Prepared for:
Elena McDonald-Buller
Project Manager
Texas Air Quality Research Program

Prepared by:
Tejas Shah, Ling Huang, Lynsey Parker, Yuge Shi, Chris Emery and Greg Yarwood
Ramboll US Consulting, Inc.

Alex Guenther
1 Cervantes Ct
Irvine, California

August 2021

Texas Urban Vegetation BVOC Emission Source Inventory

Final Report

AQRP Project No. 20-007

QA Requirements: Audits of Data Quality: 10% Required

RAMBOLL

Texas Urban Vegetation BVOC Emission Source Inventory

Ramboll
7250 Redwood Boulevard
Suite 105
Novato, CA 94945
USA

T +1 415 899 0700
https://ramboll.com

Contents

Appendices

Table of Figures

Table of Tables

ACKNOWLEDGEMENT

The preparation of this report (Project No. 20-007) was funded by a grant from the Texas Air Quality Research Program (AQRP) at The University of Texas at Austin through the Texas Emission Reduction Program (TERP) and the Texas Commission on Environmental Quality (TCEQ). The findings, opinions and conclusions are the work of the author(s) and do not necessarily represent findings, opinions, or conclusions of the AQRP or the TCEQ.

LIST OF ACRONYMS AND ABBREVIATIONS

AQRP	Air Quality Research Program
BDSNP	Berkley Dalhousie Soil NO Parameterization
BEIS	Biogenic Emission Inventory System
BELD5	Biogenic Emissions Landuse Database, Version 5
BVOC	Biogenic Volatile Organic Compound
CGL	Copernicus Global Land
CAMx	Comprehensive Air Quality Model with Extensions
CMAQ	Community Multiscale Air Quality
DBH	Diameter at Breast Height
EC	Emission Calculator
ECD	ESRI Classifier Description
EFP	Emission Factor Processor
ESA	European Space Agency
FIA	Forest Inventory Assessment
LAI	Leaf Area Index
MEGAN	Model of Emissions of Gases and Aerosols from Nature
MODIS	Moderate Resolution Imaging Spectroradiometer
MSTI	Municipal Street Tree Inventories
NAIP	National Agriculture Imagery Program
NLCD	National Land Cover Dataset
NO	Nitric Oxide
PAR	Photosynthetically Active Radiation
QA/QC	Quality Assurance/Quality Control
QAPP	Quality Assurance Project Plan
S2/MSI	Sentinel-2 satellite Multispectral Instrument
SL2P	Simplified Level 2 Product Prototype Processor
SNAP	Sentinel Application Platform
TAMU	Texas A&M University
TCEQ	Texas Commission on Environmental Quality
TERP	Texas Emission Reduction Program
TROPOMI	TROPOspheric Monitoring Instrument
uFIA	Urban Forest Inventory Assessment
USDA	US Department of Agriculture
UTC	Urban Tree Cover
VCF	Vegetation Cover Fraction
VOC	Volatile Organic Compound
VUTS	Virtual Urban Tree Surveys
WRF	Weather Research Forecast

EXECUTIVE SUMMARY

Isoprene and other biogenic volatile organic compound (BVOC) strongly influence atmospheric chemistry in Texas urban areas and can dominate the total VOC reactivity of at least some Texas urban locations. Urban areas are the most challenging for BVOC emissions estimation, due to heterogeneity and a lack of vegetation information, and yet they continue to be the least studied. Recent ground surveys of urban tree inventories and increasingly higher resolution remote sensing data products have substantially improved the potential for characterizing the landcover inputs required for biogenic emission models.

The overall goal of Texas Air Quality Research Program (AQRP) Project 20-007 was to improve numerical predictions of regional ozone and aerosol distributions in Texas by developing more accurate estimates of BVOC emissions in Texas urban areas. To accomplish this, we used urban tree inventories and aerial and satellite imagery to develop fine spatial resolution (~1 km) time-varying Leaf Area Index (LAI), total vegetation cover, and the relative abundance of high BVOC-emitting trees (e.g., live oaks, deciduous oaks, sweetgum, palms, pines, juniper) and other vegetation cover types for three Texas urban areas: San Antonio, Austin and Houston. We evaluated sensitivity of Texas urban biogenic emissions to the updates developed in this project. Results from this study were incorporated into a new version 3.2 of the Model of Emissions of Gases and Aerosols from Nature (MEGAN3.2).

The overall benefit of this project is more accurate VOC emission estimates for the Texas air quality simulations that are critical for scientific understanding and the development of regulatory control strategies that will enhance efforts to improve and maintain clean air.

1.0 INTRODUCTION

This document provides the final report for the Texas AQRP Project 20-007, "Texas urban vegetation BVOC emission source inventory". The project Co-Principal Investigators (Co-PIs) are Mr. Tejas Shah and Dr. Greg Yarwood of Ramboll and Dr. Alex Guenther. The AQRP project manager is Dr. Elena McDonald-Buller at the University of Texas, Austin. The project liaison for the Texas Commission on Environmental Quality (TCEQ) is Ms. Miranda Kosty.

The overall goal of Project 20-007 was to improve numerical predictions of regional ozone and aerosol distributions in Texas by using more accurate estimates of BVOC emissions in Texas urban areas. The overall benefit of this project is more accurate VOC emission estimates for the Texas air quality simulations that are critical for scientific understanding and the development of regulatory control strategies that will enhance efforts to improve and maintain clean air.

1.1 Background

Isoprene and other BVOC strongly influence atmospheric chemistry in Texas urban areas and can dominate the total VOC reactivity of at least some Texas urban locations. Urban areas are the most challenging for BVOC emissions estimation, due to heterogeneity and a lack of vegetation information, and yet they continue to be the least studied. Recent ground surveys of urban tree inventories and increasingly higher resolution remote sensing data products have substantially improved the potential for characterizing the landcover inputs required for biogenic emission models. The overall goal of this project was to improve numerical predictions of regional ozone and aerosol distributions in Texas by developing more accurate estimates of BVOC emissions in Texas urban areas.

1.2 Overview of Approach

The project aimed to improve the spatial representation of biogenic emissions in Texas urban areas by using fine resolution satellite imagery. This was accomplished by developing an urban LAI and high BVOC tree fraction inventory for Texas by synthesizing ground survey data with aerial and satellite imagery. Our approach captured spatial variation in urban biogenic emissions with ~1 km resolution. These data were then processed to formats needed by the MEGAN model.

Our specific objectives included:

1. To quantify high resolution (8 day, 10 m) LAI and vegetation cover fraction data for three major Texas urban areas: Houston, San Antonio, Austin.

2. To quantify high resolution (10 m) vegetation distributions of high BVOC emitting trees (e.g., live oaks, deciduous oaks, sweetgum, palms, pines, juniper) in the three urban areas.

3. To compile and assess vegetation characteristics input data for both MEGAN and Biogenic Emission Inventory System (BEIS) models, update MEGAN Emission Factor Processor to improve processing of urban and other landcover data and investigate sensitivity of Texas urban biogenic emissions to landcover inputs.

1.3 Overview of Report

In Section 2, we describe the development of high-resolution LAI and vegetation cover fraction (VCF) data for three major Texas urban areas using high resolution satellite imagery. In Section 3, we provide an overview of our approach to quantify tree cover distributions and tree species identification in the three major Texas urban areas. Section 4 describes a sensitivity analysis of Texas biogenic emissions. In Section 5, we present conclusions and recommendations for future work. Finally, Section 6 contains results of the data quality audits.

2.0 TASK 1: LAI AND VEGETATION COVER FRACTION

LAI and VCF data are used to represent the total amount of vegetation foliage for estimating BVOC emissions with the MEGAN model. LAI is defined as one half (i.e., just one side of a leaf) of the total leaf area per unit ground surface area and has units of m2 m-2. The LAI data used for MEGAN is green LAI which can be quantitatively defined as green leaf elements that have a leaf chlorophyll content > 15 ug/m2. VCF is the fraction of total area that is covered by green canopy elements as seen from the nadir (looking directly down at the canopy). VCF is the sum of all vegetation cover and is a constraint on the sum of individual vegetation cover types (growth forms) which include tree, shrub, grass/herbaceous and crop in the MEGAN model.

LAI and VCF input for MEGAN are typically derived from satellite images although they can also be predicted using dynamic vegetation models. LAI estimates based on satellite images are indirect measurements that require a retrieval method (model) which is based on empirical data and/or a radiative transfer model. For this project, we used the Simplified Level 2 Product Prototype Processor (SL2P) to estimate LAI and total VCF from the Sentinel-2 satellite Multispectral Instrument (S2/MSI) 10 m resolution data. Sentinel consists of a set of two satellites, Sentinel 2A launched in 2015 and Sentinel 2b launched in 2017, which together have about ~5-day temporal resolution. The SL2P uses a neural network approach trained with a globally representative set of simulations from the PROSAILH canopy radiative transfer model (Weiss and Baret, 2000). The S2/MSI data processing includes radiometric and geometric correction using ground control points and a digital elevation model to correct for parallax error. The broad swath width (290 km) of each S2/MSI tile spans across entire urban regions while the high resolution (10 m) can accurately capture tree cover in an urban setting (Wong et al. 2019). European Space Agency's (ESA) Sen2Cor processor was used for cloud and shadow removal and to convert Sentinel-2 level 1 (MSIL1C) data to atmospherically corrected top-of-canopy reflectance data. We selected relatively cloud free images for processing and visually inspected each image to assess quality and geolocation errors. The S2/MSI satellite data was accessed from the Copernicus Open Access Hub (https://scihub.copernicus.eu/). A full annual cycle (January to December 2019) of 10-m resolution LAI was generated for each of the three urban areas (Austin, Houston, San Antonio) by interpolating in time between the available images. The area processed was 40,000 km^2 (which resulted in 400 million locations at 10 m resolution) including 10,000 km^2 each for Austin and San Antonio 20,000 km^2 for Houston. LAI was estimated for 46 8-day periods for each location for a total of 18.4 billion (=46 x 400,000,000) LAI estimates.

A tutorial describing the approach is given in Appendix 1 to enable users to process data for additional urban areas and additional years. Appendix 1A describes how to access the data and process individual tiles using ESA's Sentinel Application Platform (SNAP) toolbox. Appendix 1B describes how to process individual tiles to generate MEGAN LAI inputs of 8-day or 10-day average data using ArcGIS and integrate SNAP-derived LAI with global LAI data.

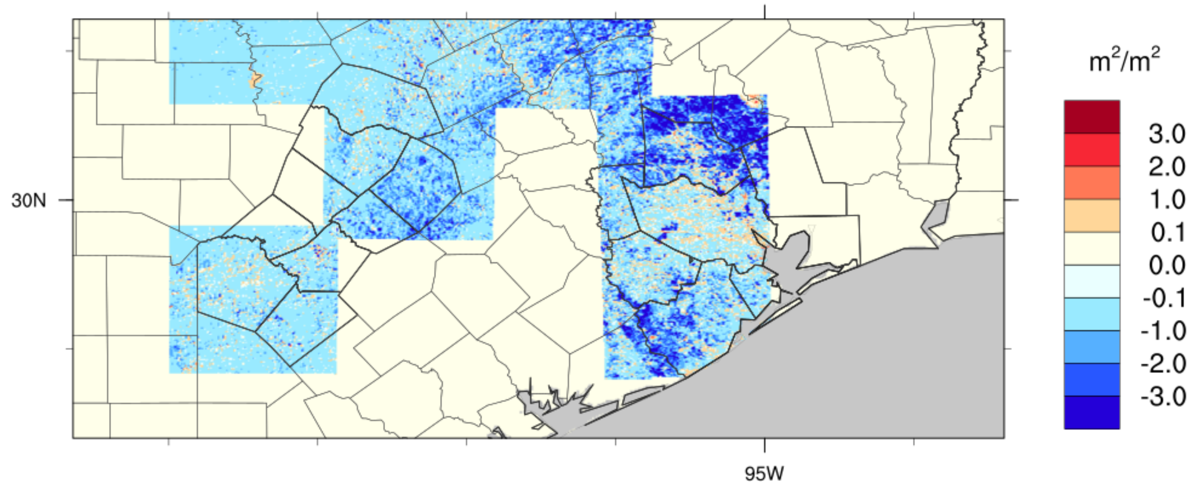


Figure 2-1 shows a spatial plot of change in LAI for the three urban areas and compares the SNAP-derived LAI from Sentinel-2 with a global LAI product. The Sentinel SNAP estimates are a bit lower than the global data but showed overall agreement with observed (in-situ) LAI (Kganyago, 2020).

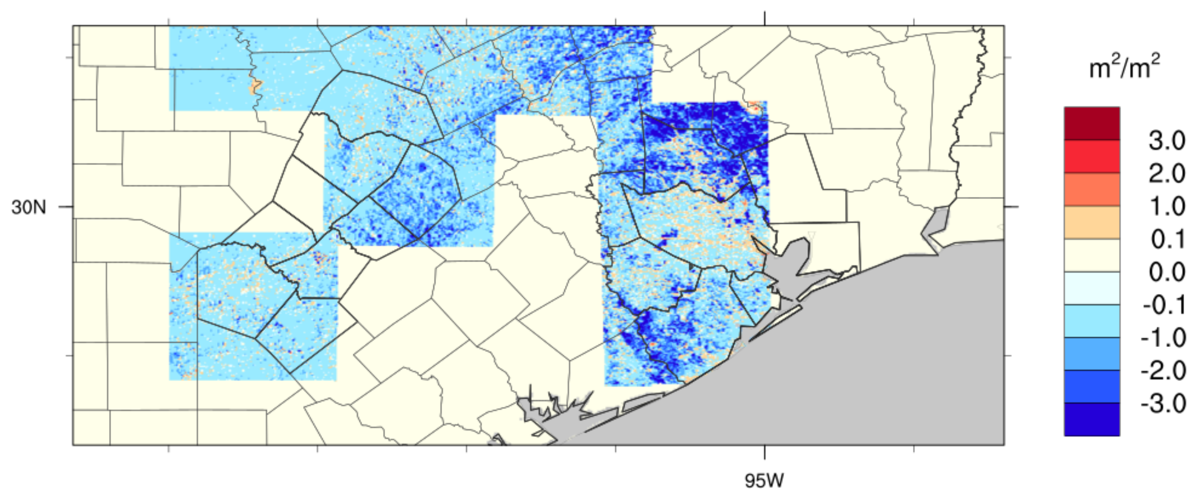


Figure 2-1. Change in LAI between SNAP-derived estimates and global LAI product.

3.0 TASK 2: URBAN TREE CHARACTERIZATION

The LAI acquired for task 1 is a necessary but insufficient step for estimating BVOC emissions because of the enormous diversity in the BVOC emission rates of different types of vegetation. There are substantial emission rate differences among different growth forms (e.g., trees, shrubs, grass, crops) and various plant species that make up a growth form category. Trees are expected to be the dominate BVOC emission source in urban areas and are the focus of this project. The MEGAN model inputs include global databases that quantify the spatial distributions of growth form fractions and landscape types and a database that contains the vegetation speciation of each growth form in every landscape. Section 3.1 describes the approach for quantifying tree cover distributions and section 3.2 describes the tree speciation methods and results.

3.1 Tree cover fraction

The fraction of the earth surface covered by general vegetation growth forms (e.g., trees, crops, grass and other herbaceous plants, shrubs) required for input to the MEGAN BVOC emission model has previously been generated using global satellite imagery with 500 to 1000 m spatial resolution (Guenther et al. 2012). We have updated the MEGAN global growth form database (generation 4) using the 100 m resolution Copernicus Global Land (CGL) Service landcover products version 3 dataset (see land.copernicus.eu/global/products/lc) based on imagery from the PROBA-V satellite. The CGL data are available as an integrated global dataset for 2015, with annual updates after that, on global scale (78.25N to 60S) at a resolution of 0.0099206 (~100 m at the equator) and with an accuracy of 80% when compared to 28,000 independent validation points (Buchhorn et al. 2020). Our assessment of the CGL tree cover fraction indicated that it is a significant improvement over the previous MEGAN global tree cover product, but tree cover was still underestimated in heterogeneous landscapes, including urban areas and arid woodlands. The higher resolution (30-m) National Land Cover Dataset (NLCD) tree cover data, available only for the contiguous US, was able capture some of these missing trees and so was integrated into the CGL data (i.e., the higher NLCD values were used for urban areas and woodlands). However, our assessment using the i-Tree tool, shown in Table 3-1, indicates that both the CGL and NLCD miss a substantial fraction of tree cover in the heterogeneous urban landscape. Since most mature trees have a crown of several meters or more, individual trees can be identified, and the crown area quantified, using imagery with a spatial resolution of 1 m or less (Figure 3-1).

Figure 3-1. Tree and other ground cover distribution at the Lions Municipal Golf Course in Austin Texas. NAIP natural color image (Left) and classified cover (Right) demonstrate that the 60 cm resolution is sufficient to quantify distributions of groves and individual trees.

This project used imagery from the National Agriculture Imagery Program (NAIP) which acquires aerial imagery by aircraft in the contiguous U.S. Detailed descriptions of the specific steps used to quantify urban tree cover are described in Appendix 2 and the approach to assess tree cover in Appendix 3.

The NAIP imagery used for this project was acquired from the ArcGIS Portal Living Atlas (https://livingatlas.arcgis.com/) for the year 2018 and has a spatial resolution of 60 cm. The areas characterized for this project include Austin (~9.3 billion pixels covering 1496 km^2), San Antonio (~7.2 billion pixels covering 1360 km^2) and Houston (~11.5 billion pixels covering 2135 km^2). The NAIP imagery can be used to generate urban tree cover data sets throughout the contiguous US for specific years starting in 2003 although some earlier NAIP imagery has lower resolution (e.g., 1 to 2 m). The "ArcGIS USA NAIP Imagery: Color Infrared product" was used which consists of three bands: near infrared and the visible green and blue bands). The ArcGIS Pro "Segment Mean Shift" geoprocessing tool was used to segment each NAIP image into discrete objects (e.g., individual features such as buildings, trees, lawns, lakes). An ESRI Classifier Description (ECD) file was then created that typically has about 12 to 15 landcover categories including several types of trees (e.g., juniper, oaks, pines, other deciduous broadleaf trees), ground covers (e.g., shrub, green grass, brown grass, bare soil, crops) and non-vegetated (e.g., water, asphalt, colored rooftops) landscapes. Appendix 2 describes how to create an ECD file. Training samples were generated according to the procedures described in Appendix 2 and used to classify the objects using the ArcGIS "Random Tree" machine learning geoprocessing tool. Images were then classified using the ArcGIS "ClassifyRaster" geoprocessing tool which assigned a specific landcover type (e.g., oak tree) to each object. The ArcGIS "RECLASS" tool was then used to assign a tree cover estimate to each 0.6 x 0.6 m location.

The resulting tree cover distribution digital maps were assessed using reported city-average tree cover (fraction of total landscape covered by tree crown) statistics and by conducting tree surveys using the i-Tree program developed by the US Department of Agriculture (USDA). The i-Tree program is based on random sampling of hundreds of points, within a specified domain such as a municipal boundary, on a google Earth image with operator identification of objects at the sample location to classify landscape features. Three or more i-Tree surveys were conducted by different members of the project team in each of the three cities. Each survey sampled 300 or more random points resulting in 900 or more sampled points in each of the three cities. The i-Tree tree cover surveys conducted by various team members agreed within 5 to 10% of the mean tree cover fraction, which was 0.408 for Austin, 0.316 for San Antonio and 0.271 for Houston. In addition to assessing the tree cover distributions generated for this project, the i-Tree observations were used to assess three other tree cover distribution databases: the Copernicus Global Landcover (CGL) tree cover data based on 100 m PROBA-V satellite imagery, the NLCD tree cover data based on 30 m Landsat satellite imagery, and the Texas A&M University (TAMU) Urban Tree Cover (UTC) tree cover distribution data, downloaded from https://texasforestinfo.tamu.edu/utc/, which was derived from NAIP imagery using an approach similar to that used for this project.

Table 3-1. Comparison of tree cover fraction estimates for 3 Texas cities estimated by i-Tree observations and by four image classification approaches.

Estimate	Year	Approach	Austin	San Antonio	Houston
i-Tree range	2020	random point survey	0.385 to 0.4423	0.301 to 0.340	0.256 to 0.303
i-Tree average	2020	random point survey	0.408	0.316	0.271
AQRP 14-016	2011	30 m imagery	0.302	0.204	0.264
CGL	2015	100 m imagery	0.255	0.190	0.178
NLCD 2016	2016	30 m imagery	0.300	0.152	0.117
UTC	2018	0.6 m imagery	0.349	0.317	Not available
This work	2018	0.6 m imagery	0.362	0.294	0.289

The uncertainty estimates associated with the i-Tree tree cover approach were typically about 5 to 10% of the mean values which agreed well with the variability we observed among values computed by different team members. In comparison to the i-Tree observations, the 100-m CGL data, which does not capture individual or small groups of trees, consistently underpredicted tree cover by 35 to 40% in all three cities. The 30-m NLCD 2016 data underpredicted tree cover by 52 to 57% in San Antonio and Houston but only 27% in Austin. A previous tree cover distribution developed for AQRP project 14-1016, based on NLCD 2011 tree cover data and other data sets, underpredicted San Antonio and Austin tree cover by 25% and 35%, respectively, but surprisingly was only 3% lower for Houston. The NAIP imagery tree cover product developed for this project was within the uncertainties (5 to 11%) of the i-Tree estimates for all three cities with results that ranged from a 6% overprediction for Houston to an 11% underprediction for Austin. The results were similar (within ~5%) for the TAMU UTC data that were derived using a similar NAIP imagery classification approach. These results indicate that although the global CGL and regional NLCD tree cover data have acceptable accuracy for relatively homogeneous (on a spatial scale of hundreds of meters) landscapes, they have a significant bias in urban areas where there are many isolated trees. As a result, the tree cover characterization approach used for this project can significantly improve tree cover estimates for Texas urban landscapes.

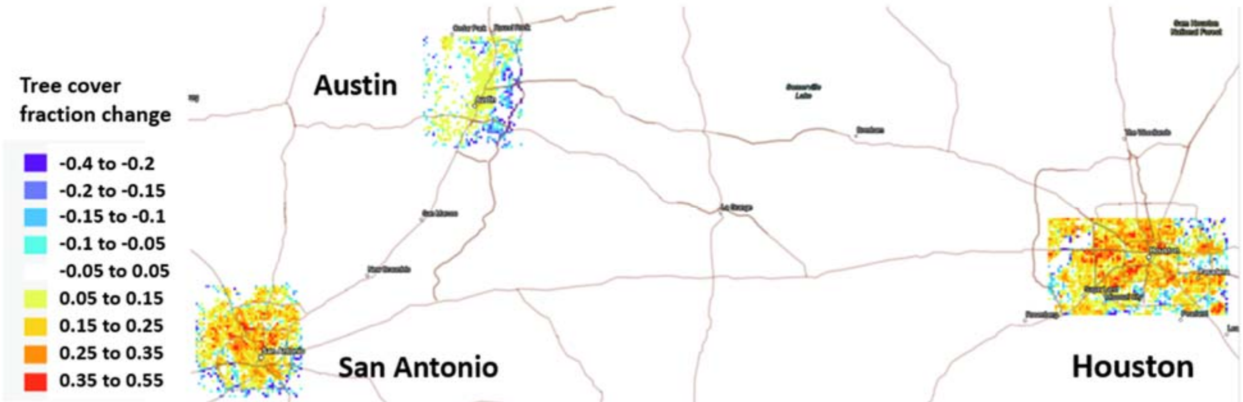


Figure 3-2. Change in percent tree cover between global/regional estimates and urban tree cover estimation approach.

The city average values compiled in Table 3-1 indicate that the NAIP (0.217 to 0.408) and CGL (0.178 to 0.255) tree cover estimates are within 40% but the differences for regions within the cities differed by 90% or more. Figure 3-2 shows that the tree cover fraction was typically 0.15 to 0.55 higher in the central core of these cities where the CGL tree cover fractions were typically ~0.05 while the NAIP imagery-based values were often 0.2 to 0.5. In contrast, the tree cover fraction estimates of the edges of the urban areas were often 0.2 to 0.4 higher in the CGL database. The higher NAIP imagery-based tree cover in the urban core likely indicates the ability of the NAIP imagery to capture individual trees and small groves. Urban development can lead to rapid losses in tree cover which may explain some of the decrease in the 2018 NAIP tree cover compared to the 2015 CGL tree cover in suburban areas.

After achieving an acceptable tree cover accuracy (which we defined as being within 12%), the 60 cm digital tree cover map was aggregated, using the ArcGIS AGGREGATE function, to the NLCD 30 m grid and integrated into the NLCD USA digital tree cover map. These data were then aggregated to ~ 1 km and inserted into the MEGAN global tree cover input file. The other global MEGAN growth form files (shrub, herbaceous, crop, barren) were adjusted so that the total cover did not exceed 100%.

3.2 Tree speciation

The emission factors of some BVOC, such as isoprene, vary by more than order of magnitude for different tree species. Oaks and Sweetgum, for example, have a high isoprene emission rate while other trees, such as Elm, Juniper and Pecan have negligible isoprene emissions. This requires quantifying tree species composition for each landscape to determine the fraction of each emission type.

Three quantitative urban tree species composition approaches were used for this study: urban Forest Inventory Assessment (uFIA) data (https://www.fs.fed.us/research/urban/fia.php) collected by the US Forest Service, Municipal Street Tree Inventories (MSTI) compiled for city governments (https://koordinates.com/layer/25245-houston-texas-street-tree-inventory/, https://koordinates.com/layer/101487-city-of-austin-texas-downtown-tree-inventory-2013/), and Virtual Urban Tree Surveys (VUTS) conducted for this project. Each of these approaches are described briefly below with additional details provided in Appendices.

- The uFIA data is available for only a few US cities but fortunately Houston, Austin and San Antonio all have available data. The uFIA approach uses ground surveys of 1/6 acre plots that are randomly selected within each of the main landuse types in the city. There was one plot selected per ~3 square miles so that Houston (640 square miles), for example, has about 200 plots. Every tree in the plot is identified and the Diameter at Breast Height (DBH) of each tree is measured. With an average of ~8 trees per plot in Houston, this represents a total of ~1,600 trees sampled. The crown cover area was estimated from these data using the equations described by Geron et al. (1994) and the species composition of the landscape was estimated from the relative crown cover area.

- MSTI inventories are conducted by cities to quantify trees on city property. Inventories were available for all of Houston (~200,000 trees throughout Houston) and part of Austin (~7,300 trees in downtown Austin). The number of trees in the MSTI data is 100 times more than the number of trees sampled for uFIA but the MSTI data is only representative of city streets and which is only ~15% of the total the urban area and can differ significantly from the city average tree species composition.

- The VUTS species composition was based on several hundred randomly selected trees within each city urban area. The resulting data is representative of each entire city but represents relatively few sampling points. The procedures for selecting the trees are described in Appendix 3 and the approach for identifying them using high resolution aerial imagery and, where available, Google Street View images is described in Appendix 4 using the Texas urban tree key in Appendix 5.

MEGANv3.0 and MEGANv3.1 have four urban classes for the US that are based on the LANDFIRE landcover database[1]. The four categories are low intensity (i.e., few built structures), medium intensity, high intensity (i.e., highly built up) and roads. The tree composition for each urban category is based on trees observed in FIA plots that fell within these LANDFIRE landcover types and are averaged over the entire US. Figure 3-3 shows that the assigned composition is nearly the same for all four urban types. The tree composition is based on a US

[1] http://www.landfire.gov/

average is generally reflective of US forests but is not representative of any specific US urban forests.

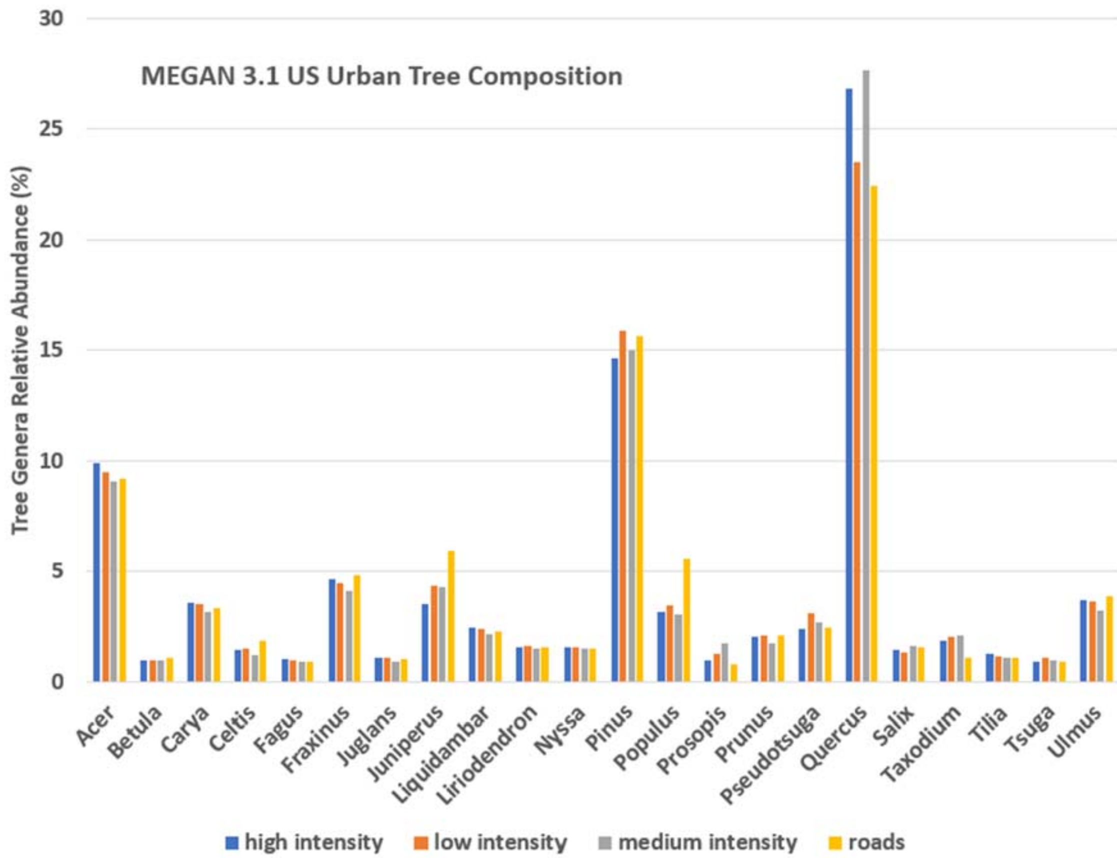


Figure 3-3. MEGANv3.1 tree composition for US urban landcover classes.

Based on the uFIA data there were 21 tree genera that comprised 94% of all trees in Houston, Austin and San Antonio. Nine of these genera comprise 86% of all trees: Oaks, Elms, Hackberries, Hickories (e.g., pecan), Ashes, Pine, Chinese Tallow, Juniper, Mesquite. This includes five genera that together comprise 57% of all trees and individually are present at 2 to 37% of the trees in the three cities: Oaks, Elms, Hackberries, Pecan and Hickories, and Ashes. Another 4 genera comprise 29% of all trees and individually can be found at high (> 8%) or low (< 0.2%) levels: Pine and Chinese Tallow are high only in Houston, Junipers are high in Austin and San Antonio and Mesquite is high only in San Antonio. A third group of genera were consistently found at 1 to 2% in all three cities: Chinaberry, Privet, and Crape Myrtle. Another seven genera were 1 to 3% in only some cities: Willow (Houston and San Antonio), Sycamore (Austin), Acacia and Persimmon (San Antonio), Sweetgum, Mulberry, Yaupon (Houston). The remaining genera were less than 0.1% in all cities: Magnolia, Palm, Blackgum.

Figure 3-4 shows that MEGAN3.1 tree genera composition overestimates the dominant genera Pines (Pinus) and Poplars (Populus) and underestimates the dominant genera Oaks (Quercus), Hackberries (Celtis), and Junipers (Juniperus). There were also six genera that were a negligible in the MEGAN3.1 urban tree composition but were found to be a small but significant component of the actual urban forests: Acacia, Ilex (Yaupon), Lagerstroemia (Crape Myrtle), Ligustrum (Privet), Platanus (Plane tree) and Triadica (Chinese Tallow). Acer (Maple) species were a major component of the MEGAN3.1 urban trees but less than 0.1% of the urban trees found for the three Texas cities. There were nine other genera that were about 1 to 2% of the MEGAN3.1 tree genera but less than 0.1% of the urban Texas trees: Betula, Fagus, Juglans, Liriodendron, Nyssa, Prunus, Pseudotsuga, Tilia, Tsuga.

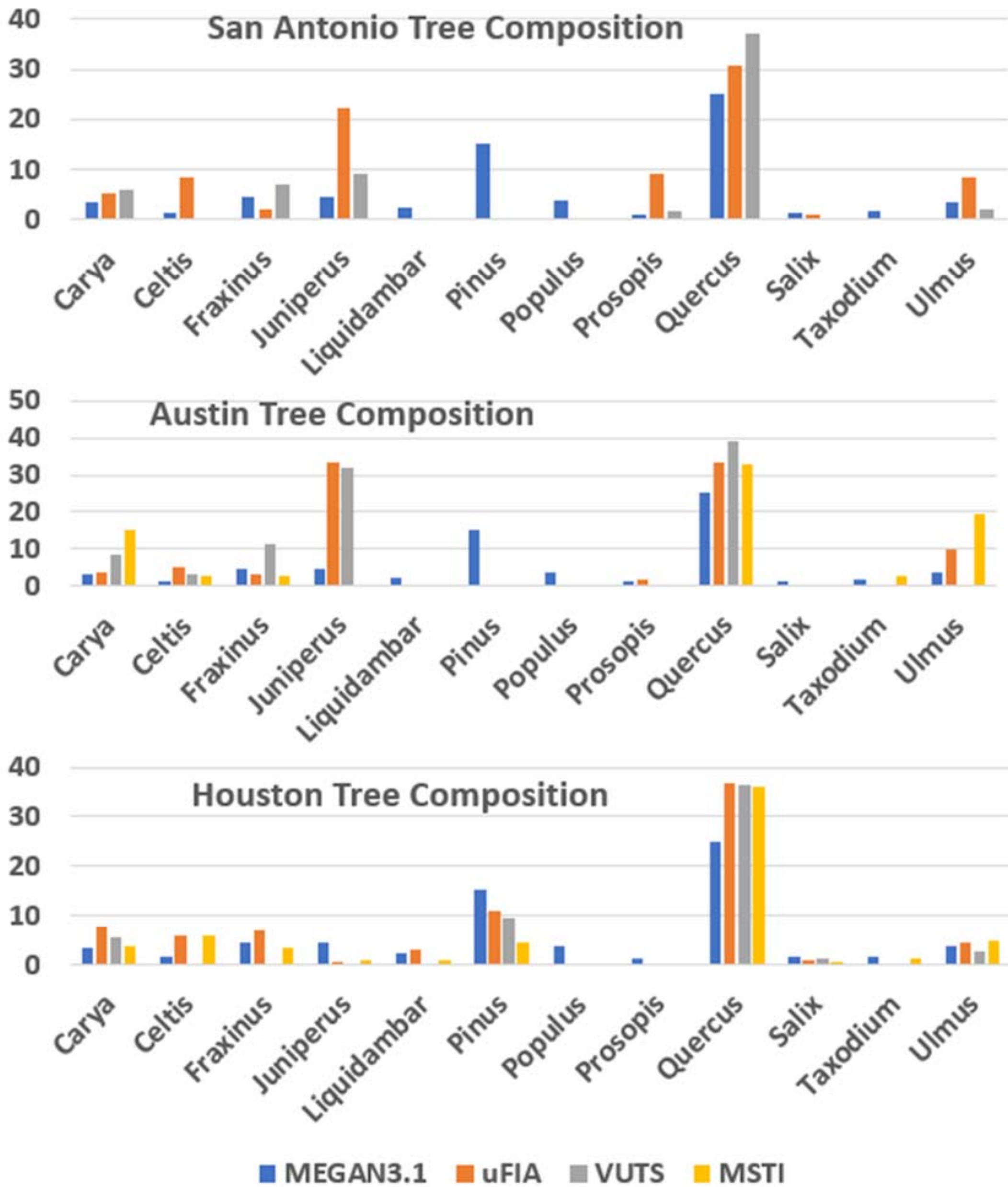


Figure 3-4. MEGAN3.1 average urban tree genera composition compared with three approaches used to estimate Houston, Austin and San Antonio urban tree composition: urban Forest Inventory and Analysis (uFIA), Virtual Urban Tree Survey (VUTS), and Municipal Street Tree Inventories (MSTI).

Table 3-2 shows that while there are broad similarities between the uFIA, VUTS and MSTI estimates of tree distributions, there are substantial differences as well. For example, there is general agreement that oak is the dominant tree genera, and the major isoprene source, in all three cities, with a range of 36.1 to 36.9% for Houston, 33 to 39.3% for Austin and 30.6 to 37.1% for San Antonio. When the four main isoprene emitting genera (Quercus, Liquidambar, Platanus, and Salix) are combined, there continues to be agreement for the sum of these isoprene emitters in Houston (37.9 to 42.2%) and Austin (33.92 to 39.3%) while San Antonio has a larger range (31.7 to 43.6%). The uFIA and VUTS estimates of pine, juniper and baldcypress contributions were similar in all three cities indicating that these approaches can accurately estimate these conifers that tend to have relatively high terpenoid emissions. The MSTI differed considerably which may just indicate a difference between private and public tree species composition. This was also observed for some broadleaf trees such as Crepe myrtle which appear to be more widely planted by the city than by private landowners. The uFIA and VUTS estimates of the sum of the three dominant non-isoprene emitting broadleaf tree genera (pecan and hickories, ashes and elms) was similar in Austin and San Antonio but the uFIA estimate was considerably higher in Houston (19.5% compared to 8.1%).

Tree species composition was determined by weighting a city tree estimate based on MSTI (15% of total) with a private tree estimate based on uFIA (85% of total). The generally good agreement between the uFIA and VUTS demonstrated consistency between these approaches and suggests that the VUTS approach is suitable for applying to cities that do not have uFIA data.

Table 3-2. Tree species composition (%) estimates for 3 Texas cities estimated by uFIA, MSTI and VUTS survey approaches.

Common Name	Genus	Houston: uFIA	Houston: VUTS	Houston: MSTI	Austin: uFIA	Austin: VUTS	Austin: MSTI	San Antonio: uFIA	San Antonio: VUTS
Acacias	Acacia	<0.01	<0.01	<0.01	0.06	<0.01	0.05	1.9	1.9
Pecan and Hickories	Carya	7.8	5.4	3.8	3.9	8.5	15.3	5.3	5.9
Hackberries	Celtis	5.8	<0.01	6.0	5.2	3.4	2.7	8.4	<0.01
Ash	Fraxinus	7.1	<0.01	3.4	3.2	11.5	2.8	2.0	7.1
Yaupon	Ilex	1.4	1.4	0.4	.06	<0.01	1.3	<0.01	<0.01
Junipers	Juniperus	0.2	<0.01	0.8	33.3	32.0	0.4	22.1	9.3
Crepe Myrtle	Lagerstroemia	0.02	1.4	16.0	0.01	<0.01	8.2	0.01	2.3
Privets	Ligustrum	2.2	<0.01	0.4	1.8	<0.01	0.8	1.4	<0.01
Sweetgum	Liquidambar	2.9	<0.01	0.7	<0.01	<0.01	0.1	<0.01	<0.01
Pines	Pinus	11.1	9.5	4.6	0.3	<0.01	0.2	<0.01	<0.01
Sycamores	Platanus	1.4	<0.01	1.8	0.8	<0.01	0.8	<0.01	6.5
Mesquite	Prosopis	<0.01	<0.01	<0.01	1.9	<0.01	0.2	9.1	1.8

Oaks	Quercus	36.9	36.5	36.1	33.6	39.3	33.0	30.6	37.1
Willows	Salix	1	1.4	0.2	<0.01	<0.01	0.02	1.1	<0.01
Baldcypress	Taxodium	<0.01	<0.01	1.2	<0.01	<0.01	2.7	<0.01	<0.01
Chinese Tallow	Triadica	8.4	1.4	4.7	0.3	0.9	0.4	0.3	<0.01
Elms	Ulmus	4.6	2.7	4.9	10.1	<0.01	19.6	8.4	2.1
Other Trees		9.2	40.3	15.0	5.5	4.4	11.4	9.4	24.3
All isoprene emitters	Quercus, Liquidambar, Platanus, and Salix	42.2	37.9	38.8	34.4	39.3	33.92	31.7	43.6
Pecan, Ash and Elm	Carya, Fraxinus and Ulmus	19.5	8.1	12.1	17.8	20.0	37.7	16.7	15.1

4.0 TASK 3: MEGAN AND BEIS INPUTS, PROCESSORS AND RESULTS

4.1 MEGAN inputs and processor

The MEGAN input files developed for this project include the landcover data described in section 2 (LAI) and section 3 (tree cover and speciation). High resolution satellite data and ground survey information were combined to generate three urban landcover products: 8-day LAI for 2019, Growth form fractions (tree, shrub, herbaceous, crop) for 2018, and tree speciation representative of 2015 to 2018. The growth form fraction and tree speciation data serve as inputs to the MEGAN-Emission Factor Processor (EFP) version 3.2 (Python code) while the growth form fraction and LAI are input to the MEGAN-Emission Calculator (EC) version 3.2 (FORTRAN code).

The MEGAN-EFP Python code and inputs were modified to generate a simpler and more effective approach that minimizes errors. The previous procedures for assigning emission factors based on quality ("j-values") and taxonomic relationships have been eliminated. Instead, the emission factors are now assigned by the database developers as is the case for BEIS and for previous versions of MEGAN. The major reason for this is that the amount of available high quality emission data is insufficient for applying the taxonomic approach which requires not only data on emitters but also for non-emitters. The assignments of emission factors to various plant species and types will instead be documented in a series of reviews of the emission measurement literature that explain the origin of the individual emission factors in the model. The new emission factor estimates are similar for most plant species that have been extensively studied (e.g., pines and oaks) but can be substantially different for other species that are not well studied. As an example, maple trees were previously assigned a moderate isoprene emission rate because the database did not include any isoprene measurements of maples simply because many investigators do not report the compounds that are not emitted but only the ones that are.

The MEGAN Preprocessor and Calculator Fortran codes and inputs were modified to improve model performance. The specific changes include:

- Rename files to clarify purpose
- Replace cantype processor. It previously read in six files: four growth forms plus needle tree fraction and tropical tree fraction. The new version reads in a single file with six canopy types.
- The LAI processor was modified to allow 8-day, 10-day or monthly time step and LAI or LAIv (i.e., LAI of vegetated surface) input to enable use of various data sets. If the input is LAI, the LAIv will be calculated within the model using total VCF which is calculated by adding four growth form cover fractions (tree, shrub, grass, crop).
- Other modifications include making namelist files more flexible, enable all "prepmegan" programs to compile at the same time, allow processing data with descending latitude, clean up source code and update README accordingly

- Allow user to specify the value to convert between total solar radiation and PAR or use default value
- Correct code errors including incorrect surrogate species for sesquiterpenes, looping ending earlier than expected in soil Nitric Oxide (NO) diurnal emissions routine and incorrect number of soil texture types in the Berkley Dalhousie Soil NO Parameterization (BDSNP) model.

4.1.1 New WRFCAMx support for MEGAN

MEGAN relies on meteorological input data that typically originate from the Weather Research Forecast (WRF) model. Historically, MEGAN has required specific meteorological input variables and formats generated by the MCIP program, which is part of the Community Multiscale Air Quality (CMAQ) modeling system. A recent MCIP update deprecated the output of certain data that MEGAN requires but CMAQ no longer requires. The Comprehensive Air Quality Model with Extensions (CAMx) interface preprocessor that corresponds to MCIP is named WRFCAMx. In this project, we updated WRFCAMx to output meteorological data required by MEGAN, thus ensuring these data remain accessible and simplifying data processing when MEGAN is used with CAMx.

Table 4-1 presents a list of WRFCAMx meteorological parameters used by MEGAN and corresponding parameters in MCIP output. It should be noted that the "soilmoist" and "soiltemp" parameters are taken from the top WRF soil layer, which is uppermost 10 cm for the NOAH Land Surface Model (LSM) and the top 1 cm for Pleim-Xiu LSM, consistent with MCIP.

CAMx meteorological input files do not follow the same dot/cross point convention as CMAQ for winds (Ramboll, 2021). The U/V vector wind components in CAMx 3-D input files are staggered to cell interfaces, so for MEGAN to use 3-D winds at point (I, J) it must average the "uwind" variable from (I-1, J) and (I, J), and the "vwind" variable from (I, J) and (I, J-1). It is not necessary to space-average for CAMx 2-D 10-m wind components (U10 and V10) because they are carried at cell centers.

Table 4-1. List of WRFCAMx parameters used by MEGAN and corresponding parameter in MCIP.

MEGAN Input Data Requirement	WRFCAMx Parameter [a]	MCIP Parameter
Pressure	"pressure" (mb)	"PRES"
Water vapor mixing ratio	"humidity" (ppm)	"QV"
U-component of wind	"uwind" (m/s): U-component (east/west) wind speed on CAMx cell interfaces	"UWIND": U-component wind speed at CMAQ dot (cell corner) points

V-component of wind	"vwind" (m/s): V-component (north/south) wind speed on CAMx cell interfaces	"VWIND": V-component wind speed at CMAQ dot (cell corner) points
Total convective plus non-convective precipitation	"preciprate" (mm/hr)	"RC" and "RN"
Snow cover	"snowewd" (m): equivalent water depth	"SNOCOV": 0 or 1 flag
Total cloud fraction	"cloudfrac" (dec): total resolved and sub-grid clouds	"CFRAC": resolved clouds only
Wind speed at 10 m	"u10" and "v10" (m/s)	"WSPD10"
Solar radiation reaching surface	"swsfc" (W/m2): affected by resolved and sub-grid cloud cover	"RGRND"
Soil texture type	"soiltype" (unitless)	"SLTYP"
Volumetric soil moisture	"soilmoist" (m^3/m^3): top LSM layer	"SOIM1": top LSM layer
Soil temperature	"soiltemp" (K): top LSM layer	"SOIT1": top LSM layer

(a) WRFCAMx outputs the parameters pressure, humidity, uwind and vwind to CAMx 3-D files and all other parameters listed to CAMx 2-D files.

4.2 BEIS inputs

EPA recently released a new version of Biogenic Emissions Landuse Database, Version 5 (BELD5) landcover database which we planned to update for BEIS with updated Texas urban landcover data developed for this project. Upon review of the BELD5 data, it appeared that BELD5 uses only Moderate Resolution Imaging Spectroradiometer (MODIS) landcover to characterize almost all of Texas including the Austin, Houston, and San Antonio urban regions. BELD5 does not appear to consider any tree species information for most of Texas and classifies large areas of the landscape as a generic woody savanna or a grassland/cropland. Upon bringing up the issue to EPA, they indicated that this issue will be rectified for urban areas in the next version of BELD (i.e., BELD6) due to land use data updates. Since BELD5 did not contain any tree data for our target region, it did not make much sense to do any adjustments to the BELD5 database. For these reasons, instead of BEIS/BELD5 updates, a review of input data between BEIS and MEGAN models was performed to better understand differences in their underlying data.

The two key input data differences between BEIS and MEGAN are (1) land use database and (2) LAI. The land cover database used by MEGAN is based on NLCD for tree growth form and breakdowns based on the ecotype estimated by Yu et al. (2017) for other growth forms. The BELD5 land cover database used by BEIS[2] is based on:

- Newer version of the Forest Inventory and Analysis (FIA) version 8.0
- Agricultural land use from the 2017 USDA crop data layer
- Global Moderate Resolution Imaging Spectroradiometer (MODIS) 20 category data with enhanced lakes

In both the models, LAI is used to adjust the isoprene emissions for the effects of Photosynthetically Active Radiation (PAR) penetrating through the leaf canopy (Vukovich, 2002). MEGAN uses actual year-specific LAI derived from satellite images, but BEIS estimates LAI using species and land use type-specific LAI factor which are not year specific.

4.3 Assess and compare MEGAN and BEIS sensitivity to urban inputs

MEGAN emissions sensitivity to updated Texas urban landcover and meteorological input data was assessed by comparing soil NO and BVOC emissions generated with (1) current BEIS (BEIS3.7), (2) current MEGAN (MEGAN3.1) and (3) updated MEGAN (MEGAN3.2).

4.3.1 MEGAN3.2 soil NO emissions

Soil NO emissions simulated using recently released version of BEIS (BEIS3.7), MEGAN3.1 (with J=4), MEGAN3.2 with MCIP and MEGAN3.2 with WRFCAMx model approaches were compared for the summer 2019 scenario (May-September period) for the TCEQ 12 km domain used in TCEQ's SIP modeling. MEGAN3.2 with two sets of meteorological input data (i.e., MCIP and WRFCAMx) predicted nearly the same soil NO emissions with identical spatial distribution which demonstrates that the newly added WRFCAMx support in MEGAN is working. MEGAN3.2 NO emission estimates are 50% higher than MEGAN3.1 primarily due to a couple of bug fixes related to (1) a looping ending earlier than expected in soil NO diurnal emissions routine and (2) incorrect number of soil texture type in the BDSNP model. Compared to BEIS, soil NO estimates by MEGAN3.2 are a factor of 3 higher (Figure 4-1). In Texas, MEGAN3.2 soil NO emissions are 40% higher than MEGAN3.1 and 200% higher than BEIS3.7 (Figure 4-2). Also notable are the higher MEGAN emissions in northern and eastern Texas agricultural landscapes (Figure 4-3 and Figure 4-4).

[2] https://www.epa.gov/air-emissions-modeling/biogenic-emission-inventory-system-beis/

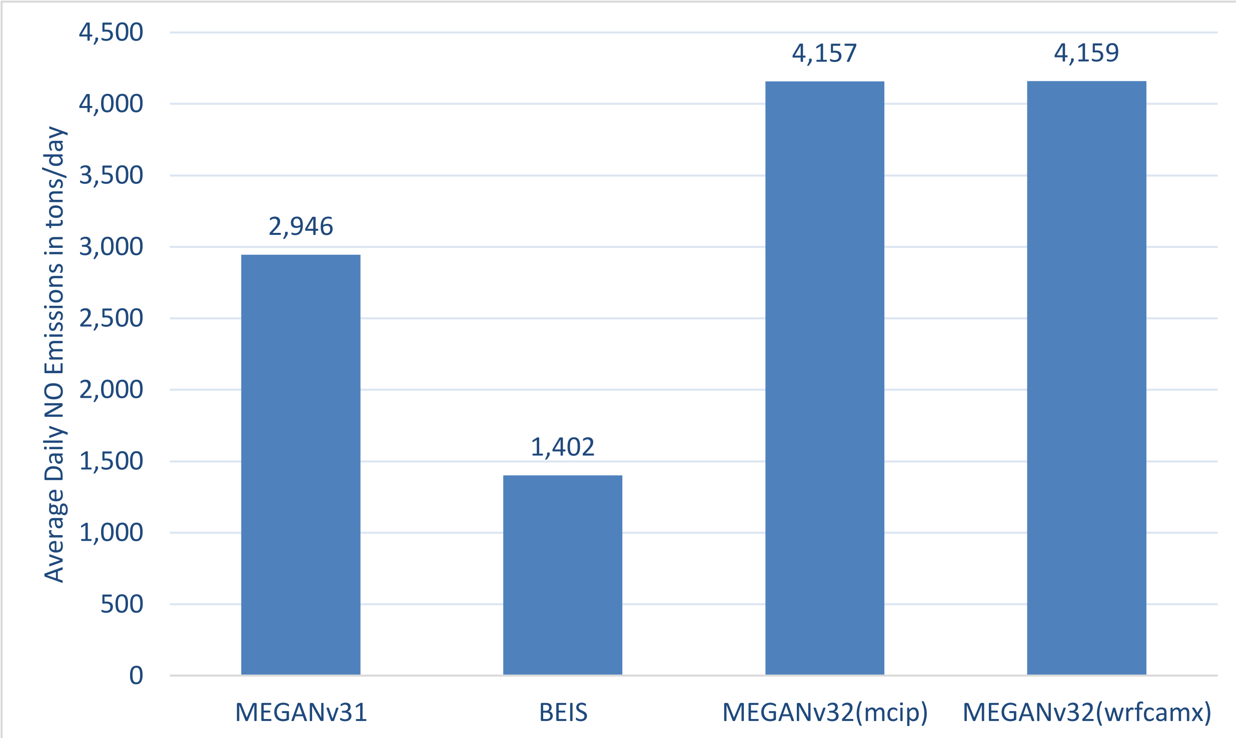


Figure 4-1. Total domain episode average NO emissions (tons/day) for summer 2019 in the TCEQ 12 km domain estimated by MEGAN3.1, BEIS, MEGAN3.2 with MCIP and MEGAN3.2 with WRFCAMx models.

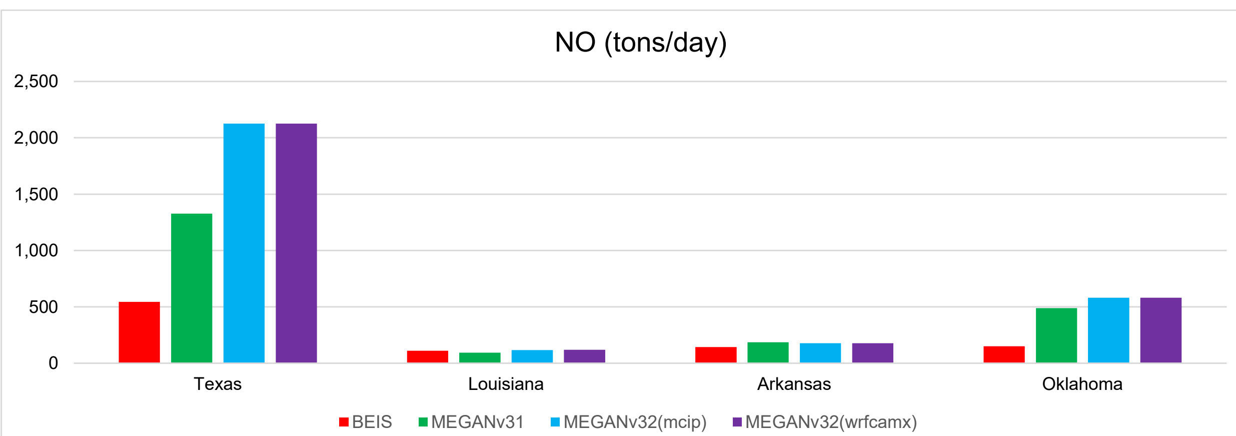


Figure 4-2. US State total soil NO emissions (tons/day) for summer 2019 in the TCEQ 12 km domain estimated by BEIS, MEGAN3.1 and MEGAN3.2 models.

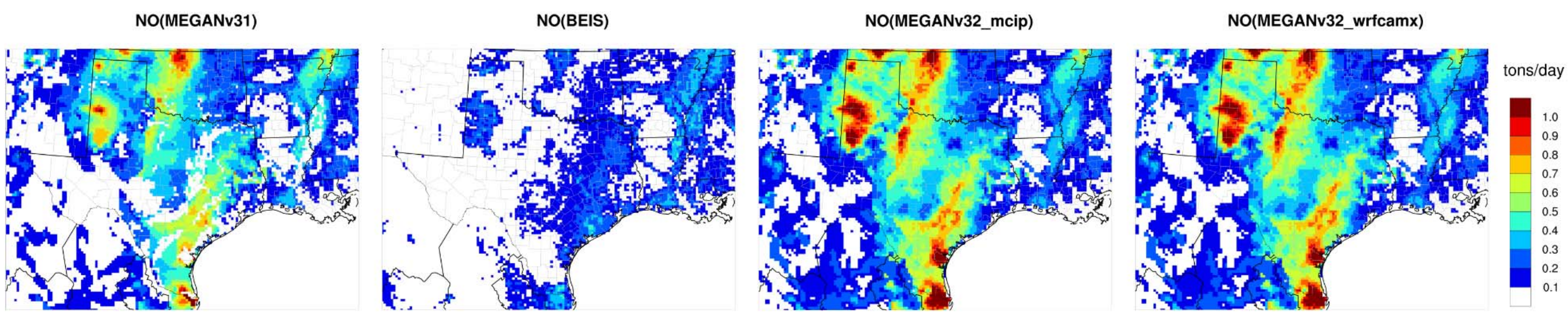


Figure 4-3. Episode average soil NO emissions (tons/day) simulated using MEGAN3.1, BEIS, MEGAN3.2 with MCIP and MEGAN3.2 with WRFCAMx models for summer 2019 in the TCEQ 12 km domain.

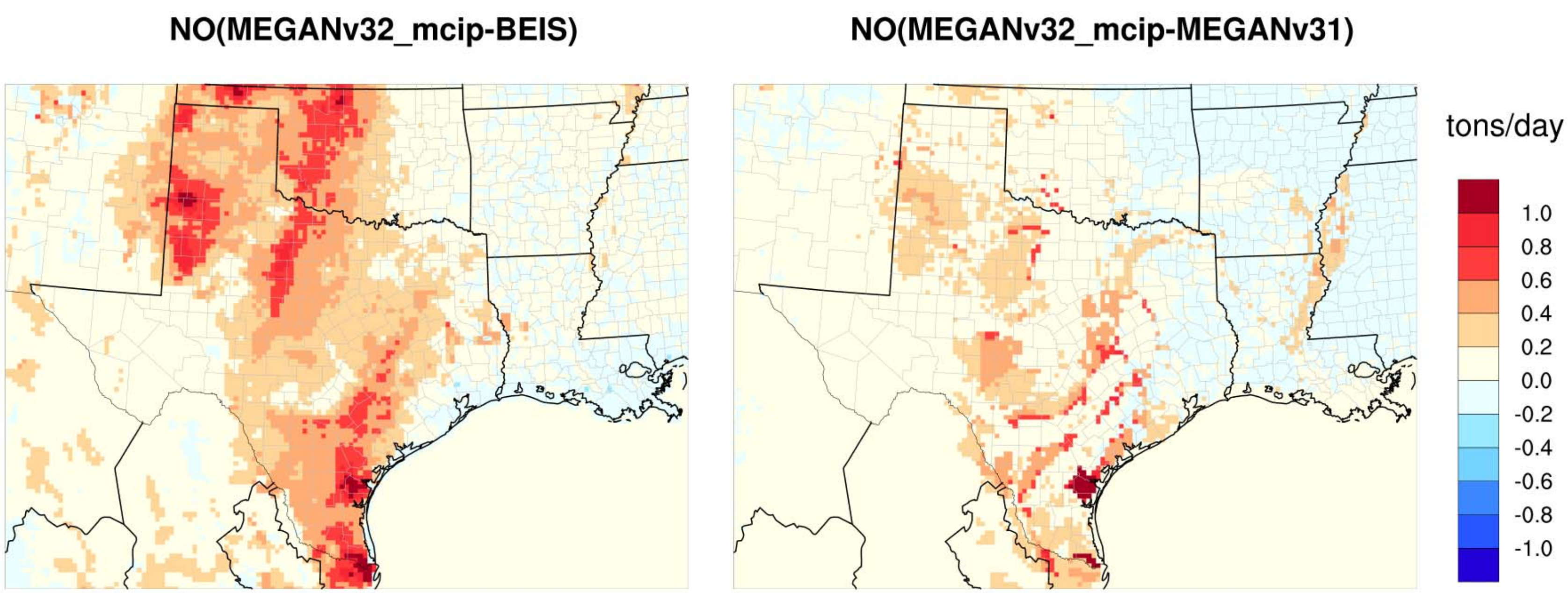


Figure 4-4. Difference (MEGAN3.2 minus MEGAN3.1 or BEIS) in soil NO emissions (tons/day) simulated, and shown in Figure 4-2, for summer 2019 in the TCEQ 12 km domain.

4.3.2 MEGAN3.2 BVOC emissions

Like soil NO, BVOC emissions simulated using BEIS3.7, MEGAN3.1 (with J=4), MEGAN3.2 with MCIP and MEGAN3.2 with WRFCAMx model approaches were compared for the summer 2019 scenario (May-September period) for the TCEQ 12 km domain. Domain-wide isoprene emissions estimated by MEGAN3.2 are slightly higher than MEGAN3.1, but much higher than BEIS (Figure 4-5). Between the two MEGAN versions, the differences in isoprene estimates are mainly due to implementing a new approach in the MEGAN EFP to assign emission factors as described in section 4.1 and using new vegetation cover data developed in this study. MEGAN 3.2 and MEGAN 3.1 isoprene and monoterpene emissions are similar when averaged across Texas but there are large regional differences (Figure 4-6 and Figure 4-11).

All four models estimate higher isoprene emissions in the oak forest of the Ozarks and southeastern US (Figure 4-7). MEGAN3.2 isoprene is higher than BEIS especially in the Piney Woods ecoregion and Ozarks (Figure 4-8). Figure 4-9 shows zoom in on the three Texas urban areas. Compared to MEGAN3.1, isoprene emissions estimated by MEGAN3.2 are lower for the three urban areas, especially in the urban core.

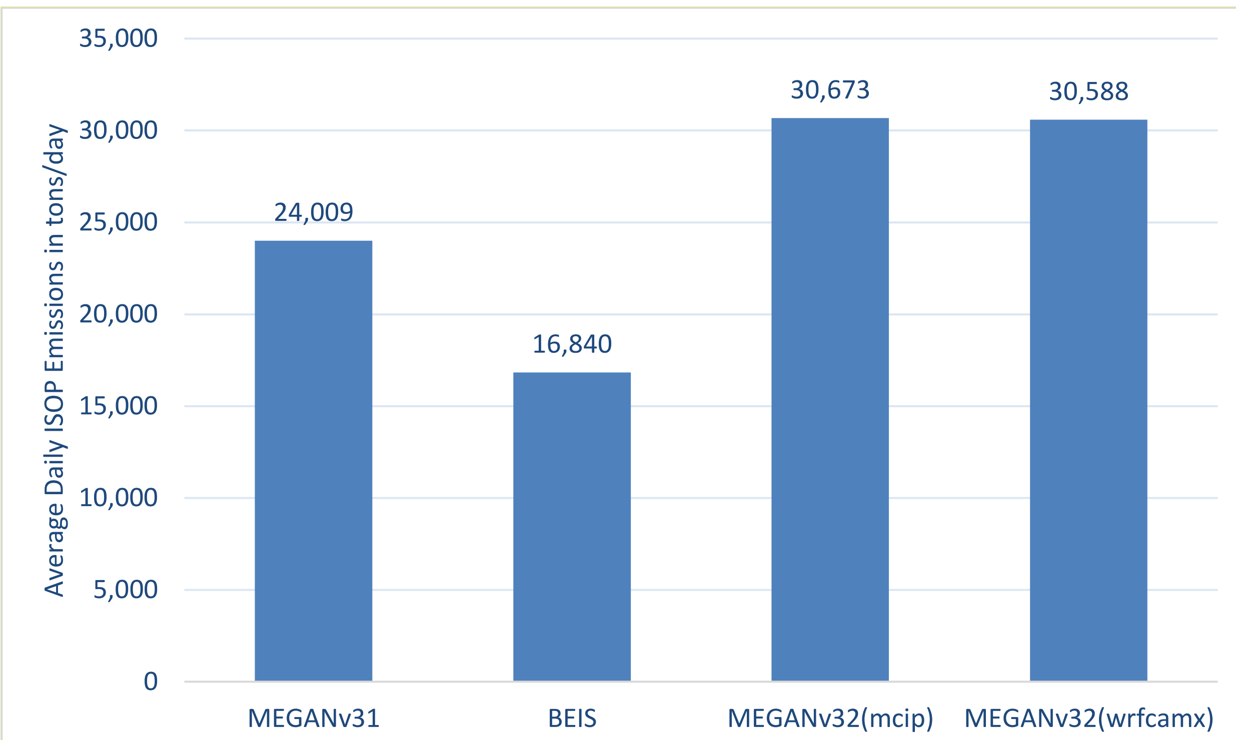


Figure 4-5. Total domain episode average isoprene emissions (tons/day) for summer 2019 in the TCEQ 12 km domain estimated by MEGAN3.1, BEIS, MEGAN3.2 with MCIP and MEGAN3.2 with WRFCAMx models.

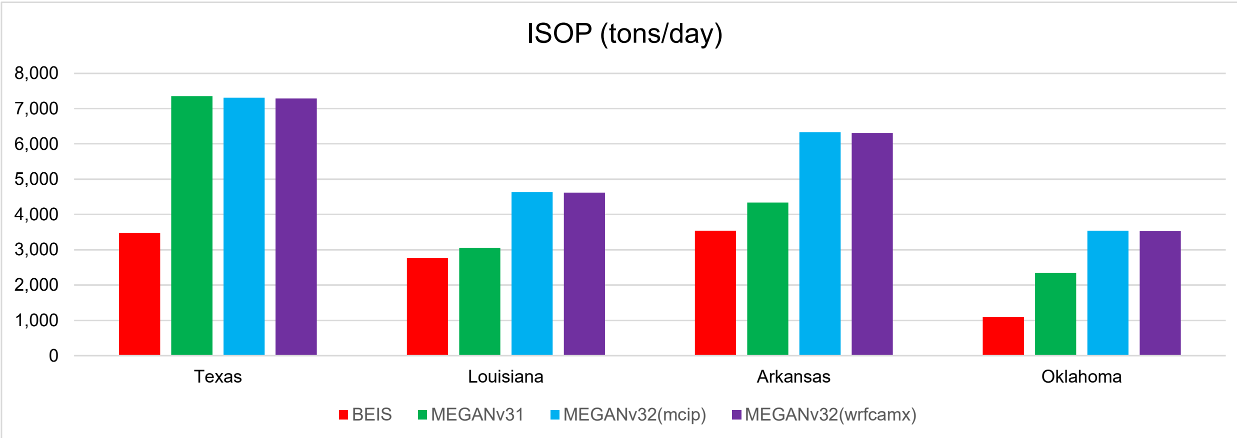


Figure 4-6. US State total ISOP emissions (tons/day) for summer 2019 in the TCEQ 12 km domain estimated by BEIS, MEGAN3.1 and MEGAN3.2 models.

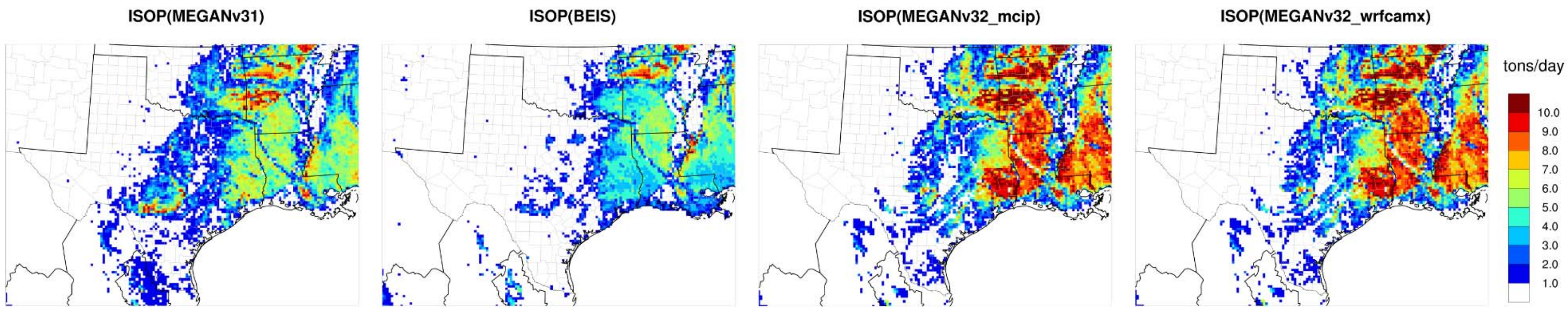


Figure 4-7. Episode average isoprene emissions (tons/day) simulated using MEGAN3.1, BEIS, MEGAN3.2 with MCIP and MEGAN3.2 with WRFCAMx models for summer 2019 in the TCEQ 12 km domain.

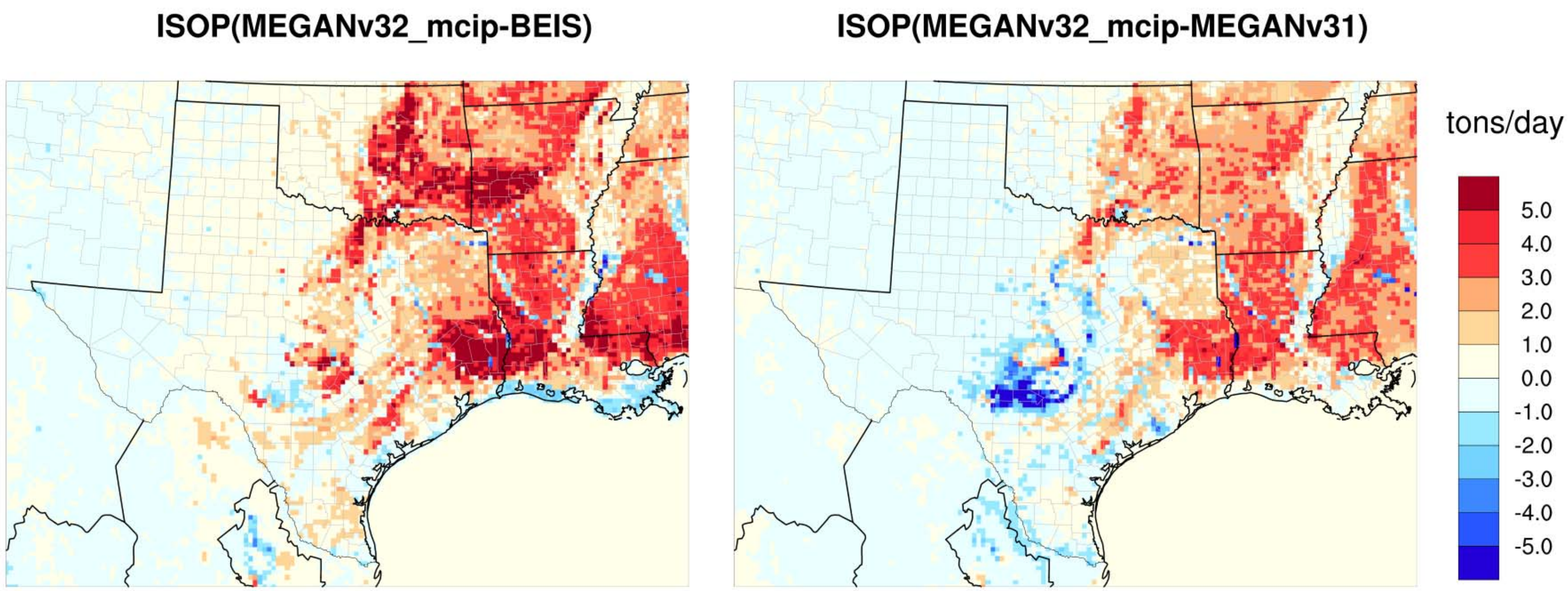


Figure 4-8. Difference (MEGAN3.2 minus MEGAN3.1 or BEIS) in isoprene emissions (tons/day) simulated, and shown in

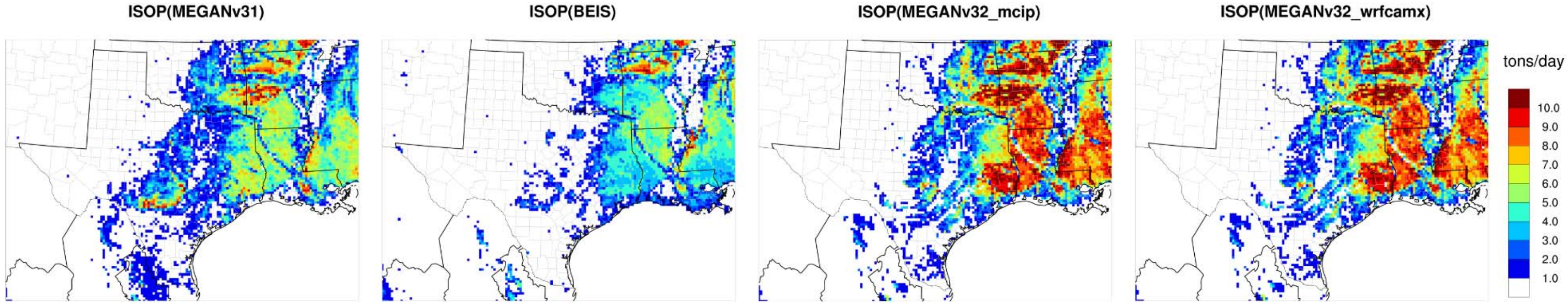


Figure 4-7, for summer 2019 in the TCEQ 12 km domain.

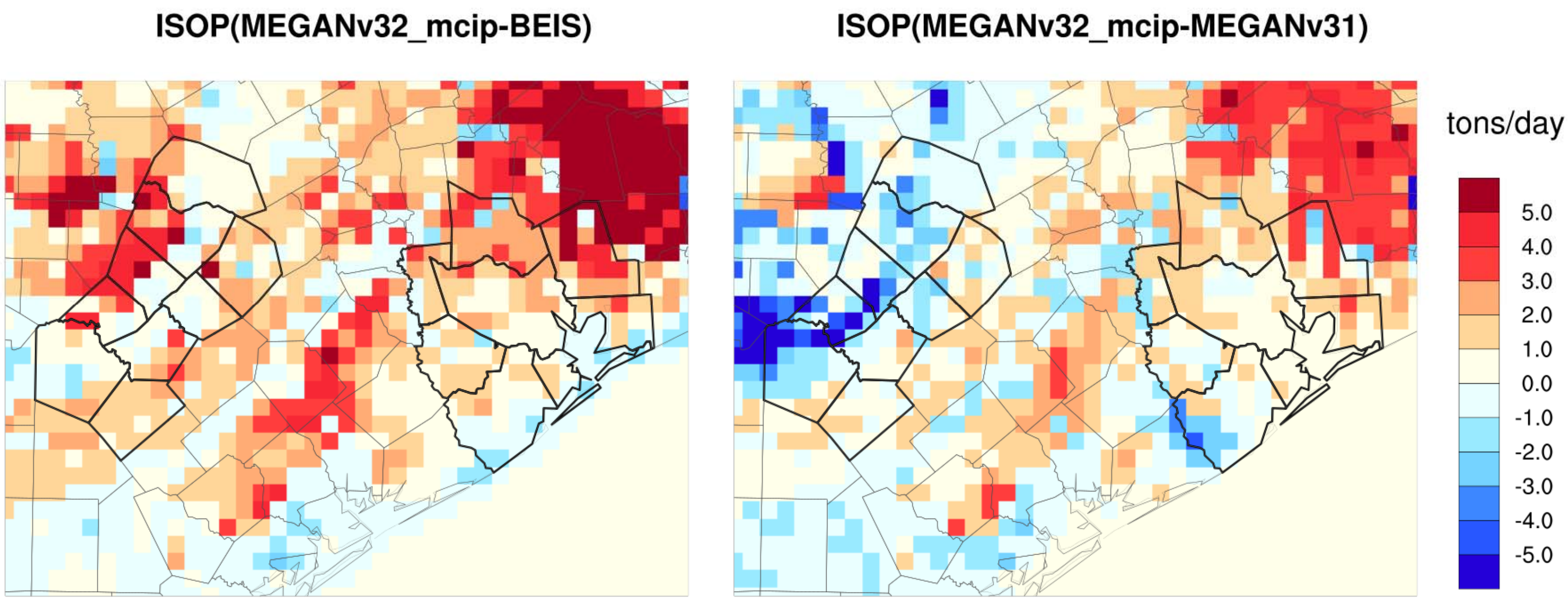


Figure 4-9. Difference (MEGAN3.2 minus MEGAN3.1 or BEIS) in isoprene emissions (tons/day) for the three Texas urban areas and summer 2019

Terpene emissions estimated by MEGAN3.2 are higher than MEGAN3.1 and BEIS for the TCEQ 12 km domain resembling isoprene (Figure 4-10). Elevated terpene emissions in the Piney Woods region are notable (Figure 4-12). In Texas, MEGAN3.2 terpene emission estimates are lower than MEGAN3.1 in shrublands and higher in East Texas (Figure 4-13Figure 4-13Figure 4-12). Like isoprene, terpene emissions estimated by MEGAN3.2 are lower than MEGAN3.1 in the urban areas (Figure 4-14).

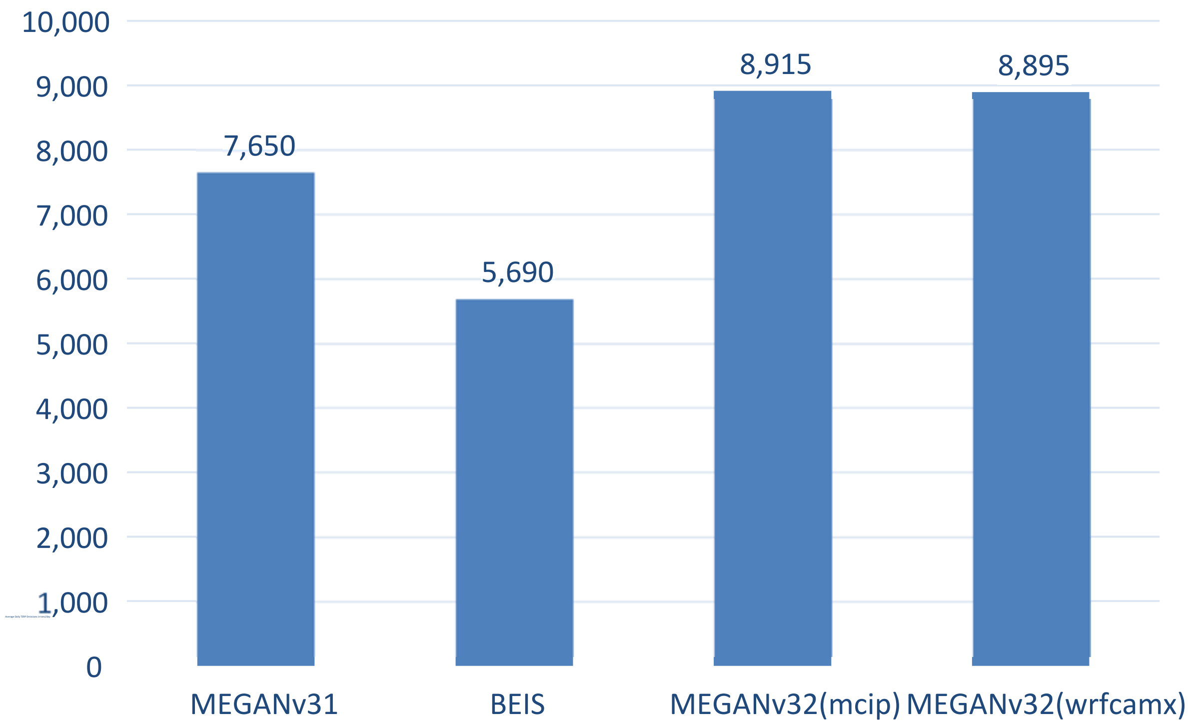


Figure 4-10. Episode average terpene emissions (tons/day) simulated using MEGAN3.1, BEIS, MEGAN3.2 with MCIP and MEGAN3.2 with WRFCAMx models for summer 2019 in the TCEQ 12 km domain.

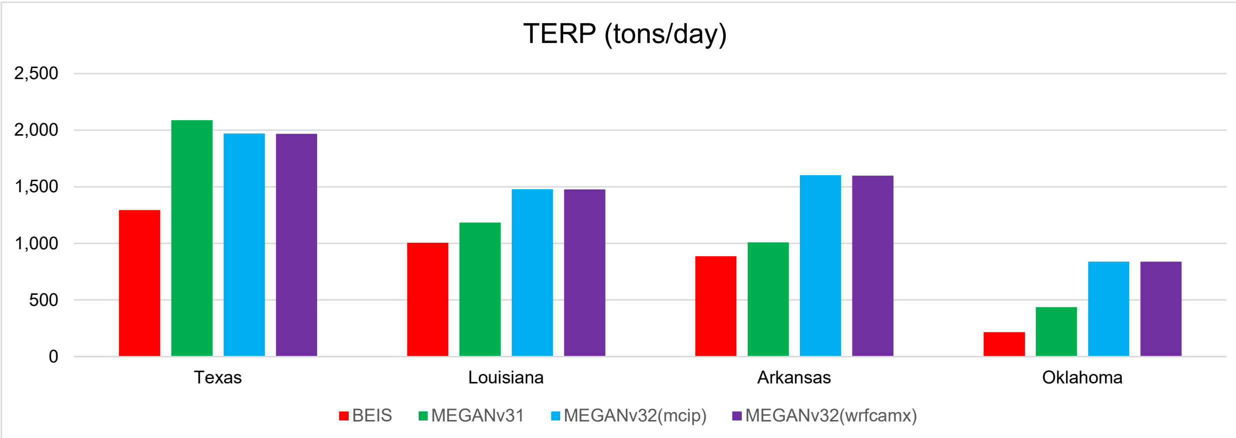


Figure 4-11. US State total TERP emissions (tons/day) for summer 2019 in the TCEQ 12 km domain estimated by BEIS, MEGAN3.1 and MEGAN3.2 models.

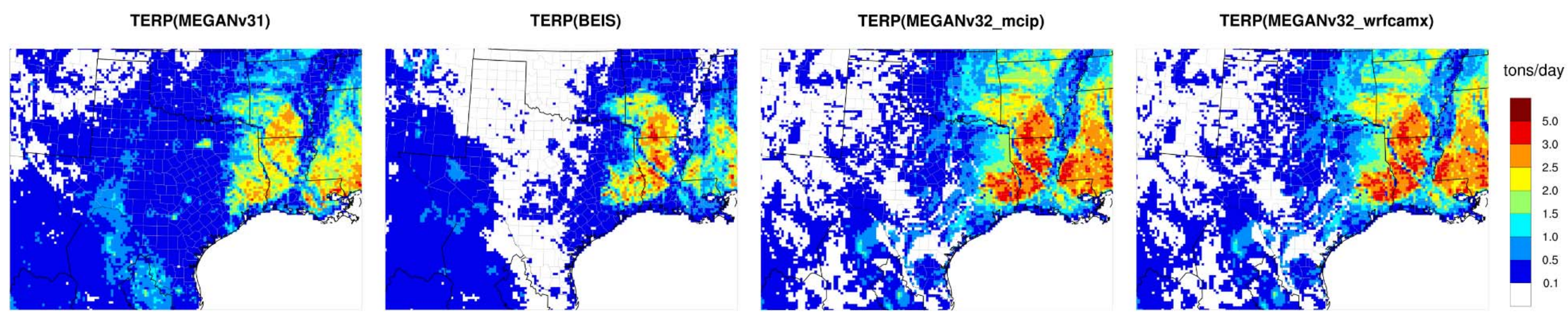


Figure 4-12. Episode average terpene emissions (tons/day) simulated using MEGAN3.1, BEIS, MEGAN3.2 with MCIP and MEGAN3.2 with WRFCAMx models for summer 2019 in the TCEQ 12 km domain.

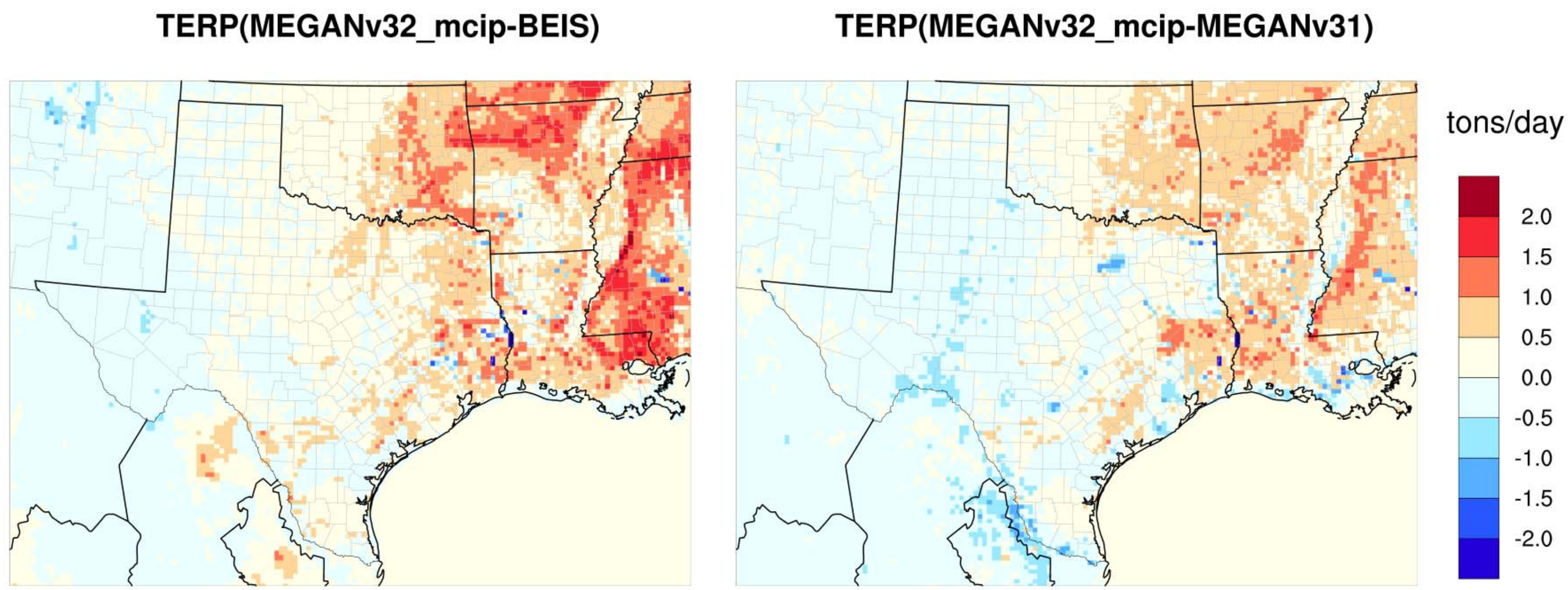


Figure 4-13. Difference (MEGAN3.2 minus MEGAN3.1 or BEIS) in terpene emissions (tons/day) simulated, and shown in

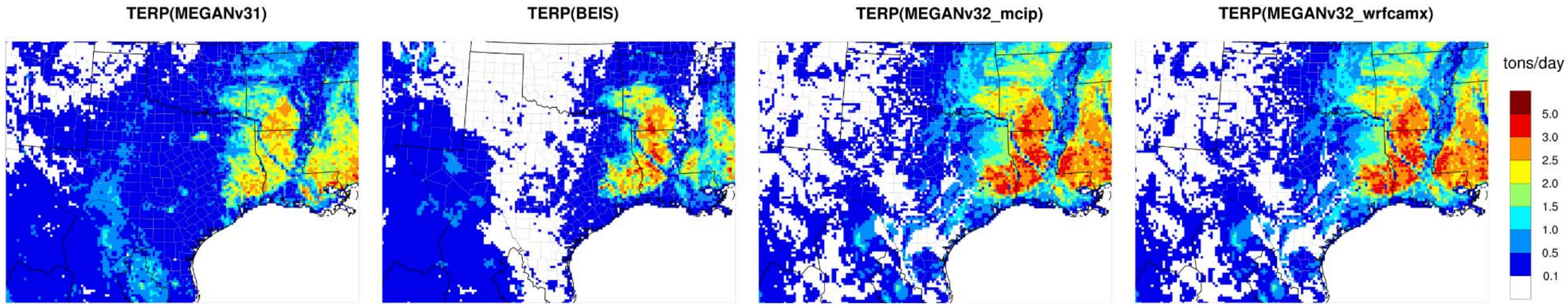


Figure 4-12, for summer 2019 in the TCEQ 12 km domain.

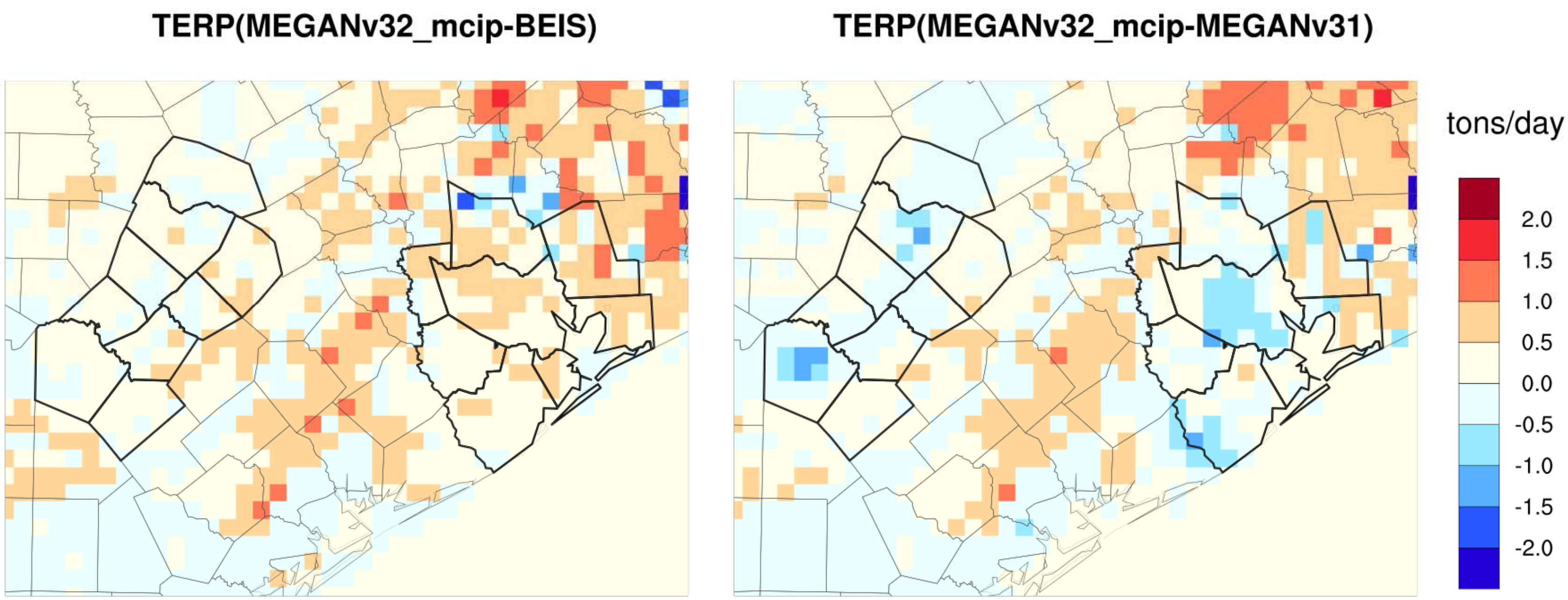


Figure 4-14. Difference (MEGAN3.2 minus MEGAN3.1 or BEIS) in terpene emissions (tons/day) for the three Texas urban areas and summer 2019

5.0 CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

Below, we provide a summary of findings of this study.

5.1 Summary of Findings

- Uncertainties in landcover data, including LAI and tree cover distributions, continue to make a significant contribution to overall uncertainties in Texas BVOC emission estimates.

- Landcover datasets based on 30-meter (and coarser) resolution imagery tend to underestimate tree cover in urban areas by 35% or more. Sub-meter resolution (e.g., NAIP) can capture tree cover associated with individual trees and is suitable for quantifying urban tree cover.

- Virtual urban tree surveys are a cost-effective approach for quantifying urban tree species composition for BVOC emission modeling.

- Oak trees are the dominant isoprene (and total BVOC) emission source in Texas urban forests. The distributions of oak trees can be accurately (typically within ~15%) characterized using the virtual urban tree survey approach.

- Conifers (e.g., pines and junipers) are an important monoterpene emission source in some Texas cities. The distributions of these trees can be accurately characterized if they are a substantial component (>5%) of an urban forest.

- The revised and simplified MEGAN3.2 Emission Factor Processor can provide more representative and accurate landscape average emission factors that are more easily assessed.

- MEGAN3.2 tends to have lower monoterpene and higher NO emissions for Texas as compared to MEGAN3.1. MEGAN 3.2 isoprene emissions are about the same as MEGAN 3.1 when averaged across Texas but there are large differences for specific regions.

5.2 Recommendations for future work

- Satellite-based isoprene: Use TROPOspheric Monitoring Instrument (TROPOMI) formaldehyde columns data to quantify and assess isoprene emission capacities for major Texas ecotypes.

- Climate control over isoprene emission capacities: Determine whether regional variations in climate produce corresponding regional variations in landscape-average isoprene emission capacities across Texas and investigate implications for isoprene response to drought and heat waves.

- Constraining terpenes: Develop and apply an optimal (for cost, time, accuracy) approach for characterizing regional monoterpene and other terpenoid emission factors for Texas using a combination of enclosure and ambient terpenoid measurements.

- MEGAN3.2: MEGAN3.2 should be used for estimating emissions of NO, isoprene, monoterpene, sesquiterpenes and other biogenic emissions in Texas while continuing to improve MEGAN inputs, especially landcover and emissions data. The main reasons to choose MEGAN are: (1) BEIS data structures don't allow urban vegetation to be characterized, and (2) MEGAN simplifies the process of updating and improving landcover and emissions data.

6.0 AUDITS OF DATA QUALITY

During this study, we performed Quality Assurance/Quality Control (QA/QC) procedures to ensure that all data and products generated are of known and acceptable quality. QA/QC procedures were performed in accordance with the Category III and IV composite Quality Assurance Project Plan (QAPP) that was completed at the beginning of the study. In a Category III and IV composite Project, data audits must be performed for at least 10% of the data sets and a report of QA findings must be given in the final report. A technical systems audit is not required. In this section, we report on the findings of our QA audits during this project.

6.1 LAI and VCF data for the three major Texas urban areas

The team of Wildland Solution and Dr. Alex Guenther reviewed more than 10% of satellite imagery data used to derive LAI and VCF data for quality assurance purposes before completing the analysis presented in Sections 4 and 5. The estimated LAI and VCF data were evaluated against the quality metrics outlined in the QAPP for AQRP Project 20-007 and were found to be of acceptable quality.

6.2 Tree distribution data for the three major Texas urban areas

The team reviewed more than 10% of processed tree cover distribution data for quality assurance purposes before completing the analysis presented in Sections 4 and 5. The tree cover data were evaluated against the quality metrics outlined in the QAPP for AQRP Project 20-007 and were found to be of acceptable quality.

6.3 MEGAN Emissions Modeling Data

The MEGAN and BEIS models were run by Ramboll using inputs developed by Wildland Solutions and Dr. Alex Guenther. The Ramboll team member who performed the MEGAN and BEIS modeling documented steps taken to obtain and process the inputs, file paths to all inputs and outputs and provided a summary of the results. The documentation was used in the subsequent MEGAN and BEIS modeling to ensure consistency in methods. Once the modeling was completed, a Ramboll team member who had not performed the MEGAN and BEIS modeling reviewed the modeling scripts for accuracy and then reviewed the model outputs. Model outputs for 10% of the episode days were reviewed using the PAVE visualization tool for completeness and compared with observations as well as the episode mean values reported in this document. All data for these days were examined for values that were outliers or otherwise unreasonable and none were found. A second Ramboll team member reviewed outputs for each day for the entire episode and found all values were reasonable. The MEGAN output data were determined to be correctly developed and complete and therefore suitable for the purposes of this study.

7.0 REFERENCES

Buchhorn, M., Lesiv, M., Tsendbazar, N. E., Herold, M., Bertels, L., & Smets, B. (2020). Copernicus global land cover layers—collection 2. *Remote Sensing*, *12*(6), 1044.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., & Wang, X. (2012). The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2. 1): an extended and updated framework for modeling biogenic emissions. Geoscientific Model Development, 5(6), 1471-1492.

Kganyago, M., Mhangara, P., Alexandridis, T., Laneve, G., Ovakoglou, G., & Mashiyi, N. (2020). Validation of sentinel-2 leaf area index (LAI) product derived from SNAP toolbox and its comparison with global LAI products in an African semi-arid agricultural landscape. Remote Sensing Letters, 11(10), 883-892.

Ramboll, 2021. User's Guide: Comprehensive Air quality Model with extensions, Version 7.10 (January 2021). http://www.camx.com/files/camxusersguide_v7.10.pdf.

Vukovich, J., & Pierce, T. (2002, April). The implementation of BEIS3 within the SMOKE modeling framework. In *Proceedings of the 11th International Emissions Inventory Conference, Atlanta, Georgia* (pp. 15-18).

Weiss, M., Baret, F., Myneni, R., Pragnère, A., & Knyazikhin, Y. (2000). Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data. *Agronomie*, *20*(1), 3-22.

Yu, H., Guenther, A., Gu, D., Warneke, C., Geron, C., Goldstein, A., ... & Yuan, B. (2017). Airborne measurements of isoprene and monoterpene emissions from southeastern US forests. *Science of the Total Environment*, *595*, 149-158.

Appendix 1: Using SNAP to estimate LAI and VCF using 10-m resolution ESA Sentinel2-MSI data

This section provides a tutorial to process Sentinel2-MSI data to develop LAI inputs for the MEGAN model. Appendix 1A describes how to access S2/MSI data and process individual tiles using the SNAP tool. Appendix 1B describes how to process individual tiles to generate MEGAN LAI inputs of 8-day or 10-day average data using ArcGIS.

Appendix 1A

Step 1. Download Sentinel 2 data
Go to https://scihub.copernicus.eu/dhus/#/home. Log in (sign up for account if you have not yet done so). Start in “Navigate mode” and navigate to the desired region of the planet and zoom into the appropriate scale. Switch to “Area mode” and select area of interest on map by holding down left button and drag polygon over area. Go to advanced search by clicking on the three bars at top left. Enter desired sensing period (start date on the left and end date on the right). Click on the box to the left of “Mission: sentinel 2”. Do not put anything in “Satellite platform”. Choose S2MSI2A for product type. Enter acceptable cloud cover, e.g. “[0 TO 10]”. Note that the “TO” must be capital letters. Click on search (right side of top bar). Download images by placing cursor over the tab and then click on arrow at far-right bottom. Note that you can have only three files downloading at a time. The time to download a file depends on the size (many are just a fraction of a full tile), your internet speed and probably how busy the server is. A 500 MB file can take from 10 seconds to 10 minutes to download. The download is a zipped file containing all of the MSI sensor bands.

Step 2. Use SNAP to calculate LAI and VCF
Open SNAP app. If you don’t have it then it can be download from http://step.esa.int/main/download/snap-download/.

Choose sentinel 2 toolbox. Click on the “open file” icon on the far left and locate the zipped file that you downloaded. It will then show up in the product explorer window.

First you need to resample the bands so they are all at 10m: Go to Raster> Geometric operation > Resampling to open the Resampling tool window. Click on “resampling parameters” and in the “define size of resampled product” box under “By reference band from source product” select “B2”. The is will set the size to 10 m and it will indicate a target height and width of 10980. Click “Run” then “OK” and then close the resampling window.

Next you will calculate LAI and veg cover fraction: Go to Optical> Thematic land processing> Biophysical Processor (LAI, fAPAR…) to open the Biophysical processor tool window.

In the “Source Product” box, select the resampled file you just created (it will say “resampled” at the end of the file name (you can widen the box to enable viewing the whole file name).

Save as BEAM-DIMAP format

Click on “Processing parameters” and deselect “FAPAR”, Cab, and CWC.

Then click on “Run”. It will take up to ~1 hour for a full tile (Windows system with 64GB RAM and 8 core processors) but most files are only a fraction of the whole tile and so take less time.

Step 3: Process the 10-m data to generate inputs for MEGAN.
The LAI and Fcover files (and their associated “flag” files) are created in a folder that is named the same as the original downloaded file but with “resampled_biophysical.data” attached. The files are ENVI image files which can be uploaded into ArcGIS or other programs where they can interpolated/resampled to the appropriate time step (8-day for MEGAN2.1, 8 or 10 day or monthly for MEGAN 3.2) and spatial resolution (1/120 of a degree). The flag files can be used to mask out cloud contamination.

Appendix 1B Processing Sentinel LAI and VCF data using ArcGIS**.**
As a first step, the LAI and fcover images (.img files) generated using the ESA SNAP tool should be visually inspected to identify and remove high cloud cover images and the SNAP output should be checked for corrupt images. For an example, see Figure 6-1 below.

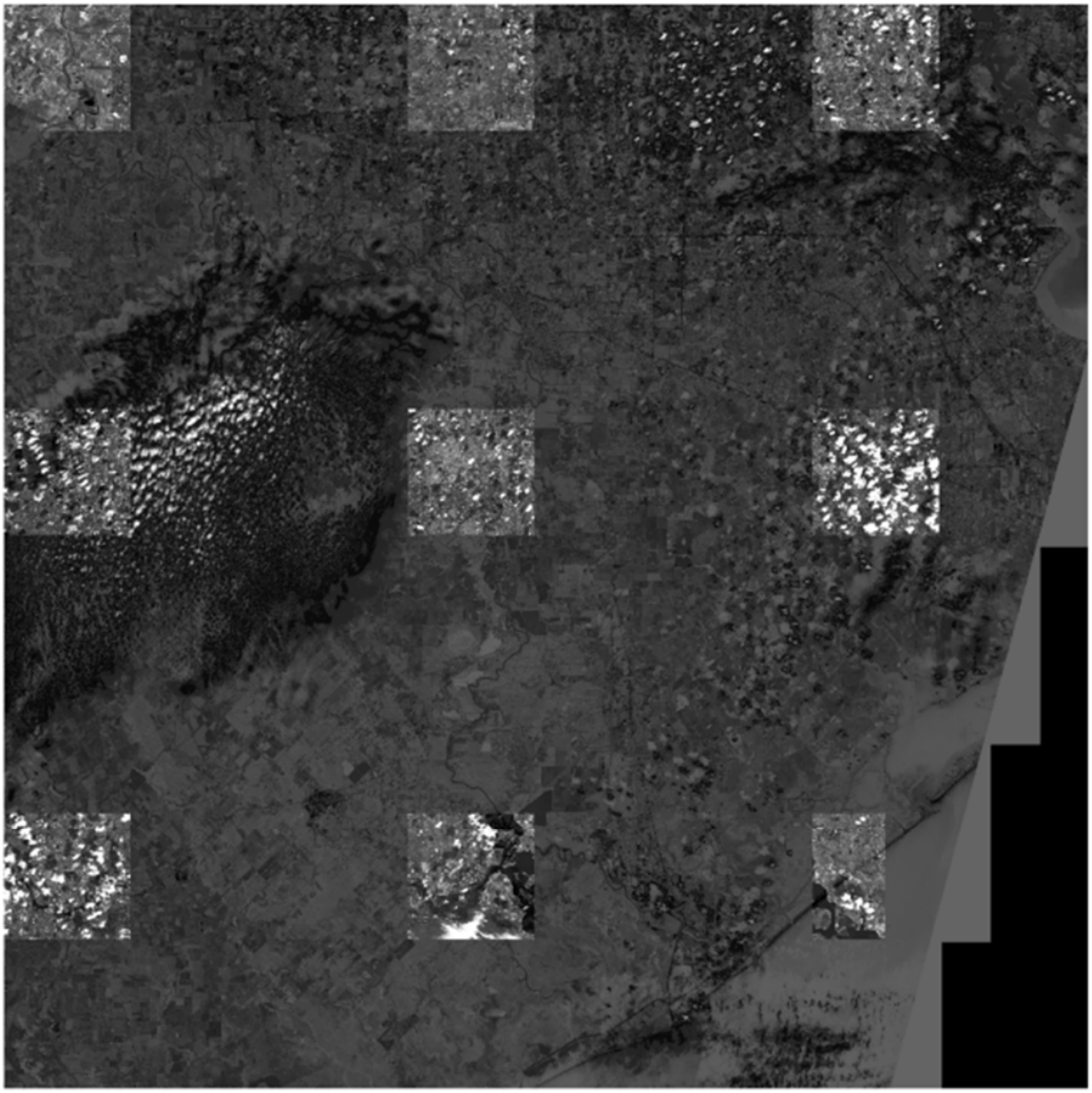
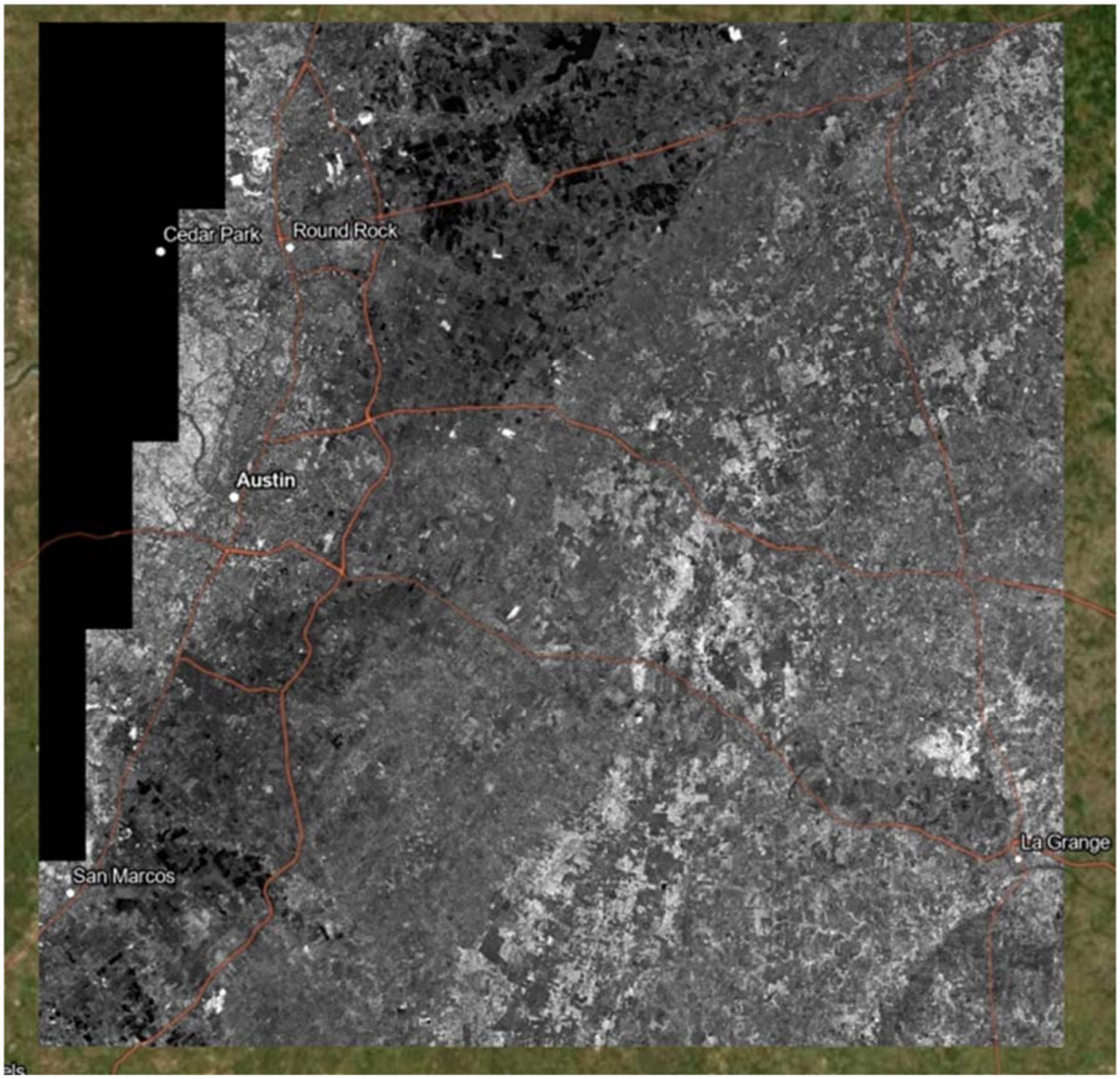


Figure 6-1. Corrupted LAI image for January 27, 2019 in Houston tile RTN

ArcGIS python can be used to process the uncorrupted images. The processing includes 5 steps that can be accomplished using the 5 ArcGIS python programs:

1. **Merge partial scenes (see SentinelRTNStep1.py).** For some locations, the 100 km x 100 km Sentinel tile is fully within the swath of both Sentinel sensors (2A and 2B) so that each satellite pass provides a full image. An example is the area around San Antonio, Texas (ESA tile T14 RNT). In that case step 1 is not required. For other locations, such as the Austin Texas example shown in Figure 6-2, some or all images includes a swath of missing data on the left (west) or right (east). The missing data in each of the scenes can be replaced by the nearest (in time) scene with data in the missing area. The missing data includes both "nan" shown in black in Figure 6-2 and constant values appearing as grey

triangles along the border of the “nan” data (see red oval in Figure 6-2, bottom). The “black” areas in Figure 6-2 can be identified by a value of zero in the Sentinel Scene classification (SCL) provided at 20 meter resolution for each image. The “grey” triangle missing data (Figure 6-2 bottom) can be identified as locations where SCL is greater than zero and Lai is greater than zero.



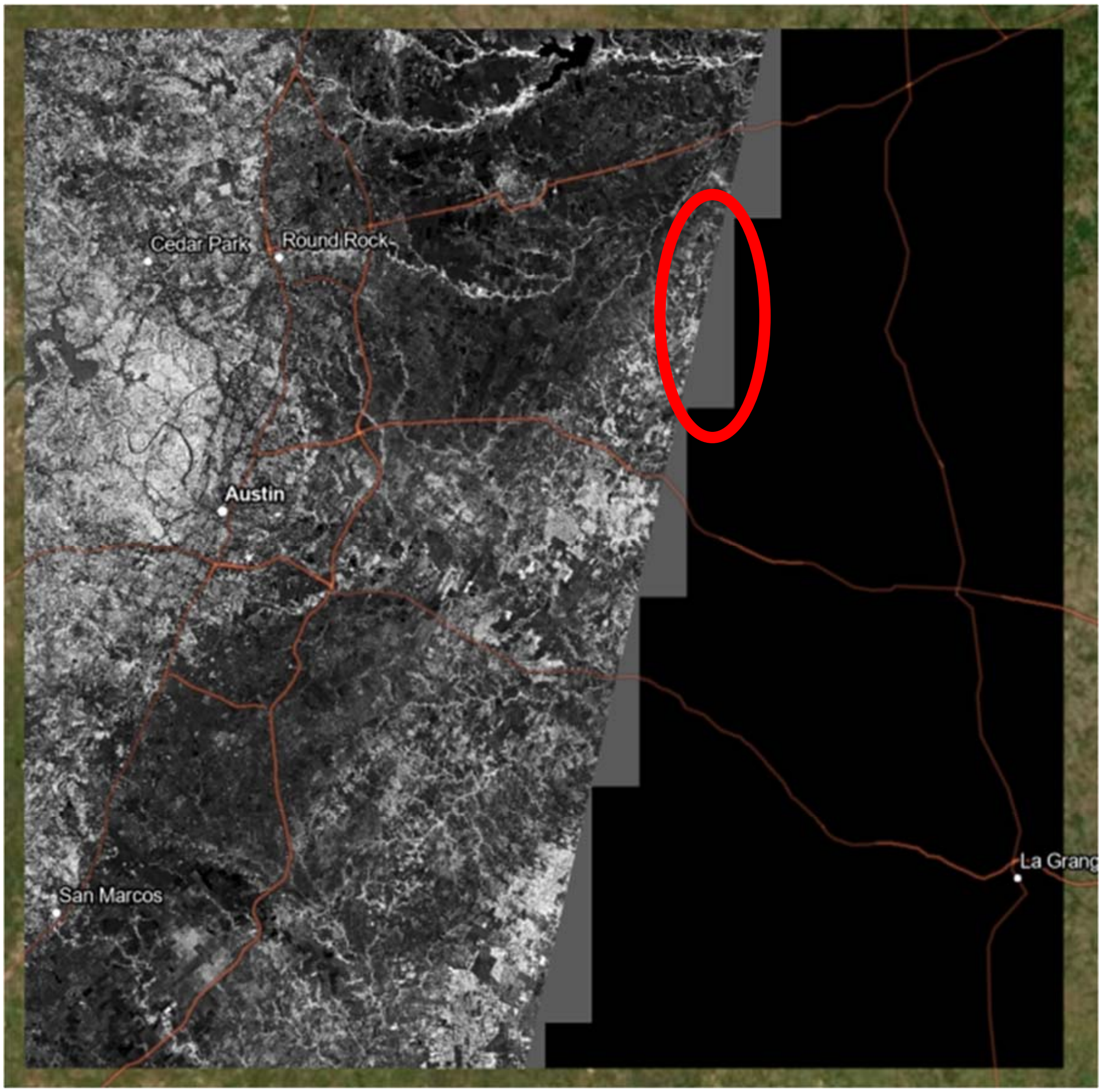


Figure 6-2. Examples of partial LAI images for Austin TX region that have missing data on the west side (top) or on the east side (bottom)

2. **Identify missing data (see SentinelRTNStep2.py).** Missing data in each image can be replaced by the data from the nearest (in time) scene. Clouds are typically flagged by SNAP with a value of "9" and "10" in the SCL image and so can be identified as missing data. Cloud shadows and the borders of some clouds are assigned values of 2, 3, 7 and 8 in the SCL images but in some cases these values are not clouds but are locations with low vegetation cover. They can be at least partially identified by checking to see if a given location is assigned these values in most scenes (in which case it is probably due to the landcover) or in only a few scenes (in which case it is probably due to clouds).

3. **Replace missing data and identify maximum vegetation cover fraction (see SentinelRTNStep3.py).** The missing vegetation cover data identified in step 2 is replaced

by the data from the nearest (in time) scene. The annual maximum vegetation cover is then determined as the maximum of all cover fractions for a location.

4. **Replace missing data and identify maximum vegetation cover fraction (see SentinelRTNStep4.py) Interpolate LAI.** The missing LAI data identified in step 2 is replaced by the data from the nearest (in time) scene. The LAI range for any location is then set with a minimum of zero (SNAP will give some negative numbers) and a maximum based on the maximum allowed LAIv (SNAP gives some unrealistically high values of LAI). MEGAN2 and MEGAN3 data preprocessors expect LAI data for 8-day time periods (46 in a year). This was designed to align with time step of MODIS, the primary source of satellite LAI and cover data. While Sentinel has the potential to have data coverage every 5 days (there are two satellites that each have a return of 10 days) cloud cover and missing scenes typically reduces the number of available images. For these cases, values for each 8-day period can be interpolated using a linear interpolation of the available images to generate a time weighted average value for each 8-day period.

5. **Calculate LAI and LAIv: vegetation covered surfaces (see SentinelRTNStep5.py).** 8-day LAIv inputs required for MEGAN2.1 and MEGAN3.1 are calculated by dividing LAI by annual maximum vegetation cover fraction. MEGAN3.2 uses just LAI and can be 8- or 10-day LAI.

6. **Integrate Texas urban LAI data into MEGAN global LAI data.** The 2021 version of the MEGAN global LAI data uses the CGL Service 2019 LAI 1km version 2 products which are provided as global, multi-band netCDF4 files with metadata according to the Climate and Forecast (CF) conventions (v1.6). The CGL LAI data are available at a temporal resolution of ~10 days and spatial resolution of 1/112 degrees (i.e., ~1 km) for the whole world.

   To merge SNAP-derived LAI with global data, the two datasets need to be aligned in temporal and spatial resolution. The global 10-day LAI data were converted into 8-day to align with the Texas urban LAI data. ArcGIS was used to remap the fine resolution LAI data from SNAP onto ~ 1km resolution global LAI data. To achieve this, the global dataset in ncf format was first converted into raster dataset for processing. ArcGIS “Zonal Statistics” was used to calculate average LAI data from the 10-m raster data within each grid cell of the global raster data. This can then be converted to a lat/lon grid with 1/112 degrees resolution (0.008-degree cell size) using ArcGIS Resample set to “bilinear interpolation”. Multiple input raster datasets were merged into one file using ArcGIS “Mosaic” tool for ease of processing. Once the SNAP-derived LAI data were merged and on the same grid as global data, the urban LAI data were merged into global data using ArcGIS “Mosaic to New Raster tool” and then converted into NCF format needed by MEGAN.

Appendix 2. Generating MEGAN growth form fraction distributions using high resolution imagery (e.g., NAIP)

Summary

This appendix section provides a tutorial on ArcGIS procedures for calculating growth form fraction distributions (e.g. tree cover, grass cover, shrub cover) using high resolution imagery than enables mapping at the level of individual trees and incorporating the results into the MEGAN growth form cover fraction landcover input files. The three main steps are:

Step 1. Generate high resolution tree cover fraction maps for selected areas
Approach: Use ArcGIS tools to process ultra-high resolution imagery (e.g. 60 cm NAIP) for selected region.
Product: High resolution growth form fraction maps for selected areas.
Step 2. Assess accuracy. Return to step 1 if necessary.
Product: Accuracy estimates.
Step 3. After achieving desired accuracy, integrate growth form fraction data into MEGAN global 1 km landcover database.
Approach: Use ArcGIS tools to reproject, aggregate and integrate growth form data into 1 km global grid.
Product: MEGAN global growth form cover fraction files (tree, shrub, crop, grass).

Step 1. Mapping individual urban trees using high resolution imagery and ArcGIS

There are four main tasks required for this approach to mapping trees and other land cover using high (60 cm or less) spatial resolution using ArcGIS. This includes 1) acquire the imagery, 2) create an object-based image (this groups neighboring pixels together based on their similarity to create objects that are then used for the classification) the objects include trees, buildings, lawns, cars, etc 3) compile a database of training samples to identify specific landcover types (e.g., oak trees, juniper trees, water bodies, laws, buildings, roads, etc) within the area, 4) classify all of the objects in the image.

1. Obtain the high-resolution imagery for the targeted region.

The example given here is for using NAIP imagery which is available for all of the contiguous US with new imagery available approximately every other year. The ArcGIS Portal Living Atlas contains NAIP data in an integrated file. Use the “USA NAIP Imagery: Color Infrared”. This is a 3 band dataset that includes two visible light (blue and green) and one near-infrared (NIR). It works well for identifying vegetation in most landscapes by may not be optimal for all landscapes so you could also try “USA NAIP: Natural Color” which is also three bands but has red visible light instead of infrared. CLIP the national NAIP database to generate an image of the target region. With 64 Gb Ram, an intel i9 chip or better, and an SSD drive you should be able to do an area of several thousand km^2 (note that this is ~10 billion points for 50 cm data). Output the clipped image (Right click on image, select data, export raster) to a TIF file that will be used for the processing described below. If a specific year and month is desired (NAIP is available about every 2 years and can be taken in different months of the year which might help with identifying certain vegetation types) the data can be accessed by right-clicking on the base "USA NAIP Imagery: Color Infrared" file and selecting individual NAIP files (~ 7 km x ~7 km). NAIP files can also be downloaded from location such as https://datagateway.nrcs.usda.gov/. You can mosaic the individual tiles into one file for processing using ArcGIS "Mosaic to new raster".

2. **Use ACRGIS Geoprocessing tool "Segment Mean Shift" to segment and create an image with discrete objects (e.g., individual features such as buildings, trees, lawns, lakes).**

Select the TIF file generated in step 1 and then go to Imagery> Classification tools > Segmentation (or Geoprocessing tool "Segment Mean Shift"). Settings that generally work well for heterogeneous urban landscapes are: spectral detail= 18, spatial detail=18, minimum segment size = 50

Other settings may be optimal for other specific landscapes.

An example of a small (800 m x 800 m) object based segmented image based on NAIP color infrared (NIR blue and 2 visible bands) is shown in Figure 6-3.

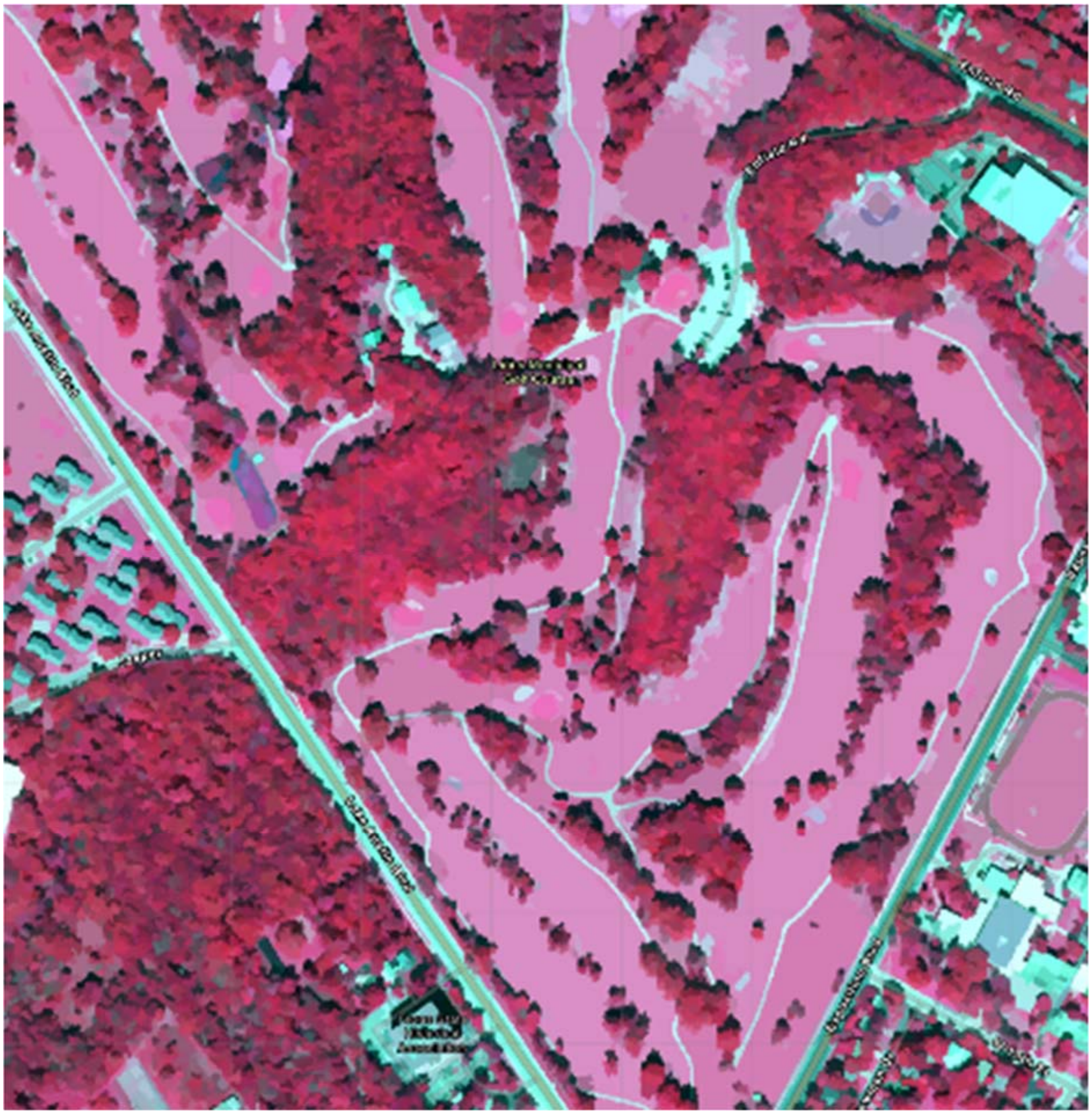


Figure 6-3. Object based segmented image based on 50 cm NAIP visible and IR bands. The figure shows the Lions Municipal Golf Course in Austin TX. Trees are dark red, grass is pink, buildings, roads and trails are grey/bluish/white in this false color image.

3. Use ArcGIS to create an ESRI Classifier Description (.ecd) file

ESRI Classifier Description files are used to classify images. To create one, first select the segmented object image file generated in step 2 and then go to Imagery> Classification tools > **Training Samples Manager**. Create a schema (or upload an existing one) that includes the targeted landcover types that are present within the image. To add classes to a new schema, right click on "new schema". This may include several types of trees (e.g., live oak, juniper, pine, pecan, deciduous oaks, other deciduous trees, etc), several ground cover types (e.g., shrub, green grass, dry grass, crops, bare soil), and several non-vegetated types (e.g., dark water, clear water, pavement, colored buildings, rock). Populate each landcover class with training samples by first selecting the class and then using the “segment picker” to select “objects” in the segmented object image that represents the landcover type (i.e, select live oak and then click on a live oak tree in the image). Ground cover and non-vegetated types will also need training samples. Breaking down total tree cover into individual species or subtypes requires obtaining geolocated features (points or polygons) representing each tree type and uploading them to ArcGIS. They can then be used within the training samples manager to identify trees representing the targeted types. The features used for training specific tree types can be obtained from geolocated municipal tree inventories (available for some cities such as Houston and Austin) or from surveys that can be conducted in situ in the field or virtually using Google Earth and Google Street Map.

After obtaining sufficient training samples (usually dozens of each landcover type but could require more) they can be used with Imagery> Classification tools > **Classify** with classifier= Random Trees (default values). The “Random Tree” classifier works well for some urban areas but other settings may work better for other landscapes. This task may require some iteration. Use ArcGIS **“Tabulate Area”** to calculate the total area covered by each landcover type. After an initial classification is generated and assessed to determine if landscapes are accurately assigned landcover types, this step can be redone to generate additional training areas for any landcover types that are incorrectly classified.

4. **Use** ArcGIS **“ClassifyRaster” Function to classify segmented object images (generated by step 2) using ESRI Classifier description (.ecd) file (generated by step 3).**

An ESRI Classifier description file and a segmented image are used to generate a classified image with various types of trees and other landcover. The RECLASS function to generate a numeric growth form fraction file at 0.5 m resolution. An example of the resulting 0.5 m resolution landcover distribution data is shown in Figure 6-4 and can be compared with the natural color image shown in Figure 6-5.

Figure 6-4. Landcover map (50 cm resolution) of Lions Municipal Golf Course in Austin showing trees (various types in different shades of green), grass (yellow) and bare soil (brown), buildings and pavement (white, grey and red).

Figure 6-5. NAIP natural color image (60 cm resolution) of Lions Municipal Golf Course in Austin.

5. Use ArcGIS "Reclassify" to set numeric values to cover types.

The landcover classes created in section 1.4 can be converted to numeric values using ArcGIS **"Reclassify"**. For example, if landcover types 1,2, and 3 are trees, then a numeric tree cover map can be created by assigning a value of 1 to each of those categories and a value of 0 to all others. Indeterminate classes (e.g., shadows) may be assigned a value between 0 and 1.

Step 2. Assess accuracy

Trees and other growth form distributions can be assessed using geolocated observations of individual trees and other vegetation types. These data can be generated in several ways including:

1) i-Tree: this program uses random sampling of hundreds of points on a google Earth image where the user identifies objects at random locations within a specified zone which can be user defined or geopolitical boundaries such as city or county boundaries.
2) Available geolocated tree inventories and other growth forms. These are becoming increasingly available for city street trees. The inventory should include latitude, longitude, growth form (and tree species) and can be uploaded into ArcGIS as a CSV file. The "Display XY data" (right click on file in standalone tables) is used to generate a layer feature class of points. Then the "extract values to points" tool is used to find the classification value for each point. The attribute table of the resulting file can be opened and copied and pasted into an excel spreadsheet.
3) Conducting field surveys. Additional geolocated tree inventories and growth form data can be generated with field surveys.
4) Available tree and vegetation cover summaries. Published reports provide average vegetation cover fraction for some cities and other areas. This includes urban FIA, city reports, and other publications.

Use ArcGIS "Zonal Statistics as Table" to calculate vegetation cover fractions for defined areas (such as city boundaries) that can be assessed with the cover fractions obtained from the above sources.

Step 3. Integrate data into MEGAN global 30 second (~1 km) grid growth form database.
The 2021 version of the MEGAN global growth form database uses the CGL Service landcover products version 3 dataset. The CGL data are available as an integrated global dataset for 2015, with annual updates after that, on global scale (78.25N to 60S) at a resolution of 0.0099206 (~100 m at the equator) and accuracy of 80% when compared to 28000 independent validation points. These data were generated using imagery from the PROBA-V satellite and released in September 2020 (Buchhorn et al. 2020; https://land.copernicus.eu/global/products/lc). The MEGAN Growth Forms include tree, shrub, crop, and herbaceous with the remaining area considered to be barren of vegetation. The CGL tree cover fraction can be underestimated in heterogeneous landscapes, including urban areas and arid woodlands. The higher resolution (30-m) NLCD tree cover data can capture some of these missing trees and has been integrated into the CGL data (i.e., the higher NLCD values are used for urban areas and woodlands). However, ultra-high resolution imagery (e.g., the 60 cm NAIP) indicates that both the CGL and NLCD also misses 10 to >50% of trees in the heterogeneous urban landscape. As shown in Figure 6-5, the NAIP imagery can capture the tree cover contributed by isolated individual trees. Urban development can also lead to rapid changes in time (e.g., deforestation for housing developments) which should be considered, for example, if comparing 2015 CGL data with 2018 NAIP data. There could also be an increase from tree planting efforts although that would typically take longer.

MEGAN (30 second resolution = 0.008333333 degree cell size) and CGL (3.57 second resolution = 0.000992063 degree cell size) data are on a lat/lon grid while NLCD (30 m resolution) and NAIP (50 cm resolution) are on the Albers_Conical_Equal_Area grid. The 50-cm NAIP data can be set to the NLCD 30-m grid using ArcGIS “Aggregate” (cell factor of 30). This can then be converted to a lat/lon grid with 1.2 second resolution (00033333333 degree cell size) using ArcGIS Resample set to “nearest neighbor”. The data can then be set to the MEGAN 30 second resolution (0.008333333 degree cell size) using the ArcGIS Aggregate tool with a Cellfactor of 25. The domain (in ArcGIS environment) needs to be set so that it covers a latitude and longitude that starts and ends on a 30 second grid value so that the grids will line up. This 30 second file can then be integrated into the MEGAN global 30 second grid. This can be accomplished by integrating the entire image or a subset such as within the boundaries of a city or for all of the urban locations within the image.

Appendix 3. Estimating tree cover fraction and compiling random tree locations using i-Tree

This appendix describes how to use the i-Tree tool to generate a random tree location database and to estimate area average tree cover fraction (and fractions of other cover types) for individual cities and other regions. The results can be used to assess the tree cover estimates generated using the approach described in Appendix 2 and to generate a database of random tree geolocations that can be used for the Virtual Urban Tree Survey approach described in Appendix 4.

Step 1. Initiate project

The i-Tree tool can be accessed at canopy.itreetools.org. Detailed tutorials are provided at www.itreetools.org/support. The first step is to either start a new (i.e. for a new city) project ("Project" > "File" > "New/Start/over") or continue with an existing project by loading the project from "project" > "file" > "open" and skip down to step 2. A project area can be defined by select an administrative region such as a city or county or you can draw an area- either within i-Tree or created in another program (e.g., ArcGIS) and imported as a shapefile.

You will need to create a landcover scheme. A typical one would include the following seven classes: Deciduous tree, Evergreen tree, Shrub, Grass and Bare Soil, Built (e.g., roads, buildings), Water, Barren" or import one that you have previously created and saved. Grass and Bare soil may be combined since what appears as bare soil in a winter image may appear as grass in a summer image. Barren indicates a location containing bare rock, desert sand or other surface that appears unlikely to have vegetation at any season.

Step 2. Identify landcover types and generate tree cover percentage and database of randomly located trees

The i-Tree tool will randomly choose points within the selected domain and the operator will classify the point as one of the landcover types in the landcover scheme. A typical minimum number of points is 300. This can be tested by seeing if the percent tree cover values change if you add an additional 100 points. Save the project file often so you don't lose your data and always save when you finish. After completing all data points necessary and saving your project you can transfer the results (mean and standard error for each cover class) to a spreadsheet. All of the data is stored in a project with a name that you give it with a file type ".itrcnpy". This file can be used to create a database of randomly located trees that can be used for the Virtual Urban Tree Survey described in Appendix 4. This can be accomplished by outputting all of the points into a spreadsheet and then removing all of the points that are not trees. The file can then be used to create a KML file that can be imported into Google Earth for the Virtual Urban Tree Survey.

Appendix 4. Virtual Urban Tree Survey

This appendix describes procedures for quantifying city average urban tree species composition that is referred to here as the Virtual Urban Tree Survey (VUTS) approach. VUTS takes a dataset of trees that are randomly distributed throughout a city, generated using the procedures described in Appendix 3, and identifies them using a virtual tree identification key, described in Appendix 5.

Traditional approaches for characterizing tree species composition include random plots, tree inventories, and tree community surveys. An example of the random plot method is the USFS Forest Inventory Assessment (FIA) approach that uses ground surveys of 1/6 acre plots that are randomly selected within each of the main vegetation types in a landscape that could be an urban area. FIA selects one randomly located plot per ~3 square miles. Every tree in the plot is identified and the diameter at breast height (DBH) of each tree is measured which enables an estimate of crown cover. The FIA approach has been applied to most US forests including a few urban forests such as in Houston, Austin and San Antonio. Tree inventories involve identifying every tree in an area and often include DBH measurements. Urban tree inventories are becoming increasingly common, but they typically include only trees on city property which generally is only ~15% of all trees in a city. Tree community surveys identify the dominant trees in a landscape and so can provide a qualitative estimate of the dominant trees in a city. All of these approaches require a survey team to travel to the landscape being characterized. The VUTS approach was developed for this project as a low cost alternative that does not require on site measurements. The VUTS approach was assessed by comparison with random plot and tree inventory approaches.

VUTS is an extension of the i-Tree procedures (see Appendix 3) which are designed to quantify tree cover and potentially broad tree type categories such as evergreen and deciduous trees. VUTS takes this to the next level of BVOC emission types that can be identified using widely available Google Earth imagery including aerial views and street views.

Appendix 5. Virtual Urban Tree Survey (VUTS) Texas Tree Keys

This appendix contains the taxonomic keys for identify trees using the VUTS tree identification approach described in Appendix 4. A taxonomic key is a simple tool used to identify a specific object, in this case to identify a tree. It begins by looking at large, important features that can divide the possible answers into a few large groups, thus quickly ruling out most of them. The key shown in Appendix 5A is designed for identifying trees at locations where a Google Street View image of the tree is available. It considers a number of features that can be seen in a Google Street View image. If a Google Street View image is not available for a specific location, then the aerial view key shown in Appendix 5B can be used with a Google aerial image view. These are available for nearly all locations, but image quality (resolution) can vary and the number of seasons with imagery can vary. This key relies on features that can be seen from aerial images. The accuracy achieved with both keys can be improved with availability of images from more than one season.

Appendix 5A. Key for Texas trees for which Google Street View imagery is available (i.e. the tree location is near a road with Google Street View imagery).

1 leaves are narrow (< 3mm); needles or scale like
 2 leaves are needles……………………………………………………………**Pine**
 2 leaves are scale like……………………………………………………………**Juniper**
1 leaves are wider(> 5 mm)
 3 leaves are very large and fan shaped……………………………………………**Palm**
 3 leaves not as above
  4 leaves are compound (with leaflets) ……………………………………**Group 1**
  4 leaves are simple
   5 large leaves (> 10 cm wide) that are lobed; almost as wide as long
    6 leaves star shaped, with deep lobes; bark grey to brown……………………**Sweetgum**
    6 leave with 3 to 5 shallow lobes; bark pale white………………………**Sycamore**
   5 leaves not as wide as above and/or longer than wide
    7 leaves have lobes……………………………………………**Deciduous Oak**
    7 leave are entire or toothed but do not have lobes
   8 leaves are linear, length is > than 3 times the width
    9 leaf edge finely toothed; upright branching habit……………………**Willow**
    9 leaf entire with bristled tip, often appearing clumped;
      branching more horizontal ………………………**D. Oak (Willow and Sandjack Oak)**
   8 leaf length is < than 3 times the width
    10 leaf widest at base
      11 leaf about as wide as long (triangular); bark dark brown……………………**Tallow**
     11 leaf obviously longer than wide; bark pale gray……………………**Hackberry**
   10 leaf widest at middle or near the top end
    12 leaf widest near the top end……………………………………………**Group 2**

12 leaf widest near the middle……………………………………………………………………………………………..…..**Group 3**

Group 1 compound leaves

1 compound leaves are oppositely attached to twig with 5 to 9 leaflets…………………………………………….….**Ash**
1 compound leaves are alternately attached to twig.
 2 leaves are once compound- largest leaflet at tip……..…………………………………………………………..……**Pecan**
 2 leaves are multi compound
  3 leaves are bi-pinnate………………………………………………………………………………………..…………………..**Mesquite**
  3 leaves are double compound
   4 leaf has a terminal leaflet……………………………………………………………………………………….……………C**hinaberry**
   4 leaf has no terminal leaflet…………………………………………………………………………………………………**Acacia**

once compound leaf
Ash
Pecan

bi-pinnate
Mesquite

terminal leaflet
double compound
Chinaberry

no terminal leaflet
double compound
Acacia

Group 2 – Leaves widest at top end
1 large tree with main trunk; deciduous leaves with blunt lobes at tip;.**D. Oak (Water and Blackjack Oak**)
1 small tree or shrub, often with many trunks leaves usually evergreen, with no lobes at tip
 2 leaves always evergreen, with teeth…………………………………………………………………………
………………...**Yaupon**
 2 leaves evergreen or deciduous, entire
…………………..……………………………………………………..……….**Persimmon**

blunt lobes
Water Oak

blunt lobes
Blackjack Oak

no lobes,teeth
Yaupon

no lobes, no teeth
Persimmon

Group 3 leave widest around the middle

1 evergreen leaves
 2 leaves opposite...**Privet**
 2 small tree or bush with spiny tips or bluntly toothed leaves...................................**Yaupon (Holly)**
 2 large tree
  3 large entire leaves, > 15 cm long and greater than 6 cm wide;
   branches angle upward ....**Magnolia**
  3 leaves < 14 cm long and < 3 cm wide; branches more horizontal.................................**Live Oak**
1 deciduous leaves
 2 leaves coarsely toothed, often appearing clumped;
  branching often horizontal......................................**D. Oak (Emory, Chinkapin, Chestnut Oak)**
 2 leaves entire to finely toothed; branching more vertical
  3 leaves finely toothed; large trees..**Elm**
  3 leaves entire
   4 shrub or small tree; usually multi trunked; showy flowers;.....................................C**rape Myrtle**
   4 large tree; usually one trunk...**Blackgum**

opposite leaves
Privet

spiny tips
Yaupon (Holly)

bluntly toothed
Yaupon

coarsely toothed
Emory Oak

coarsely toothed
Chinkapin Oak

coarsely toothed
Swamp Chestnut Oak

entire leaves
Blackgum

entire leaves
Crape Myrtle

finely toothed leaves
Elm

Appendix 5B. Key for Texas trees for which Google Street View imagery is not available (i.e. the tree location is not near a road with Google Street View imagery). Identification is based on Google Aerial View imagery.

1) Tree leaves present in late January/February – **Group A**
1) Tree leaves absent in late January/February – **Group B**

Group A
1) Tree with long single trunk with no lateral branches, dense crown of large leaves……………...….Palm
1) Tree with branched crown
 2) Tree leaf color/density appears the same in February, March, November, and December
  3) Mature trees are > 7 Meters in diameter
   4) Irregular crown, color light yellowish green long needles often visible in shadows………………Pine
   4) Symmetrical crown with wide leaves, color dark green….…………………………………..Magnolia
  3) Mature tree is < 7 Meters in diameter…………………………………………………….....Juniper
 2) Tree aerial appearance is not the same during these months. March is lighter than January/February
  5) Mature trees < 7 Meters in diameter
   6) Tree with bare branches exposed in Nov./Dec. usually near houses……………………………..Privet
   6) Tree with few branches exposed in Nov./Dec.,
    7) Tree often away from houses, irregular crown………...………………………...…………Yaupon
    7) Tree usually near house, rounded crown…….………………………………….……….Persimmon
  5) Mature tree > 7 Meters in diameter….…………………………………………..……….Live Oak

Group B
1) Small tree < 7 Meters in diameter, crown often circular
 2) No large branches visible in Nov./Dec., often pinkish color in middle, circular crown…...Crape Myrtle
 2) Some large branches in crown, not pinkish,
  3) Irregular crown, native……………………………..………….……………………………...Acacia
  3) Symmetrical crown, non-native, invasive………………..……………………...…………..Chinaberry
1) Mature tree > 7 Meters in diameter, crown often irregular

4) Tree usually with some leaves in November or December
5) Trees with some large white branches exposed in crown in Nov. or Dec…………………….Sycamore
5) Tree without large white branches exposed in crown in Nov. or Dec.
6) Most of the leaves left on tree in Nov./Dec, often colored……………………..……Deciduous Oak
6) Over ½ of the leaves missing in Nov./Dec
7)Tree has many leaves in March, many large branches visible in January…………...…………….Elm
7) Tree has fewer leaves in March, tree has smaller branches producing a fuzzy image……...Mesquite
4) Tree usually without leaves in November or December
8) Tree with abundant small branches and minimal bare space between branches. Tree may or may not have large branches visible
9)Tree mostly in natural riparian areas – usually not near houses…………….....……...……...Blackgum
9) Tree in urban areas near houses or in natural areas, less often in riparian areas
10) Tree with some large branches intermixed with small branches…..………………..……..…..…Ash
10) Tree without large branches visible, all branches similar in size…..………………..…….Sweetgum
8) Tree with large branches visible and bare ground visible between branches
11) More than 50% bare ground visible under tree in Nov./Dec, little green up in March
12) Tree dark green in October, common…………………..…………………...…………..……Pecan
12) Tree lighter green in October, not as common
13) Tree with large and many medium size branches, native……………………..…...……Hackberry
13) Tree with large branches and few medium size branches……………………..…...Willow or Tallow
11) Less than 50% bare ground visible under tree in Nov./Dec……………………………………....Elm