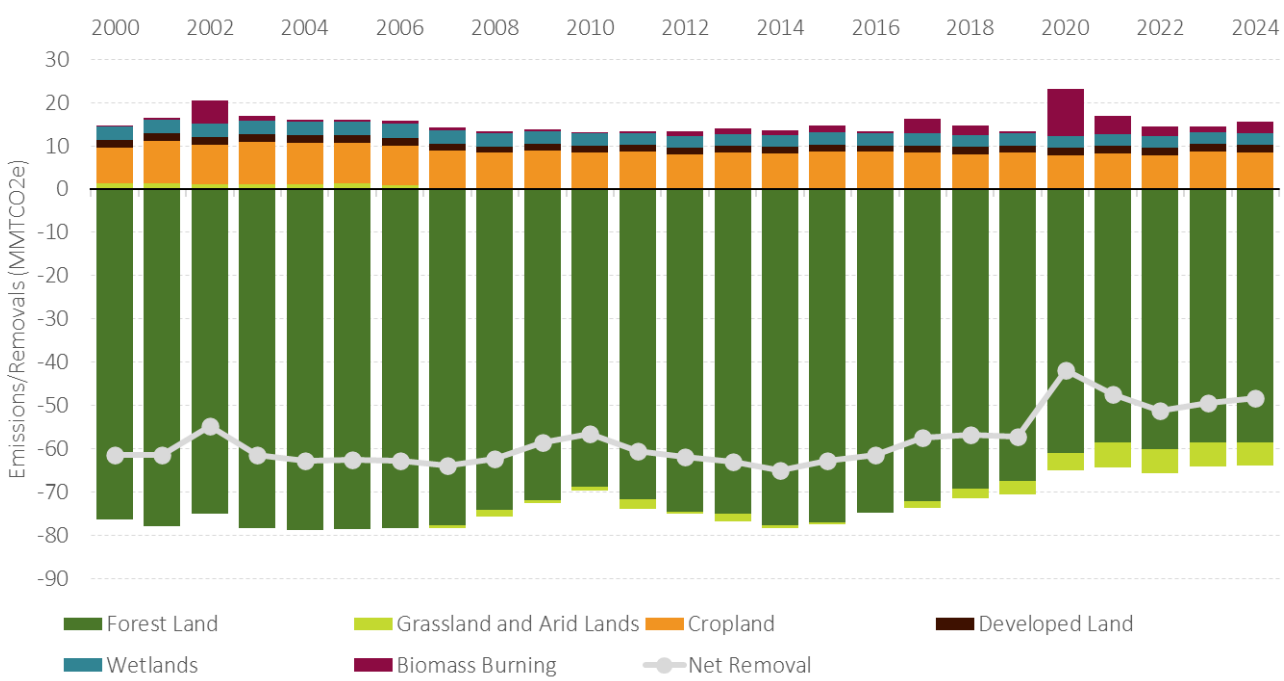
Appendix B – Analysis of Trends in the Land-based Net Carbon Inventory

This appendix provides detailed analysis of trends in land-based emissions and removals using data generated from the 2025 Land-based Net Carbon Inventory (Inventory), prepared by the Oregon Department of Energy. This appendix also provides guidance on methods selection, including testing data to ensure that they meet the assumptions for using the method(s) chosen, and testing for significant trends. The trend analysis results in this appendix are distinct from the results in the 2025 Inventory and its appendices, particularly Appendix F – Detailed Results, which did not include tests for significant trends. ODOE does not propose reasons for the trend data at this time. It is important to note that determining the reasons behind the trends may or may not be possible given historical data quality, sampling procedures, policy, and management decisions, among other reasons.

Analysis of Trends in the Land-based Net Carbon Inventory

The Land-based Net Carbon Inventory (Inventory) estimates the emissions and removals of all of Oregon's land regardless of ownership (Figure B1). These emissions (positive values) and removals (negative values) are reported by six categories in the Inventory: Forest Land, Developed Land, Wetlands, Grassland (including arid lands), Cropland, and Biomass Burning. Oregon's land is a net sink, meaning that, on net, it removes more carbon than it emits. Net removal is calculated as the sum of the emissions and removals of all categories. Oregon's net removal may change over time due to changes in category emissions, removals, or both. For example, an increase in net removals (greater negative numbers) may be the result of more sequestration from forest land or less emission from wildfires, or both. Analysis of trends for all categories and individual categories is provided in the subsections below and use data reflected in Figure B1.

Figure B1. Emissions and removals by category in the period 2000–2024 from the Land-based Net Carbon Inventory. Net removal (gray dots and line) is calculated as the sum of emissions and removals of all categories.



Trends by Category

This section contains analysis of significant trends, where present, in net flux of by category in the Inventory. Additionally, this section includes discussion of trends for net flux across all categories except for biomass burning in the sub-section titled "All Categories without Biomass Burning."

Each subsection includes a brief discussion of whether the data met the statistical assumptions necessary for use in two tests of the significance of trends: linear regression and Mann-Kendall. Data were tested for whether they met the assumptions of normality, homoscedasticity, and auto-correlation. *It is critical that data meet both the assumptions required for the statistical methods chosen and that the tests indicate a significant trend.* If either condition is not met, conclusions about trends may be erroneous. More discussion of assumptions, significance of trends, and calculating direction and magnitude of trends can be found toward the end of this document in Statistical Analysis of Trends in Time Series Data.

Each subsection also includes a discussion of the analysis of significant trends in the data, including whether any trend was significant and the direction (i.e., increase or decline) and magnitude of any trend, separated by the form of analysis. Equations of trend lines were determined using Ordinary Least Squares analysis for linear regression, or Theil-Sen analysis for Mann-Kendall. Notes on the underlying data of the categories and their trends are provided where such analysis was possible at this time. Fuller analysis of trends is planned for future updates to the Inventory.

Trends in All Categories

The data in this category are labeled as Net Removal (gray dots and lines) in the legend of the stacked bar chart showing full inventory results shown in Figure B1 of this appendix. The values for All Categories are calculated as the sum of emissions and removals of all categories (Forest Land, Developed Land, Wetlands, Grassland, Cropland, and Biomass Burning).

Data Assumptions

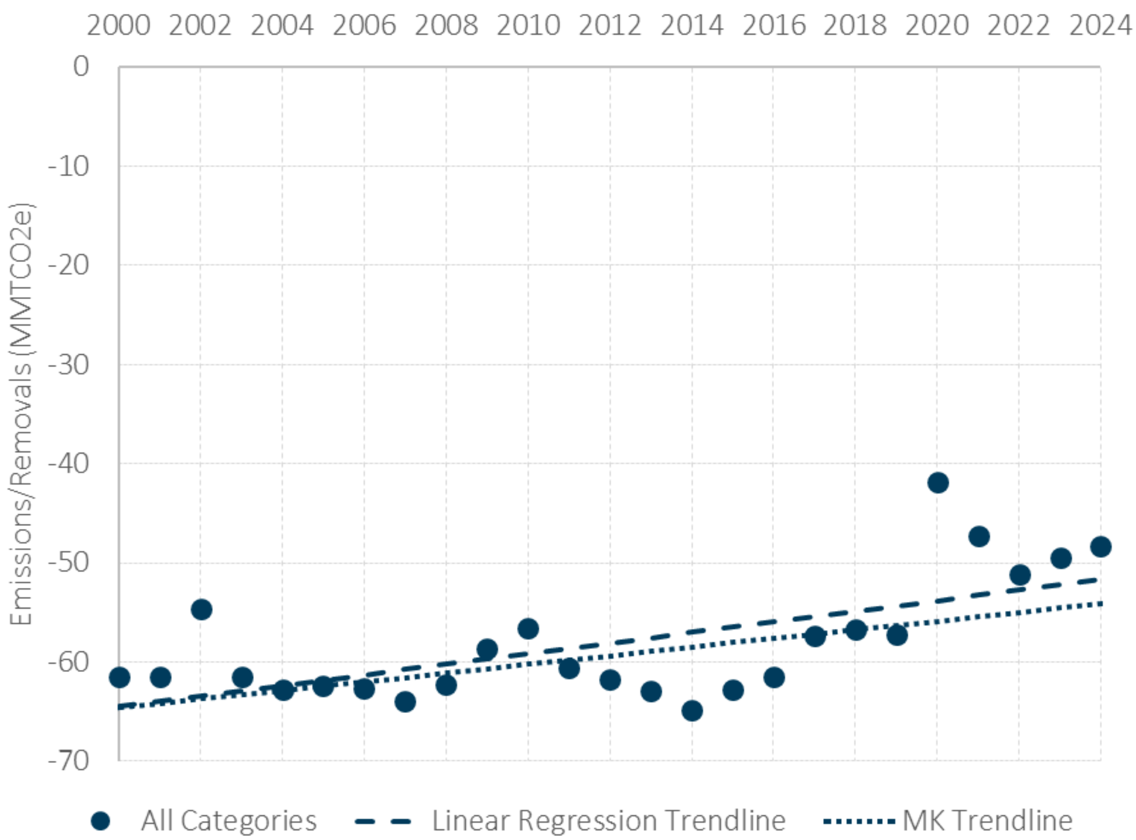
Data for All Categories satisfied the assumptions for both linear regression and Mann-Kendall analysis, so both forms of analysis are appropriate for use with these data.

Analysis of Trends

Per linear regression analysis, removals of All Categories has significantly declined over time at a rate of 0.54 MMTCO2e/year.

Per Mann-Kendall analysis, removals of All Categories have significantly declined over time at a rate of 0.58 MMTCO2e/year.

Figure B2. Removals of All Categories. Negative Y-values indicate removals. Presence of trendlines indicates that trends are significant ($p \leq 0.05$). Trendlines increasing over time indicate a significant decline in removals. Linear Regression Trendline indicates the trend determined using linear regression with Ordinary Least Squares analysis; MK Trendline indicates the trend determined using Mann-Kendall with Theil-Sen analysis.



Trends in All Categories without Biomass Burning

The data in this category are calculated by adding the data for all categories except for biomass burning. As wildfires have been substantial sources of emissions in recent years, excluding them from the analysis allows us to determine whether net removals of Oregon's lands are declining irrespective of the direct contribution of wildfires to emissions.

Data Assumptions

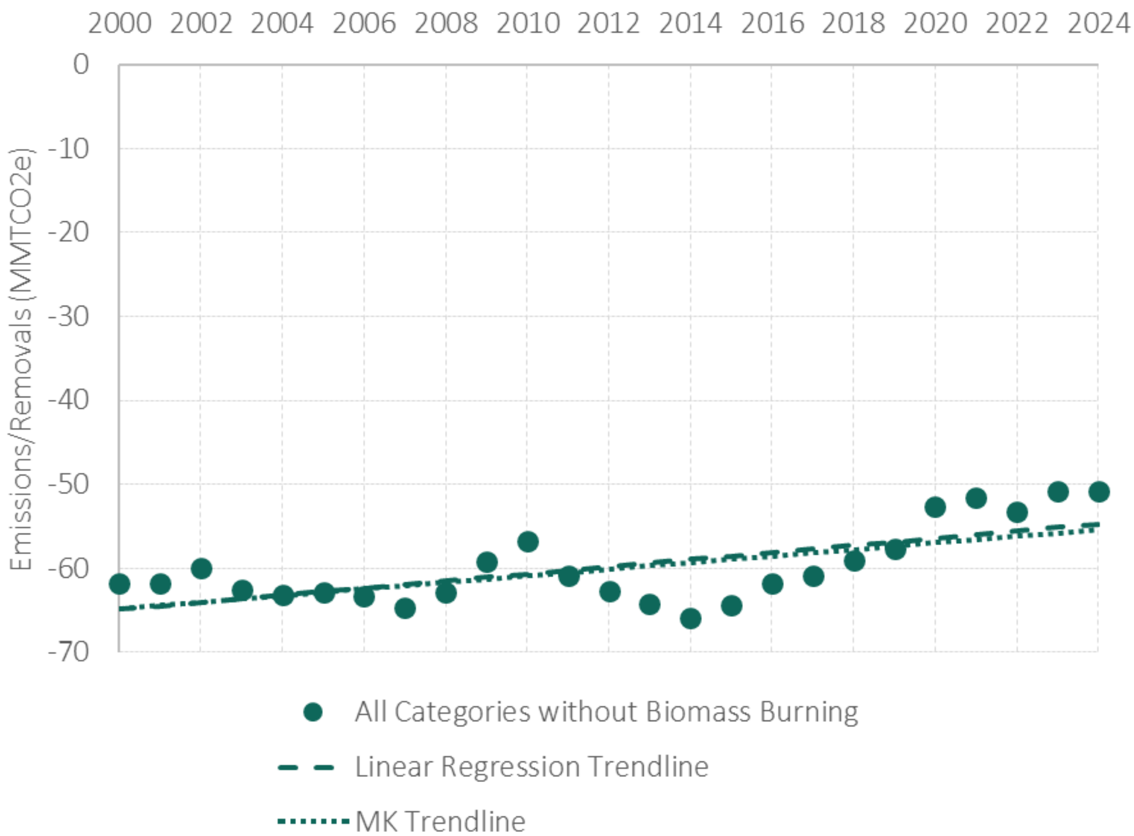
Data for All Categories without Biomass Burning satisfied the assumptions for both linear regression and Mann-Kendall analysis, so both forms of analysis are appropriate for use with these data.

Analysis of Trends

Per linear regression analysis, removals of All Categories without Biomass Burning have significantly declined at a rate of 0.42 MMTCO2e/year. This decline in net sequestration is 0.12 MMTCO2e/year less than for All Categories (with biomass burning), indicating that wildfires have contributed substantially to the trend.

Per Mann-Kendall analysis with Theil-Sen estimator, removals of All Categories without Biomass Burning have significantly declined at a rate of 0.39 MMTCO2e/year. This decline in net sequestration is 0.19 MMTCO2e/year less than for All Categories (with biomass burning), indicating that wildfires have contributed substantially to the trend.

Figure B3. Net removals of All Categories without Biomass Burning. Negative Y-values indicate net sequestration. Presence of trendlines indicates that trends are significant ($p \leq 0.05$). Trendlines increasing over time indicate a significant decline in net removals over time. Linear Regression Trendline indicates the trend determined using linear regression with Ordinary Least Squares analysis; MK Trendline indicates the trend determined using Mann-Kendall with Theil-Sen analysis.



Trends in Forest Land

Forest Land includes carbon transferred from forest biomass to harvested wood products. In the report body, Harvested Wood Products is shown as a separate subcategory in the figure, but in the Inventory and in this analysis they are reported together.

Data Assumptions

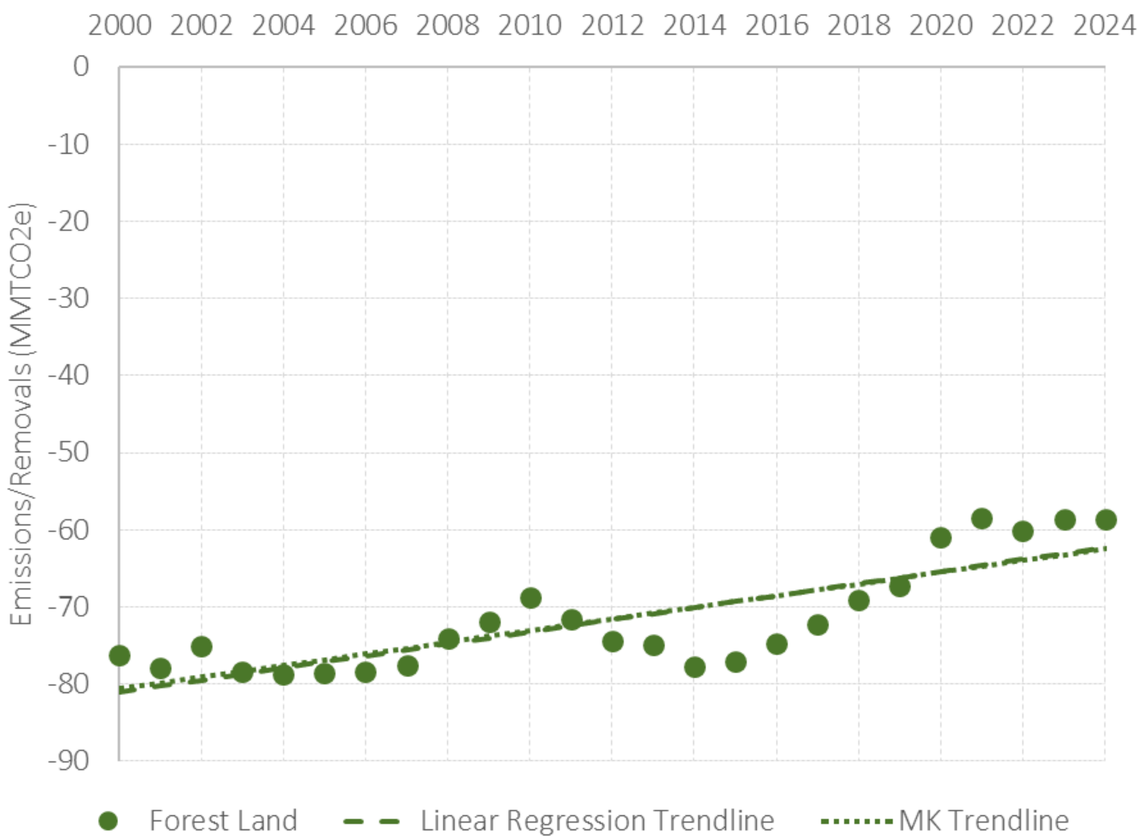
Data for Forest Land satisfied the assumptions for both linear regression and Mann-Kendall analysis, so both forms of analysis are appropriate for use with these data.

Analysis of Trends

Per linear regression analysis, removals of Forest Land have significantly declined over time at a rate of 0.78 MMTCO2e/year.

Per Mann-Kendall analysis with Theil-Sen estimator, removals of Forest Land have significantly declined over time at a rate of 0.76 MMTCO2e/year.

Figure B4. Net removals of Forest Land. Negative Y-values indicate net sequestration. Presence of trendlines indicates that trends are significant (p ≤ 0.05). Trendlines increasing over time indicate a significant decline in removals over time. Linear Regression Trendline indicates the trend determined using linear regression with Ordinary Least Squares analysis; MK Trendline indicates the trend determined using Mann-Kendall with Theil-Sen analysis.



Trends in Developed Land

Data for Developed Land is complicated by underlying Developed Land information that only allowed estimates to be produced every five years, with linear interpolation between these five-year periods. As such, the data and results of the tests of trends presented here should be used with caution.

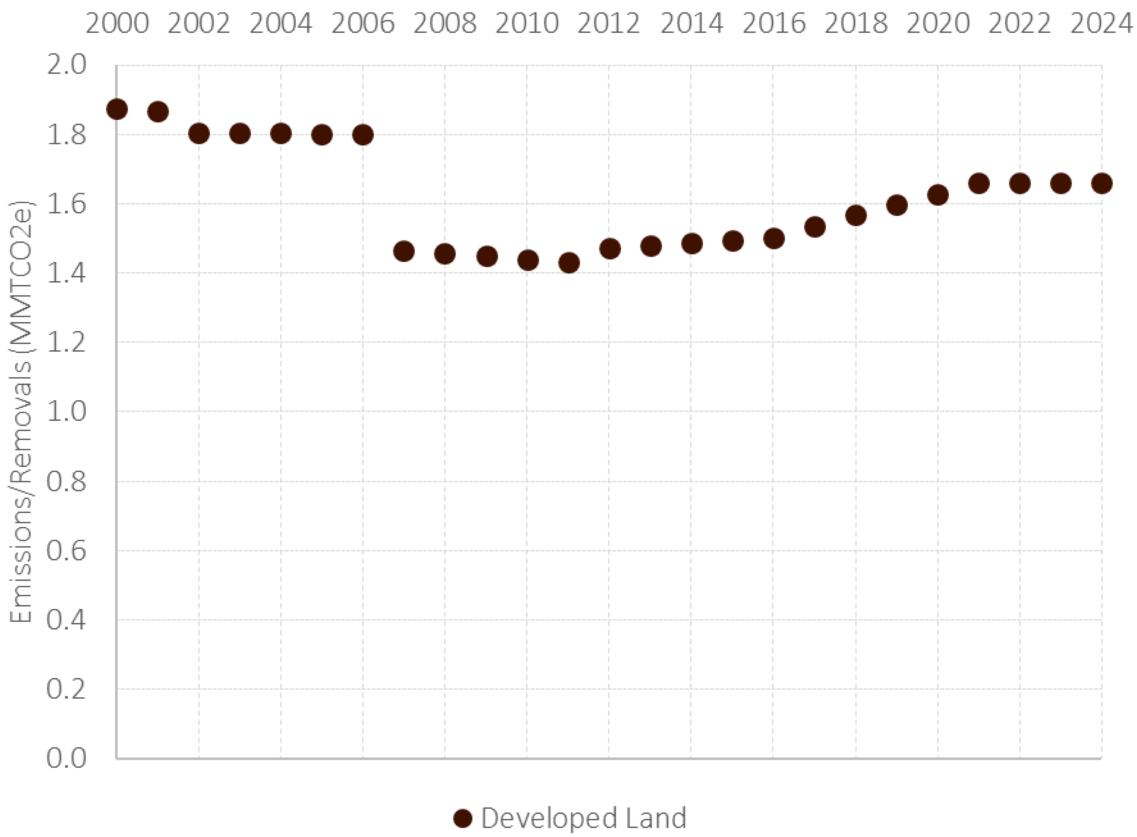
Data Assumptions

Data for Developed Land do not satisfy the assumptions for linear regression; they do satisfy the assumptions for Mann-Kendall analysis. Only Mann-Kendall analysis is appropriate for these data.

Analysis of Trends

Per Mann-Kendall analysis with Theil-Sen estimator, emissions of Developed Land showed no significant trend. The sharp change in value between 2006 and 2007 has been noted by ODOE and the causes will be investigated and addressed where appropriate in future updates to the Inventory.

Figure B5. Net emission of Developed land. Positive Y-values indicate net emission. No trends in the data were found to be significant.



Trends in Wetlands

Data for Wetlands is complicated by underlying Wetlands information that only allowed estimates to be produced every five years, with linear interpolation between these five-year periods. As such, the data and trends presented here should be used with caution.

Data Assumptions

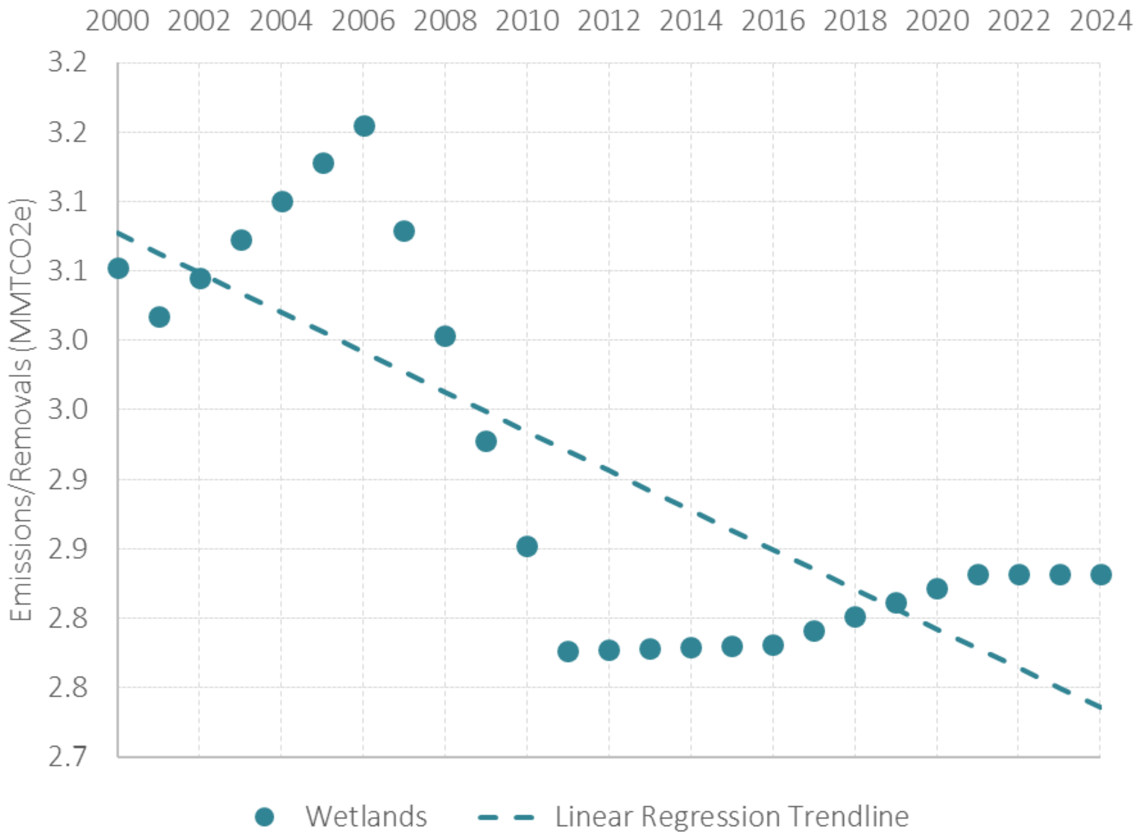
Data for Wetlands satisfied the assumptions for both linear regression and Mann-Kendall analysis, so both forms of analysis are appropriate for use with these data.

Analysis of Trends

Per linear regression analysis, emissions of Wetlands have significantly declined over time at a rate of 0.01 MMTCO2e/year.

Per Mann-Kendall analysis with Theil-Sen estimator, there was no significant trend in emissions of Wetlands. Trend results of linear regression analysis are acceptable for reporting and use.

Figure B6. Net emissions of Wetlands. Positive Y-values indicate emissions. Presence of trendlines indicates that trends are significant (p ≤ 0.05). Trendlines decreasing over time indicate a significant decline in emissions over time. Linear Regression Trendline indicates the trend determined using linear regression with Ordinary Least Squares analysis.



Trends in Grassland

Data Assumptions

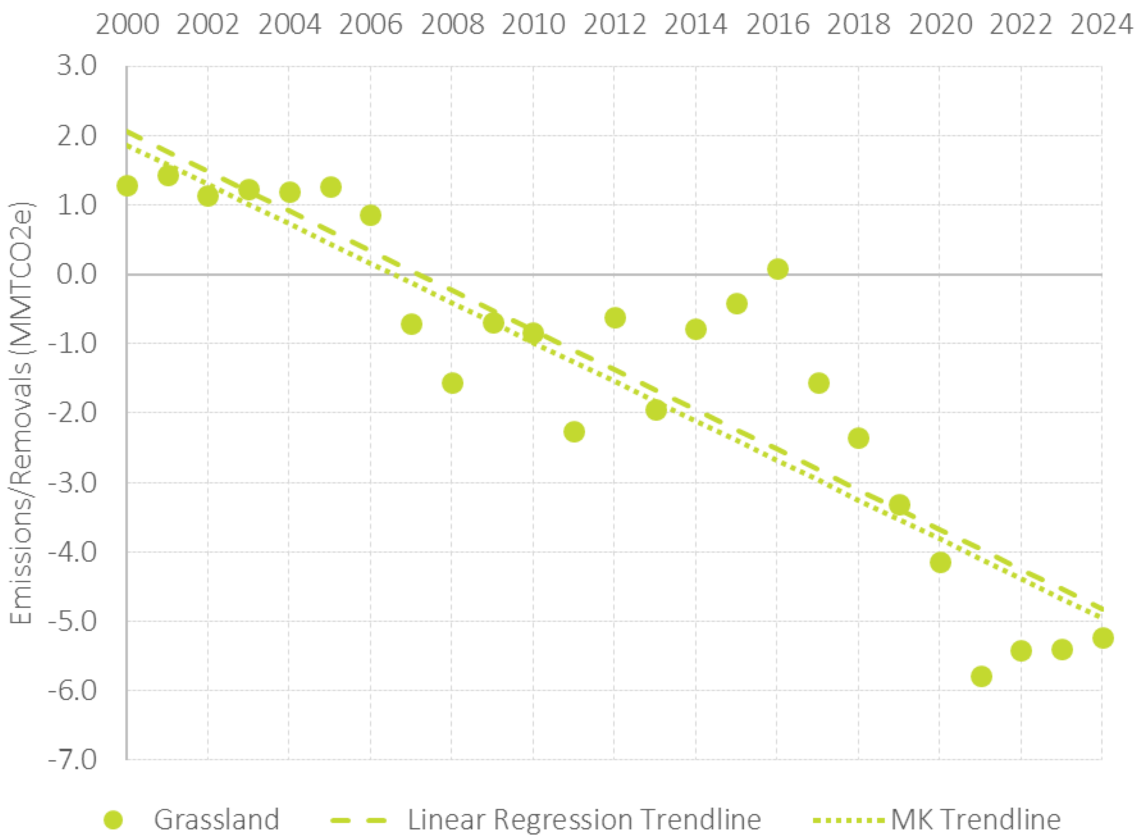
Data for Grassland satisfied the assumptions for both linear regression and Mann-Kendall analysis, so both forms of analysis are appropriate for use with these data.

Analysis of Trends

Per linear regression analysis, emissions of Grassland has significantly declined and removals have increased over time, such that Grassland has converted from a net source to a net sink, at a rate of 0.29 MMTCO2e/year.

Per Mann-Kendall analysis with Theil-Sen estimator, emissions of Grassland has significantly declined and removals have increased over time, such that Grassland has converted from a net source to a net sink, at a rate of 0.28 MMTCO2e/year.

Figure B7. Net emissions and removals of Grassland. Positive y values indicate emissions, negative y values indicate removals. Presence of trendlines indicates that trends are significant ($p \leq 0.05$). Trendlines decreasing over time across the 0 line indicate a significant decline in emissions and increase in removals. Linear Regression Trendline indicates the trend determined using linear regression with Ordinary Least Squares analysis; MK Trendline indicates the trend determined using Mann-Kendall with Theil-Sen analysis.



Trends in Cropland

Data Assumptions

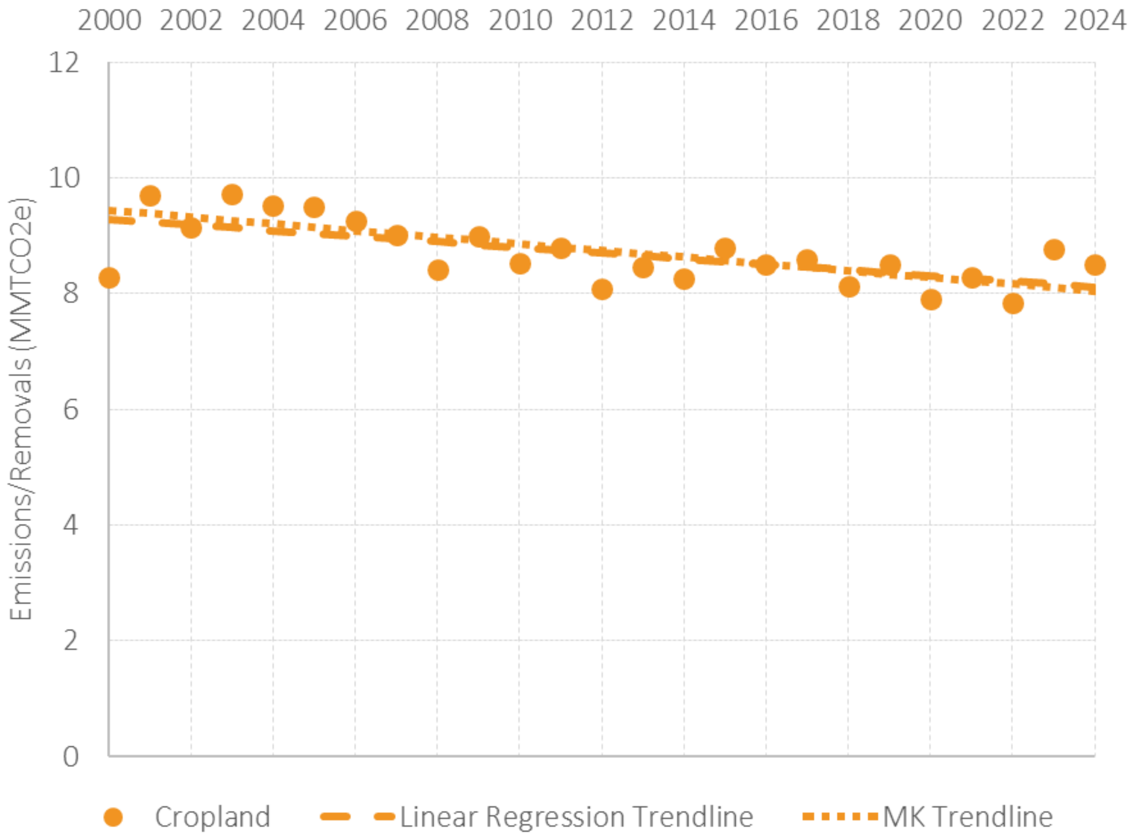
Data for Cropland satisfied the assumptions for both linear regression and Mann-Kendall analysis, so both forms of analysis are appropriate for use with these data.

Analysis of Trends

Per linear regression analysis, net emissions of Cropland have significantly declined over time at a rate of 0.05 MMTCO2e/year.

Per Mann-Kendall analysis with Theil-Sen estimator, net emissions of Cropland have significantly declined over time at a rate of 0.06 MMTCO2e/year.

Figure B8. Net emissions of Cropland. Positive Y-values indicate emissions. Presence of trendlines indicates that trends are significant ($p \leq 0.05$). Trendlines decreasing over time indicate a significant decline in emissions over time. Linear Regression Trendline indicates the trend determined using linear regression with Ordinary Least Squares analysis; MK Trendline indicates the trend determined using Mann-Kendall with Theil-Sen analysis.



Trends in Biomass Burning

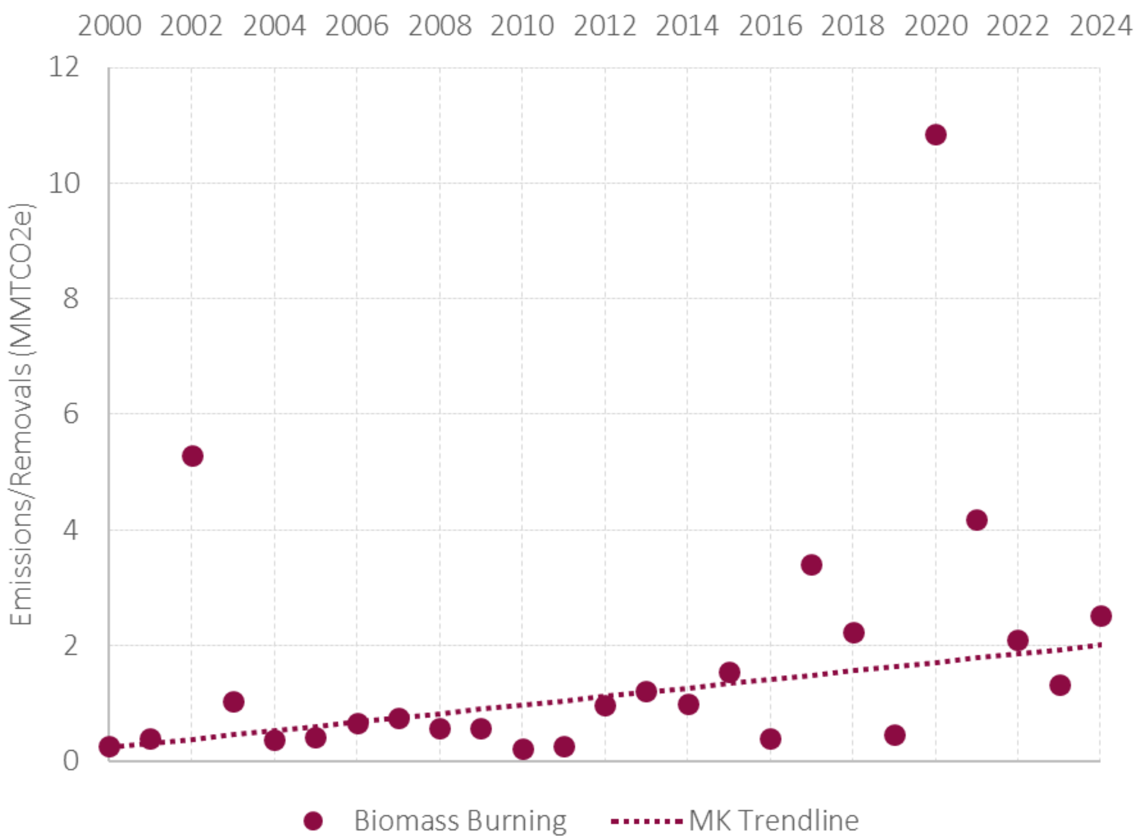
Data Assumptions

Data for Biomass Burning did not satisfy assumptions for linear regression because they are heteroscedastic, specifically that emissions from biomass burning have become increasingly variable over time. The data did satisfy assumptions for Mann-Kendall analysis. Only Mann-Kendall analysis is appropriate for use with these data.

Analysis of Trends

Per Mann-Kendall analysis with Theil-Sen estimator, emissions from Biomass Burning has increased at a rate of 0.07 MMTCO2e/year.

Figure B9. Net emissions of Biomass Burning. Positive Y-values indicate emissions. Presence of trendlines indicates that trends are significant (p ≤ 0.05). Trendlines increasing over time indicate a significant increase in net emission over time. MK Trendline indicates the trend determined using Mann-Kendall with Theil-Sen analysis.



Statistical Analysis of Trends in Time Series Data

Two common statistical methods for testing for trends in time-series data related to emissions and removals are (1) simple linear regressions and (2) Mann-Kendall tests of significance of the trend[i,ii]. Summarily, while some data in the Inventory are appropriate for use in linear regression, other data are not; where data do not meet assumptions for linear regression, findings of significance may be erroneous; Mann-Kendall analysis is generally safer for use in detecting significant trends in data from the Inventory. While other tests for trends exist, and more complex models to estimate trends may be possible to construct (i.e., non-linear models), the analysis of trends was limited to these two methods for lack of better guidance from experts contacted and documents made available by authoritative groups like the U.S. Environmental Protection Agency and the Intergovernmental Panel on Climate Change.

[i] Mann, H.B. (1945). Non-parametric tests against trend. *Econometrica*, *13*(3), 245–259. https://www.jstor.org/stable/1907187

[ii] Kendall, M.G. (1975). *Rank Correlation Methods*, 4th edition, Charles Griffin, London. Rank correlation methods

Linear Regression with Ordinary Least Squares Analysis

Linear regression is a parametric statistical test that requires that several assumptions about the underlying data are met. These assumptions are:

- Linearity: The relationship between the dependent and independent variables is linear.
- Independence: The residuals of the model are independent/not auto-correlated.
- Normality: The residuals of the model should be normally distributed.
- Homoscedasticity: The variance of the residuals is constant across all levels of the independent variable.

All of these assumptions should be met to justify the use of linear regression. This section will focus on two assumptions, normality and homoscedasticity, which are commonly misunderstood and not met before linear regression is applied.

Calculating Residuals

To test for normality and homoscedasticity, residuals must first be calculated. Residuals are distinct from raw values in a dataset. Rather, residuals are calculated using the expected value according to the model being used (in the case of this appendix, a linear model) and the dataset values such that:

$$\textit{Residual}_i = \textit{Expected Value}_i - \textit{Observed Value}_i$$

Where:

- *i* is a unique index used to identify the items in a list of variables;
- Expected Value_i is the value calculated using a model using the *i*th value of the independent variable as input; and
- Observed Value_i is the value in the input dataset, which is the value of the *i*th dependent variable for the corresponding *i*th independent variable.

For example, if the dataset is x = [2000, 2001, 2002, 2003] in units of years, and y = [45, 47, 46, 49] in units of million metric tons of carbon dioxide equivalent (MMTCO2e per year), the best fit linear model for these data is:

$$\text{Expected Value}_i = 1.1 \times x_i - 2154.9$$

This is the equation for the "line of best fit" that software such as Microsoft Excel will automatically generate when asked to produce a linear trendline of these data. Such software, or statistical analysis packages in coding languages like R and Python, determine the equation for this line using a method called Ordinary Least Squares, the details of which will not be reviewed in this Appendix.

In this equation, Expected Value_i is the value calculated using the linear model. It is the same Expected Value_i as in the equation for the residuals, and x_i is the value of the *i*th independent variable.

Using the linear model and example datasets provided above, if $i = 1$:

$$\text{Expected Value}_1 = 1.1 \times (2000) - 2154.9 = 45.1$$

Where 2000 is the 1st value of the independent variable (i.e., the list of years provided in the example dataset) and 45.1 is the calculated Expected Value_1. Using this value of 45.1 as Expected Value_1 in the equation for calculating the 1st residual, we would calculate the 1st residual as:

$$\text{Residual}_1 = 45.1 - 45 = 0.1$$

Where 45 is the 1st value of the dependent variables, which is also Observed Value$_1$. Thus, the predicted value from the model (45.1) for this 1st independent variable is off by 0.1 from the corresponding 1st Observed Value (45). The values of second, third, and fourth residuals can be similarly calculated by using the second, third, and fourth values of the independent variable (2001, 2002, and 2003, respectively) and observed values in the equations above.

Normal Distribution of Residuals

Linear regression assumes that the residuals of the observed values are normally distributed. A normal distribution is also commonly referred to as a "bell curve," where:

- There is roughly an equal number of residuals above and below the mean of the residuals;
- Roughly 68% of the residuals should be within one standard deviation of the mean of the residuals;
- Roughly 95% of the residuals should be within two standard deviations of the mean of the residuals; and
- Other characteristics not listed here.

Use of linear regression is not appropriate when residuals are not normally distributed. As demonstrated, the residuals are not the same as the observed values, so *it is not appropriate to use linear regression when the observed values themselves are normally distributed.*

The Shapiro-Wilk test is commonly used to test for significant deviation of residuals from a normal distribution.[iii] The null hypothesis of the Shapiro-Wilk test is that the input data (i.e., the residuals) are normally distributed; if $p > 0.05$ the test "fails" to reject the null hypothesis and the residuals do not significantly deviate from a normal distribution; if $p \leq 0.05$, then the null hypothesis is rejected and the data significantly deviate from a normal distribution. Thus, if $p \leq 0.05$ in the Shapiro-Wilk test, the dataset used may be appropriate for use of linear regression and Ordinary Least Squares. The full details of conducting a Shapiro-Wilk test are beyond the scope of this document but it can be calculated manually or using packages in coding languages commonly used for statistics like R or Python.

Homoscedasticity of Residuals

Another key assumption to meet before using linear regression with Ordinary Least Squares is homoscedasticity of the residuals, or the assumption that the variance of the residuals is roughly the same for all values of the independent variable. In contrast to *homo*scedasticity, *hetero*scedasticity describes a situation where residuals have more or less variance for some values of the independent variable compared to others. In terms relevant to the analysis of the Inventory results, *homo*scedasticity might be observed in datasets of annual wildfire emissions during a period where wildfire occurrence and intensity is stable over time, steadily increasing, or steadily decreasing; *hetero*scedasticity might be observed in annual wildfire emissions during a period when where wildfire becomes increasingly erratic over time (e.g., early years with low emissions, later years with extremely high or extremely low emissions).

Commonly, homoscedasticity is assessed visually by examining plots showing the distribution of residuals against a "fitted line" of the expected values. Homoscedasticity is observed when residuals are randomly distributed; heteroscedasticity is observed when residuals are distributed with clear patterns like clumping or spreading out over parts of the data. As an alternative to visual assessment, the

[iii] Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika, 52*, 591-611.
https://doi.org/10.1093/biomet/52.3-4.591

Breusch-Pagan test may be used to test for the presence of homoscedasticity using packages in coding languages commonly used for statistics like R or Python, but the details of this test are beyond the scope of this Appendix.[iv] The Breusch-Pagan test was used in this analysis to test that the assumption of homoscedasticity of datasets was met.

If residuals are not normally distributed and homoscedastic, findings of significant trends in the observed values via linear regression and Ordinary Least Squares may be erroneous. Later in this document, analysis of these assumptions is provided for the net flux of all categories in the Inventory as well as individual categories.

Mann-Kendall with Theil-Sen Analysis

The Mann-Kendall test is a non-parametric statistical test of the presence and direction of monotonic trends, which occur when there is a consistent increase or decrease of the dependent variable in a dataset. The Mann-Kendall test offers some advantages to general use with Inventory data over linear regression and Ordinary Least Squares Unlike linear regression, primarily due to not having to meet many of the same assumptions:

- Linearity: The Mann-Kendall test *does not* assume the relationship between independent and dependent variables is linear—observed values may increase or decrease in any way, such as exponentially or logarithmically, and thus does not require the assumption of linearity be met.
- Normality: The Mann-Kendall test does not assume normality of residuals.
- Homoscedasticity: The Mann-Kendall test *does not* assume homoscedasticity.

Similar to the assumption of independence in linear regression and Ordinary Least Squares, the Mann-Kendall test assumes independence of data, or that there is no auto-correlation in the observations. Broadly, this means that an observed value should not have strong dependence on the observed values around it. Auto-correlation is commonly a concern when observations are collected across seasons, where for example higher rates of sequestration would be expected in summer than in winter. However, this is not a concern in the annual data in the Inventory used to analyze trends in categories.

The Mann-Kendall test is generally more robust than linear regression for use in detecting trends in Inventory data because the underlying data do not need to meet an assumption of normality and homoscedasticity of residuals. However, there are instances where linear regression may be both preferable and provide better analysis that are not reviewed here.

As Mann-Kendall only detects the presence of significant trends and their general direction (i.e., increase and decline), follow-up statistical analysis such as the Theil-Sen estimator is commonly used for determining a line of best fit for the data, similar to a line produced using Ordinary Least Squares analysis after linear regression detects a significant trend.[v,vi] Thiel-Sen analysis provides an estimate of the *amount* of decline or increase in the data (e.g., an increase of 0.5 MMTCO2e/year). The Mann-Kendall test with Theil-Sen estimator can be calculated manually or using packages in coding languages commonly used for statistics like R or Python, but the detailed methods are beyond the scope of this Appendix.

[iv] Breusch, T. S., and Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, *47*(5), 1287–1249. https://doi.org/10.2307/1911963

[v] Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis.

[vi] Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, *63*(324), 1379-1389. https://doi.org/10.1080/01621459.1968.10480934