**Estimation of Gross Ecosystem Production for the Reynolds Creek Critical Observatory Zone Derived from Landsat 5 TM 30-Meter Imagery**

**Jeremy Facer 1, Keith Weber 1, Susan Parsons 2, Kathleen Lohse 2, Gerald Flerchinger 3, Aaron Fellows 3**

1. GIS Center, Idaho State University, Pocatello, ID 83201, USA; facejere@isu.edu (J.F.); webekeit@isu.edu (K.W.)
2. Geosciences, Idaho State University, Pocatello, ID 83201, USA; parssusa@isu.edu (S.P.); klohse@isu.edu (K.L.)
3. Agricultural Research Service, U.S. Dept. of Agricultural Services, ID 83201, USA; Gerald.Flerchinger@ars.usda.gov (G.F.); Aaron.Fellows@ars.usda.gov (A.F.)

**Abstract:** Measurements of gross primary productivity (GPP) can help explain ecosystem structure and function. Direct measurement of GPP, however is difficult or impossible. As a result, indirect measurements via satellite imagery and vegetation indices have been developed to aid in our understanding of ecosystem structure and function. This study compares annual maximum MSAVI2 against measured CO2 levels recorded throughout each water year at the Reynolds Creek Experimental Watershed (RCEW) in southwest Idaho between 2002 and 2011. A linear regression was calculated from four individual datasets and resulting best-fit equations used to predict gross ecosystem productivity (GEP). Predicted GEP values were compared against measured GEP values. Average GEP measurements from three individual tower sites across 2002-2011 resulted in the highest correlation (R2 = 0.87) with a percent error of residuals ranging from -7% to 13.5%. Resulting regression analysis with all data points showed an R2 value of 0.47 with a percent error ranging from -26% to 33%. Spatial predictions of 2007 GEP from MSAVI2 data suggests GEP levels will be highest in central and southern areas of the RCEW. Low sample sizes for each dataset must be taken into consideration when interpreting these results and further research including more data points and multiple regression analysis may improve GEP predictions.

**Introduction**

Gross primary productivity (GPP) at an ecosystem scale is described by the level of photosynthesis that occurs within an ecosystem (Chapinet al., 2012). GPP is a crucial measurement when studying global alterations and plant community health (Aaltoet al., n.d.), where higher levels of GPP indicate healthy ecosystems and vice-versa. However, direct measurement of productivity levels within an ecosystem are often extremely difficult or logistically impossible, which warrants further research into indirect GPP measurement methods. One such measurement involves the quantification of gross ecosystem productivity (GEP). GEP is quantified as the fluctuation differences of CO2 between an ecosystem and the atmosphere via gross ecosystem photosynthesis (Biederman et al., 2016). By measuring the fluctuation difference of CO2, insight into photosynthetic activities and, therefore, plant community health can be achieved. Contemporary technological advances allow for such indirect GEP measurement methods in the form of remote sensing. The Normalized Difference Vegetation Index (NDVI) is one of the most widely used spectral indices to monitor the Earth’s surface, with over 121,000 scientific articles using NDVI (Robinson et al., 2017). While this vegetation index (VI) proves useful in most scenarios, NDVI studies are limited when applied to areas containing high areas of exposed soil (Qi et al., 1994). The Modified Soil-Adjusted Vegetation Index (MSAVI2) was implemented to correct vegetation index values in areas of high soil surface exposure.

MSAVI2 is defined by equation 1 below:

Where is the reflectance in near-infrared (NIR) and is the reflectance in red (1994).

While NDVI and MSAVI2 serve as estimators of primary production, both biotic and abiotic factors such as bare soil reflectance, water reflectance, cloud cover reflectance and species differences in leaf structure are known to influence vegetation index values (2012). Because of these factors, vegetation index values determined solely by remote sensing techniques warrant further study to confirm their accuracy. A 2018 study by Fellows et al. compared maximum NDVI values against GEP levels in an area within the Reynolds Creek Critical Zone Observatory (RC-CZO) and found the R2 value between these variables to be 0.75. This association was determined using satellite imagery drawn across Landsat 5, Landsat 7 and Landsat 8 sensors between 1985 and 2016. This study follows the research conducted by Fellow et al. (2018) and compares the calculated MSAVI2 values via remote sensing techniques against measured CO2 fluctuation levels (GEP) at the RC-CZO.

The Reynolds Creek Critical Observatory Zone is located in southwestern Idaho (figure 1) and is an experimental watershed primarily focused on soil carbon research. The watershed is approximately 239 km2 and contains a wide variety of climate conditions (“Detailed Introduction,” n.d.). To determine carbon balance, eddy covariance towers were placed within the RC-CZO and GEP was monitored. These towers rapidly measure vertical wind speed and CO2 content of upward and downward air exchange between the environment and atmosphere which can then be used to determine CO2 fluctuations (2012). Data from these eddy covariance towers were used in this study.

1. **Materials and Methods**

*2.1 Data*

Landsat 5 TM collection 1 level-1 imagery of Reynolds Creek CZO was obtained via USGS Earth Explorer. Scenes from 2002-2011 where visually analyzed within Earth Explorer and scenes containing cloud cover presence over the CZO area were excluded while all other scenes were imported for VI composites. The spatial resolution for all imported scenes were 30 x 30 meters projected in a WGS 1984 UTM Zone 11N coordinate system.

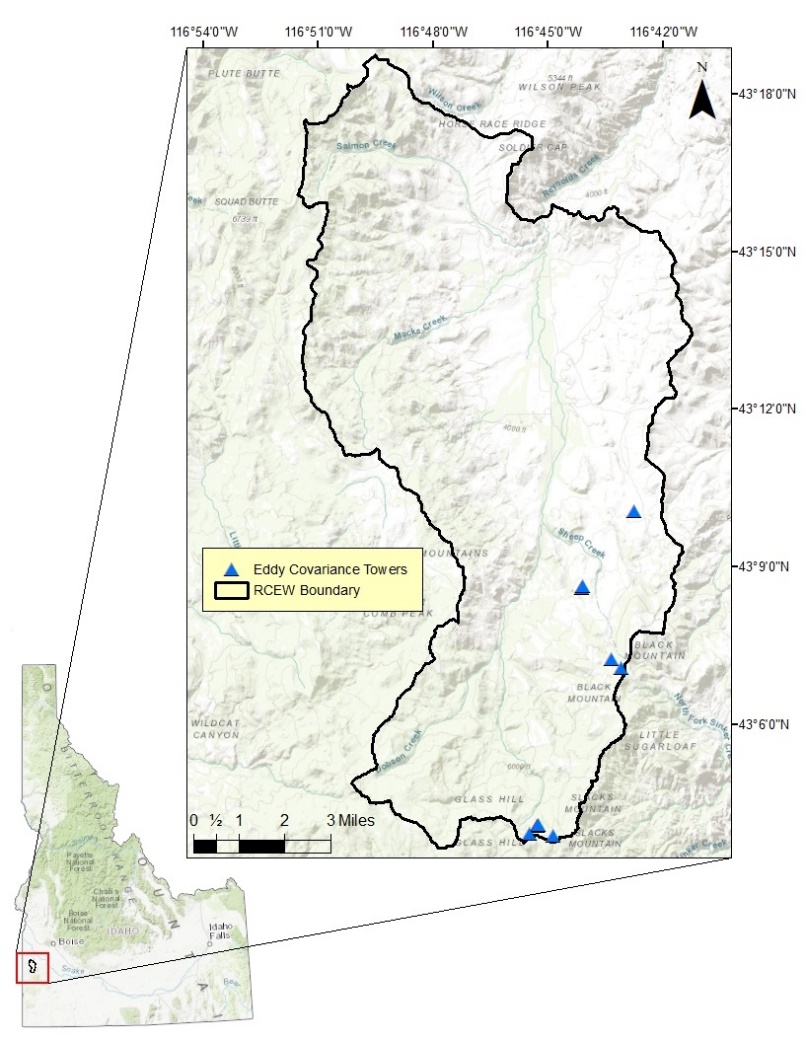
GEP data was collected via 10 eddy covariance towers found within the Reynolds Creek CZO (figure 1). Of the 10 towers, three towers contained GEP measurements within the same 2002-2011 timeframe for each water year. Because of the Landsat 5 active timeframe being between 2002-2011, GEP data collected outside this timeframe were not used in this study. Furthermore, Landsat data from 2009 contained only one usable scene and was therefore excluded from this study.

Figure 1. Reynolds Creek Critical Zone Observatory located in southwestern Idaho.

A 1-meter LiDAR-derived Digital Elevation Model (DEM) was imported into ArcMap software via the Reynolds Creek CZO database server and was used to characterize the topography at each tower site. A separate 10 x 10-meter topography dataset was imported from Idaho State University’s GIS Center and slope and aspect calculated from these data. These two respective datasets were then used to extract slope (in degrees) and aspect data for each tower. Finally, a 30 x 30-meter vegetation raster dataset from ISU’s GIS Center was used to extract vegetation cover at each tower location.

* 1. *Compositing*

A total of 72 scenes containing band 3 and band 4 data (red and near infrared bands) were imported into IDRISI TerrSet software to produce composite maximum NDVI and MSAVI2 images for each calendar year. Of the 72 scenes, 59 were automatically atmospherically corrected by IDRISI using the *Cos(t)* model correction. Five scenes were atmospherically corrected manually using information derived from ISU GIS Center sources (Table 1).

Table 1. Parameter values used for manual atmospheric correction. This information was derived using Landsat 5 TM specifications provided from Idaho State University.

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| GMT | 18 |
| DN Haze | 0 |
| Lmin for band 3 scenes | -0.117 |
| Lmax for band 3 scenes | 26.4 |
| Lmin for band 4 scenes | -0.151 |
| Lmax for band 4 scenes | 22.1 |
| Band center for band 3 scenes | 0.66 |
| Band center for band 4 scenes | 0.83 |
| Satellite viewing angle | 0 |
| Sun elevation | 60 |

Atmospherically corrected scenes were then used to create NDVI and MSAVI2 scenes using the IDRISI “VEGINDEX” command. Band 3 and 4 from each scene were used to create separate NDVI and MSAVI2 scenes. The smallest common scene extent size was found using ArcMap software and each scene was clipped to this extent to successfully composite yearly maximum scenes in IDRISI. No scene data was lost during this process. The clipped NDVI and MSAVI2 scenes were imported into IDRISI and the “NDVIComp” (composite NDVI) tool was then used to generate yearly maximum NDVI and MSAVI2 scenes between 2002 and 2011.

* 1. *Data Compilation*

Yearly maximum NDVI and MSAVI2 scenes were imported into ArcMap software for data extraction. Footprint information of each tower’s measurement influence was used to extract average VI values for each tower. Footprints for 4 of the 10 towers were provided from USDA ARS sources containing information of tower measurement influential areas that were discarded from the overall footprint areas, while the remaining 6 footprints were generated using a 90-meter buffer creation. The “Zonal Statistics as Table (Batch)” tool was used to extract each of the 18-total maximum NDVI and MSAVI2 yearly scenes for each respective tower. The table output was then imported into Excel for further analysis.

Elevation, slope, aspect and vegetation cover data for each tower location was extracted using the ArcMap “Sample” tool. The tower point feature class was reprojected to each of the four respective raster datasets before the data extraction was conducted. Slope extraction was done in degrees rather than percent. These data were copied into an Excel spreadsheet containing maximum yearly NDVI and MSAVI2 values along with respective GEP data.

Of the 10 towers, 6 towers did not match the date range of Landsat 5 TM data collection and 1 tower did not contain usable GEP data. Because of this, only three towers were statistically analyzed. These 3 towers include: (1) Big Sage tower, (2) Reynolds Mountain Exposed Ridge tower and (3) Reynolds Mountain East Above Aspen Canopy tower. The year 2007 was the only individual year where descriptive statistical analysis could be conducted. Between each of the three towers, the Reynolds Mountain Exposed Ridge tower contained the most usable observations (n = 5).

Descriptive statistical analysis regarding GEP and maximum MSAVI2 values were conducted in Excel using the Analysis ToolPak add-in. Statistics were calculated for each of the following datasets: (1) 2007 only, (2) tower averages between 2002 and 2011, (3) Reynolds Mountain Exposed Ridge tower only and (4) all data points in which GEP and MSAVI2 values were available. Each dataset was individually graphed using Excel software and R2 values determined.

* 1. *GEP Predictability*

A best-fit linear trend line for measured annual GEP levels and yearly maximum MSAVI2 values was individually determined for all four datasets using Excel software. An R2 value and linear best-fit line was determined for each individual data pool. The percent error for each dataset was calculated using the respective linear best-fit equation to determine the predicted GEP values against measured GEP values. Visual representation of predicted GEP levels for the 2007 year was generated using ArcMap software. The linear best-fit equation was applied to the maximum MSAVI2 raster image using Idrisi’s raster calculator tool where the input raster image was used as the x-variable in the linear best-fit equation.

**Results**

*3.1 Simple Linear Regression*

Simple linear regression analysis revealed that in 2007 annual maximum MSAVI2 values explained approximately 57% of the observed variation in annual GEP values at the three tower sites (n = 3). When site GEP and maximum MSAVI2 values were averaged across all years, GEP variance explained by maximum MSAVI2 improved by a factor of 1.5 (R2 of 0.57 vs 0.87, respectively). Linear regression analysis of data from the Reynolds Mountain Exposed Ridge towerresulted in an R2 value of 0.70 (n = 5). Analysis of all usable data points resulted in an R2 value of 0.47 (n = 14). GEP and maximum MSAVI2 values averaged across all years for each tower revealed the highest R2 value of 0.87 (figure 2).

*a*

*b*

*3.2 2007 Dataset*

*d*

*c*

Figure 2. Graph representations of each simple linear regression analysis conducted on the four datasets, a) 2007 GEP vs maximum MSAVI2, b) Total average GEP vs MSAVI2 values, c) Data from the Reynolds Mountain Exposed Ridge tower only, and d) All usable data points

GEP measurements for the three towers in 2007 ranged from 379 to 678 gC/m2/year (table 2). The Reynolds Mountain East Above Aspen Canopy tower contained the highest average GEP level in 2007 while the Reynolds Mountain Exposed Ridge tower contained the lowest average GEP level in 2007. Average GEP between all three towers in 2007 was 511 ± 152 gC/m2/year. The average maximum MSAVI2 values between all three towers in 2007 was 0.21 ± 0.022.

Table 2. Average maximum MSAVI2 values and average annual GEP values for three towers within 2007.

|  |  |  |
| --- | --- | --- |
| Site | Avg. Max.  MSAVI2 | GEP  (gC/m2/year) |
| Big Sage tower | 0.1868 | 475.4 |
| Reynolds Mountain Exposed Ridge tower | 0.2040 | 379.0 |
| Reynolds Mountain East Above Aspen Canopy tower | 0.2313 | 677.9 |

The equation of the linear best-fit line for the 2007 data pool was

Predicted values of GEP compared against actual GEP measurements contained a percent error ranging from -14% to 30% (table 3).

Table 3. Measured GEP values and predicted GEP values calculated using equation 2. The percent error is listed for each tower site.

|  |  |  |  |
| --- | --- | --- | --- |
| Site | Actual GEP (gC/m2/year) | Predicted GEP  (gC/m2/year) | Percent Error |
| Big Sage tower | 475.4 | 405.1 | -14.8 |
| Reynolds Mountain Exposed Ridge tower | 379.0 | 493.6 | 30.2 |
| Reynolds Mountain East Above Aspen Canopy | 677.9 | 633.5 | -6.5 |

Visual representation of the predicted GEP values across the RCEW for the 2007 year show negative GEP predictions for the majority of the RCEW. Moderate levels were predicted in the eastern and southern areas of the RCEW. The highest predicted GEP values were found in the center area and southern end of the RCEW (figure 3).

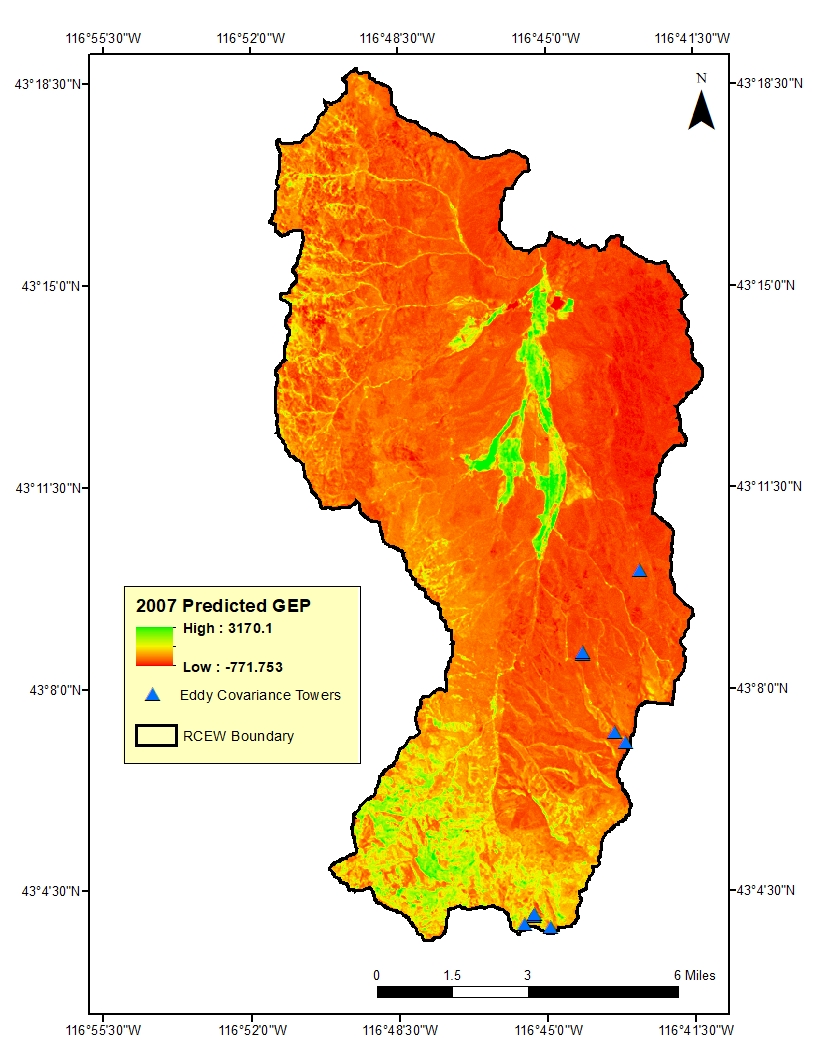
*3.3 Tower Averages Between 2002 and 2011 Dataset*

Figure 3. Predicted GEP levels across the RCEW for 2007. This prediction was determined using ArcMap's raster calculator tool, where the linear best-fit equation (equation 2) was applied.

Average annual GEP levels for the three towers across 2002 and 2011 ranged from 489 and 840 gC/m2/year (table 4). The Reynolds Mountain Exposed Ridge tower contained the lowest average GEP value of 489 gC/m2/year and the Reynolds Mountain East Above Aspen Canopy tower contained the largest average GEP value of 840 gC/m2/year. The average GEP of all three towers between 2002 and 2011 was 632 ± 184 gC/m2/year. Average maximum MSAVI2 values ranged from 0.2590 to 0.2824. The average maximum MSAVI2 values of all three towers across 2002 and 2011 was 0.26 ± 0.026.

Table 4. Average maximum MSVI2 values and average annual GEP values for the three towers across 2002 and 2011.

|  |  |  |
| --- | --- | --- |
| Site | Avg. Max. MSAVI2 | GEP (gC/m2/year) |
| Big Sage tower | 0.2590 | 568.0 |
| Reynolds Mountain Exposed Ridge tower | 0.2299 | 489.0 |
| Reynolds Mountain East Above Aspen Canopy | 0.2824 | 840.1 |

The equation of the linear best-fit line of the average values across 2002 to 2011 data pool was

The percent error of the predicted GEP values compared against the actual GEP values ranged from -7% to 13.5% (table 5). Two of the three towers contained predicted GEP values that were lower than the actual GEP values.

Table 5. Measured GEP values and predicted GEP values based off equation 3. The percent error is listed for each tower site.

|  |  |  |  |
| --- | --- | --- | --- |
| Site | Actual GEP  (gC/m2/year) | Predicted GEP  (gC/m2/year) | Percent Error |
| Big Sage tower | 568.0 | 644.7 | 13.5 |
| Reynolds Mountain Exposed Ridge tower | 489.0 | 454.8 | -7.0 |
| Reynolds Mountain East Above Aspen Canopy | 840.1 | 797.6 | -5.1 |

*3.4 Reynolds Mountain Exposed Ridge Tower Dataset*

GEP measurements for the Reynolds Mountain Exposed Ridge tower ranged from 379 to 549 gC/m2/year (table 6). The average GEP measurements between 2003 and 2007 for the Reynolds Mountain Exposed Ridge tower was 489 ± 83 gC/m2/year. The average maximum MSAVI2 value between 2003 and 2007 was 0.22 ± 0.022.

Table 6. Average maximum MSAVI2 values and average annual GEP values for the Reynolds Mountain Exposed Ridge tower between 2003 and 2007.

|  |  |  |
| --- | --- | --- |
| Year | Avg. Max. MSAVI2 | GEP (gC/m2/year) |
| 2003 | 0.1983 | 421.1 |
| 2004 | 0.2187 | 549.2 |
| 2005 | 0.2458 | 547.9 |
| 2006 | 0.2423 | 547.9 |
| 2007 | 0.2040 | 379.0 |

The equation of the linear best-fit line for this dataset was

Resulting predicted values of GEP compared against actual GEP measurements contained a percent error ranging from -12% to 14% (table 7). The average percent error was 0.8% for all five data points.

Table 7. Measured GEP values and predicted GEP values based off equation 4. The percent error is listed for each year.

|  |  |  |  |
| --- | --- | --- | --- |
| Year | Actual GEP  (gC/m2/year) | Measured GEP  (gC/m2/year) | Percent Error |
| 2003 | 421.1 | 413.8 | -1.7 |
| 2004 | 549.2 | 479.0 | -12.8 |
| 2005 | 547.9 | 565.5 | 3.2 |
| 2006 | 547.9 | 554.6 | 1.2 |
| 2007 | 379.0 | 432.2 | 14.0 |

*3.5 All Usable Data Points*

GEP measurements for all data points in which measured GEP and average maximum MSAVI2 values were available ranged from 379 to 941 gC/m2/year (table 8). The average GEP for all data points was 617 ± 178 gC/m2/year with a skewness value of 0.66 (n = 14). The average maximum MSAVI2 value for all usable data points was 0.26 ± 0.06 with a skewness value of 0.92 and a range of 0.19 to 0.39 (n = 14).

The equation of the linear best-fit line of all usable data points was

Predicted values of GEP compared against actual GEP measurements contained a percent error ranging from -26% to 33% (table 9). The average percent error was 3.8% ± 19.3% with a skewness value of -0.045 (n = 14).

Table 8. Average maximum MSAVI2 values and average annual GEP values for all complete data points between 2002 and 2011.

|  |  |  |  |
| --- | --- | --- | --- |
| Site | Avg. Max. MSAVI2 | GEP (gC/m2/year) | Year |
| Reynolds Mountain Exposed Ridge tower | 0.1983 | 421.1 | 2003 |
| Reynolds Mountain Exposed Ridge tower | 0.2187 | 549.2 | 2004 |
| Reynolds Mountain Exposed Ridge tower | 0.2458 | 547.9 | 2005 |
| Reynolds Mountain Exposed Ridge tower | 0.2423 | 547.9 | 2006 |
| Big Sage tower | 0.2392 | 540.3 | 2006 |
| Big Sage tower | 0.1868 | 475.4 | 2007 |
| Reynolds Mountain Exposed Ridge tower | 0.2040 | 379.0 | 2007 |
| Reynolds Mountain East Above Aspen Canopy tower | 0.2313 | 677.9 | 2007 |
| Big Sage tower | 0.2387 | 469.3 | 2008 |
| Reynolds Mountain East Above Aspen Canopy tower | 0.2906 | 835.0 | 2008 |
| Big Sage tower | 0.3558 | 608.3 | 2010 |
| Reynolds Mountain East Above Aspen Canopy tower | 0.2908 | 907.0 | 2010 |
| Big Sage tower | 0.3933 | 746.6 | 2011 |
| Reynolds Mountain East Above Aspen Canopy tower | 0.3301 | 940.7 | 2011 |

Table 9. Measured GEP values and predicted GEP values based off equation 5. The percent error is listed for each tower site and year.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Site | Actual GEP (gC/m2/year) | Predicted GEP (gC/m2/year) | Percent Error | Year |
| Reynolds Mountain Exposed Ridge tower | 421.1 | 492.1 | 16.9 | 2003 |
| Reynolds Mountain Exposed Ridge tower | 549.2 | 532.4 | -3.1 | 2004 |
| Reynolds Mountain Exposed Ridge tower | 547.9 | 585.8 | 6.9 | 2005 |
| Reynolds Mountain Exposed Ridge tower | 547.9 | 579.1 | 5.7 | 2006 |
| Big Sage tower | 540.3 | 572.9 | 6.0 | 2006 |
| Big Sage tower | 475.4 | 469.4 | -1.3 | 2007 |
| Reynolds Mountain Exposed Ridge tower | 379.0 | 503.5 | 32.8 | 2007 |
| Reynolds Mountain East Above Aspen | 677.9 | 557.4 | -17.8 | 2007 |
| Big Sage tower | 469.3 | 571.9 | 21.9 | 2008 |
| Reynolds Mountain East Above Aspen | 835.0 | 674.4 | -19.2 | 2008 |
| Big Sage tower | 608.3 | 802.9 | 32.0 | 2010 |
| Reynolds Mountain East Above Aspen | 907.0 | 674.6 | -25.6 | 2010 |
| Big Sage tower | 746.6 | 876.9 | 17.4 | 2011 |
| Reynolds Mountain East Above Aspen | 940.7 | 752.2 | -20.0 | 2011 |

**Discussion/Conclusion**

The cumulative maximum MSAVI2 values used in this study were able to explain more variation of measured GEP relative to cumulative maximum NDVI values. The majority of the RCEW is classified as a semi-arid ecosystem (Peel et al., 2007) and, as such, MSAVI2 predictions were expected to perform better than NDVI predictions. This may be explained by the soil adjustment factor that is used in the soil-adjusted vegetation index (SAVI) family (1994), which includes the MSAVI2 methodology.

The visual representation of predicted GEP determined by the 2007 linear best-fit equation predicted values ranging from -772 gC/m2/year to 3170 gC/m2/year with high levels of predicted GEP in central and southern zones of the RCEW. The vegetation map found on the RCEW website indicates that cultivated land and Douglas Fir were the prominent vegetation types for these areas (“Variation in Soils and Geology,” n.d.). Measured GEP values for the average 2002 to 2011 dataset ranged from 379 gC/m2/year to 678 gC/m2/year. Because of the small sample size (n = 3), predicted GEP values that are extrapolated beyond the measured GEP range may not contain accurate GEP predictions.

Of the four datasets, the dataset of average values between 2002 and 2011 had the lowest range of percent error and highest R2 value. This may be a result of the underrepresentation of data outliers due to the calculation and use of the averages for each tower rather than individual data points. While the percent error and R2 value suggest a strong association between predicted GEP and actual GEP values, a low sample size for this dataset (n = 3) must be taken into consideration when interpreting these values. Statistically, a low sample size may not warrant statistical significance with regard to this model. However, the practicality of eddy covariance towers must also be taken into consideration when conducting studies such as this. The low sample size may result in statistically insignificant results, albeit biological significance of these findings is debatable.

While linear regression of MSAVI2 values and measured GEP values provided a basic model to predict GEP values, further studies using multiple regression analysis including other biotic and abiotic factors is needed to strengthen these predictions. Alongside this, spatial prediction of GEP values across the RCEW may require more point samples for accurate GEP predictions. This may be accomplished using data from other eddy covariance towers found in the RCEW. In order to do this, however, cross calibration techniques must be developed for the various Landsat satellites due to the temporal variation of measured GEP values in the RCEW. Overall, these findings may not hold statistical significance in predicting GEP values, but both practicality of the eddy covariance towers and biological significance of the findings must be considered.

**Works Cited**

Aalto, Emil, et al. “Net Primary Production.” *Net Primary Production*, Sacramento River Watershed Program, indicators.ucdavis.edu/waf/model/indicator/net-primary-production.

Biederman, Joel A., et al. “Terrestrial Carbon Balance in a Drier World: the Effects of Water Availability in Southwestern North America.” *Global Change Biology*, vol. 22, no. 5, 2016, pp. 1867–1879., doi:10.1111/gcb.13222.

Chapin, F. Stuart, et al. *Principles of Terrestrial Ecosystem Ecology*. 2nd ed., Springer, 2012.

“Detailed Introduction to the Reynolds Creek Critical Zone Observatory (RC CZO*).” Reynolds Creek Critical Zone Observatory*, National Science Foundation, criticalzone.org/reynolds/about/.

Fellows, Aaron W., et al. “Rapid Recovery of Gross Production and Respiration in a Mesic Mountain Big Sagebrush Ecosystem Following Prescribed Fire.” *Ecosystems*, 2018, doi:10.1007/s10021-017-0218-9.

Peel, M C, et al. “Updated World Map of the Koppen-Geiger Climate Classification.” *Hydrology and Earth System Sciences,* vol. 11, Jan. 2007, pp. 1633–1644.

Robinson, Nathaniel, et al. “A Dynamic Landsat Derived Normalized Difference Vegetation Index (NDVI) Product for the Conterminous United States.” *Remote Sensing*, vol. 9, no. 8, 2017, p. 863., doi:10.3390/rs9080863.

Qi, J., et al. “A Modified Soil Adjusted Vegetation Index.” *Remote Sensing*, 1994.

“Variation in Soils and Geology.” Reynolds Creek Critical Zone Observatory, *Critical Zone Observatories,* criticalzone.org/reynolds/about/.