# Interim Progress Report:
# A New Time Series Remote Sensing Approach to Mapping Fine Fuels in Sonoran Desert Ecosystems

Project ID: 10-1-04-7
September 30, 2011

Steven E. Sesnie, Ph.D.
U.S. Fish & Wildlife Service
500 Gold Avenue SW, Rm 4127
Albuquerque, NM 87102
(505) 248-6631
Steven_Sesnie@fws.gov

## Contents

## I. Project objectives

The objectives of this project are to 1) develop an efficient means to measure Sonoran Desert fine fuels on the ground and 2) construct and validate time series and remote sensing-based models capable of repeatedly mapping seasonal and annual fine fuel production.

This interim report describes project activities and accomplishments during the performance period since project initiation on August 1st, 2010 to September 31st 2011.

A minor change to the originally proposed field study was determined necessary to obtain valid estimates of native and non-native herbaceous plant composition and abundance for fine fuel and vegetation modeling. Plot sampling was modified to improve biomass data collection and scalability with available remotely sensed data. A stratified plot and subplot design, described below, was elected to allow for vegetation modeling using both Landsat Thematic Mapper (TM) and Moderate Resolution Imaging Spectroradiometer (MODIS) image pixels of 30m (0.0009 $km^2$) and 250m (0.0625 $km^2$) respectively.

PI Sesnie started a new position with US Fish and Wildlife Service Southwest Region Office in Albuquerque, New Mexico in June of 2011. He remains project PI as an affiliate faculty member in the School of Earth Sciences and Environmental Sustainability at Northern Arizona University.

## II. Project progress

### a. Workshop I. Western Wildland Environmental Threat Assessment Center (WWETAC) ArcFuels and fire modeling

To initiate the fine fuels mapping and modeling project, our Bureau of Land Management (BLM) cooperator Erica Falkner, met with project PIs for a three day fire modeling workshop held by the NAU Lab of Landscape Ecology and Conservation in Flagstaff, AZ September 27th – 29th, 2010 (**Appendix A**). Fire modeling experts Nicole Valiant (WWETAC) and Lauren Miller (Deschutes National Forest) were principal instructors of fire modeling techniques using the ArcFuels fire model interface. The workshop was also used to plan for the preliminary 2010 field data collection and spring 2011 data collection on BLM and other lands in the study area. The workshop facilitated technical training in new fire modeling techniques, data and tools that will be used for other aspects of this project and matched Department of Defense Strategic Environmental Research and Development Program (DoD-SERDP) funded project "Integrated Spatial Models of Non-native Plant Invasion, Fire Risk, and Wildlife Habitat to Support Conservation on Military and Adjacent Lands in the Sonoran Desert". The DoD-SERDP project will utilize biomass model estimates to select custom fuel models appropriate for the Sonora study area.

b. **Preliminary field study**

A principal objective of this research is to investigate the efficacy MODIS and TM satellite image time series data to estimate and map fine fuels in the Sonoran Desert. In addition, we seek to determine the relative contribution of native and non-native herbaceous plants to fine fuel accumulation that occur seasonally, annually and internannually. Each sensor has different spatial and temporal characteristics important to estimating vegetation parameters for desert ecosystems. Therefore, a preliminary field study was conducted to develop a standard plot design and data collection protocols consistent with the spatial resolution of MODIS (250m pixels) and TM (30m pixels) image data. The preliminary study also assessed field equipment and techniques being considered for obtaining plant cover and above ground biomass measurements across a diversity of site conditions. A two-week field study took place during early October of 2010 within Sonoran Desert uplands, thornscrub and creosote/white bursage vegetation with known populations of invasive annual and perennial grasses on BLM, National Park Service (NPS) and US Forest Service (USFS) lands close to Tucson, AZ.

An evaluation of field measurement equipment and time costs resulted in a stratified plot design used to co-locate plots and subplots within MODIS and Landsat TM pixels for biomass model development (**Fig. 1A**). Plot measurements subsample across a MODIS pixel (7%) and cover the majority of a Landsat pixel (**Fig. 1B**). The preliminary field study also helped to streamline plant measurements and biomass data collection techniques. The initially proposed use of a laser point frame and field spectrometer proved cumbersome on rugged terrain with highly varied vegetation and life forms consisting of grasses, forbs, shrubs, trees and cacti. A diversity of vegetation height also precluded the use of infrared photography as a viable option for estimating herbaceous plant cover and the above ground fraction of biomass production from annual and perennial grasses and forbs. However, spectrometer measurements were taken from a subset of plots to estimate the potential for fuel bed and biomass characterization using site -level spectral reflectance measurements. Spectrometer measurement protocols and methods are outlined below.

The above efforts successfully identified a practical multi-scaled sampling design that used a point intercept method to measure vegetation composition, cover and height at 5-m intervals within five 25 x 25m subplots (**Fig.1B**). We adopted point intercept method implemented across the entire TM pixel area from five locations within a MODIS pixel as a rapid and efficient method to sample Sonoran vegetation, consistent with the two senor types. A crew of two individuals could measure approximately two plots consisting of 10 sub-plots per day, depending on travel time between plot locations.

At 25 point-intercept locations within each subplot, we recorded the presence and height of individual plants (at the species level), and dominant substrate type (i.e. litter, rock, sand, biological crust). A streamlined approach was adopted to collect herbaceous biomass of native and non-native plants within a 0.33 $m^2$ circular plot at 9 of the 25 point intercepts to estimate biomass from MODIS and TM pixel data. Biomass samples collected within subplots were placed into separate bags containing invasive plants from the current year's production, target invasive plants from the previous year, native plants from the current year, and native plants from the previous year. Field separated biomass also facilitated plant drying, weighing, and data entry for each category.

A spectrometer sampling protocol was developed to collect reflectance data using an ASD Inc. FieldSpec Max3 (350nm - 2,500nm range) from each of the 9 point intercepts on a subplot where biomass samples were collected. A pistol grip and fiberoptic cable assembly were mounted on a specialized non-reflective (black) pole and leveling device to obtain un-shadowed spectral reflectance measurements from each 0.33 $m^2$ circular biomass collection point. To measure reflectance from only the clipped area, spectrometer measurements were recorded from 1.3 m above the ground with the bare fiber cable end equivalent to a 25º field of view. All spectral measurements were taken prior to biomass clipping and calibrated to field illumination conditions with a white reference spectralon disk. Spectral samples were collected at point locations averaging 20 measurements for each of five separate spectra, taken in less than five seconds per each intercept once equipment and foreoptics were in position.

### c. Field study design, implementation, and data collection

Full deployment of the field study within a 103,000-$km^2$ consolidated study area occurred during the period of peak productivity for annual perennial grasses and forbs in the Sonoran Desert between January and April of 2011 (**Fig. 2**).

Prior to data collection we identified five target non-native invasive annual grasses and forbs associated with increased fire occurrence in Sonoran Desert Ecosystems: African buffelgrass (*Pennisetum ciliare*), red brome (*Bromus rubens*), Sahara mustard (*Brassica tournefortii*), Mediterranean grass (*Schismus* spp.), and arugula (*Eruca vesicaria sativa*). As part of a stratified random design to determine prospective sampling locations, we used existing regional occurrence data for each target species and the Maxent software package (Phillips et al. 2006) to develop probabilistic models and maps of habitat suitability for the study area (**Fig. 3A, B**). These models considered ≤ 14 covariates, including topographic, edaphic, and climatic factors. In order to determine and spatially balance plot locations for each species, we coupled the 90$^{th}$ percentile of predicted habitat suitability with slope conditions (< 30 degrees), road proximity (250-2,000 m from roads), and land accessibility (private, state, and most tribal lands were excluded). Each plot was spatially registered to a MODIS image pixel (250 m; 0.0625 $Km^2$) that encompassed five subplots, each matched with an interior Landsat TM image pixel (30 m; 0.0009 $Km^2$) (**Fig. 1A, B**). A hierarchical sampling framework was used to model invasive plant occurrence and biomass with each sensor type and image time series described below.

Fine fuel production in the Sonoran Desert is an ephemeral and varied event creating logistical challenges for field sampling. Plot establishment and measurement was conducted by 12 temporary field technicians hired for this and the DoD-SERDP project, with assistance from project PI, research associates and NAU faculty. All technicians were experienced botanists and biologist who were trained to implement the field sampling protocol in a consistent fashion. Multiple sampling teams (6) were be rotated into the field at 8-day intervals to cover 8 geographic units within study area. A minimum of 4 field crews were in the field conducting sampling during an 8-day period in order to cover as much of the study area as feasible during the spring plant production period. Sampling during the spring production period was essential to measuring and identifying native and non-native plants in addition to safely working in hot desert environments.

We established 239 plots, 1,174 subplots, and 29,350 point-intercept locations across the study area during the 2011 field season (**Fig. 2**). A total of 158 plots (66%) occurred on BLM lands and an additional 81 plots (34%) were measured on land jurisdictions such as military land, Indian reservations, US Forest Service, National Wildlife Refuge, and State and National Parks. To maximize efficiency in detecting targeted invasive plants (e.g., drawing on methods described by Sesnie et al. in press), we re-evaluated and prioritized our plot locations in April 2011, and focused efforts on sandy and loose textured soils more likely to support target invasive plants. Sandy soils were discriminated from other substrates in Landsat TM imagery based on field-measured reflectance of sand (see also below). With these approaches, we detected at least one of our target non-native species in 173 plots (642 subplots). Mediterranean grass (plot-level $n$ = 130) and Sahara mustard ($n$ = 105) were detected most frequently (**Table 1**). Detections of red brome ($n$ = 15), arugula ($n$ = 9), and African buffelgrass ($n$ = 3) were less common.

High resolution (350nm to 2,500nm) field spectrometer measurements were collected from a total of 10 plots and 50 subplots. These resulted in 315 point intercepts with biomass and spectral reflectance data. The lower number of point intercepts with reflectance data is a result of discarding points with shading from surrounding or overhead vegetation. Point spectral data were consolidated into a single average measurement for each intercept location and analyzed using Unscrambler X (Camo 2010) multivariate statistical software package V. 10.0.1 and partial least squares regression (PLSR), that is well suited to regression analysis with high resolution spectral reflectance data and multicollinearity among numerous predictor variables.

### d. Data processing and preliminary analyses

***Field data processing***

A Microsoft Access data entry form was developed to facilitate data entry and database development. All vegetation data collected in the field was entered into the project database and summarized by target invasive species occurrence and percent cover for native and non-native plants as a measure of abundance. Biomass samples collected in the field were first oven dried and then weighed to record biomass for each native and non-native plant category and entered into the database. As anticipated, biomass data collection, processing, data entry and summarization was the most labor intensive data to collect. Most phases of biomass data development have been accomplished and are currently being summarized by category to model fuel bed composition and structure. Preliminary models and model comparisons will be developed by the end of November.

***Satellite image processing***

To develop preliminary occurrence and abundance model runs for target invasive plants, we identified two image tiles in the Landsat World Reference System 2 (WRS2: path 38/row 37 and path 37/row 37 overlapping a majority of 2011 field plots. We acquired and processed 208 cloud-free TM scenes (Jan 1, 2000 – Feb 28, 2011) within these two path/rows acquired from the USGS Global Visualization Viewer (http://glovis.usgs.gov/). All seven TM bands, including visible (blue, green, and red corresponding to bands 1-3), near-infrared (band 4), shortwave-infrared (bands 5 and 7), and a thermal band (band 6) were used and resampled to a 30-m resolution. For each scene, we used custom-developed and automated routines to calculate two vegetation indices and a water index. Vegetation indices used were Normalized Difference Vegetation Index (NDVI) and near-infrared scaled NDVI (NDVI/NIR), where NIR is the reflected radiance in the near-infrared region. A Normalized Water Difference Index ((NIR-SWIR)/(NIR+SWIR)) was also used, where SWIR represents the reflected radiance in the short-wave infrared wavelength region (TM band 5). We considered that the changes in plant phenology induced by invasive species would also be reflected in the soil moisture (Gao 1996). All indices were based on the raw (i.e., digital number) pixel values. In addition, a principal components rotation based on the correlation matrix was applied to Landsat TM bands and the first three principal components were used in our analysis.

We also acquired all MODIS 16-day NDVI composites images (MOD13Q1, version v05; Jan 1, 2000 – Jan 31, 2011) overlapping the study area. These composites included 20 MODIS tiles from 2000, 23 tiles in each year from 2001 to 2010, and five scenes from 2011. The MODIS composites were re-projected to two coordinate systems that matched the Landsat TM scenes using an automated processing routine.

***Plant phenology metrics***

Time series MODIS and TM data were used to derive an initial set of satellite image-based metrics to characterize plant phenology within the study area. MODIS-derived variables included 16-day NDVI composite scenes from 2000 to 2011 and NDVI derivatives such as maximum NDVI values, date of maximum NDVI and the trend in NDVI over one, two, and three periods (roughly 16, 32, and 48-days). Trend, calculated as the change in NDVI divided by the number of days between dates, reflected the rate of increase or decrease in greenness for a given imagery pixel. We calculated the trend both as an absolute change (change in NDVI over time) and relative change (change in NDVI over time in proportion to NDVI of the earlier date). For TM imagery, we calculated the maximum spring and fall NDVI and the date at which they occurred. In addition to the single year spring and fall time series metrics, we calculated the mean, maximum, minimum, and standard deviation of these maxima over the image time series.

In total, we calculated 108 phenology metrics for MODIS and 80 metrics each for two Landsat tiles that intersected our study area, based on 104 Landsat TM scenes for each tile (path 37/row 37 and path 38/row 37).

***Preliminary models***

We selected two invasive species, *B. tournefortii* and *Schismus* spp., to focus our preliminary distribution and abundance modeling efforts using Random Forest classification and regression trees (Breiman 2001). Random Forest trees generate robust categorical and continuous model predictions that can enhance conditional relationships between predictor and response variables (Sesnie et al. 2008, 2010). Random Forest models allowed us to evaluate the importance of time series phenology data derived from MODIS and Landsat TM spectral bands, NDVI, and NDWI. We used 1,512 phenology-based covariates for MODIS-level model predictions and 1,432 covariates for both sets of TM-level models. Multiple ($n$ = 1,500) classification trees were grown and aggregated to make model predictions and estimate variable importance. To determine spatial and temporal relationships between our target invasive species and model covariates, we used a Monte Carlo approach to estimate the performance of binary (presence/absence) MODIS-based models at using randomly selected plots increasing geographic distances across the study area. We refer to these models as ‘local-scale’ models. The local model analysis was performed using target species presence data and employed a binary modeling approach fitted to MODIS-based covariates. Models developed using multiple subsamples of the data were used to estimate how physiographic differences may impact plant phenology and model prediction error. We refer to these models as ‘regional-scale’ models.

***Preliminary model performance***

We initially identified (*a priori*) 80% accuracy as a performance metric (i.e., benchmark) for evaluating our models, as per Congalton and Green (1999). We identified additional performance metrics during the course of this project because any single metric may simultaneously highlight a single aspect of model performance while ignoring another. Therefore, we calculated an additional set of metrics for each continuous and binary model (Jensen 2005). For binary models, we calculated classification error rates, Cohen’s kappa, overall accuracy, and the true positive rate (TPR). Cohen’s kappa is a “chance corrected” measure of classification accuracy between two or more classes that is less sensitive to unevenness in the proportion of samples in each class than the basic accuracy calculation (Congalton and Green 1999). Kappa measurements above 0.4 have been considered fair to good while kappa statistics above 0.75 have been considered almost perfect (Fleiss 1981). For continuous models, we calculated the percentage of variance explained for target species abundance as well as the area under the receiver operating characteristic curve (AUC). The AUC is a diagnostic accuracy assessment for continuous models that scales between 0.5 (no better than random) to 1.0 (perfect discrimination). The random forest algorithm provides robust error estimation for both class and continuous variables. Variance explained is calculated as the mean $R^2$ value of 'all trees in the forest' based on data that are set aside. Model validation statistics for random forests were calculated by extracting multiple bootstrapped training samples from plot or subplot data and leaving one-third of the sample data aside for error testing at each model iteration (Breiman 2001). Error and variance statistics were then aggregated and averaged as a final step. Each performance metric provided a complimentary indicator of the potential to discriminate target non-native from native vegetation. Initial performance metrics and thresholds are listed in **Table 2**.

Performance metrics for the regional-scale models are summarized in **Tables 3** and **4**, and the performance of local models is presented in **Figure 4** (MODIS). Classification accuracies of the binary models were high, exceeding our benchmark of 80%, as evidenced by accuracies above 86% for MODIS-based models and above 92% for Landsat-based models (**Table 2**). In contrast, overall classification accuracy was poor to fair based on Cohen's kappa results. Cohen's kappa was 0.33 for the *B. tournefortii* binary MODIS model and the *S.* spp. binary Landsat model (path 37, row 37). The best models according to Cohen's kappa were the regional binary MODIS model for *S.* spp. (κ=0.59) and the binary TM model of *B. tournefortii* for path 38, row 37 (κ=0.60). While the regional MODIS-based *S.* spp. model was better than *B. tournefortii* models, the local-scale Landsat TM-based *B. tournefortii* models were better than *S.* spp. models for both Landsat tiles. Inspection of the TPRs for each of these models reveals that the models with higher κ also had the highest TPRs and better predicted the presence of the target invasive species. Error rate and accuracy did not always reflect the same model ranking as κ and TPR. We suspect that our detection rate was low because data collection occurred during a below-average precipitation year, discussed below, which can artificially inflate the number of falsely predicted absences.

Continuous region-scale models of target species abundance had good discriminatory power with AUC values in the range of 0.74 to 0.82 (MODIS-based models) and 0.66 to 0.80 (Landsat-based models). However, these models poorly characterized the abundance of our target invasive species as evidenced by low values for variance explained: 8% to 9% (MODIS) and 23% to 36% (Landsat).

For local-scale analyses, we report only kappa to demonstrate the strength of spatial proximity in describing the presence of our invasive target species. Local-scale binary models demonstrated a distinct relationship between the size of the neighborhood and quality of the model (**Fig. 4**). At very small neighborhoods (e.g., < 25 neighbors), model prediction was inconsistent and κ values varied widely as a result. For both *B. tournefortii* and *Schismus* spp., predictions were good at smaller neighborhood sizes (e.g., 25-35 neighbors), and the 95% confidence interval of κ exceeded 0.4. However, mean κ decreased and was not significantly greater than 0.4 for progressively larger neighborhoods, suggesting that seasonal weather and invasive plant phenology patterns may differ across increasingly larger spatial extents.

***Field spectrometer data analysis***

Combined old and new herbaceous biomass (total biomass) was used as the principle response variable and spectral reflectance values as predictor variables with PLSR models. All spectral values highly impacted by water vapor absorption and suspended solids such as dust, were removed prior to analysis. Nevertheless, field spectrometer measurements were impacted by low biomass productivity on nearly all sites. Herbaceous biomass collected from plots ranged from 0 to 110 g, with a majority the data points showing no herbaceous biomass (**Fig. 5**). As a result, very low variation was explained (29% maximum) by PLSR models with a maximum of 9 orthogonal factors and a maximum of 12% of the variance explained (6 factors) using 10 fold cross validation (**Fig. 6**). Most samples were dominated by surface material such as rock, sand, soil crusts or dry woody material. Further analyses will be conducted to explore narrow band (3 – 10nm) vegetation indices such as cellulose absorption indices (CAI) that are sensitive to plant litter, cellulose and lignin (Nagler et al. 2003).

Field spectral reflectance measurements proved to be valuable for substrate modeling and spectral mixture analysis of TM imagery to map general substrate categories such as sandy soil versus extrusive basalt which supports very limited plant production. Plots overlaying sandy soils were given priority during the later sampling period to increase the number of plots within areas likely to be colonized by target invasive plants. Substrate maps will also aid future sample site selection and modeling efforts.

## III. Conclusions & next steps

***Biomass modeling***

Biomass data summaries were finalized by the end of September 2011 to initiate modeling efforts. Biomass modeling will utilize similar methods and image data sources, but include vegetation indices sensitive to green and senesced vegetation such as the soil adjusted total vegetation index (SATVI) for investigating fine fuel accumulation from annual and perennial grasses and forbs (Marsett et al. 2006). We anticipate preliminary models to be developed by the end of November 2011, however further model testing and comparisons with alternative VI will be evaluated until January of 2012.

***Field sampling***

Although we identified the optimal sampling period and locations to detect and measure attributes for each of our five target plant species in the field, annual plant productivity was low across the study area in 2011. Therefore individual plants were small in stature and cover was low, or no herbaceous production was measured on a plot. Indeed, total precipitation in Yuma County from December, 2010, to April, 2011, was 37% below the previous ten year average (2000-2010) and 76% below the wettest year on record (2005) (WRCC 2011). Consequently, maximum spring NDVI in 2011 was 19% below the average peak NDVI calculated over the previous 10 years. Thus, we were unable to relate detections and measurements of each species to contemporary (i.e., 2011) images and associated spectral characteristics with high accuracy and precision; much of the signal we anticipated and hypothesized would be present was minimized by the noise of a dry year.

Our spring 2012 field campaign will be focused in areas with a greater abundance of target invasive species, as indicated by our 2011 data. We intend to prioritize and locate field plots in hotspots for red brome, arugula, and African buffelgrass in order to better model the distribution and abundance of these species. Based on our initial stratified random sampling effort, these species typically occur within the study area in sparse populations and very low abundances during a year with low rainfall. Relatively large stands of African buffelgrass and Sahara mustard, however, were been observed in the eastern and western portions of our study area.

Spectral mixture analysis will also be used to identify common geologic substrates for use in a stratified sampling design. For example, contiguous pixels with a high proportion of sand (e.g., > 75%) will provide additional strata for targeted sampling for *Brassica tournefortii*. Soil substrate data layers may also play a role as model covariates pending field verification.

***Satellite image processing***

The exclusion of phenology metrics from the best distribution models for *Schismus* spp. suggested that key features that distinguished these species from the surrounding vegetation may not be well characterized by TM imagery. Sesnie et al. (in press) concluded that there can be too few TM image-acquisition dates in a given year to adequately characterize plant phenology except at a very coarse level. MODIS 16-day VI, on the other hand, provided more consistent time series imagery for approximately 23 dates a year that give a bimonthly estimate of plant phenology. The 8-day gap-filled MODIS NDVI data and phenology products that are currently being developed by the Goddard Space Flight Center National for the North American Carbon Program (NACP; http://accweb.nascom.nasa.gov/index.html) could greatly enhance phenology characterization with as many as four image observations in a given month.

We continue to acquire and process MODIS and Landsat imagery and will continue to apply automated routines to efficiently build and maintain a phenological database that extends through our next sampling effort. These products and their derivatives (e.g., new 8- and 16-day MODIS NDVI composite images and other VI) also will be considered as covariates for the purposes of estimating plant biomass, invasion risk, and fire risk.

***Deriving temporal and spatial patterns of plant phenology***

To derive 11 additional seasonality parameters that represent the start, mid, end, peak, amplitude, and level of photosynthetic activity in the vegetation canopy, we have initiated an effort to use the time-series analysis program, Timesat (Jönsson and Eklundh 2004), in combination with our existing MODIS time-series data. We anticipate that additional Timesat phenology metrics, based on more refined model fitting functions, will shed more light on how phenological differences between our target invasive species and the surrounding vegetation are anticipated to increase model predictive power and assist in differentiating plant communities dominated by non-native vs. native species. We also anticipate that the start of season, rate of green-up, end of season, and rate of senescence from multiple years may be among the most discriminating variables for identifying phenological signatures of invasive species in the Sonoran Desert.

Related to this task, our recent efforts and published work (e.g., Sesnie et al. in press) indicate that time-series MODIS NDVI can be an exceptional product for evaluating plant (e.g., animal forage) phenology in our study area, particularly in areas of variable topography. Other vegetation indices may provide additional explanatory power for modeling low productivity years where soil background effects are substantial. Vegetation indices with a soil adjustment factor will be considered with future modeling efforts, particularly for target species less prone to establishing on steep rocky terrain. Future targeted field sampling efforts will incorporate soil substrate maps derived from spectral mixture analysis of high resolution field spectrometer data and satellite image classification with Landsat TM.

***Modeling the occurrence and distribution of target non-native invasive plants***

Over the coming months, we will focus on improving our models of invasive species distribution, abundance and biomass using the following four approaches:

1) Derive additional phenology metrics, including start of and end of season dates and NDVI values, which we anticipate will be more discriminating than NDVI amplitude. Our results thus far have found little difference in the peak greenness between native vegetation and highly invaded plots. Further analyses will seek to test for potential asynchronies in the timing of vegetation greenup and senescence of invasive species that may differ from un-invaded areas.
2) Investigate pattern of *Schismus* spp. with respect to SWIR and NDWI Landsat TM time series.
3) Incorporate community and soil substrate data into the model structure to account for variation in vegetation cover, abundance of native life forms, and site-level differences.

4) Evaluate alternative statistical model structures to Random Forest, including maximum likelihood-based and Bayesian hierarchical models developed for modeling low prevalence rates, and that may be more sensitive to detecting relationships with our phenological metrics and other remotely sensed covariates.

## IV. References

Breiman, L. 2001. Random Forests. Machine Learning 45:5–32.

Congalton, R.G. and K. Green. 1999. Assessing the Accuracy of Remotely Sensed Data: Principals and Practices. Lewis Publishers, Boca Raton. FL. Pp. 137

Fleiss, J.L., 1981. Statistical methods for rates and proportions (2nd ed.). New York: John Wiley.

Gao, B.C., 1996: NDWI - a normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment 58:257–266.

Jensen, J.R., 2005. Introductory Digital Image Processing: A Remote Sensing Perspective, 3rd edition. Upper Saddle River, NJ 07458. Pp. 495-515.

Jönsson, P. and L. Eklundh. 2004. TIMESAT - a program for analyzing time-series of satellite sensor data. Computers and Geosciences 30:833–845.

Marsett, R. C., Q. Jiaguo, P. Heilman, S.H. Biedenbender, M.C. Watson, S. Amer, M. Weltz, D. Goodrich, and R. Marsett. 2006. Remote sensing for grassland management in the arid southwest. Rangeland Ecology and Management 59:530-540.

Nagler, P.L., Y. Inoue, E.P. Glenn, A.L. Russ, and C.S.T. Daughtry. 2003. Cellulose absorption index (CAI) to quantify mixed soil-plant litter scenes. Remote Sensing of Environment 87: 310–325.

Phillips, S.J., R.P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231–259.

Sesnie, S.E., P. Gessler, B. Finegan, and S. Thessler. 2008. Integrating Landsat TM and SRTM-DEM derived variables with decision trees for habitat classification and change detection in complex neotropical environments. Remote Sensing of Environment 112:2145–2159.

Sesnie, S.E**.**, B. Finegan, P. Gessler, S. Thessler, Z. Ramos and A. Smith. 2010. The multispectral separability of Costa Rican rain forest types with Random Forest decision trees and Support Vector Machines. International Journal of Remote Sensing 31:2885–2909.

Sesnie, S.E., B.G. Dickson, S. Rosenstock, and J.M. Rundall. In press. A comparison of Landsat TM and MODIS vegetation indices for estimating forage phenology in desert bighorn sheep (*Ovis canadensis nelsoni*) habitat in the Sonoran Desert, USA. International Journal of Remote Sensing.

WRCC (2010) Arizona Climate Summaries. Western Regional Climate Center.
http://www.wrcc.dri.edu/summary/climsmaz.html. Accessed 7/1/2011.

## V. Tables

**Table 1**. Frequency of detection for target invasive species at three nested sampling levels.

| | Number of detections | | |
|---|---|---|---|
| Species | Plot | Subplot | Point intercept |
| *Schismuss* spp. (Mediterranean grass) | 130 | 488 | 17,657 |
| *Brassica tournefortii* (Sahara mustard) | 105 | 305 | 11,405 |
| *Bromus rubens* (red brome) | 15 | 52 | 2,253 |
| *Eruca vesicaria sativa* (arugula) | 9 | 24 | 1,175 |
| *Pennisetum ciliare* (African buffelgrass ) | 3 | 7 | 241 |

**Table 2**. Model performance metrics used to evaluate models for this Go/No-Go step.

| Performance metric | Model type | Benchmark threshold |
|---|---|---|
| Cohen's kappa (κ) | Binary | 0.40 |
| Accuracy | Binary | 80% |
| Error rate | Binary | 20% |
| Variance explained | Continuous | 60% |
| AUC | Continuous | 75% |
| True Positive Rate (TPR) | Continuous | 50% |

**Table 3**. Model performance metrics (ER = error rate, Acc = accuracy, κ = Cohen's kappa, TPR = true positive rate) for binary models of *Schismus* spp. and *Brassica tournefortii* abundance based on MODIS and Landsat TM data.

| | MODIS | | | | Landsat TM | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | path 37, row 37 | | | | path 38, row 37 | | | |
| Species | ER | Acc | κ | TPR | ER | Acc | κ | TPR | ER | Acc | κ | TPR |
| *S*. spp. | 15.4 | 86.2 | 0.587 | 52.9 | 6.7 | 98.2 | 0.327 | 21.1 | 7.5 | 92.4 | 0.449 | 36.2 |
| *B. tournefortii* | 10.8 | 89.2 | 0.330 | 22.2 | 3.1 | 97.0 | 0.446 | 31.4 | 5.1 | 95.0 | 0.599 | 50.0 |

## VI. Figures

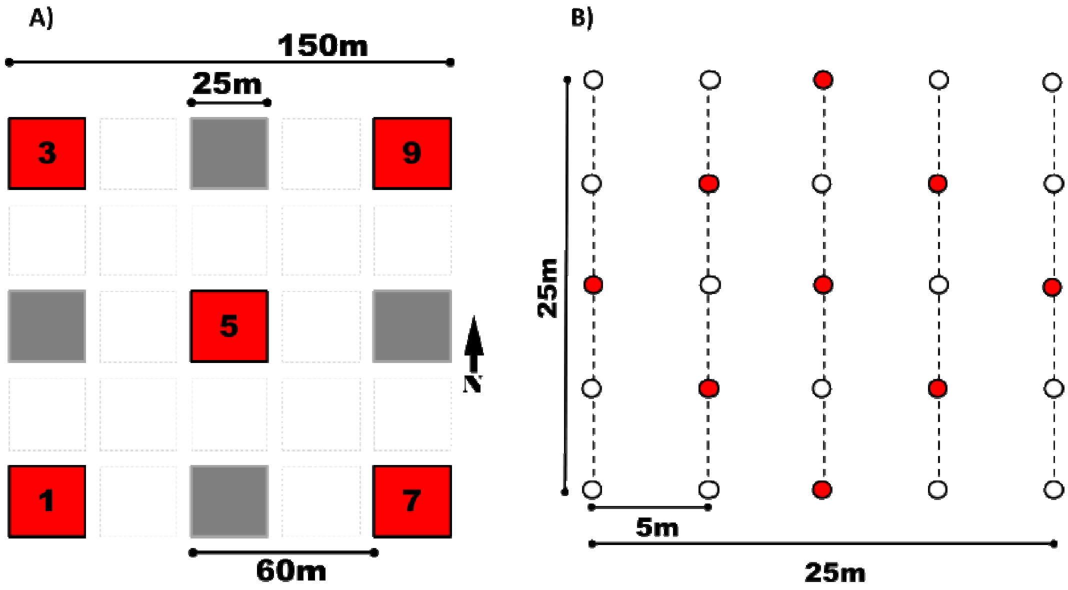



**Figure 1.** Revised field sampling design used to detect and measure herbaceous native and non-native invasive plants in the Sonoran Desert. Plots and subplots were developed for sampling across **A**) a 250-m MODIS image pixel that includes five nested 25-m subplots (Red) and **B**) a Landsat image pixel and 25 point-intercepts used to measure vegetation composition, cover, and height at 5-m intervals. The red circles ($0.33m^2$) are locations where biomass samples and field spectrometer measurements[1] were taken.

[1] A limited number of spectrometer measurements on subplots and points overlapping biomass samples were taken to determine how high resolution field spectral can be used to estimate herbaceous biomass.

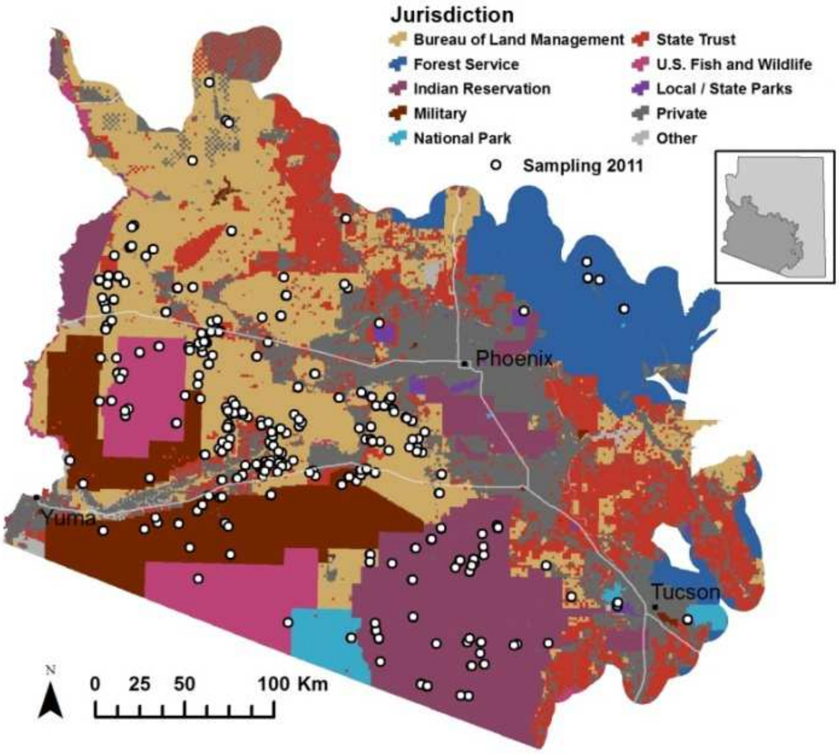



**Figure 2**. The 103,000-km$^2$ extent that encompasses the Sonoran Desert JFSP/SERDP study area in southwestern Arizona, including 239 plot locations sampled from Jan-April, 2011.

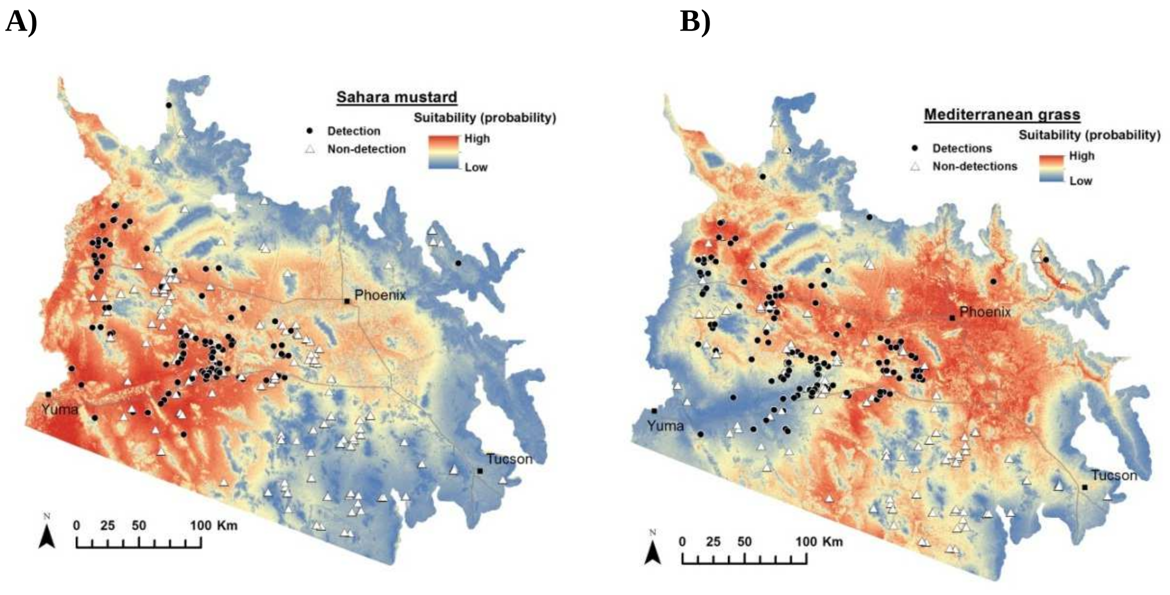
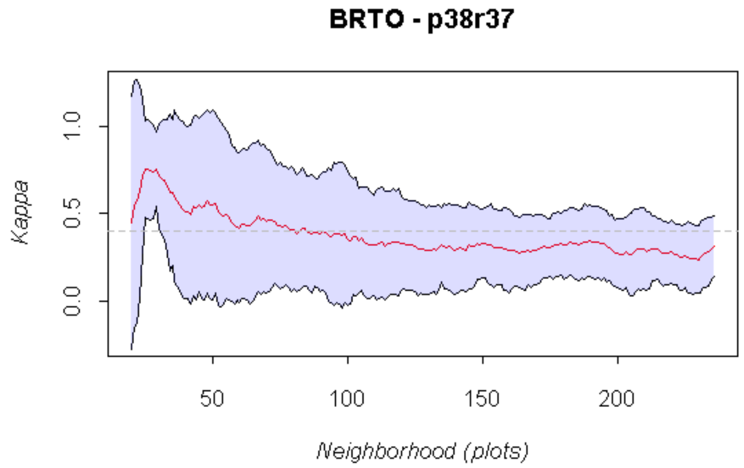



**Figure. 3**. Detection locations and probabilistic models of habitat suitability for target invasive species **A**) Sahara mustard and **B**) Mediterranean grass across the study area.

A)




B)

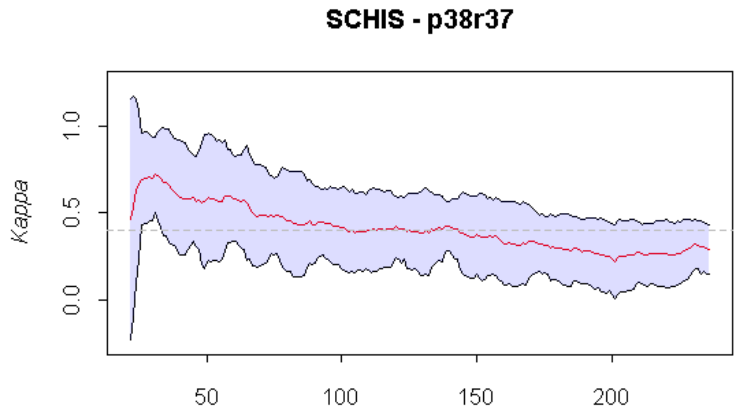



**Figure 4**. Kappa values for local-scale MODIS model runs (*n* = 1,000) of increasing neighborhood sizes and random seed plots for A) *Brassica tournefortii* (BRTO) and B) *Schismus* spp. (SCHIS, bottom) in Landsat tile path 38, row 37. The running means (red line) and 95% confidence intervals (black lines) also are presented.

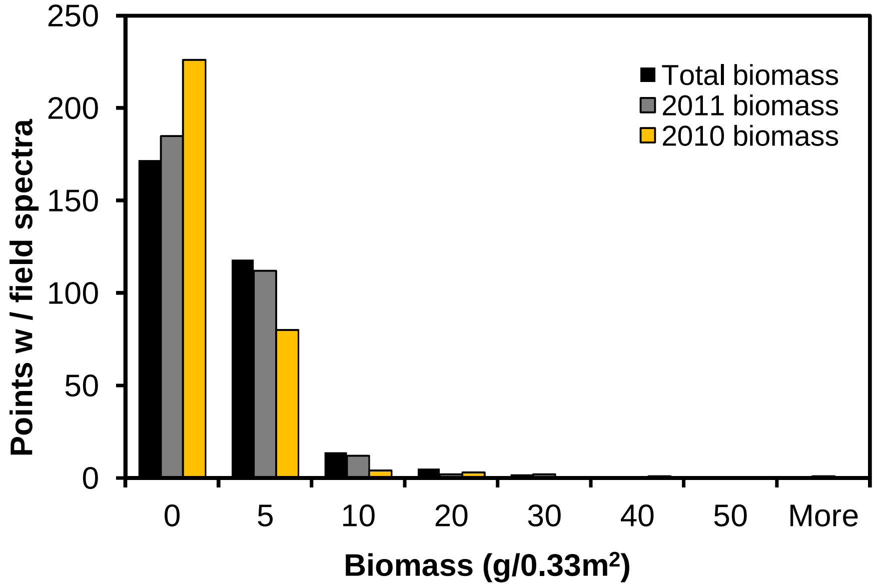



**Figure 5.** Herbaceous plant biomass (g/0.33 $m^2$, dry weight) distribution for total, current (2011) and the previous year's (2010) production for point intercepts with spectral reflectance measurements ($n$ = 315), obtained with an ASD Inc. field spectrometer.

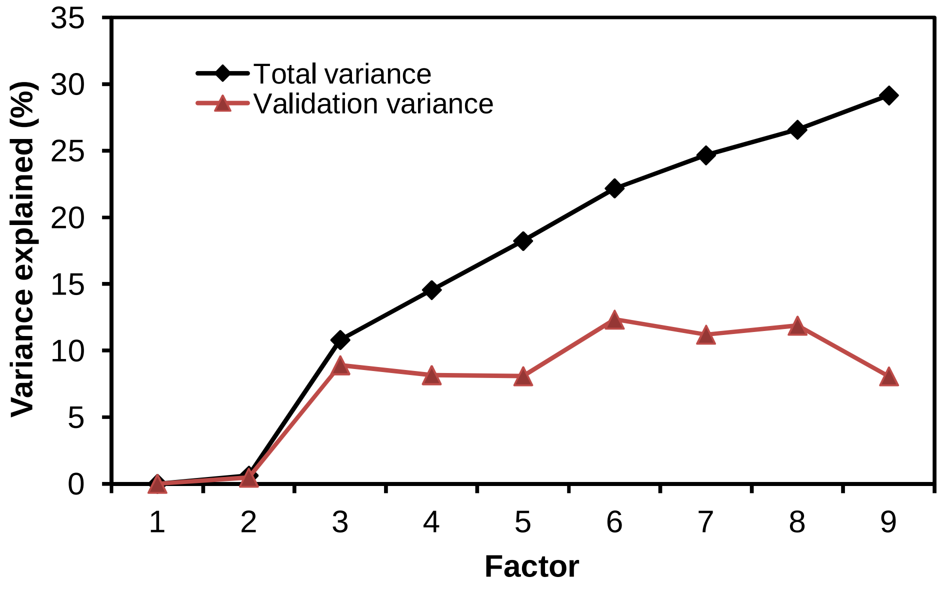



**Figure 6.** Partial least squared regression results indicating total amount of variance explained by the model and variance explained from 10-fold cross validation with an increased number of orthogonal factors.

**Appendix A.** Fire modeling workshop and technical training for fire and fuels specialists, land managers and stakeholders.

### FIRE MODEL TRAINING AND APPLICATIONS

**Instructors:**
**Nicole Valiant** (Western Wildland Environmental Threat Assessment Center: WWETAC[1]): Nicole received her BS from University of California Davis in Evolution and Ecology, and MS and PhD from University of California Berkeley in Environmental Science Policy and Management where her research centered around fuel treatment effectiveness at multiple scales. She has worked for the Forest Service since 2000 first as a seasonal fire fighter including one season on the Redding IHC and is currently a Fire Ecologist with WWETAC.

**Lauren Miller** (Deschutes National Forest): Lauren received her BS from University of Michigan in Ecology and MS from University of Idaho in Forestry, studying 20th century fire history methodologies. Since 2000, she has worked as an environmental consultant in the private sector, for the state of Arizona (NAU-ERI), and for federal land management agencies, primarily in fire effects/fire ecology. She currently works for the Deschutes National Forest, providing technical support to complex fire and fuels planning efforts.

**Location:**
Northern Arizona Applied Research and Development, building #56 (See attached map)

**Training summary:**
Wildland fire is a prominent forest disturbance in the western US and a principal driver of a number of important ecosystem processes. However, fire regime disruption over the last century has contributed to unprecedented changes in fire behavior. Fire behavior modeling is quickly becoming prerequisite to implementing forest management and restoration activities on federal and state managed forest land.

The **first two days** of this intensive training is specifically for GIS analysts with a forestry and fire ecology background to become skilled at effectively using the ArcFuel interface for fire modeling and forest growth simulation models such as the Forest Vegetation Simulator. Analysts will participate in hands-on training to parameterize, customize and interpret fire models and their associated outputs. **A third day** will be dedicated to both GIS technicians and fire managers on how fire model outputs are interpreted and integrated into fire management in addition to identifying how field managers can inform the fire modeling process.

**Day 1 (09/27/10)**
Location: GRAIL Computer Lab, ARD 226
Schedule: (***Technical Group***[2]) 8:30am to 12:00pm / Lunch / 1:00pm to 5:00pm

**Day 2 (09/28/10)**
Location: GRAIL Computer Lab, ARD 226
**Schedule:** (***Technical Group***) 8:30am to 12:00pm / Lunch / 1:00pm to 5:00pm

**Day 3 – (09/29/10)**
**Location:** Large Pod Conference Room, ARD 1st floor
**Schedule:** (***General Audience***[3]) 9:00am to 12:00pm

**Fire Model Training Modules**

| Date | Training module | Module Description |
|---|---|---|
| 09-27-10 | Fire modeling techniques (***Technical Group 15 participants***) | • Intro. to modeling and applications (FlamMap, FVS, treatment minimizer/rTOM, and others)<br>• Working w/ArcFuels interface<br>• Fire model data and inputs<br>• Fire model parameterization<br>• Preliminary model runs |
| 09-28-10 | Fire model techniques (***Technical Group 15 participants***) | • Fire model implementation<br>• Interpreting results<br>• Output and GIS analysis techniques |
| 09-29-10 | Fire model applications & land management planning (***General Audience 40 participants***) | • Discussion of fire model outputs, understanding how managers inform the models; interpretation, landscape planning applications and future fire model developments |

**For information contact:**

**USFS/Agency personnel: Mary Lata: mlata@usfs.fed.us**
**Non-Agency personnel: Steven Sesnie: steven.sesnie@nau.edu**

[1]WWETAC - http://www.fs.fed.us/wwetac/
[2]Approximately 16 individuals with technical expertise in GIS will be accommodated for this training
[3]In addition to the technical group, approximately 25 individuals with a forestry and fire management background will be accommodated in this group