Hotspot Analysis Reveals Large Landscape Controls Over Cheatgrass (*Bromus tectorum*) Persistence Across Arid Landscapes

Bishop, Tara B.B.[1], Munson, S.[3], Gill, R.A.[1], Belnap, J.[2], St. Clair, S.B.[1]

[1]Brigham Young University, [2] Southwest Biological Science Center, USGS, [3] Colorado Plateau Research Station, USGS

INTRODUCTION

The invasion of exotic annual grasses during the last century has transformed plant habitats and communities worldwide. Cheatgrass (*Bromus tectorum*) is a winter annual grass that has invaded over 100 million acres of the western United States (Pellant and Hall, 1994. Pellant, 1996). Cheatgrass quickly utilizes available resources especially after a disturbance to the landscape. A major impact of invasion is the increased frequency in fires (D'Antonio and Vitousek, 1992). As cheatgrass is highly successful at invading open and disturbed landscapes at a rapid pace it increases the frequency and severity of fires in arid landscapes (Brooks, 2005). Cheatgrass' prolific seed production and flammability allows it to competitively exclude native plant species (Seabloom et al., 2003). The successful life strategy of cheatgrass gives a unique spectral image reflectance that can allow the use of remote sensing platforms to track and locate invasions.

Cheatgrass invasion is particularly worrisome in eastern and southern Utah as it spreads and degrades much of Utah's wildlands. Utah has 13 national parks and monuments with over 10 million visitors annually. Within those parks there are over 18 threatened and endangered species and pristine habitat for over 200 endemic plant species. With an economic benefit of over $725,00,000 annually (National Park Service, 2014) the increasing invasion of

cheatgrass puts all national parks at risk of altering valuable visitor experiences and economic benefit in the future.

Increasing invasion, and thus potential and actual fire frequency, also has serious ecological impacts as the native plants have a decreased ability to re-establish after a fire. This leads to the degradation of the native plant community as the cheatgrass continues to replace the native perennials and/or shrubs (Zouhar, 2003). This change in the native plant community can lead to negative impacts on the surrounding wildlife habitat and changes in the surrounding physical environment.

Restoration and rehabilitation of areas that have been invaded are a top priority of land managers. But large-scale surveying of the land is timely and can have high cost association. Using a geographic information system modeling (Hotspot Analysis; GIS, ESRI) with Detection of Early Season Invasive (DESI) software (Kokaly, 2011) landscape level analysis can be done of invasive annual grasses. Understanding landscape controls and the temporal dynamics of large, full scale invasions may be critical to controlling, managing and even preventing loss of natural habitat to the conversion of invasive grasslands. Our primary objectives to achieve this understanding are to (1) Identify areas that have spatially significant cheatgrass invasion; (2) Develop and interpret a statistical model that explains the landscape controls over the spatial and temporal distribution of cheatgrass and (3) Identify areas that are currently free of cheatgrass but are sensitive to cheatgrass expansion in the future.

METHODS

Detection of Early Season Invasives

The study sites, in collaboration with USGS, will be conducted at seven national parks and monuments: Arches National Park, Bryce Canyon National Park, Canyonlands National Park, Capitol Reef National Park, Dinosaur National Monument, Glen Canyon National Recreation Area, and Natural Bridges National Monument all located in the state of Utah.

The United States Geological Survey (USGS) developed a software algorithm that uses remote sensing data from Landsat TM/ETM imagery to detect locations and populations of cheatgrass called Detection of Early Season Invasives (DESI) (Kokaly, 2011).

Using ENVI (Exelis Visual Information Solutions) software the Normalized

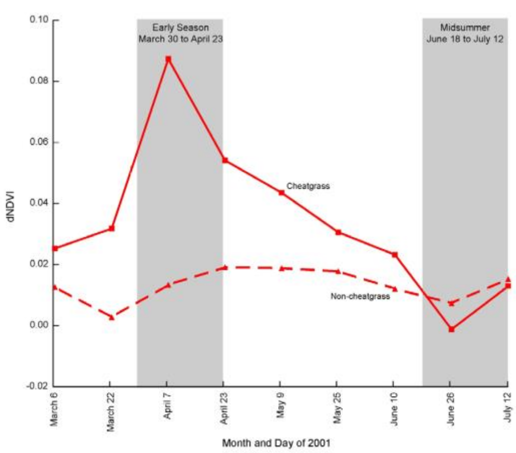

Figure 1: Seasonal trends of dNDVI for plots in Canyonlands National Park in 2001.

Difference Vegetation Index (NDVI) values for reflectance of red and near-infrared radiation by plants are extracted from the Landsat TM/ETM images. NDVI is taken for early spring and summer to capture the senescence of early season invasives. By taking the difference of NDVI (dNDVI) values in early spring and summer (Figure 1), and including masks for cloud cover and other climatic conditions, the software can detect locations for early seasons invasives, specifically cheatgrass. The image produced is a map with 30m x 30m pixels designating locations at which cheatgrass meets high and low thresholds. The thresholds are determined by examining the value (minimum dNDVI values) of a specific pixel and then the surrounding pixels and their corresponding value.

DESI images were produced for each study site for years 1999-2009 (Figure 2). Not all parks had a complete 10-year data set as some images were not acceptable for proper analysis due to cloud cover and other environmental factors.

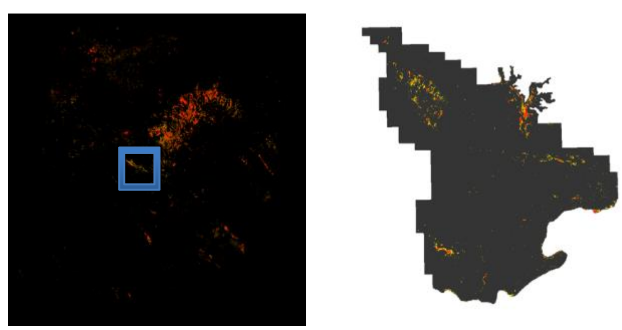

Figure 2: Left is a DESI output image for Landsat imagery encompassing central and southeast Utah. Right is Arches National Park DESI output (clipped from larger image). Red indicates the high threshold for cheatgrass growth and yellow indicates the low threshold for cheatgrass growth.

Producing Final DESI Image

Analysis of the DESI images required building models in GIS software, ArcMap 10x (ESRI, 2011). All of the DESI images for each individual management unit were overlain each other. Then using the Raster Calculator tool, syntax was derived to add all pixel values at each location together.

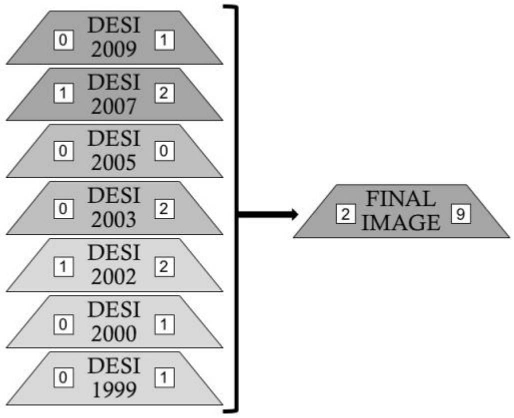

Figure 3: Final DESI image. All available years are added together using the raster calculator to produce one image for the entire park area where each point counts the individual year's presence of cheatgrass

(Figure 3). The end result was a raster layer where each pixel represented all years added together. Higher numbers then signify where cheatgrass perseveres and is there most years, whereas lower numbers indicate areas where cheatgrass is not present with much consistency.

Hotspot Analysis

Because cheatgrass is so widespread, it is important to be able to identify areas that may be central in the seed bank production. HotSpot analysis (ESRI) provides a means to statistically evaluate a DESI output image. Using ArcGIS 10x (ESRI) tool "Hotspot Analysis" gives an output feature of statistical analysis of spatial clustering in a point image. The final image of all combined DESI years was converted from raster to vector data. Where the centroid of each pixel becomes a point with the associated value. Hotspot Analysis calculates the Getis-Ord Gi* (pronounced G-i=star) (Getis and Ord 1992, Ord and Getis 1995) which evaluates the sum of value of an individual point of all

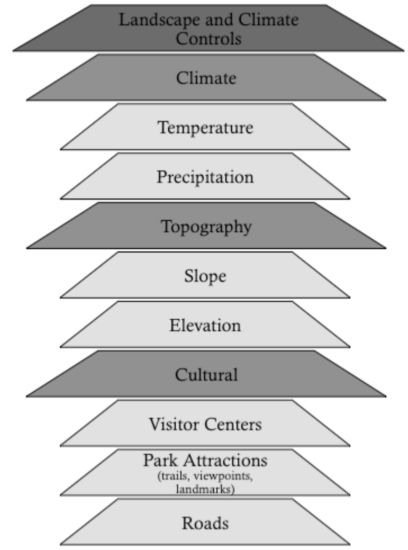

Figure 4: Data layers used for DESI output analysis acquired by remote sensing and satellite imagery

surrounding points in relation proportionally to the sum of all points. Z-scores and p-values are then calculated for each point. If a point has a resulting large z-score and points surrounding it also have a large z-score it will be significant spatial clustering called a hotspot. The larger the positive z-score the more intense the spatial clustering of high occurrence points it will be. High occurrence points represents persistent populations of cheatgrass. If a point has a resulting small negative z-score with significant p-value it falls in the category of coldspot, which is significant spatial clustering of low occurrence points. Low occurrence points represent populations of cheatgrass that have high inter-annual presence variability. If the z-score is close to zero it becomes statistically insignificant for spatial clustering.

Landscape and Climate Models

We initially began our work by focusing on Arches National Park, evaluating lags between precipitation in preceding seasons and DESI estimates of annual grass cover. We found inconclusive results, leading us to believe that there are other heavily weighing factors that will determine the locations and predict growth or decline in certain areas of the park. Factors that are currently being considered are climate, topographic, and cultural in nature (Figure 4).

Topographic data include: DEM (digital elevation model), slope, and soil texture and percent clay.

Digital Elevation Model (DEM) for elevation (USGS) has been collected from Utah Automated Geographic Reference Center (UT AGRC). Tiles were mosaicked using ArcMap 10x to encompass all areas of each park. Slope was calculated using the ArcMap 10x Slope Tool with the DEM layer.

Soil texture and percent clay data was downloaded from the NRCS Web Soil Survey (Soil Survey Staff, 2015).

Climate data include precipitation for the preceding fall DESI year. If the DESI output image was for 2003 then fall 2002 precipitation was used. Maximum and minimum temperatures for the preceding fall year were also included with the climate data. Climate data has been collected from PRISM climate datasets (PRISM Climate Group, 2004). PRISM data was resampled from a 90m x 90m pixel using a cubic convolution to match the 30m x 30m pixel size of the DESI output image.

The cultural data was gathered from existing GIS databases as well as digitizing trail maps and other sources gathered from the National Park Service. Trails, visitor centers, campgrounds and any other high traffic use areas were located and combined into one layer. A 100m buffer zone was created (ESRI) around all locations. This buffer zone is used as an error buffer as well as to account for growth that may occur near but not on these specific locations. For instance, cheatgrass would not grow on a road but on the shoulder or adjacent land to the road.

All data preparation was done using ArcMap 10x to ensure quality and compatibility of the multiple data layers. An example model that was used for these adjustments can be seen in Figure 5.

Transformations were required to ensure accuracy of the data for spatial and statistical analysis. Not all databases were found to be in the same datum or geographic coordinate system. Once the data was aligned, extrapolation was done using Multi-Value to Point Tool to build a statistical model that explains the control over the spatial and temporal distribution of cheatgrass. This statistical modeling and analysis will be done using Program R (R Core Team, 2012).

Statistical Modeling

PCA

To identify whether parks could be grouped together for easier analysis, general characteristics for precipitation, temperature, elevation, slope, and soil characteristics were pooled for each park. Mean fall temperature was left out of the principal component analysis (Pearson's correlation >.85) due to uneven loading potential. Using a benchmark cumulative Eigenvalue of 70% the parks were categorized into like groups. This statistical tool was performed using JMP 13 pro.

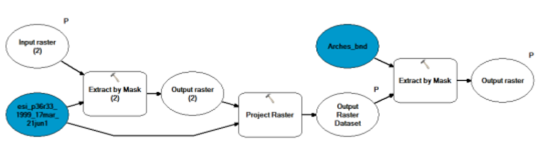

Figure 5: One of the models used to adjust and transform the various data layers to all align with the DESI output images. Input raster (2) will be target data layer for manipulation. Output raster will be the new data transformed and clipped to

Stepwise Discriminant Analysis

To test which biophysical factors were significant in explaining cheatgrass occurrence and persistence from the Hotspot analysis a stepwise discriminant analysis was performed using SAS [CIT]. Data exploration and preparation was done using methods from ZURR TEXT BOOK [cit] Due to the incredibly low proportion of coldspots (Table 1) causing a violation of the general 9:1 ratios assumption, coldspots were dropped from the analysis.

National Park Unit	Park Characteristics				Cheatgrass % of Total Park Area	Hotspot Analysis of Total Cheatgrass		
	Average annual precipitation (mm)	Average Tmax C	Average Tmin C	Elevation Range (m)		Hotspot	Ephemeral	Coldspot
Arches National Park	209.6	22.2	6.3		2.6%	10.3%	89.7%	0.0%
Bryce Canyon National Park	369.5	13.3	-1.9		1.5%	2.6%	97.4%	0.0%
Canyonlands National Park	212.5	19.0	5.5		5.9%	12.1%	87.9%	0.0%
Captiol Reef National Park	189.0	18.6	5.5		0.8%	2.7%	97.3%	0.0%
Dinosaur National Monument	227.8	16.3	2.7		24.8%	16.8%	82.9%	0.3%
Glen Canyon National Recreation Area	181.7	21.8	9.5		5.7%	11.3%	86.6%	2.1%
Natural Bridges National Monument	299.6	17.2	3.7		14.8%	13.0%	86.9%	0.0%
All Parks	241.4	18.3	4.5		3.8%	12.9%	85.9%	0.04%
AN Park Group	218.5	19.8	6.1		5.0%	11.3%	87.2%	1.50%
BD Park Group	298.6	14.8	0.4		2.5%	16.6%	83.1%	0.13%

Table 1: Summarizing park characteristics and results of hotspot analysis and coverage of cheatgrass in each park unit and park group

Mean fall temperature was also dropped due to high correlation to mean fall precipitation, mean winter precipitation and temperature, and mean spring precipitation (Pearson's correlation value >.9). The decision to drop mean fall temperature rather than the other climate variables was because of studies showing fall precipitation to largely affect *Bromus spp.* growth and reproduction [1] as well as the implication of winter temperature in seedling survival and the effect of competition with native plants [2]. SPEI data will hopefully alleviate any problems with dropping fall temperature as SPEI takes into account the temperature and precipitation to create the index. To avoid violating the assumption of independence we performed a repeated simulation of randomly selecting 1000 points from each Hotspot analysis category and ran the stepwise discriminant analysis 1000 times. The order in which variables were pulled in for the discriminant functions were recorded in summary tables to identify which biophysical attributes best discriminated against hotspot categories in a weighted frequency table.

Weighted frequency was calculated by taking into account the order in the discriminant function (1^{st}, 2^{nd}, 3^{rd}, etc) and how frequent the variable was brought in at in that order. This was done for all parks combined and then for each park group categorized by the PCA. Once those biophysical variables are identified in the weighted frequency table, one discriminant function was made for each grouping.

RESULTS

PCA

Using a scree plot and cumulative Eigenvalues, two components were used in determining park groups. Component one (Eigenvalue=3.1033 and 51.293%) largely consisted of mean fall precipitation, mean winter temperature and precipitation, and spring precipitation, and DEM with fairly equal contribution, while the second component (Eigenvalue=1.5862 and 19.828%) largely consisted of the soil characteristics percent clay and sand and depth of plant water supply (cm) with equal spread. Component one clustered the parks into two main groups which we have labelled the AN group (Arches, Canyonlands, Capitol Reef, Glen Canyon, and Natural Bridges) and the BD group (Bryce Canyon and Dinosaur).

Hotspot Analysis

Hotspot analysis has clearly shown areas of cheatgrass that are occurring at high density and are spatially significant. The only park to have Coldspots (spatially significant low occurrence points) was Dinosaur National Monument (see Figure 6).

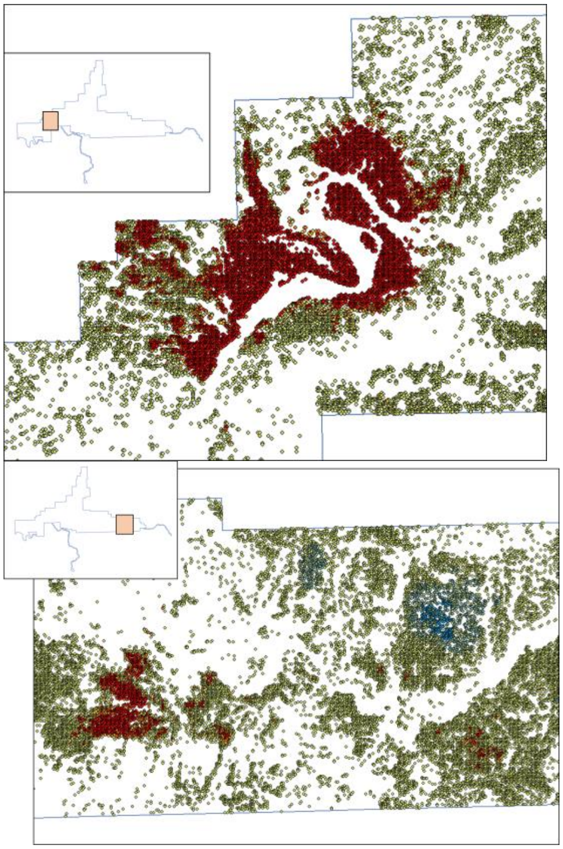

Figure 6: Dinosaur Naitonal Monument. Top: Hotspot analysis showing areas of spatially significant high occurrence cheatgrass growth (red), tan/green color are spatially insignificant. Bottom: Area of DNM showing hotspots (red) and coldspots (blue; spatially significant low occurring cheatgrass growth

Stepwise analysis

All biophysical variables (Table 2) used were deemed significant in the discriminant analysis ($p < 0.001$).

	Weighted frequency percent by park group		
	All Parks	AN Group	BD group
Aspect	0.9%	3.3%	0.4%
Depth of plant available water (cm)	41.5%	8.0%	11.3%
Distance to human populated area	0.5%	3.8%	3.6%
Distance to park boundary	0.3%	5.0%	0.8%
Elevation (DEM)	22.1%	23.5%	4.9%
Mean fall precipitation (mm)	1.7%	1.2%	6.8%
Mean spring precipitation (mm)	7.9%	11.4%	3.1%
Mean winter temperature (C)	12.9%	2.5%	25.4%
Mean SPEI	0.4%	1.2%	4.9%
Percent clay in top 20 cm of soil	3.4%	35.1%	3.7%
Percent sand in top 20 cm of soil	0.8%	2.8%	5.3%
Slope	7.6%	2.1%	29.8%

Table 2: The listed weighted frequency percent indicate the importance a certain variable has in discriminating between the analysis groups of hotspots, ephemeral populations, and areas with no cheatgrass.

However, some variables came in consistently as the most heavily weighted factors. To give proper weight to what place a variable was pulled into the discriminant function, a weighted frequency table was created showing the results of the repeated measures stepwise discriminant analysis (Table 2). For all parks combined elevation and depth of plant available water supply were the two most heavily weighted variables for significance in discriminant functions followed by mean winter temperature and spring precipitation. Using the weighted values, a final discriminant function was made for hypothesis testing. The final discriminant function for all parks combined used five variables in the following order: depth of plant available water, elevation, mean winter temperature, mean spring precipitation and slope, and was statistically significant ($\Lambda = 0.888755$, $F = 121.36$, $p < 0.001$). When the parks were split up into their respective groups, slope and mean winter temperature were most important for the BD group, followed by mean fall precipitation and depth of plant available water supply where a steady decrease in weighted frequency can be seen. The final discriminant function for the BD group used seven variables in the following order: slope, mean winter temperature, depth of plant available water, mean fall precipitation, percent sand in the top 20 cm of soil, mean SPEI, elevation, and was statistically significant ($\Lambda = 0.73112$, $F = 32$, $p < .001$). For the AN group, percent of clay content in the top 20 cm of soil and elevation were the most important. A sharp decrease in weighted frequency then leads to depth of water supply and mean spring precipitation being brought into the functions. The final discriminant function for the AN group used four variables: percent clay, elevation, mean spring precipitation, depth of plant available water, and was statistically significant ($\Lambda = 0.8925$, $F = 331$, $p < 0.001$).

DISCUSSION

Cheatgrass is highly dynamic and temporally variable from year to year [3]. Based on preliminary results we know that there are other factors other than preceding fall precipitation and clay content in soils that will determine whether cheatgrass will become established or not.

Visualization of the hotspot maps along with other topographic and cultural data show patterns across the landscape. Preliminary results showed that distance to human features is negatively correlated with hotspot Z score. However, this was not considered one of the most important variables in the discriminant analysis. Previous work in trying to identify core populations of cheatgrass were insignificant (based on this visual validation it was determined that the core populations were in fact, ecologically irrelevant) thus showing the largely dynamic nature of annual invasive grasses. It is important that this large landscape level work be ecologically relevant as the primary goal of this work is to be useful in land management and conservation goals. There was no spatial clustering of the core population pixels to indicate that there was a large scale invasion that is well established enough to be present every year within the datasets. Hotspot analysis allowed us to analyze spatially significant areas of cheatgrass persistence rather than continual presence.

Continuing research and analysis is being done to define criteria for areas that are sensitive to and conditions that will promote cheatgrass expansion. This information will be used to then identify what could be considered sensitive but cheatgrass has not yet established. Once the research has been completed, this set of criteria will be used to model as control factors that indicate locales that either are sensitive to or promote the invasion of cheatgrass.

Characterizing conditions and locations of cheatgrass populations will give land managers insight into areas that should be of high priority for conservation. It will also give NPS an understanding if the park has been fully invaded or if there are areas that meet the criteria for invasion but have not yet experience large-scale cheatgrass growth. If the factors that control cheatgrass invasion are controllable then these areas would mostly likely set apart has high priority for conservation. Understanding this system will also make it possible to focus the restorative efforts on areas that have an increased likelihood of success in those endeavors.

The novelty of this work is it will give the scientific community, including land managers, the ability to monitor invasions at an unprecedented landscape scale using remote sensing technology that is available at minor cost, reducing time and overall processing cost.

Since it is known that annual alien grass species contribute to increasing fire cycles and is an aggressive invader it will be crucial to maintain the integrity of the wildlands in Utah on a large scale. This technology of using free open source Landsat imagery will allow for this continued large scale monitoring to occur.

All data collected for this project utilizes existing data and remote sensing platforms and is available in free, open-source databases, reducing the costs directly to land managers. It reduces the need for large field crews to be extensively sampling remote areas and reduces human bias in the collection process based on conditions of the landscape (Peterson, 2008).

References

1. Bradley, B.A. and J.F. Mustard, *Comparison of phenology trends by land cover class: a case study in the Great Basin, USA.* Global Change Biology, 2008. **14**(2): p. 334-346.
2. Brooks, M. L., D'Antonio, C. M., Richardson, D. M., Grace, J. B., Keeley, J. E., DiTomaso, J. M., ... & Pyke, D. (2004). Effects of invasive alien plants on fire regimes. *BioScience*, *54*(7), 677-688.
3. D'Antonio, C.M., Vitousek, P. (1992). Biological invasions by exotic grasses, the grass/firecycle, and global change. *Annual Review of Ecological Systems.* 25, 63-87
4. ESRI 2011. ArcGIS Desktop: Release 10.3. Redlands, CA: Environmental Systems Research Institute.
5. Exelis Visual Information Solutions, Boulder, Colorado
6. Horn, K.J., T.B.B. Bishop, and S.B. St. Clair, *Precipitation timing and soil heterogeneity regulate the growth and seed production of the invasive grass red brome.* Biological Invasions, 2017. **19**(4): p. 1339-1350.
7. Kokaly, R.F..(2011). Detecting cheatgrass on the Colorado Plateau using Landsat data: A tutorial for the DESI software: *U.S. Geological Survey Open-File Report 2010–1327*, 88
8. Melgoza, G., R.S. Nowak, and R.J. Tausch, *SOIL-WATER EXPLOITATION AFTER FIRE - COMPETITION BETWEEN BROMUS-TECTORUM (CHEATGRASS) AND 2 NATIVE SPECIES.* Oecologia, 1990. **83**(1): p. 7-13.
9. NASA (2014) Strategic Plan. http://www.nasa.gov/sites/default/files/files/FY2014_NASA_SP_508c.pdf
10. National Park Service (2014) Working with Utah: By the numbers.
11. Pellant, M.. (1996) Cheatgrass: The invader that won the west. *Bureau of Land Management Interior Columbia Basin Ecosystem Management Project*
12. Peterson, E. B. (2008). Invasion of the exotic grasses: mapping their progression via satellite. *Shrublands under fire: disturbance and recovery in a changing world. Usda Forest Service Rocky Mountain research station proceedings*, *52*, 33-39.
13. PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu, created 4 Feb 2004.
14. R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/
15. Seabloom, E.W., Harpole, W.S., Reichman, O.J., Tilman, D.. (2003). Invasion, competitive dominance and resources use and native California grassland species. *Proceedings of the National Academy of Sciences of the United States of America.* 100(23), 13384-13389
16. Sieberg, M., Howery, L., Thwaits, A.. (2007) Western rangeland weeds. *Learning Modules; University of Arizona College of Agriculture and Life Sciences.*
17. Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online at http://websoilsurvey.nrcs.usda.gov/. Accessed [07/30/2015].
18. Zouhar, K. 2003. Bromus tectorum. In: Fire Effects Information System, [Online]. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory (Producer). Available: http://www.fs.fed.us/database/feis/ [2015, April 20].