recently eliminated Russian knapweed stands. These patches were superimposed on climate and soils maps to identify 1 km grid cells that were known to contain Russian knapweed. The status of Russian knapweed within a cell was used as the dependent variable in a logistic regression model to define the environmental envelope for this species. At the scale of our analysis, Russian knapweed was most prevalent on fine-textured soils (clay and clay loam), and in warmer, drier regions of Colorado (precipitation 18–73 cm/yr, mean annual temperature 1–12° C). June precipitation was the most important single factor, although nearly all environmental, annual, and monthly climatic factors were significantly related to Russian knapweed occurrence. The multivariate logistic regression model we developed was used to predict the probability of occurrence of Russian knapweed for the entire state of Colorado. Our predictions matched the areas of highest abundance of Russian knapweed from a new field survey, and also indicated areas of high risk that were not identified by the field survey.**

**Key Words: generalized additive model, geographic information system, logistic regression**

## Resumen

**El "Russian knapweed" (*Acroptilon repens* (L.) DC.) se introdujo al oeste de los Estados Unidos a inicio del siglo pasado, en los primeros años de 1900's. Esta especie perenne invasora fue un contaminante de la semilla de alfalfa y se distribuyo ampliamente a través de Colorado. Así que, su distribución actual refleja la tolerancia del "Russian knapweed" al clima y suelos e historial de manejo mas que el proceso de dispersión. Inspeccionamos agentes de extensión y de control de maleza del estado de Colorado y fuimos capaces de localizar 528 poblaciones de "Russian knapweed" actual o recientemente eliminadas. Estos parches fueron sobrepuestos en mapas de suelo y clima para identificar en una cuadricula celdas de 1 km$^2$ en las que sabíamos que el "Russian knapweed" estaba presente. El posición del "Russian knapweed" dentro de la celda se uso como variable dependiente en un modelo de regresión logística para definir el rango ambiental de esta especie. En la escala de nuestro estudio, el "Russian knapweed" prevaleció mas en suelos de textura fina (arcillosos o franco arcillosos) y en las regiones cálidas y secas del estado de Colorado (precipitación de 18—73 cm/año, temperatura media anual de 1-12° C). La precipitación de Junio fue el factor individual mas importante, aunque casi todos los factores ambientales climáticos, anuales y mensuales, estuvieron significativamente relacionados a la ocurrencia del "Russian knapweed". El modelo multivariado de regresión logística que desarrollamos se uso para predecir la probabilidad de ocurrencia del "Russian knapweed" en todo el estado de Colorado. Nuestras predicciones concordaron con las áreas de mayor abundancia de "Russian knapweed" identificadas en un nuevo muestreo de campo y también indicaron las áreas con mayor riesgo que no fueron identificadas por el reconocimiento de campo.**

Invasive species are a severe problem world-wide. While control or eradication are usually the ultimate goals, in many cases scientists lack even basic information about the biology and ecology of invasive species. Knowledge of the distribution of a species with respect to environmental and climatic factors can be used to prioritize management efforts for those regions where an invasive is most likely to succeed (Panetta and Dodd 1987, Higgins et al. 1999). An understanding of climatic controls on species distribution also makes it possible to predict changes in distribution under potential climate change scenarios (Franklin 1998).

Russian knapweed (*Acroptilon repens* (L.) DC.) is a long-lived (> 75 yr), highly aggressive $C_3$ perennial forb. It was introduced to the US in contaminated alfalfa seed during the early 1900s (Rogers 1928, Watson 1980). This invasive weed has large economic and ecological impacts, including reduction of forage quality and poisoning of horses, increased soil erosion, and reduced wildlife populations (Watson 1980, Roché and Roché 1988, 1991). It is a legislated noxious weed in Colorado (State of Colorado 1996), and is a problem species in every western state (Roché and Roché 1991). It was introduced to Colorado in contaminated alfalfa seed, and has been widely distributed across the state (Rogers 1928).

This research was funded by the USDA National Research Initiative Competitive Grants Program award #9600468. We greatly appreciate the cooperation of the many Colorado County Extension Agents and Weed Supervisors who assisted us in identifying the locations of Russian knapweed patches.
Manuscript accepted 27 Jul. 02.

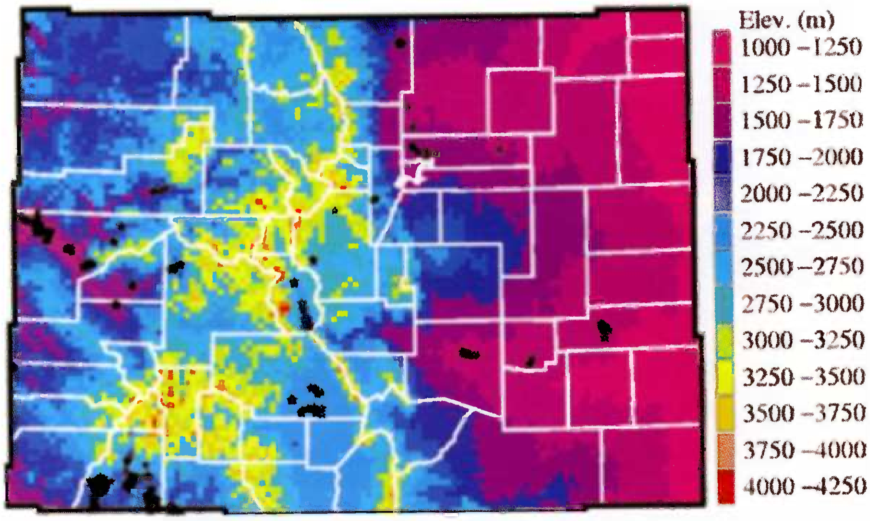


**Fig. 1. Locations of known Russian knapweed stands (black stars) in relation to elevation (USGS 1990). County boundaries are shown as white lines.**

Both field observations and simulation results show that Russian knapweed does not do equally well on all soil types and under all climatic regimes found in Colorado (Goslee et al. 2001, Goslee, S.C., K.G. Beck, and D.P.C. Peters, unpublished data). If we can identify the environmental distribution limits of this species, that information can be used to a) identify areas most vulnerable to Russian knapweed invasion under current conditions; and b) predict where Russian knapweed may become a problem under proposed climate change scenarios. Our objectives were to map known Russian knapweed stands across Colorado, and to use the accumulated data to describe the environmental and climatic variables important in determining Russian knapweed distribution. Russian knapweed is a particularly good candidate for this type of analysis since seeds were widely distributed and the location of current stands should be related to environmental conditions rather than dispersal constraints. Likewise, Colorado spans a relatively wide range of climatic conditions, so that it should be possible to identify both maxima and minima for temperature and precipitation descriptors.

## Methods

To obtain the locations of patches of Russian knapweed, we surveyed cooperative extension agents and weed supervisors across the state of Colorado. All patch locations were entered into a geographic information system (GIS) database using ArcInfo (ESRI, Inc.). We extracted soil texture data (percentage sand, silt, and clay) from STATSGO soils maps (USDA 1991). Elevation data were obtained from 1:250,000-scale digital elevation models (USGS 1990). Temperature and precipitation maps for Colorado from the PRISM database were used as a source for minimum and maximum monthly temperature and monthly precipitation, which were used to calculate annual averages and totals (Daly et al. 1994). All GIS data were rescaled to a consistent pixel size (1 km), resulting in 267310 grid cells within Colorado. The Russian knapweed patch locations were superimposed on the gridded state map to create another GIS layer indicating presence or absence of Russian knapweed in each grid cell. Although it is highly unlikely that we were able to identify every location where Russian knapweed is currently found, the survey methodology allowed us to obtain a representative sample across the state. We compared the distribution of soils and climate values for the patches to the distribution of these variables for Colorado as a whole, not to the areas where no Russian knapweed was reported since we cannot be certain that those areas are actually free of Russian knapweed.

Logistic regression is commonly used to fit species presence/absence to environmental factors, and to predict probability of species occurrence in new areas (Franklin 1995). The form of the regression is based on assumptions about the way that a species responds to environmental factors. Ecologists often assume that species follow a linear or a normal distribution, but in many cases this assumption is incorrect. Generalized additive models (GAMs) provide an alternative way to fit a curve describing species responses without making any assumptions about distribution shape, and thus should more closely fit data of any form (Frescino et al. 2001). The drawback to GAMs is that it is much more computationally challenging to fit and interpret these models, and to use them for predictive purposes.

We fit 3 separate curves using S-Plus (Version 3.4, MathSoft, Inc.): linear logistic regression (assuming linear distribution), second order polynomial logistic regression (assuming normal distribution), and a spline-smoothed generalized additive model (no distributional assumptions), to identify the most appropriate functional form for each variable (Chambers and Hastie 1992, Venables and Ripley 1994). Models were initially fit for each environmental variable separately. To identify the best-fitting models, we examined the

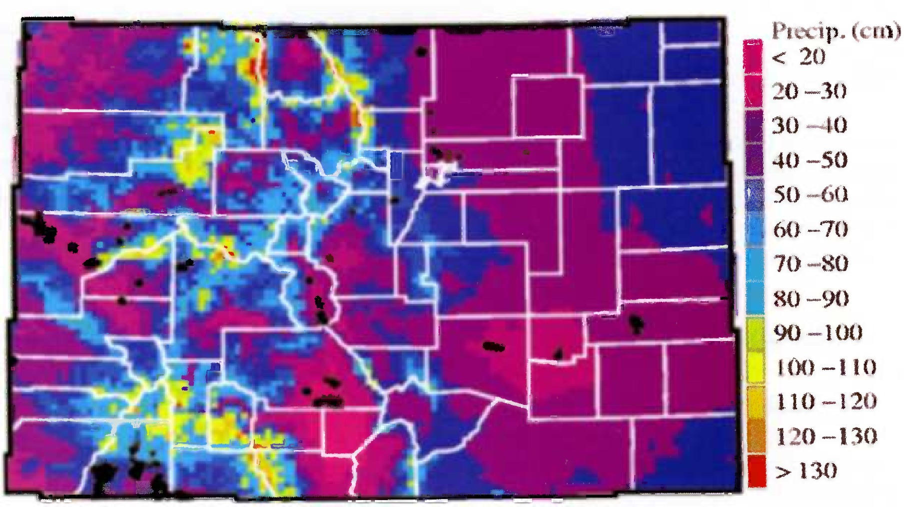


**Fig. 2. Locations of known Russian knapweed stands (black stars) in relation to annual precipitation (Daly et al. 1994). County boundaries are shown as white lines.**

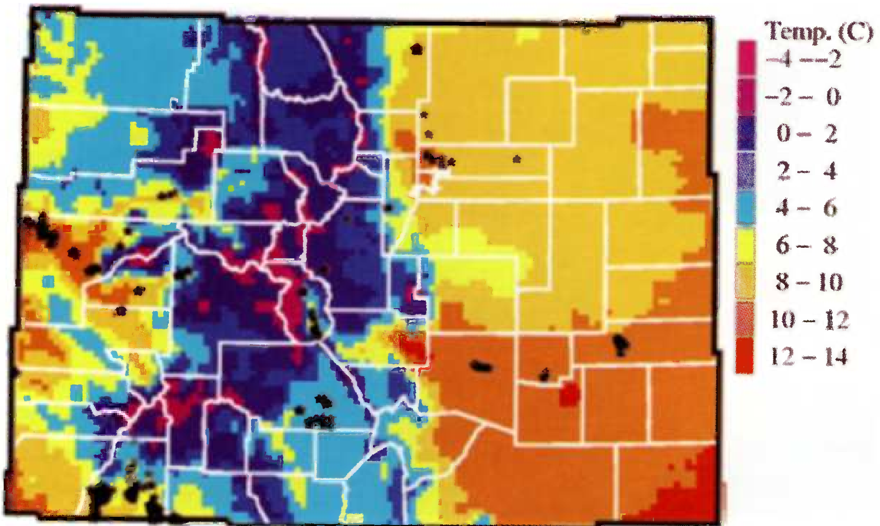


**Fig. 3. Locations of known Russian knapweed stands (black stars) in relation to mean annual temperature (Daly et al. 1994). County boundaries are shown as white lines.**

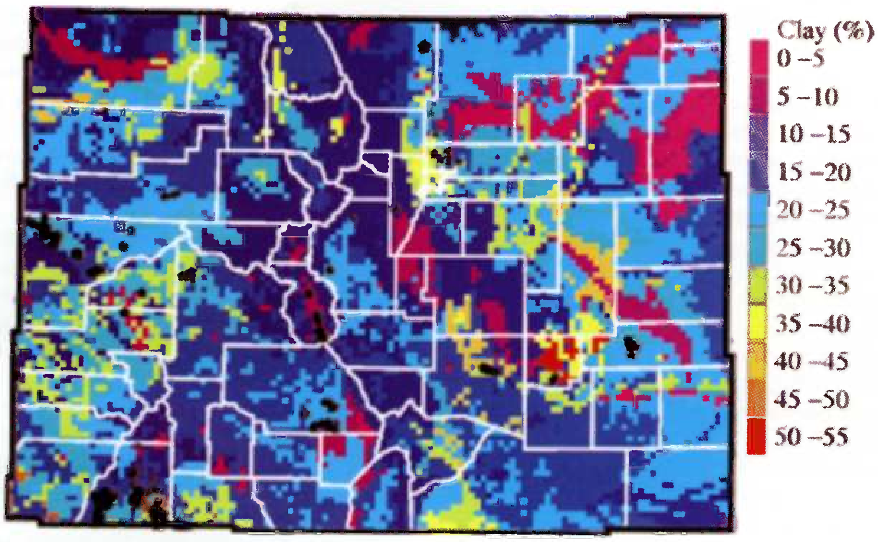


**Fig. 4. Locations of known Russian knapweed stands (black stars) in relation to soil texture (USDA 1991). County boundaries are shown as white lines.**

stands were located around reservoirs or in river valleys. Russian knapweed patches were found mostly in areas of low to moderate annual precipitation (18—73 cm/yr; Fig. 2) and moderate to high mean annual temperature (1—12 C; Fig. 3). Many sites were found on fine-textured soils. Nearly 40% of stands were found on clay loam or clay soils, although those types make up less than 16% of Colorado soils (Fig. 4). A map and list of coordinates for Russian knapweed patches are available from the authors. A previous survey identified location only to the county level, but general regions correspond well with our reported infestations (Maddox et al. 1985).

Summary plots for the entire state of Colorado and for identified Russian knapweed locations for annual totals and for a winter and a summer month are shown in Fig. 5. The precipitation variables (annual, maximum monthly, minimum monthly, June and December) showed the greatest differences between Russian knapweed sites and the state as a whole, indicating that precipitation is an important variable for this species. Certain temperature variables (annual, maximum monthly, and June maximum) showed a trend of Russian knapweed towards warmer sites. These summaries also suggest that these patches span most of its climatic and environmental envelope since Russian knapweed was rarely found in the extremes of any of the variables selected.

We tested models of Russian knapweed distribution in relation to elevation, soil texture (percentages sand, silt, and clay), and 6 annual climate summary parameters: minimum and maximum monthly precipitation, total annual precipitation, mean minimum and maximum temperature, and deviance explained by each, and also compared the models formally using a Chi-squared test (Venables and Ripley 1994). Those explanatory variables with the strongest relationships to Russian knapweed presence were combined in a multivariate logistic regression. Computer limitations prevented an in-depth investigation of multivariate and interaction models.

## Results and Discussion

We were able to identify the locations of 528 Russian knapweed stands from our survey. These patches were found throughout the state, except for the northeastern plains and high elevations in the mountains (Fig. 1). Large clusters of

**Table 1. Polynomial regression equations for individual environmental variables of the form $y = Intercept + Coef1 * x + Coef2 * x^2$. The dependent variable (y) is the natural log of the probability of Russian knapweed presence. For 2 variables, soil clay and maximum monthly precipitation, the polynomial term was not significant, and the linear equation is given. The last column gives the percentage of the residual deviance explained by the appropriate model.**

| | Intercept | Coef1 | Coef2 | %Dev. |
|---|---|---|---|---|
| Elevation (m) | –17.45 | 0.013 | –3.25E-06 | 4.44 |
| Soil clay (%) | –7.75 | 0.072 | | 3.27 |
| Annual precipitation (cm) | –2.24 | –0.127 | 6.09E-04 | 4.77 |
| Minimum monthly precip. (cm) | –9.03 | 4.691 | –1.47E+00 | 4.62 |
| Maximum monthly precip. (cm) | –2.56 | –0.661 | | 7.68 |
| Mean annual temperature (C) | –10.87 | 1.338 | –8.36E-02 | 3.56 |
| Minimum monthly temp. (C) | –13.92 | –1.242 | –4.76E-02 | 1.07 |
| Maximum monthly temp. (C) | –32.64 | 1.751 | –2.85E-02 | 2.66 |
| June precip. (cm) | –1.85 | –1.945 | 1.37E-01 | 13.79 |
| June minimum temp. (C) | –8.63 | 0.821 | –5.32E.02 | 2.71 |
| June maximum temp. (C) | –28.32 | 1.578 | –2.75E-02 | 3.08 |
| December precip. (cm) | –7.88 | 1.550 | –2.36E-01 | 3.73 |
| December minimum temp. (C) | –9.58 | –0.719 | –3.50E-02 | 1.19 |
| December maximum temp. (C) | –7.81 | 1.220 | –1.49E-01 | 5.77 |

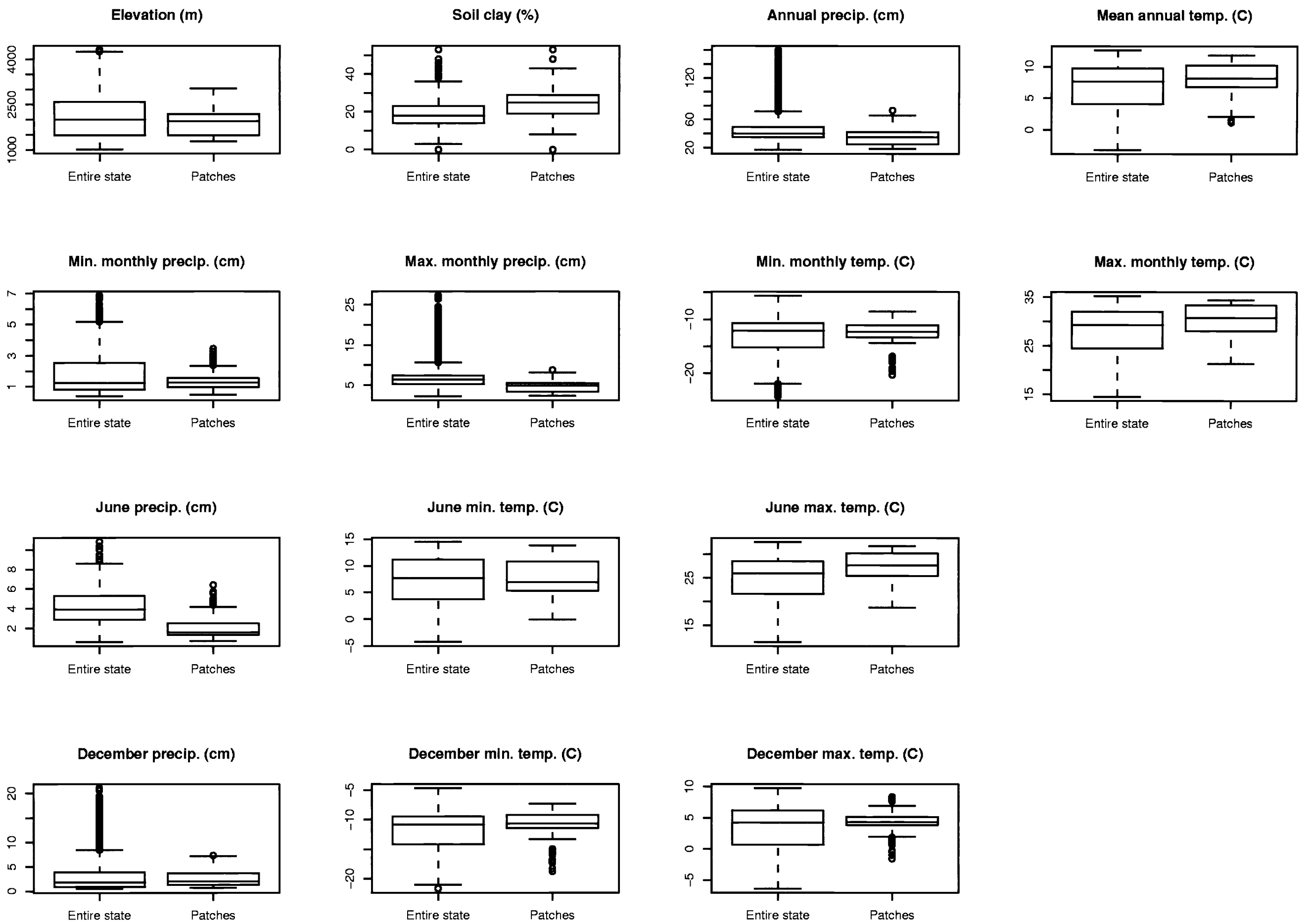


**Fig. 5. Comparison of major environmental and climatic variables across the entire state of Colorado and for locations where Russian knapweed was found. The box area contains the middle 50% of the data, and the median is voided from the box so that skewness can be assessed. Outliers are shown as lines beyond the whiskers, which indicate 1.5 times the interquartile range (S-Plus Version 3.4, MathSoft, Inc.).**

mean annual temperature. We also tested 3 climate parameters for each month: minimum and maximum temperature and total precipitation. Each of these 46 parameters were highly significant explanatory variables for both the polynomial and spline models ($P < 0.001$), and all but October precipitation were also significant for the linear model. June precipitation was the strongest single predictor, explaining nearly 15% of the total deviance in Russian knapweed occurrence (Fig. 6). Maximum monthly and annual precipitation, winter maximum temperature, and elevation were also strong predictors.

As we expected, the spline-fitting model most closely matched the data in all cases. The shape of the polynomial model was surprisingly similar to that of the spline model for all strong predictor variables, and the linear model provided a generally poorer fit (Fig. 7). Although the polynomial model was not quite as good a fit as the spline model, it is more straightforward conceptually and it is easier to parameterize and to use for prediction. Because of these considerations, and the similarity between the shapes of the spline and polynomial models, we have chosen to focus on the latter.

Regression equations for the strongest single-variable polynomial models are given in Table 1. Two of the variables, soil clay and maximum monthly precipitation, were better fit by linear than polynomial relationships, so the coefficients for

**Table 2. Coefficients for the multivariate logistic polynomial model containing elevation and the strongest soil texture, precipitation and temperature variables. The dependent variable (y) is the natural log of the probability of Russian knapweed presence. NS indicates terms which were not significant in the full model, and are not included in the model presented here. This model explained 19% of the total deviance.**

| | Coefficient |
|---|---|
| Intercept | –2.43E+000 |
| Elevation | –1.97E-003 |
| $(\text{Elevation})^2$ | 2.30E-007 |
| Clay | 3.37E-002 |
| $(\text{Clay})^2$ | **NS** |
| June precipitation | –9.04E-001 |
| $(\text{June precipitation})^2$ | **NS** |
| December maximum temp. | 7.28E-001 |
| $(\text{December maximum temp.})^2$ | –7.61E-002 |

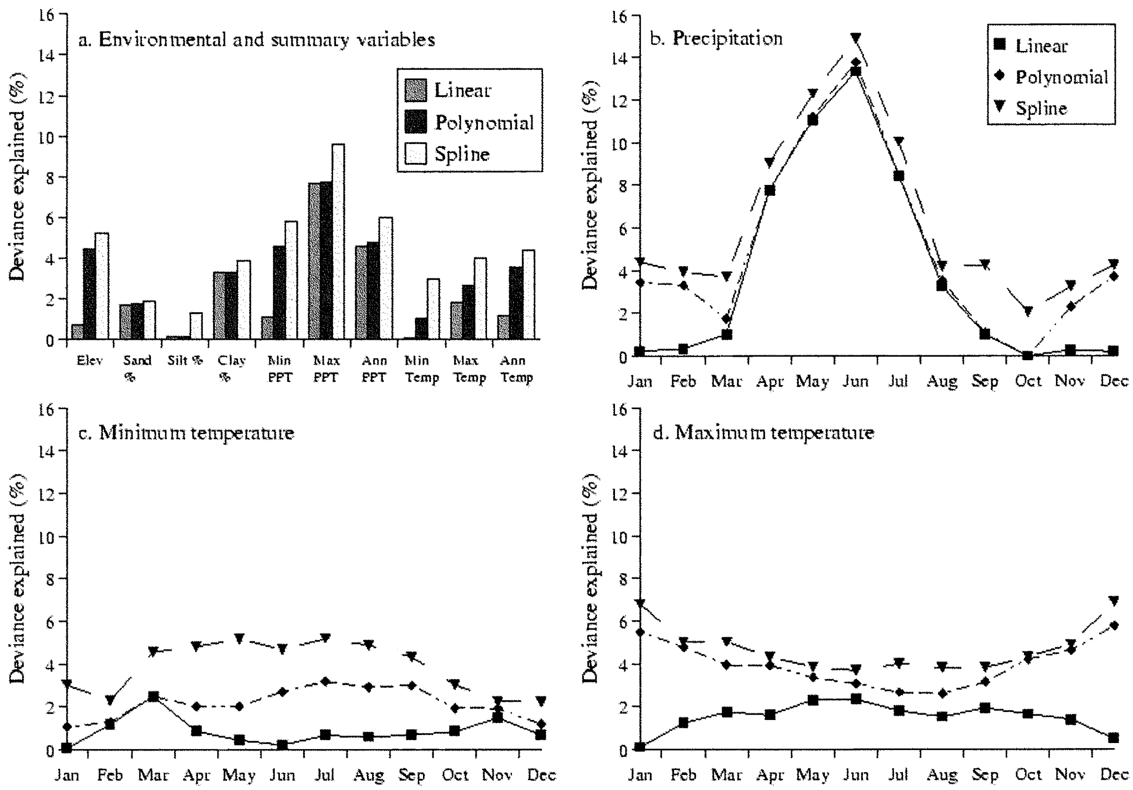


**Fig. 6. Percentage of deviance in Russian knapweed occurrence explained by single-variable linear logistic regression, polynomial logistic regression and spline-smoothed generalized additive models.**

that form are shown. These equations can be used to predict the probability of Russian knapweed occurrence in a novel area within our realm of inference. A polynomial model including elevation, clay, June precipitation, maximum monthly precipitation and December minimum temperature (the most important environmental, soil, precipitation and temperature variables) explained 19% of the total deviance (Table 2).

The fitted values from the multivariate polynomial model were used to draw a map of Colorado showing the predicted probability of occurrence for Russian knapweed in each grid cell (Fig. 8). To assess the validity of our model, we compared our predictions to a recently-released map of Russian knapweed abundance developed by the Colorado Department of Agriculture (2002). While the major areas of infestation identified by the Colorado Department of Agriculture correspond closely to areas identified by the logistic regression model (primarily in river valleys), the regression model highlighted a greater portion of the state as potential Russian knapweed habitat. Most of these mismatches occur in counties where Russian knapweed is known to exist, but has not yet been reported to the Colorado Department of Agriculture. The close similarity of the independently-derived maps suggests that the multivariate logistic regression model adequately describes potential Russian knapweed habitat.

Our results demonstrated that Russian knapweed was found more often on sites with low June precipitation, low elevation, high percentage soil clay, and high December maximum temperatures. The environmental and climatic factors most strongly related to Russian knapweed occurrence were probably determined by its phenology and physiology. Russian knapweed is a $C_3$ species, and begins to grow in the spring as soon as the soil no longer freezes (Watson 1980). Sites with warm winter temperatures may encourage earlier growth, allowing this perennial weed to become active before its competitors. Russian knapweed has a deep root system, which likely allows it to do well in dry areas (Selleck 1964, Watson 1980). Its relationship to precipitation patterns supports that assumption. In sites with low spring precipitation, this deep-rooted perennial may be able to establish an advantage over annuals or shallow-rooted perennials and this advantage could persist throughout the growing season. Information in the literature conflicts over the role of soil texture on Russian knapweed success. Within its native range, Russian knapweed can be found in a wide range of environmental conditions, including sandy, stony and clay soils (Komarov 1963). In the western US and Canada, where Russian knapweed has been introduced, early studies claimed that soil texture is unimportant (Rogers 1928, Watson 1980), but more recent research suggests that this species invades more successfully on fine-textured than on coarser soils (Goslee et al. 2001, Goslee, S.C., K.G. Beck, and D.P.C. Peters, unpublished data). Our analysis supports the observation that Russian knapweed invasion occurs more often on soils with a high clay content, although that relationship was not as strong as with other environmental variables.

Even when including several variables in the model, we were able to explain only a small portion of the total variance in Russian knapweed occurrence. This could be due to the relatively small number of known stands, or because information on other biologically important variables was unavailable. The weak relationships are most likely due to the coarse grain (1 km) of the available environmental data. The distributions of weeds, including Russian knapweed, are often related to distur-

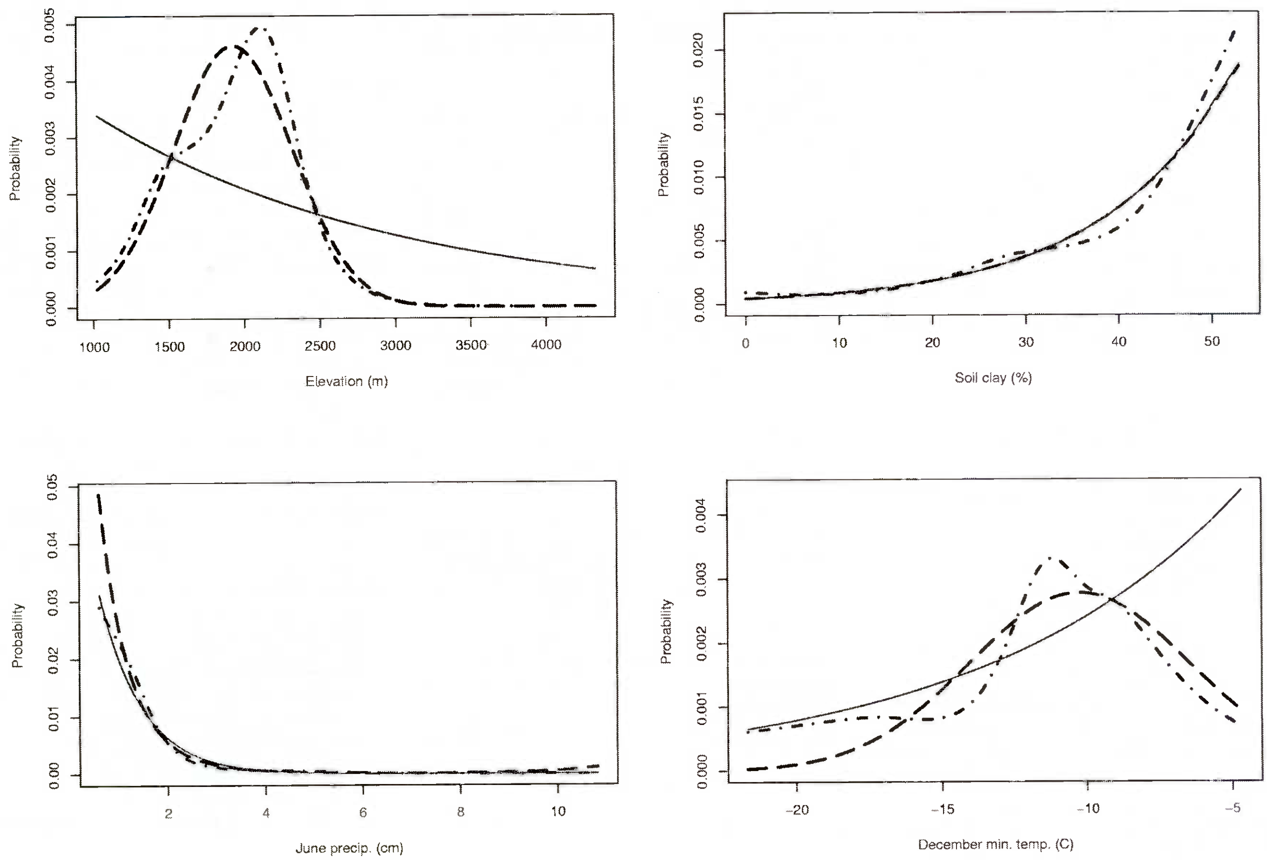


**Fig. 7. Comparison of the fitted curves for linear logistic regression (solid line), polynomial logistic regression (dashed line) and spline - smoothed generalized additive model (dot-dashed line) for four predictor variables. Each shows a different response of the probability of Russian knapweed presence to that variable.**

bances that allow them to become established. The location of the stands identified with respect to roadsides, reservoirs and other disturbances is not known. Although environmental conditions must be correct for long-term survival, disturbance could determine where Russian knapweed became established, and thus affect the strength of the regression relationship by reducing the number of potential sites where it was actually present. Competitive effects of the original vegeta-

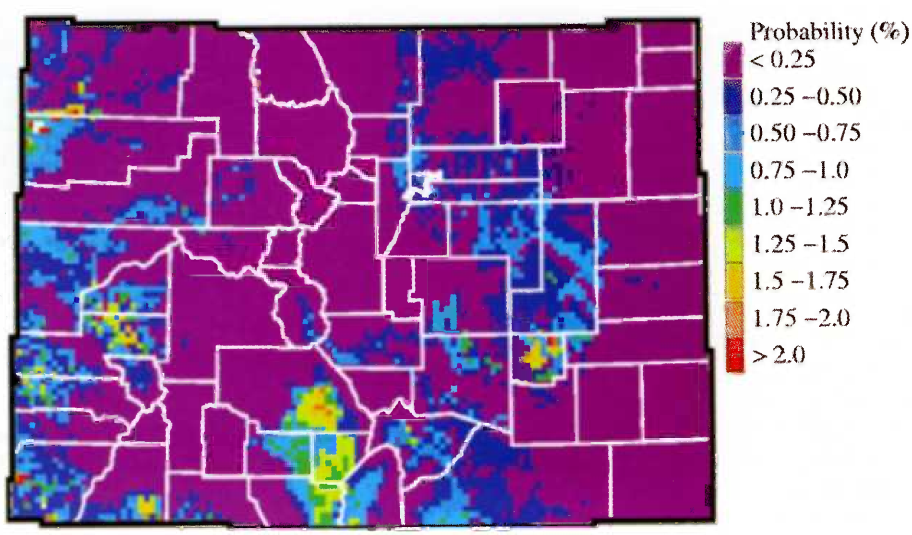


**Fig. 8. Probability of occurrence of Russian knapweed across Colorado derived from the multivariate logistic regression model in Table 2.**

tion could also be important (Panetta and Mitchell 1991).

Our survey of Russian knapweed locations across Colorado identified several important variables, especially June precipitation, elevation, soil clay, and December maximum temperatures. The physiology of Russian knapweed makes these variables important for its success. The relationships identified here can be used for prediction of Russian knapweed success in novel environments, and under novel climatic conditions.

## Literature Cited


**Chambers, J.M. and T.J. Hastie. 1992.** Statistical models in S. Wadsworth and Brooks/Cole, Pacific Grove, Calif.

**Colorado Department of Agriculture. 2002.** Russian knapweed 2001 distribution and abundance in Colorado. http://www.ag.state.co.us/dpi/weeds/mapping/russianknapweed.gif. (Accessed 5 April, 2002.)

**Daly, C., R.P. Nielson, and D.L. Phillips. 1994.** A statistical-topographic model for mapping climatological precipitation over mountainous terrain. J. Appl. Meteorol. 33:140–158.

**Franklin, J. 1995.** Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. Prog. Physical Geog. 19:474–499.

**Franklin, J. 1998.** Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. J. Veg. Sci. 9:733–748.

**Frescino, T.S., T.C. Edwards, Jr., and G.G. Moisen. 2001.** Modeling spatially explicit forest structural attributes using Generalized Additive Models. J. Veg. Sci. 12:15–26.

**Goslee, S.C., D.P.C. Peters, and K.G. Beck. 2001.** Modeling invasive weeds in grasslands: the role of allelopathy in *Acroptilon repens* invasion. Ecol. Model. 139:31-45.

**Higgins, S.I., D.M. Richardson, R.M. Cowling, and T.H. Trinder-Smith. 1999.** Predicting the landscape-scale distribution of alien plants and their threat to plant diversity. Conserv. Biol. 13:303–313.

**Komarov, V.L. 1963.** Flora of the U.S.S.R. Published for the Smithsonian Institution and the NSF by the Israel Program for Scientific Translations. Jerusalem, Israel.

**Maddox, D.M., A. Mayfield, and N.H. Poritz. 1985.** Distribution of yellow starthistle (*Centaurea solstitialis*) and Russian knapweed (*Centaurea repens*). Weed Sci. 33:315–327.

**Panetta, F.D. and J. Dodd. 1987.** Bioclimatic prediction of the potential distribution of skeleton weed *Chondrilla juncea* L. in western Australia. J. Australian Inst. Agricultural Sci. 53:11-16.

**Panetta, F.D. and N.D. Mitchell. 1991.** Homocline analysis and the prediction of weediness. Weed Res. 31:273–284.

**Roché, B.F., Jr. and C.T. Roché. 1991.** Identification, introduction, distribution, ecology and economics of *Centaurea* species. p. 275–291 *In:* L.F. James and J.O. Evans. (eds.) Noxious Range Weeds. Westview Press, Boulder, Colo.

**Roché, C.T. and B.F. Roché, Jr. 1988.** Distribution and amount of four knapweed (*Centaurea* L.) species in eastern Washington. Northwest Sci. 62:242–253.

**Rogers, C.F. 1928.** Canada thistle and Russian knapweed and their control. Colorado Exp. Sta. Bull. 348, Fort Collins, Colo.

**Selleck, G.W. 1964.** A competition study of *Cardaria* spp. and *Centaurea repens*. Proc. Seventh Brit. Weed Control Conf. 7:569–576.

**State of Colorado. 1996.** Colorado noxious weed act. Colorado House Bill 96-108, Denver, Colo.

**USDA Soil Conservation Service. 1991.** State Soil Geographic Data Base (STATSGO): Data Users Guide. Misc. Publ. No. 1492.

**USGS. 1990.** Digital Elevation Models, National Mapping Program Technical Instructions, Data Users Guide 5, Second Printing (Revised), Reston, Virg.

**Venables, W.N. and B.D. Ripley. 1994.** Modern applied statistics with S-plus. Springer-Verlag, New York, N.Y.

**Watson, A.K. 1980.** The biology of Canadian weeds. 43. *Acroptilon (Centaurea) repens* (L.) DC. Can. J. Plant Sci. 60:993–1004.