1 **Sensitivity of biogenic volatile organic compounds (BVOCs) to land**

2 **surface parameterizations and vegetation distributions in California**

3 [1]Chun Zhao, [1]Maoyi Huang, [1]Jerome D. Fast, [1]Larry K. Berg, [1]Yun Qian, [2]Alex

4 Guenther, [2]Dasa Gu, [1]Manish Shrivastava, [1]Ying Liu, [3]Stacy Walters, [3]Gabriele Pfister,

5 [4]Jiming Jin, [1]John E. Shilling, [5,6]Carsten Warneke

6

7

8 [1]*Atmospheric Science and Global Change Division, Pacific Northwest National*

9 *Laboratory, Richland, WA, USA*

10 [2]*Department of Earth System Science, University of California, Irvine, California, USA*

11 [3]*National Center for Atmospheric Research, CO, USA*

12 [4]*Departments of Watershed Sciences and Plants, Soils, and Climate, Utah State*

13 *University, Logan, UT, USA*

14 [5]*National Oceanic and Atmospheric Administration, Earth System Research Laboratory,*

15 *USA*

16 [6]*CIRES, University of Colorado, Boulder, CO, USA*

17

18

20

21 *Corresponding author: Chun Zhao, phone: (509) 371-6372; email: chun.zhao@pnnl.gov

22

23

Abstract. Current climate models still have large uncertainties in estimating biogenic

trace gases, which can significantly affect atmospheric chemistry and secondary aerosol

formation that ultimately influences air quality and aerosol radiative forcing. These

uncertainties result from many factors, including uncertainties in land-surface processes

and specification of vegetation types, both of which can affect the simulated near-surface

fluxes of biogenic volatile organic compounds (BVOCs). In this study, the latest version

of Model of Emissions of Gases and Aerosols from Nature (MEGAN v2.1) is coupled

within the land surface scheme CLM4 in the Weather Research and Forecasting model

with chemistry (WRF-Chem). In this implementation, MEGAN v2.1 shares a consistent

vegetation map with CLM4 for estimating BVOC emissions. This is unlike MEGAN v2.0

in the public version of WRF-Chem that uses a standalone vegetation map that differs

from what is used by land surface schemes. This improved modeling framework is used

to investigate the impact of two land surface schemes, CLM4 and Noah, on BVOCs and

examine the sensitivity of BVOCs to vegetation distributions in California. The

measurements collected during the Carbonaceous Aerosol and Radiative Effects Study

(CARES) and the California Nexus of Air Quality and Climate Experiment (CalNex)

conducted in June of 2010 provide an opportunity to evaluate the simulated BVOCs.

Sensitivity experiments show that land surface schemes do influence the simulated

BVOCs, but the impact is much smaller than that of vegetation distributions. This study

indicates that more effort is needed to obtain the most appropriate and accurate land

cover datasets for climate and air quality models in terms of simulating BVOCs, oxidant

chemistry, and consequently secondary organic aerosol formation.

1. Introduction

Volatile organic compounds (VOCs) in the atmosphere play an important role in atmospheric chemistry, and therefore can significantly affect ozone and secondary organic aerosol (SOA) formation and ultimately air quality and climate [e.g., Chameides et al., 1992; Fehsenfeld et al., 1992; Andreae and Crutzen, 1997; Pierce et al., 1998; Poisson et al., 2000; Sanderson et al., 2003; Claeys et al., 2004; Arneth et al., 2010]. Significant effort has been made on obtaining accurate predictions of atmospheric VOC concentrations; however, there remain large differences between observed and simulated values. These uncertainties result from many factors, including biogenic emission rates that are influenced by near-surface meteorological processes, sub-surface processes, representation of vegetation distributions, and plant biology [Guenther et al., 2013].

Biogenic emissions are a major source of VOCs [e.g., Zimmerman et al., 1978; Mueller, 1992] in the atmosphere. In particular, isoprenoids (consisting mainly of isoprene and monoterpenes) that dominate biogenic VOCs (BVOCs) have been extensively investigated during the last five decades [e.g., Went, 1960; Rasmussen, 1972; Zimmerman et al., 1979; Lamb et al., 1987; Pierce et al., 1998; Niinemets et al., 1999 and 2002; Arneth et al., 2007; Schurgers et al., 2009; Guenther et al., 1995 and 2012]. BVOC emissions were originally computed offline, producing prescribed emission inventories used by regional and global models [e.g., Huang et al., 2011]. However, emissions of BVOCs depend on diurnal, multi-day, and seasonal variations in light intensity, temperature, soil moisture, vegetation type, and leaf area index (LAI) [e.g., Pierce et al., 1998; Niinemets et al., 1999 and 2002; Arneth et al., 2007; Schurgers et al., 2009; Guenther et al., 2012]. Therefore, various BVOC emission algorithms have been

69   proposed that extrapolate limited laboratory and field measurements to prescribed

70   regional and global ecosystems [e.g., Pierce et al., 1998; Niinemets et al., 1999 and 2002;

71   Arneth et al., 2007; Schurgers et al., 2009; Guenther et al., 1995 and 2012]. The

72   uncertainties in biogenic emission schemes are mainly due to the scarcity of observations

73   of BVOC fluxes and vegetation distributions over regional scales. Inappropriate coupling

74   strategies between biogenic emission and land-surface schemes may also introduce errors

75   in estimating atmospheric BVOCs. For example, some models specify different

76   vegetation distributions for biogenic emissions and land-atmosphere interaction processes

77   as applied in different parts of models.

78   BVOCs play a significant role in affecting the air quality and regional climate

79   over California, where there have been many studies, such as the Carbonaceous Aerosol

80   and Radiative Effects Study (CARES) [Zaveri et al., 2012] and the California Nexus of

81   Air Quality and Climate Experiment (CalNex) [Ryerson et al., 2013], investigating the

82   impacts of BVOCs and their interaction with anthropogenic pollutants. In the past 20

83   years, California's economy has grown rapidly and the population has increased by 33%

84   [Cox et al., 2009]. Although California has reduced the emissions of most primary

85   pollutants, poor air quality still affects the well-being of millions of people. Nearly all

86   Californians live in areas that are designated as nonattainment for the state (about 99%)

87   and national (about 93%) health-based O_3 and/or PM standards. Accurate predictions of

88   O_3 and PM concentrations are needed to develop effective attainment strategies, but this

89   is complicated, in part, due to uncertainties associated with long-range transport of

90   pollutants and local natural emission sources such as BVOCs.

91   In California, the complex topography and distribution of vegetation makes it

92   difficult for models to capture the variability of BVOCs at regional and local scales. For

93   example, Fast et al. [2014] showed that simulated biogenic emissions varied by as much

94   as a factor of 2 within 8 km of an observation site in Cool, California. They also found

95   that daytime mixing ratios of isoprene and monoterpenes from a regional simulation

96   using the Weather Research and Forecasting model with chemistry (WRF-Chem) [Grell

97   et al., 2005; Fast et al., 2006] are usually a factor of two smaller than the observations

98   collected both at the rural Cool site and an urban Sacramento site. Conversely, simulated

99   monoterpene mixing ratios were similar to observations during the day but by a factor of

100   three too high at night at the observation site in Cool. They suggested that the biogenic

101   emission rates calculated based on the Model of Emissions of Gases and Aerosols from

102   Nature version 2.0 (MEGAN v2.0) might contribute to major biases in their simulations.

103   Knote et al. [2014] also found that their simulations using WRF-Chem with MEGAN

104   v2.0 produced BVOC concentrations that were too small over Los Angeles, and

105   suggested that there might be deficiencies in the description of vegetation in urban areas.

106   Thus, it is evident that uncertainties in simulated atmospheric BVOCs can arise from how

107   well vegetation is represented in models. Furthermore, to our knowledge, none of the

108   numerous chemical transport modeling studies for California have investigated the

109   sensitivity of BVOC simulations to land surface schemes and vegetation distributions.

110   To better understand the uncertainties in simulating BVOCs associated with land

111   surface schemes and vegetation distributions in California, the latest version of MEGAN

112   (MEGAN v2.1) is coupled into the CLM4 land surface scheme of WRF-Chem in this

113   study. Multiple sensitivity experiments are conducted using this improved modeling

114   framework at a relatively high spatial resolution to capture the region's complex

115 topography and vegetation distribution. Simulations are conducted using WRF-Chem

116 with "fully" coupled version of CLM4 and MEGAN v2.1 (i.e., CLM4 and MEGAN share

117 a consistent vegetation dataset) and compared with the measurements collected during

118 CARES and CalNex conducted in June 2010. This new coupling also adds the capability

119 of quantifying the impact of different vegetation distributions on simulating BVOCs.

120 Simulations are also performed using two land surface schemes (Noah and CLM4)

121 coupled with MEGAN v2.0. As with previous studies using WRF-Chem, MEGAN v2.0

122 uses a different vegetation dataset from the land surface schemes. The WRF-Chem

123 experiments with MEGAN v2.0 and MEGAN v2.1 are included together here as a

124 reference for future studies in the community and for users interested in migrating from

125 the widely used v2.0 to v2.1.

126   The rest of manuscript is organized as follows. Sections 2 and 3 describe the

127 WRF-Chem model and the observations used in this study, respectively. The sensitivity

128 of modeling BVOCs to the land surface schemes and the vegetation distributions are

129 analyzed in section 4. The findings are then summarized and discussed in section 5.

130

131 **2. Model Description and Experimental Design**

132 **2.1 WRF-Chem**

133   The WRF-Chem (v3.5.1) configuration is similar to that used by Fast et al. [2014]

134 for studying aerosol evolution over California, except that this study excludes aerosols

135 and focuses on simulated BVOCs. The model includes numerous options for the

136 treatment of physics and chemistry processes. In this study, the SAPRC-99

137 photochemical mechanism [Carter, 2000a,b] is selected to simulate gas-phase chemistry,

138 and the Fast-J parameterization [Wild et al., 2000] for photolysis rates. For all the

139 simulations in this study, we use the Yonsei University (YSU) parameterization [Hong et

140 al., 2006] for the planetary boundary layer (PBL), the Monin-Obukhov similarity theory

141 [Paulson, 1970] to represent the surface layer, the Morrison two-moment

142 parameterization [Morrison et al., 2009] for cloud microphysics, the Kain-Fritsch

143 parameterization [Kain 2004] for sub-grid scale clouds and precipitation, the rapid

144 radiative transfer parameterization (RRTMG) for longwave and shortwave radiation

145 [Iacono et al., 2008]. Since Fast et al. [2014] has evaluated the simulated meteorological

146 fields and gases and aerosols with a similar model configuration, this study will focus

147 primarily on the BVOC simulation.

148 **2.2 Land surface schemes**

149   Two land surface schemes, Noah and CLM4.0, are used to quantify how

150 differences in the treatment of land surface processes, including latent and sensible heat

151 fluxes, soil moisture, and surface albedo, affect near-surface meteorological conditions

152 and consequently simulated BVOC emissions and concentrations. The Noah land surface

153 scheme, described by Barlage et al. [2010] and LeMone et al. [2010a, 2010b], has been

154 used in numerous studies with WRF-Chem. Noah has four soil layers, with a total depth

155 of two meters and a single slab snow layer that is lumped with the top-soil layer, which is

156 set to a combined depth of 10 cm. It uses the 24 United States Geological Survey (USGS)

157 land-use types, and does not treat sub-grid scale variability within a model grid cell.

158   The CLM4 (Community Land Model version 4.0) [Lawrence et al. 2011; Jin et

159 al., 2012] was recently coupled and released with WRF (since v3.5) as one of the land

160 surface scheme options. CLM4 in global and region applications has been shown to be

161 accurate in describing snow, soil, and vegetation processes [Zeng et al., 2002; Jin and

162 Miller, 2007; Zhao et al., 2014]. CLM4 includes five layers for snow, 10 layers for soil,

163 and a single-layer for vegetation. The soil is divided into 19 categories defined according

164 to percentages of sand and clay. The two-stream approximation [Dickinson, 1983] is

165 applied to vegetation when calculating solar radiation reflected and absorbed by the

166 canopy as well as radiation transfer within the canopy. Each model grid cell can be

167 divided into a maximum of 10 smaller cells to account for sub-grid scale heterogeneity

168 and its impact on the land surface processes. The 24 USGS land-use types are mapped to

169 the 16 plant functional types (PFTs) in CLM4 based on a lookup table derived from

170 Bonan et al. [1996]. Additional technical details of CLM4 are provided in Oleson et al.

171 [2004].

172 **2.3 MEGAN and coupling with CLM4**

173 MEGAN is a modeling framework for estimating fluxes of biogenic compounds

174 between terrestrial ecosystems and the atmosphere using simple mechanistic algorithms

175 to account for the major known processes controlling biogenic emissions [Guenther et al.,

176 2006, 2012]. Two versions (v2.0 and v2.1) of MEGAN are used in this study. MEGAN

177 v2.1 is an update from MEGAN v2.0 [Guenther et al., 2006; Sakulyanontvittaya et al.,

178 2008] that includes additional compounds, emission types, and controlling processes.

179 MEGAN v2.1 estimates emissions (F_i) for 19 compound classes (*i*) from terrestrial

180 landscapes based on emission factors ($\varepsilon_{i,j}$) at standard conditions for vegetation type *j*

181 with fractional grid box areal coverage χ_j, i.e., $F_i = \gamma_i \Sigma \varepsilon_{i,j} \chi_j$, where γ_i is emission activity

182 factor from the processes controlling emission responses to environmental and

183 phenological conditions [Guenther et al., 2006, 2012].

184   For emission factors, MEGAN v2.0 enabled users to customize vegetation

185   emission type schemes ranging from detailed (e.g. individual plant species or sub species)

186   to generic (e.g. a few broad vegetation categories). MEGAN2.1 emission factors can be

187   specified from gridded maps based on species composition and species-specific emission

188   factors or by using PFT distributions and the PFT specific emission factors. MEGAN2.0

189   defines emission factors as the net flux of a compound into the atmosphere, while

190   MEGAN2.1 emission factor represents the net primary emission that escapes into the

191   atmosphere but is not the net flux because it does not include the downward flux of

192   chemicals from above canopy. The difference in the definition (net flux versus primary

193   emission) of emission factors affects the emission factors of compounds with

194   bidirectional exchange but does not impact MEGAN isoprene and monoterpene emission

195   factors because they have small deposition rates relative to emission rates. In this study,

196   both MEGAN v2.0 and v2.1 estimate biogenic species emissions based on the PFT

197   distributions and the PFT specific emission factors. MEGAN v2.0 and v2.1 use 4 and 16

198   PFTs, respectively, as described below in Section 2.4.

199   The publically available version of WRF-Chem includes the MEGAN v2.0

200   scheme for calculating BVOC emission fluxes (WRF-Chem user guide:

201   http://ruc.noaa.gov/wrf/WG11/Users guide.pdf). It has been widely used for gas and

202   aerosol simulations [e.g., Shrivastava et al., 2011, 2013; Gao et al., 2011, 2014; Knote et

203   al., 2014; Fast et al., 2014]. In the released version, MEGAN v2.0 can be used with any

204   land surface scheme available in WRF-Chem including Noah and CLM4. However,

205   MEGAN v2.0 was originally not coupled into the land surface scheme in WRF-Chem

206   (since v3.1). The biogenic emission calculation in MEGAN uses both instantaneous and

207   the past-days' surface air temperature and solar radiation. MEGAN v2.0 obtains the

208   instantaneous value from the land surface scheme and the past-days' value from the

209   climatological monthly mean dataset. In contrast, MEGAN v2.1 obtains both values

210   directly from CLM. Figure 1 shows the example of the comparison between the input

211   climatological and model simulated monthly mean surface air temperature in June. It is

212   apparent that the monthly-averaged simulated surface air temperature is much different

213   from the climatology value. In addition, the vegetation dataset (referred to as VEG-M,

214   will be discussed in Section 2.4) used in MEGAN v2.0 for calculating BVOC emission

215   fluxes is also different from the one used by the land surface scheme, which allows

216   MEGAN v2.0 to be used with any of the available land surface schemes (e.g., Noah and

217   CLM4) in WRF-Chem. This inconsistency in vegetation distributions may introduce

218   errors in simulating emissions and concentrations of BVOC. To avoid this inconsistency,

219   we have coupled MEGAN v2.1 with WRF-Chem embedded in the CLM4 land surface

220   scheme. Therefore, the coupling of MEGAN v2.1 and CLM4 in WRF-Chem now has the

221   same functionality as CLM4 in the Community Earth System Model (CESM) [Lawrence

222   et al. 2011]. With this coupling strategy, MEGAN v2.1 also uses the same vegetation

223   dataset (i.e., 16 PFTs converted from the USGS dataset as discussed in Section 2.2) that

224   CLM4 uses for all other land surface processes; this means, however, that MEGAN v2.1

225   can only be used with CLM4 in WRF-Chem. In addition, MEGAN v2.1 can compute

226   BVOC emissions that account for the sub-grid variability of vegetation distributions

227   within CLM4.

228   **2.4 Vegetation datasets**

229   As mentioned previously, the first 16-PFT dataset (referred to as USGS hereafter)

230   used by CLM4 is converted from the default 24 USGS land cover dataset used by WRF-

231   Chem based on a lookup table derived from Bonan et al. [1996]. This method is also

232   applied to three other 16-PFTs datasets (referred to as VEG1, VEG2, and VEG3) used by

233   CLM4 in WRF-Chem. The sensitivity of simulating BVOC emissions by CLM4 to these

234   four 16-PFTs datasets is quantified. The VEG1, VEG2, and VEG3 datasets are derived

235   from different sources as described next.

236   The VEG1 dataset is from the PFT fractional cover product by Ke et al. [2012],

237   which was developed from the Moderate Resolution Imaging Spectroradiometer

238   (MODIS) PFT classifications for the year 2005 for determining seven PFTs including

239   needleleaf evergreen trees, needleleaf deciduous trees, broadleaf evergreen trees,

240   broadleaf deciduous trees, shrub, grass and crop for each 500 m pixel. The WorldClim 5

241   arc-minute (0.0833°) [Hijmans et al., 2005] climatological global monthly surface air

242   temperature and precipitation data was interpolated to a 500 m grid and used to further

243   reclassify the PFTs into 15 PFTs, and fractions of crop grasses were mapped based on the

244   method presented in Still et al. [2003]. Pixels with barren land and urban areas were

245   reassigned to the bare soil class. The bare soil and the 15 PFTs from the 500-m grid were

246   then aggregated to a 0.05° grid.

247   The VEG2 dataset is obtained from the NCAR CESM data repository [Oleson et

248   al., 2010], available on a 0.05° grid and derived using a combination of the 2001 MODIS

249   Vegetation Continuous Field (VCF), MODIS land cover product for year 2000

250   [Lawrence and Chase, 2006; Lawrence and Chase, 2007], and 1992-1993 AVHRR

251   Continuous Field Tree Cover Project data [Lawrence and Chase, 2007; Lawrence et al.,

252   2011]. The monthly surface air temperature and precipitation data from Willmott and

253   Matsuura [2001] was used to further reclassify the seven PFTs into bare soil and 15 PFTs

254   in the tropical, temperate and boreal climate groups based on climate rules described by

255   Bonan et al. [2002]. Fractions of crop grasses were mapped based on the method

256   presented in Still et al. [2003].

257   The VEG3 dataset is derived from a high-resolution (30 arc-second) dataset over

258   the U.S. with 16 PFT classifications for the year 2008. The dataset was created by

259   combining the National Land Cover Dataset (NLCD, Homer et al., 2004) and the

260   Cropland Data Layer (see http://nassgeodata.gmu.edu/CropScape/), both of which were

261   based on the 30-m LANDSAT-TM satellite data. Vegetation species composition

262   information   was   obtained   from   the   Forest   Inventory   and   Analysis   (see

263   http://www.fia.fs.fed.us) and the soil data from the Natural Resources Conservation

264   Services (see http://sdmdataaccess. nrcs.usda.gov/). The processing included adjusting

265   the NLCD tree cover estimates in urban areas to account for the substantial

266   underestimation of trees in the LANDSAT-TM data [Duhl et al., 2012]. This was

267   accomplished using the regionally specific adjustment factors for urban NLCD developed

268   by Greenfield et al. [2009] using the high-resolution imagery.

269   Figure 2 shows the spatial distributions of the dominant PFT in each 4×4 km^2 grid

270   cell of the simulation domain from each of the four datasets. Not only are the grid-

271   dominant PFTs very different among the four datasets, but the sub-grid distributions of

272   PFTs are different as well (not shown). The domain-averaged fractions of 16 PFTs from

273   the four datasets listed in Table 1 also illustrate the differences in PFT distributions. For

274   example, the fraction of temperate broadleaf deciduous tree ranges from 0.4% in VEG1

275   to 1.8% in VEG2 and the fraction of temperate broadleaf deciduous shrub ranges from

276   10.8% in VEG3 to 37.5% in VEG1. In MEGAN v2.0 of WRF-Chem, only four PFTs

277   (refer to VEG-M) that are broadleaf tree, needleleaf tree, shrub, and herbaceous

278   vegetation categories, are considered for the biogenic emission calculation because they

279   are the only ones included in the MEGAN v2.0 PFT scheme. As discussed previously,

280   these are different from the USGS vegetation distribution used by Noah and CLM4 and

281   may cause additional biases. The distributions of the four PFTs used by MEGAN v2.0 are

282   shown in Figure 3. This difference in PFT distributions can affect the BVOC emission

283   calculations primarily through determining distributions of PFT specific emission factors

284   and leaf area indices (LAI) that are prescribed with PFTs in this study. For example,

285   Figure 4 shows the biogenic isoprene emission factor for each PFT prescribed in

286   MEGAN v2.0 and MEGAN v2.1 in CLM4. In MEGAN v2.1, it shows that temperate

287   broadleaf deciduous tree (PFT 7 listed in Table 1) has a large isoprene emission factor,

288   while temperate needleleaf evergreen tree (PFT 1 listed in Table 1) has a small isoprene

289   emission factor. A similar difference between broadleaf trees and needleleaf trees is

290   indicated for MEGAN v2.0. Figure 5 shows the spatial distributions of averaged biogenic

291   isoprene emission factor used in MEGAN v2.0 and v2.1 with different PFTs. It is evident

292   that the difference in the distributions of PFTs results in a significant difference in spatial

293   distributions of the isoprene emission factor. Figure 6 shows the spatial distributions of

294   LAI used in MEGAN v2.0 and v2.1. The differences in the spatial distributions of LAI

295   can significantly affect the biogenic emission calculation in MEGAN. It should be noted

296   that in MEGAN v2.0 used in WRF-Chem, the LAI used for the calculation of the

297   biogenic emissions is prescribed using the 4 PFTs, which is different than the land

298   scheme that uses the LAI derived from the 24 USGS land categories.

13

299

2.5 Numerical experiments

The simulations are performed using a domain encompassing California (Fig. 1) with a horizontal grid spacing of 4 km and 279×279 grid cells (113°W-128°W, 32°N-43°N) and 51 vertical layers up to 100 hPa with about 35 layers below 2 km. The simulation period is from May 25 to June 30 2010, but only the results in June are used for analysis to allow for the model to "spin-up" realistic distributions of trace gases. The initial and boundary conditions are prescribed by large-scale meteorological fields obtained from the North American Regional Reanalysis (NARR) data with updates provided at 6-h intervals, which also provide the prescribed sea surface temperature (SST) for the simulations. The modeled u and v wind components and temperature in the free atmosphere above the planetary boundary layer are nudged towards the NARR reanalysis data with a time scale of 6 hours [Stauffer and Seaman, 1990]. Chemical lateral boundary conditions are from the default profiles in WRF-Chem, which are based on the averages of mid-latitude aircraft profiles from several field studies over the eastern Pacific Ocean [McKeen et al., 2002].

Anthropogenic emissions were obtained from the CARB 2008 ARCTAS emission inventory developed for the NASA Arctic Research of the Composition of the Troposphere from Aircraft and Satellite (ARCTAS) mission over California [Pfister et al., 2011]. The CARB inventory contains hourly emissions for a 13-day period using a 4-km grid spacing over California. We created diurnally averaged emissions from 5 of the weekdays and 2 of the weekend days and used those averages for all weekdays and weekends and applied these over the entire simulation period. Anthropogenic emissions

322   from the 2005 National Emissions Inventory (NEI) (WRF-Chem user guide from

323   http://ruc.noaa.gov/wrf/WG11/Users guide.pdf) were used for regions outside of

324   California. Biomass burning is not considered in the present study, because satellite

325   detection methods indicated that there were very few fires in California during the

326   simulation period. Biogenic emissions were computed on-line using the MEGAN model

327   and lumped into isoprene, terpenes, and sesquiterpenes for the SAPRC-99 photochemical

328   mechanism.

329   As discussed previously, multiple numerical experiments summarized in Table 2

330   are conducted with different combinations of land surface schemes and vegetation

331   datasets to investigate the sensitivity of BVOC simulation to land surface schemes and

332   vegetation distributions. First, we conduct two experiments using MEGAN v2.0 coupled

333   with the Noah (Mv20Noah) and CLM4 (Mv20CLM) land surface schemes, respectively.

334   The Noah land surface scheme is only coupled with MEGAN v2.0 in WRF-Chem. In

335   these two experiments, the two land surface schemes use the USGS vegetation

336   distributions while MEGAN v2.0 uses a separate vegetation map (VEG-M) to estimate

337   BVOC emissions. By comparing these two experiments, the impact of land surface

338   schemes on simulated BVOC concentrations are examined. Second, we conduct four

339   experiments using MEGAN v2.1 embedded in the CLM4 land surface scheme with four

340   different vegetation datasets, i.e., USGS (Mv21USGS), VEG1 (Mv21V1), VEG2

341   (Mv21V2), and VEG3 (Mv21V3). The differences among these four experiments show

342   the impact of vegetation distributions on simulated BVOC concentrations.

343   We note that MEGAN v2.0 and v2.1 use different vegetation datasets and are

344   implemented in WRF-Chem in different ways, but the objective of this study is not to

15

345 explore how the formulations of these two versions of MEGAN affect BVOC

346 concentrations. The better way for exploring the version difference of MEGAN is to

347 implement both versions in the same way and use the same vegetation dataset. The

348 simulated BVOC emissions and concentrations by WRF-Chem with MEGAN v2.0 and

349 MEGAN v2.1 are included together here as a reference for future studies in the

350 community and for users interested in migrating from the widely used v2.0 to v2.1.

351

352 3. Observations

353   Measurements of VOCs collected by proton transfer reaction mass spectrometer

354 (PTR-MS) instruments [Lindinger et al., 1998] and a gas chromatography instrument

355 [Gentner et al., 2012] over California during June of 2010 as part of the CARES and

356 CalNex campaigns are used to evaluate the simulated isoprene and monoterpene

357 concentrations. CARES was designed to address science issues associated with the

358 interactions of biogenic and anthropogenic precursors on SOA, black carbon mixing state,

359 and the effects of organic species and aerosol mixing state on optical properties and the

360 activation of cloud condensation nuclei [Zaveri et al., 2012]. As shown in Figure 7,

361 ground-based instruments were deployed at two sites (T0 and T1) in northern California:

362 T0 in Sacramento (38.649 °N, -121.349°W, ~ 30 m m.s.l., denoted by red upward

363 triangle) and T1 in Cool (38.889°N, -120.974°W, ~ 450 m m.s.l., denoted by red

364 downward triangle), a small town located about 40 km northeast of Sacramento. The U.S.

365 Department of Energy (DOE) Gulfstream 1 (G-1) research aircraft sampled

366 meteorological, trace gas, and aerosol quantities aloft in the vicinity of the T0 and T1

367 sites, denoted by black lines in Figure 8. Zaveri et al. [2012] described the

16

368 instrumentation for each of the surface sites and Shilling et al. [2013] described VOC

369 measurements on the G-1. Most of the sampling during CARES occurred between 2 and

370 28 June, and only the aircraft sampling within 1 km of the surface is used to evaluate

371 model simulations because G-1 sampled below 1 km for the majority of time.

372   CalNex was designed to address science issues relevant to emission inventories,

373 dispersion of trace gases and aerosols, atmospheric chemistry, and the interactions of

374 aerosols, clouds, and radiation [Ryerson et al., 2013]. Ground-based instruments were

375 deployed at two sites in southern California as shown in Figure 7: one in Pasadena

376 (34.141°N, -118.112°W, ~240 m m.s.l., denoted by the red circle) and one in Bakersfield

377 (35.346°N, −118.965°W, ~ 123 m m.s.l., denoted by the red square). The NOAA WP-3D

378 research aircraft sampled meteorological, trace gas, and aerosol quantities aloft along

379 flight paths shown in Figure 7 (denoted by blue lines). While most of the CalNex aircraft

380 tracks below an altitude of 1 km were conducted in southern California in the vicinity of

381 the Los Angeles basin, the WP-3D also flew within the Central Valley and in the vicinity

382 of Sacramento on some days. A detailed description of the instrumentation for each of the

383 CalNex surface sites and mobile platforms is given by Ryerson et al. [2013]. Most of the

384 sampling during CalNex was conducted before June 16 and only the aircraft sampling

385 below 1 km is used to evaluate the model simulations.

386

387 **4. Results**

388 **4.1 Impact of land surface schemes**

389 4.1.1 Biogenic isoprene and monoterpene emissions

390   Figure 7 shows the spatial distributions of biogenic isoprene emissions averaged

391   over June for the six simulations listed in Table 2. Biogenic isoprene emissions occur in

392   vegetated regions of California with the highest emission rates along the foothills of the

393   Sierra Nevada where oak trees are the dominant plant species. To show the difference in

394   biogenic isoprene emissions among the cases more clearly, Figure 8a and 8b zoom in on

395   the CARES (northern California) and CalNex (southern California) sampling regions,

396   respectively. In both regions the differences in land surface schemes had a relatively

397   small impact on the biogenic isoprene emissions over California in terms of both spatial

398   distribution and magnitude, although the emissions from Mv20CLM were a little larger

399   than those from Mv20Noah. The domain summed biogenic isoprene emissions for the

400   entire month of June from Mv20Noah and Mv20CLM are 1.4×10^9 and 1.6×10^9 mole,

401   respectively. Figure 9a and 9b are similar to Figure 8a and 8b, except that biogenic

402   monoterpene emission fluxes are shown. In general, the spatial patterns of emissions of

403   the two biogenic species are similar, except that the peak areas of monoterpene emissions

404   are shifted slightly. For example, the peak monoterpene emissions in northern California

405   occur further northeast at higher elevations of the Sierra Nevada that are dominated by

406   needleleaf evergreen trees. The impact of land surface schemes on biogenic monoterpene

407   emissions is also small over California in terms of both spatial patterns and magnitudes,

408   although the emissions from Mv20CLM are a little larger than those from Mv20Noah.

409   The domain summed biogenic monoterpene emissions for the entire month of June from

410   Mv20Noah and Mv20CLM are 1.0×10^8 and 1.1×10^8 mole, respectively.

411   The similarity in estimating biogenic emissions between the experiments with two

412   land surface schemes is also summarized in Figures 10 and 11, which show the average

413   diurnal biogenic isoprene and monoterpene emission rates at the four observation sites.

414   The similarity between Mv20Noah and Mv20CLM (red and orange lines) is likely due to

415   the same vegetation map in MEGAN v2.0 to estimate biogenic emissions. Although the

416   two land surface schemes produce slightly different values of surface temperature (Fig.

417   1), soil moisture (not shown), and net solar radiation near the surface (not shown), their

418   impact on the biogenic emissions was small. Both BVOC species have peak emission

419   rates in the early afternoon. One noteworthy difference in diurnal variation of the two

420   biogenic species emission rates is that there is no isoprene emitted during the night while

421   the amount of monoterpenes emitted during the night is small but not negligible. This can

422   contribute to differences in the diurnal variation of the mixing ratios of two biogenic

423   species, as will be discussed next.

424   4.1.2 Isoprene and monoterpene mixing ratios

425   Figures 12a,b and 13a,b show the spatial distributions of monthly-averaged

426   surface mixing ratios of isoprene+MVK(methyl-vinylketone)+MACR(methacrolein) and

427   monoterpenes, respectively, around the CARES (northern California) and the CalNex

428   (central and southern California) sampling regions simulated by the six experiments

429   listed in Table 2. Due to the fast chemical transition from isoprene to MVK and MACR,

430   the sum of isoprene+MVK+MACR mixing ratios can better reflect the impact of

431   biogenic isoprene emissions than isoprene mixing ratio alone [Shilling et al., 2013]. In

432   general, the spatial patterns and magnitudes of surface isoprene+MVK+MACR and

433   monoterpene mixing ratios over the two regions are similar from the two MEGAN v2.0

434   experiments with the Noah and CLM4 land surface schemes, respectively. The spatial

435   patterns of surface mixing ratios of isoprene+MVK+MACR and monoterpenes are

436   similar to the spatial variability in the emission rates.

19

437   There is difference between the two experiments at specific locations, which is

438   partly reflected in the comparison of average diurnal variations of surface mixing ratios

439   of isoprene+MVK+MACR and monoterpenes at the four observation sites shown in

440   Figure 14 and Figure 15. At the Bakersfield site, only isoprene mixing ratios were

441   reported so that the comparison is for isoprene only. Note that the values for the

442   Bakersfield and Pasadena sites are averaged over the first two weeks of June to be

443   consistent with the observations. Although both experiments with Noah and CLM4 (red

444   and orange lines, respectively) simulate similar isoprene emission fluxes with the

445   maximum in the afternoon (Fig. 10), their respective isoprene+MVK+MACR mixing

446   ratios are different at the four sites, particularly at site T0, where the Mv20CLM

447   simulated isoprene+MVK+MACR mixing ratios during the daytime are about a factor of

448   2 larger than those from Mv20Noah. This inconsistence mainly results from the

449   differences in the near surface meteorology, such as net surface radiation and

450   temperature, between the two experiments (not shown) that affects photochemistry, but

451   this impact of surface meteorology occurs only at limited locations. When compared to

452   the observations, both experiments significantly underestimate the

453   isoprene+MVK+MACR mixing ratios except at the Bakersfield site. Figure 15 is

454   identical to Figure 14, except for surface monoterpene mixing ratios. Note that there were

455   no monoterpene data reported for the Bakersfield and Pasadena sites, so only the

456   simulation results are shown. In contrast to isoprene+MVK+MACR, monoterpenes

457   exhibit peak surface mixing ratios during the nighttime due to the strong photolysis

458   activity that makes the lifetime of monoterpenes short during the daytime and the small

459   emissions into a shallow boundary layer during the nighttime (Fig. 11). In general, the

460   difference between the Mv20Noah and MV20CLM experiments in monoterpene mixing

461   ratios is relatively small at these four sites, particularly during the daytime. When

462   compared to the observations, both experiments overestimate the diurnal variation and

463   the nighttime surface monoterpene mixing ratios at the T0 and T1 sites.

464   Figures 16 and 17 show the comparison of the observed and simulated mixing

465   ratios of isoprene+MVK+MACR and monoterpenes, respectively, along the G-1 and

466   WP-3D flight tracks below 1 km. Model results are sampled along the flight tracks. As

467   shown in Figure 7, the G-1 flight mainly flew over northern California around the T0 and

468   T1 sites, while the WP-3D flew over a larger area covering both southern California and

469   the Central Valley. To better reflect the spatial variability in the BVOCs, the flight tracks

470   of both flights are separated into two regions as indicated by the black lines in Figure

471   12a,b and Figure 13a,b. For the G-1, the flight paths are divided into regions of southwest

472   and northeast of the black line shown in Figures 12a and 13a that is parallel to the Sierra

473   Nevada. The two regions have significantly different vegetation (Fig. 2) resulting in large

474   differences in biogenic emissions. For the WP-3D, the flight paths are divided into

475   regions of south and north of the black line shown in Figures 12b and 13b to separate

476   southern California and the Central Valley. Over southern California, the measured

477   isoprene+MVK+MACR mixing ratios by the PTR-MS over the WP-3D are the upper

478   limit since the PTR-MS may have a small interference in urban areas for isoprene and

479   MVK+MACR.

480   In Figure 16, it is interesting to note that both experiments Mv20Noah and

481   Mv20CLM reasonably capture the variability seen in the G-1 isoprene+MVK+MACR

482   measurements over the southwest region even though they underestimate the surface

483 observations by as much as a factor of 2 at the T0 site (Fig. 14). While both experiment

484 mixing ratios are slightly smaller than observed, the Mv20CLM simulated mixing ratios

485 are a little larger than those from Mv20Noah and closer to the observations. Over the

486 northeast region, both experiments produced similar mixing ratios that were significantly

487 smaller than the observations, which is consistent with the comparison between the

488 simulated and observed isoprene+MVK+MACR at the T1 site (Fig. 14). As shown in

489 Figure 16, the Mv20CLM simulation produced somewhat larger isoprene+MVK+MACR

490 mixing ratios than Mv20Noah in both southern California and the Central Valley. This is

491 consistent with the comparison at the Bakersfield and Pasadena surface sites. Both

492 simulations also underestimate and overestimate the isoprene+MVK+MACR mixing

493 ratios over southern California and the Central Valley, respectively. The comparison of

494 isoprene+MVK+MACR with aircraft observations may suggest that both experiments

495 underestimate biogenic isoprene emissions over the forested foothills of Sierra Nevada

496 and southern California around Los Angeles, but overestimate the emissions over the

497 Central Valley. The model biases may also be affected, to some extent, by anthropogenic

498 emissions with large uncertainties and the associated non-linear chemistry due to the

499 mixing of anthropogenic and biogenic plumes [Fast et al., 2014].

500   Figure 17 shows that both experiments Mv20Noah and Mv20CLM significantly

501 underestimate the monoterpene mixing ratios over all the regions sampled by the G-1 and

502 WP-3D aircraft and that the differences between the simulations were negligible. The

503 average monoterpene mixing ratios sampled by the G-1 below 1 km was comparable to

504 the surface measurement at the T0 site during the daytime, but somewhat higher than the

505 observations at the T1 site. The simulated mixing ratios averaged along the flight tracks

506 were much smaller than those at the two surface sites, suggesting that it may be difficult

507 for model to simulate the large spatial heterogeneity of the monoterpene mixing ratios.

508 This could result from the biases in biogenic monoterpene emissions and/or the chemical

509 mechanism for monoterpene oxidation and how chemistry is coupled with turbulent

510 mixing within the simulated convective boundary layer. It also needs to be noted that the

511 G-1 and WP-3D measured monoterpene mixing ratios are generally below the Limit Of

512 Detection (LOD) of instruments (0.1-0.3 ppbv). Therefore, the true monoterpene mixing

513 ratios could be range between 0 ~ 0.1-0.3 ppbv, which may also contribute to the

514 discrepancy between observations and simulations.

515

516 **4.2 Impact of vegetation distributions**

517 4.2.1 Biogenic isoprene and monoterpene emissions

518   Figures 8a,b and 9a,b show that the differences in biogenic isoprene and

519 monoterpene emission distributions due to using the various vegetation datasets are larger

520 than the differences resulting from the two land surface schemes. The domain summed

521 biogenic isoprene emissions for the entire month of June are 2.3, 0.76, 1.7, and 0.92

522 ($\times 10^9$ mole) from the experiments of Mv21USGS, Mv21V1, Mv21V2, and Mv21V3,

523 respectively, and biogenic monoterpene emissions are 2.5, 1.7, 1.9, and 1.1 ($\times 10^8$ mole)

524 from the four experiments, respectively. Each of the four simulations produces high

525 biogenic isoprene and monoterpene emission rates along the Sierra Nevada that is

526 covered mainly by oak and pine forests. However, the different forest classifications and

527 their coverage (Table 1) produce different biogenic isoprene and monoterpene emission

528 rates along the Sierra Nevada. Another distinct difference among these four simulations

529   is found over the Central Valley, where the Mv21V1 and Mv21V3 experiments produce

530   significantly lower biogenic isoprene and monoterpene emissions than the Mv21USGS

531   and Mv21V2 experiments. This results from their different spatial distributions of

532   vegetation types. For example, the vegetation dataset in MV21USGS assigns a relatively

533   larger fraction of vegetation over the Central Valley to broadleaf trees, which are biggest

534   contributors of isoprene emissions (Fig. 4).

535   The differences in the spatial distributions of biogenic isoprene and monoterpene

536   emissions due to various vegetation distributions is also illustrated by the average diurnal

537   biogenic isoprene emission rates at the four observation sites shown in Figures 10 and 11.

538   For example, the Mv21V3 simulation produces the largest biogenic isoprene and

539   monoterpene emissions at three of the sites. At the T1 site over the forested foothills of

540   the Sierra Nevada, the Mv21USGS and Mv21V3 simulations produce much larger

541   biogenic isoprene emissions than Mv21V1 and Mv21V2. Even though forest is the

542   dominant vegetation type along the foothills of the Sierra Nevada in all four vegetation

543   datasets (Fig. 2), their different forest classifications and coverage result in biogenic

544   isoprene emission rates that differ by as much as a factor of 8 at the T1 site. Similar to

545   isoprene emissions, the Mv21USGS simulation produces the largest monoterpene

546   emissions at the T1 site. However, the differences in monoterpene emissions among the

547   four vegetation dataset experiments are smaller overall than that for biogenic isoprene

548   emissions. Different vegetation distributions for a typical urban area can also lead to

549   differences in biogenic isoprene and monoterpene emissions. For example at the urban

550   T0 and Pasadena sites, biogenic isoprene and monoterpene emission rates are almost 0 in

551   the Mv21USGS and Mv21V1 experiments, while the rates were significant larger in the

552   Mv21V3 experiment. This could have profound implications on local oxidant chemistry

553   influencing urban air quality.

554   4.2.2 Isoprene+MVK+MACR and monoterpene mixing ratios

555   As expected, the differences in biogenic isoprene and monoterpene emissions

556   among the four different vegetation distribution experiments lead to large differences in

557   the simulated surface isoprene+MVK+MACR and monoterpene mixing ratios (Figs.

558   12a,b and 13a,b). Although all the four experiments simulate highest biogenic

559   isoprene+MVK+MACR and monoterpene mixing ratios along the forested foothills of

560   Sierra Nevada, the Mv21V1 and Mv21V3 experiments have the lowest

561   isoprene+MVK+MACR and monoterpene mixing ratios, respectively, corresponding to

562   their lowest biogenic emission rates. Over the Central Valley, Mv21USGS and Mv21V2

563   experiments produce significantly higher isoprene+MVK+MACR mixing ratios than the

564   other two experiments, while Mv21V3 simulates the lowest monoterpene mixing ratios

565   among all the experiments.

566   At the T1 site located in the forested foothills of Sierra Nevada, the Mv21V1

567   simulation produces the lowest isoprene+MVK+MACR mixing ratios (Fig. 14),

568   significantly underestimating the peak concentrations during the day. In contrast, the

569   Mv21USGS and Mv21V3 simulations reasonably capture the observed

570   isoprene+MVK+MACR mixing ratios during the daytime. All four experiments

571   underestimate the isoprene+MVK+MACR mixing ratios by about a factor of 2 during the

572   night. This may indicate that the transported isoprene+MVK+MACR from the

573   surrounding areas of T1 was too low. The negative biases of simulated

574   isoprene+MVK+MACR mixing ratios over the areas surrounding T1 can be reflected by

575   Figure 16 that shows all the four experiments significantly underestimate the observed

576   isoprene+MVK+MACR mixing ratios below 1 km in the northeast area around the T1

577   site (Fig. 12a). Figure 16 also shows that Mv21USGS and MV21V3 simulate larger

578   isoprene+MVK+MACR mixing ratios averaged over the northeast region of northern

579   California than Mv21V1 and Mv21V2. All four experiments produce similar surface

580   monoterpene mixing ratios, which are smaller than that from the Mv20Noah and

581   Mv20CLM with MEGAN v2.0 and are closer to the observed values particularly during

582   the night. This is consistent with their much lower biogenic monoterpene emissions

583   during the night (Fig. 11). The four experiments with MEGAN v2.1 simulate higher

584   daytime monoterpene mixing ratios averaged along the flight tracks below 1 km than the

585   two experiments with MEGAN v2.0. The simulated mixing ratios are still much lower

586   than the aircraft observations, although the simulated surface mixing ratios are higher

587   than the observations at the T1 site (Fig. 15). However, the aircraft measured

588   monoterpene mixing ratios may also be higher than the true values due to the LOD of

589   instruments (0.1-0.3 ppbv).

590   At the T0 site, an urban site, the vegetation coverage in both the Mv21USGS and

591   Mv21V1 experiments is small so that the isoprene+MVK+MACR and monoterpene

592   mixing ratios are significantly lower than observed during the daytime. The Mv21V2 and

593   Mv21V3 experiments reasonably simulate isoprene+MVK+MACR mixing ratios during

594   the daytime. Over the area surrounding the T0 site (i.e., the southwest area in Fig. 12a), it

595   is interesting to note that the Mv21USGS and Mv21V2 simulations produced larger

596   isoprene+MVK+MACR mixing ratios than Mv21V1 and Mv21V3 and closer to the

597   observations (Fig. 16). This is mainly due to the relatively large isoprene+MVK+MACR

26

598 mixing ratios over the northwest corner of CARES sampling region (Fig. 12a) in the

599 Mv21USGS and Mv21V2 simulations, consistent with the distributions of biogenic

600 isoprene emissions over the region. The Mv21V2 and Mv21V3 simulations produced

601 higher monoterpene mixing ratios than Mv21USGS and Mv21V1, but are still smaller

602 than the observed values during the daytime not only for the T0 site but also for the

603 region surrounding T0 as shown in Figure 17.

604   At the Bakersfield site, the experiments often simulate significantly larger

605 isoprene mixing ratios than the observations, except for the Mv21V1 simulation that was

606 always too small. The Mv21V3 simulation produced the highest isoprene mixing ratios

607 among the experiments. This is consistent with its biogenic isoprene emission rates (Fig.

608 10). In addition, the observed surface isoprene mixing ratios show negligible diurnal

609 variation in contrast to the experiments that produced larger diurnal variations. The

610 Mv21V3 simulation produced peak isoprene mixing ratios during the daytime that were

611 likely controlled by its large daytime local biogenic isoprene emission rates (Fig. 10).

612 The Mv21USGS and Mv21V2 simulations produced peak isoprene mixing ratios during

613 the early evening, possibly the result of chemistry and transport from regions with higher

614 biogenic emissions. All four experiments produce small diurnal variation of surface

615 monoterpene mixing ratios. The Mv21USGS and Mv21V3 simulations produce larger

616 monoterpene mixing ratios than the other two, consistent with their local emission rates

617 (Fig. 11).

618   At the Pasadena site, the Mv21V3 simulation reproduces the observed diurnal

619 variation of isoprene+MVK+MACR mixing ratios reasonably well. This is consistent

620 with the area surrounding the Pasadena site, in which the Mv21V3 simulation produces

621   the largest mixing ratios of isoprene+MVK+MACR both at the surface (Fig. 12b) and

622   aloft (Fig. 16) in the vicinity of Los Angeles. The other three experiments simulated

623   significantly smaller mixing ratios of isoprene+MVK+MACR. Although the values from

624   the other three experiments are still smaller than the observations, they are much closer to

625   the aircraft measurements (within a factor of 2) than at the Pasadena site (Fig. 14).

626   Among the four vegetation sensitivity simulations, Mv21V3 produces higher surface

627   monoterpene mixing ratios than the other three experiments, consistent with their

628   emission rates (Fig. 11). All four vegetation sensitivity experiments produced much

629   lower monoterpene mixing ratios below 1 km (Fig. 17), compared to the aircraft

630   measurements over southern California that may overestimate the true values due to the

631   LOD of instruments (0.1-0.3 ppbv).

632   As discussed previously, all four experiments simulate significantly different

633   isoprene+MVK+MACR and monoterpene mixing ratios over the Central Valley (Figs.

634   12a,b and 13a,b). The Mv21USGS and Mv21V2 simulations produce much larger

635   isoprene+MVK+MACR mixing ratios (0.6 ppbV and 0.5 ppbV, respectively) over the

636   Central Valley than the observed values (~0.1 ppbV). The Mv21V1 and Mv21V3

637   simulations produce monoterpene mixing ratios much closer to observed values. This

638   may indicate that the fraction of broadleaf trees (the main emitter over the region) over

639   the Central Valley from the vegetation datasets USGS and VEG2 are overestimated or

640   the biogenic emission factors estimated for the broadleaf trees are overestimated for this

641   area. For monoterpenes, the Mv21V3 simulation was much smaller than observed, while

642   the mixing ratios from the other three experiments were more comparable. This suggests

643   that the fraction of vegetation emitting monoterpenes is significantly underestimated over

644   this area in the VEG3 dataset.

645

5. Summary and discussion

647   In this study, the latest version of MEGAN (v2.1) is coupled within the CLM4

648   land scheme as part of WRF-Chem. Specifically, MEGAN v2.1 is implemented into the

649   CLM4 scheme so that a consistent vegetation map can be used for estimating biogenic

650   VOC emissions as well as surface fluxes. This is unlike the older version of MEGAN

651   (v2.0) in the public-released WRF-Chem that uses a standalone vegetation map that

652   differs from what is used in land surface schemes. With this improved WRF-Chem

653   modeling framework coupled with CLM4-MEGAN v2.1, the sensitivity of biogenic VOC

654   emissions and hence of atmospheric VOC mixing ratios to vegetation distributions is

655   investigated. The WRF-Chem simulations are also conducted with the two land surface

656   schemes, Noah and CLM4, with the MEGAN v2.0 scheme for biogenic emissions in each

657   case. The comparison between the Noah and CLM4 driven MEGAN v2.0 biogenic

658   emissions not only serves for investigating the impact of different land surface schemes

659   on the emissions but also provides a reference for all previous studies that used the Noah

660   land surface scheme. Experiments are conducted for June 2010 over California,

661   compared with the measurements from the CARES and CalNex campaigns. The main

662   findings about the modeling sensitivity to the land surface schemes and vegetation

663   distributions include:

664   •   The WRF-Chem simulation with the CLM4 land surface scheme and the MEGAN

665   v2.0 module (Mv20CLM) produces similar biogenic isoprene and monoterpene

666   emissions in terms of spatial patterns, magnitudes, and diurnal variations as the one

667   with the Noah land surface scheme (Mv20Noah) in June over California. The

668   similarity in the biogenic emissions between the experiments using two different land

669   schemes is primarily because of using MEGAN v2.0 and the same vegetation map in

670   the two experiments. The spatial patterns and magnitudes of surface

671   isoprene+MVK+MACR and monoterpene mixing ratios are generally similar

672   between the two experiments with the Noah and CLM4 land surface schemes,

673   although there are significant differences at some specific locations due to their

674   differences in the near surface meteorology such as surface net radiation and

675   temperature. Compared with surface and aircraft measurements, both experiments

676   generally underestimate the daytime mixing ratios of isoprene+MVK+MACR but

677   overestimate the nighttime mixing ratios of monoterpenes.

678 • The experiments with the four vegetation datasets result in much larger differences in

679   biogenic isoprene and monoterpene emissions than the ones with the two land surface

680   schemes. The simulated total biogenic isoprene and monoterpene emissions over

681   California can differ by a factor of 3 among the experiments and the difference can be

682   even larger over specific locations. The comparison of mixing ratios of

683   isoprene+MVK+MACR and monoterpenes with the observations indicates the

684   simulation biases can be largely reduced with accurate vegetation distributions over

685   some regions of California. For example, at an observation site at the forested

686   foothills of Sierra Nevada, two experiments with the vegetation distributions from the

687   USGS and VEG3 datasets capture the observed daytime surface mixing ratios of

688   isoprene+MVK+MACR well, with values that are much larger than the experiments

689   with the other two vegetation datasets.

690   • Although vegetation distributions from some datasets do significantly improve the

691   model performance in simulating BVOC mixing ratios more than others, the optimal

692   vegetation dataset cannot be determined, because the improvement by vegetation

693   datasets has dependence on both the region and BVOC species of interest. For

694   example, over the Central Valley, the experiments with the VEG1 and VEG3

695   vegetation datasets simulate isoprene+MVK+MACR mixing ratios that are much

696   closer to observations than the USGS and VEG2 datasets, while the VEG3 dataset

697   significantly underestimates the observed monoterpene mixing ratios. Large biases

698   over some regions of California in all the experiments with current vegetation

699   datasets imply that more effort is needed to improve land cover datasets and/or

700   biogenic emission factors.

701   There are still some large biases existing over some regions of California

702   regardless of the vegetation distributions. For example, all the experiments significantly

703   underestimate the observed isoprene+MVK+MACR mixing ratios below an altitude of 1

704   km over the forest-covered Sierra Nevada. Over the Pasadena area, all the experiments

705   simulate significantly smaller monoterpene mixing ratios than observed. The biases in

706   BVOCs identified in this study may be partly due to inaccurate vegetation distributions in

707   all the vegetation distribution datasets. The biases can also result from the uncertainties in

708   BVOC emission factors for the individual types of vegetation commonly found in

709   California. The constraints on BVOC emission factors applied in models are limited due

710   to sparse measurements of BVOC emission fluxes. The MEGAN scheme in WRF-Chem

711   uses the global averaged emission factors for BVOC emissions for each PFT. Over

712   California, the broadleaf temperate trees are primarily oaks that have relatively higher

31

713   BVOC emission factors compared to the global mean values for temperate broadleaf

714   trees. In addition, the needleleaf trees are pines that have relatively larger monoterpene

715   emission factors compared to global mean values. These biases in emission factors may

716   partly explain why all the experiments generally underestimate mixing ratios of

717   isoprene+MVK+MACR and monoterpenes over the regions with large amounts of trees.

718   The MEGAN scheme using the location-specified emission factor maps that accounts for

719   species composition of trees may provide a better estimate on regional scales.

720   This study demonstrates large difference between the experiments with the two

721   versions of MEGAN (v2.0 versus v2.1), and that MEGAN v2.1 results in a better

722   comparison with the observations over some parts of the study domain. However, this

723   difference should not be fully attributed to the improvement of MEGAN between the two

724   versions, because the two versions also use different vegetation distributions. The results

725   highlight the importance of sub-grid vegetation distributions in simulating biogenic

726   emissions even at a relatively high horizontal grid spacing (e.g., 4 km in this study). The

727   biogenic emissions can be significantly different even though the dominant vegetation

728   within a model grid box is similar. The comparison of the simulations and the

729   observations at the surface sites and along the aircraft tracks reflects the large spatial

730   variability of biogenic emissions and BVOC mixing ratios over California. It is

731   challenging for model to capture such a spatial heterogeneity of BVOCs if the vegetation

732   distributions are not appropriately represented in the simulation. The relatively large

733   LOD of instruments on the aircrafts for monoterpenes compared to the true

734   concentrations also make the evaluation of simulated monoterpenes difficult. Over a

735   region with relatively low monoterpene concentrations, an instrument with lower LOD is

736   needed. It is also noteworthy that this study is in a relatively dry and warm season;

737   therefore the impact of biogenic emission treatments may change for other seasons and

738   during periods with higher cloudiness. A multiple-season investigation may be needed in

739   future. Finally, it is also noteworthy that factors other than biogenic emissions can

740   influence the simulated BVOC mixing ratios over California, such as anthropogenic

741   emissions and the oxidation mechanism of BVOCs used in simulations. Therefore,

742   additional direct measurements of biogenic emission fluxes are needed for a better

743   evaluation of simulated BVOC fluxes.

744

Code availability

746   The   WRF-Chem   version   3.5.1   release   can   be   obtained   at

747   http://www2.mmm.ucar.edu/wrf/users/download/get_source.html. Code modifications for

748   implementing MEGANv2.1 into CLM are available upon request by contacting the

749   corresponding author and will be released to public WRF-Chem version.

750

764

765
766

767 **Reference**

768 Andreae, M. and P. J. Crutzen: Atmospheric Aerosols : Biogeochemical Sources and

769   Role in Atmospheric Chemistry. Science, 276, 5315, 1052-1058,

770   doi:10.1126/science.276.5315.1052, 1997.

771 Arneth, A., S. P. Harrison, S. Zaehle, K. Tsigaridis, S. Menon, P. J. Bartlein, J. Feichter,

772   A. Korhola, M. Kulmala, D. O'Donnell, G. Schurgers, S. Sorvari, and T. Vesala:

773   Terrestrial biogeochemical feedbacks in the climate system, Nature Geoscience, 3,

774   525-532, 2000.

775 Arneth, A., Niinemets, Ü., Pressley, S., Bäck, J., Hari, P., Karl, T., Noe, S., Prentice, I.

776   C., Serça, D., Hickler, T., Wolf, A., and Smith, B.: Process‑based estimates of

777   terrestrial ecosystem isoprene emissions: Incorporating the effects of a direct CO_2‑

778   isoprene interaction, Atmos. Chem. Phys., 7, 31–53, 2007.

779 Barlage, M., F. Chen, M. Tewari, K. Ikeda, D. Gochis, J. Dudhia, R. Rasmussen, B.

780   Livneh, M. Ek, and K. Mitchell: Noah land surface model modifications to improve

781   snowpack prediction in the Colorado Rocky Mountains, J. Geophys. Res., 115,

782   D22101, doi:10.1029/2009JD013470, 2010.

783 Bonan, G. B.: A land surface model (LSM ver. 1.0) for ecological, hydrological, and

784   atmospheric studies: Technical description and user's guide. NCAR Tech. Note

785   4171STR, 150 pp, 1996.

786 Bonan, G. B., K. W. Oleson, M. Vertenstein, S. Levis, X. Zeng, Y. Dai, R. E. Dickinson,

787   and Z.-L. Yang: The Land Surface Climatology of the Community Land Model

788   Coupled to the NCAR Community Climate Model. J. Climate, 15, 3123–3149, doi:

789   http://dx.doi.org/10.1175/1520-0442(2002)015<3123:TLSCOT>2.0.CO;2, 2002.

790   Carter,W. P. L.: Documentation of the SAPRC-99 Chemical Mechanism for VOC

791   Reactivity Assessment, Draft report to the California Air Resources Board, Contracts

792   92–329 and 95–308, 8 May, available at: http://www.cert.ucr.edu/~carter/pubs/ (last

793   access:1 December 2015), 2000a.

794   Carter,W. P. L.: Implementation of the SAPRC-99 Chemical Mechanism into the

795   Models-3 Framework, Report to the United States Environmental Protection Agency,

796   29 January, available at: http://www.cert.ucr.edu/~carter/pubs/ (last access: 1

797   December 2015), 2000b.

798   Chameides, W. L., F. Fehsenfeld, M. O. Rodgers, C. Cardelino, J. Martinez, D. Parrish,

799   W. Lonneman, D. R. Lawson, R. A. Rasmussen, P. Zimmerman, J. Greenberg, P.

800   Mlddleton, T. Wang: Ozone precursor relationships in the ambient atmosphere, J.

801   Geophys. Res., 97(D5), 6037–6055, doi:10.1029/91JD03014, 1992.

802   Claeys, M., B. Graham, G. Vas, W. Wang, R. Vermeylen, V. Pashynska, J. Cafmeyer, P.

803   Guyon, M. O. Andreae, P. Artaxo, W. Maenhaut: Formation of Secondary Organic

804   Aerosols Through Photooxidation of Isoprene, Vol. 303 no. 5661 pp. 1173-1176, doi:

805   10.1126/science.1092805, 2004.

806   Cox, P., Delao, A., Komorniczak, A., andWeller, R.: The California almanac of

807   emissions   and   air   quality   2009   edition,

808   http://www.arb.ca.gov/aqd/almanac/almanac09/almanac2009all.pdf, 2009.

809   Dickinson R. E.: Land Surface Processes and Climate Surface Albedos and Energy-

810   Balance. Adv Geophys 25:305-353. doi: 10.1016/S0065-2687(08)60176-4, 1983.

811   Duhl, T. R., Guenther, A., and Helmig, D.: Estimating urban vegetation cover fraction

812   using   Google   Earth®images,   J.   Land   Use   Sci.,   7,   311–329,

813   doi:10.1080/1747423X.2011.587207, 2012.

814   Fast, J. D, Gustafson Jr., W. I., Easter, R. C., Zaveri, R. A., Barnard, J. C., Chapman, E.

815   G., and. Grell, G. A.: Evolution of ozone, particulates, and aerosol direct forcing in an

816   urban area using a new fully-coupled meteorology, chemistry, and aerosol model, J.

817   Geophys. Res., 111, D21305, doi:10.1029/2005JD006721, 2006.

818   Fast, J. D., Allan, J., Bahreini, R., Craven, J., Emmons, L., Ferrare, R., Hayes, P. L.,

819   Hodzic, A., Holloway, J., Hostetler, C., Jimenez, J. L., Jonsson, H., Liu, S., Liu, Y.,

820   Metcalf, A., Middlebrook, A., Nowak, J., Pekour, M., Perring, A., Russell, L.,

821   Sedlacek, A., Seinfeld, J., Setyan, A., Shilling, J., Shrivastava, M., Springston, S.,

822   Song, C., Subramanian, R., Taylor, J. W., Vinoj, V., Yang, Q., Zaveri, R. A., and

823   Zhang, Q.: Modeling regional aerosol and aerosol precursor variability over

824   California and its sensitivity to emissions and long-range transport during the 2010

825   CalNex   and   CARES   campaigns,   Atmos.   Chem.   Phys.,   14,   10013-10060,

826   doi:10.5194/acp-14-10013-2014, 2014.

827   Fehsenfeld, F., Calvert, J., Fall, R., Goldan, P., Guenther, A.B., Hewitt, C.N., Lamb, B.,

828   Liu, S., Trainer, M., Westberg, H. and Zimmerman, P.: Emissions of volatile organic

829   compounds from vegetation and the implications for atmospheric chemistry. Global

830   Biogeochemical Cycles, 6: doi: 10.1029/92GB02125. issn: 0886-6236, 1992.

831   Gao, Y., X. Liu, C. Zhao, and M. Zhang: Emission controls versus meteorological

832   conditions in determining aerosol concentrations in Beijing during the 2008 Olympic

833   Games, Atmos. Chem. Phys., 11, 12437-12451, 2011.

834   Gao, Y., C. Zhao, X. Liu, M. Zhang, and L. R. Leung: Regional modeling of aerosol and

835   its radiative forcing over East Asia using WRF-Chem, Atmos. Environ., 92, 250-266,

836   2014.

837   Gentner, D. R., Isaacman, G., Worton, D. R., Chan, A. W. H., Dallmann, T. R., Davis, L.,

838   Liu, S., Day, D. A., Russell, L. M., Wilson, K. R., Weber, R., Guha, A., Harley, R.

839   A., and Goldstein, A. H.: Elucidating secondary organic aerosol from diesel and

840   gasoline vehicles through detailed characterization of organic carbon emissions, P.

841   Natl. Acad. Sci. USA, 109, 18318–18323, doi:10.1073/pnas.1212272109, 2012.

842   Greenfield, E.J., D.J. Nowak, and J.T. Walton: Assessment of 2001 NLCD percent tree

843   and impervious cover estimates, Photogrammetric Engineering & Remote Sensing ,

844   75(11): 1279–1286, 2009.

845   Grell, G. A., Peckham, S. E., Schmitz, R., and McKeen, S. A., Frost, G., Skamarock, W.

846   C., and Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmos.

847   Environ., 39, 6957–6976, 2005.

848   Guenther, A. B., P. R. Zimmerman, P. C. Harley, R. K. Monson, and R. Fall: Isoprene

849   and monoterpene emission rate variability: model evaluations and sensitivity

850   analyses. Journal of Geophysical Research, vol. 98, no. 7, pp. 12–617, 1993.

851   Guenther, A. B., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., Harley, P.,

852   Klinger, L., Lerdau, M., McKay, W. A., Pierce, T., Scholes, B., Steinbrecher, R.,

853   Tallamraju, R., Taylor, J., and Zimmerman, P.: A global model of natural volatile

854   organic compound emissions, J. Geophys. Res.-Atmos., 100, 8873–8892, 1995.

855   Guenther, A. B., W. Baugh, K. Davis, Gary Hampton, P. Harley, L. Klinger, L. Vierling,

856   P. Zimmerman, E. Allwine, S. Dilts, B. Lamb, H. Westberg, D. Baldocchi, C. Geron,

857   T. Pierce: Isoprene fluxes measured by enclosure, relaxed eddy accumulation, surface

858   layer gradient, mixed layer gradient, and mixed layer mass balance techniques.

859   Journal of Geophysical Research D, vol. 101, no. 13, pp. 18555–18567, 1996.

860 Guenther, A., T. Karl, P. Harley, C. Wiedinmyer, P. I. Palmer, and C. Geron: Estimates

861   of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases

862   and Aerosols from Nature). Atmospheric Chemistry and Physics, vol. 6, no. 11, pp.

863   3181–3210, 2006.

864 Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L.

865   K., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature

866   version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic

867   emissions, Geosci. Model Dev., 5, 1471-1492, doi:10.5194/gmd-5-1471-2012, 2012.

868 Guenther, A. B.: Biological and chemical diversity of biogenic volatile organic emissions

869   into the atmosphere, ISRN Atmospheric Sciences, 2013, 786290, 1-27, 2013.

870 Hong, Song-You, Yign Noh, Jimy Dudhia: A new vertical diffusion package with an

871   explicit treatment of entrainment processes. Mon. Wea. Rev., 134, 2318–2341, 2006.

872 Huang, M., Carmichael, G. R., Spak, S. N., Adhikary, B., Kulkarni, S., Cheng, Y., Wei,

873   C., Tang, Y., D'Allura, A., Wennberg, P. O., Huey, G. L., Dibb, J. E., Jimenez, J. L.,

874   Cubison, M. J., Weinheimer, A. J., Kaduwela, A., Cai, C., Wong, M., Bradley Pierce,

875   R., Al-Saadi, J. A., Streets, D. G., and Zhang, Q.: Multi-scale modeling study of the

876   source contributions to near-surface ozone and sulfur oxides levels over California

877   during the ARCTAS-CARB period, Atmos. Chem. Phys., 11, 3173-3194,

878   doi:10.5194/acp-11-3173-2011, 2011.

879 Hijmas, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis: Very high resolution

880   interpolated climate surfaces for global land areas. Int. J. Climatol. 25: 1965–1978

881   (2005), DOI: 10.1002/joc.1276, 2005.

882 Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D.

883   Collins: Radiative forcing by long–lived greenhouse gases: Calculations with the

884   AER radiative transfer models. J. Geophys. Res., 113, D13103, 2008.

885 Jin, J., and N. L. Miller (2007), Analysis of the impact of snow on dailyweather

886   variability in mountainous regions using MM5, J. Hydrometeorol.,8, 245 –258,

887   doi:10.1175/JHM565.1.

888 Jin, J., and L. Wen: Evaluation of snowmelt simulation in the Weather Research and

889   Forecasting model, J. Geophys. Res., 117, D10110, doi:10.1029/2011JD016980,

890   2012.

891 Kain, John S.: The Kain–Fritsch convective parameterization: An update. J. Appl.

892   Meteor., 43, 170–181, 2004.

893 Ke Y, LYR Leung, M Huang, AM Coleman, H Li, and MS Wigmosta: Development of

894   High Resolution Land Surface Parameters for the Community Land Model.

895   Geoscientific Model Development 5(6):1341-1362. doi:10.5194/gmd-5-1341-2012,

896   2012.

897 Knote, C., Hodzic, A., Jimenez, J. L., Volkamer, R., Orlando, J. J., Baidar, S., Brioude,

898   J., Fast, J., Gentner, D. R., Goldstein, A. H., Hayes, P. L., Knighton, W. B., Oetjen,

899   H., Setyan, A., Stark, H., Thalman, R., Tyndall, G., Washenfelder, R., Waxman, E.,

900   and Zhang, Q.: Simulation of semi-explicit mechanisms of SOA formation from

901   glyoxal in aerosol in a 3-D model, Atmos. Chem. Phys., 14, 6213-6239,

902   doi:10.5194/acp-14-6213-2014, 2014.

903   Lamb, B., A. Guenther, D. Gay, H. Westberg: A national inventory of biogenic

904   hydrocarbon emissions. Atmos. Environ., 21, 8, 1695-1705, doi:10.1016/0004-

905   6981(87)90108-9, 1987.

906   Lawrence, P.J. and T.N. Chase: Climate Impacts. Our earth's changing land: an

907   encyclopedia of land-use and land-cover change, H. Geist, Greenwood Press,

908   Westport, 115-124, 2006.

909   Lawrence, P. J. and T. N. Chase: Representing a new MODIS consistent land surface in

910   the Community Land Model (CLM 3.0). J. Geophys. Res.-Biogeosci.: Vol. 112,

911   2007.

912   Lawrence, D.M., K.W. Oleson, M.G. Flanner, P.E. Thornton, S.C. Swenson, P.J.

913   Lawrence, X. Zeng, Z.-L. Yang, S. Levis, K. Sakaguchi, G.B. Bonan, and A.G.

914   Slater: Parameterization improvements and functional and structural advances in

915   version 4 of the Community Land Model. J. Adv. Model. Earth Sys., 3, DOI:

916   10.1029/2011MS000045, 2011.

917   Lemone, M. A., F. Chen, M. Tewari, J. Dudhia, B. Geerts, Q. Miao, R. L. Coulter, and R.

918   L. Grossman (2010a), Simulating the IHOP_2002 Fair-Weather CBL with the WRF-

919   ARW-Noah modeling system. Part I: Surface fluxes and CBL structure and evolution

920   along the eastern track, Mon. Weather Rev., 138(3), 722–744,

921   doi:10.1175/2009mwr3003.1.

922   Lemone, M. A., F. Chen, M. Tewari, J. Dudhia, B. Geerts, Q. Mia, R. L. Coulter, and R.

923   L. Grossman (2010b), Simulating the IHOP_2002 Fair-Weather CBL with the WRF-

924   ARW Noah modeling system. Part II: Structures from a few kilometers to 100 km

925   across, Mon. Weather Rev., 138(3), 745–764, doi:10.1175/2009mwr3004.1.

926   Lindinger, W., Hansel, A., and Jordan, A.: On-line monitoring of volatile organic

927   compounds at pptv levels by means of protontransfer-reaction mass spectrometry

928   (PTR-MS) – medical applications, food control and environmental research, Int. J.

929   Mass Spectrom., 173, 191–241, 1998.

930   McKeen, S. A., G. Wotawa, D. D. Parrish, J. S. Holloway, M. P. Buhr, G. Hubler, F. C.

931   Fehsenfeld, and J. F. Meagher: Ozone production from Canadian wildfires during

932   June and July of 1995, J. Geophys. Res., 107(D14), 4192,

933   doi:10.1029/2001JD000697, 2002.

934   Morrison, H., G. Thompson, V. Tatarskii: Impact of Cloud Microphysics on the

935   Development of Trailing Stratiform Precipitation in a Simulated Squall Line:

936   Comparison of One– and Two–Moment Schemes. Mon. Wea. Rev., 137, 991–1007,

937   2009.

938   Müller J.-F.: Geographical distribution and seasonal variation of surface emissions and

939   deposition velocities of atmospheric trace gases. J. Geophys. Res. 97(D4): 3787-

940   3804, 1992.

941   Niinemets, Ü., J. D. Tenhunen, P. C. Harley, and R. Steinbrecher: A model of isoprene

942   emission based on energetic requirements for isoprene synthesis and leaf

943   photosynthetic properties for Liquidambar and Quercus, Plant Cell Environ., 22,

944   1319–1335, 1999.

945   Niinemets, Ü., G. Seufert, R. Steinbrecher, and J. D. Tenhunen: A model coupling foliar

946   monoterpene emissions to leaf photosynthetic characteristics in Mediterranean

947   evergreen Quercus species, New Phytol., 153, 257–275, 2002.

948   Oleson, K. W., Lawrence, D. M., Bonan, G. B., Flanner, M. G., Kluzek, E., Lawrence, P.

949   J., Levis, S., Swenson, S. C., Thornton, P. E., Dai, A., Decker, M., Dickinson, R.,

950   Feddema, J., Heald, C. L., Hoffman, F., Lamarque, J.-F., Mahowald, N., Niu, G.-Y.,

951   Qian, T., Randerson, J., Running, S., Sakaguchi, K., Slater, A., Sto ̈ckli, R., Wang,

952   A., Yang, Z.-L., Zeng, X., and Zeng, X.: Tech- nical Description of version 4.0 of the

953   Community Land Model (CLM), Tech. Rep. NCAR/TN-478 + STR, National Center

954   for Atmospheric Research, 2010.

955   Oliver Wild, Xin Zhu, Michael J. Prather: Fast-J: Accurate Simulation of In- and Below-

956   Cloud Photolysis in Tropospheric Chemical Models, J. Atmos. Chem., Volume 37,

957   Issue 3, pp 245-282, 2000.

958   Paulson, C. A.: The mathematical representation of wind speed and temperature profiles

959   in the unstable atmospheric surface layer. J. Appl. Meteor., 9, 857–861, 1970.

960   Pfister, G. G., Parrish, D. D.,Worden, H., Emmons, L. K., Edwards, D. P., Wiedinmyer,

961   C., Diskin, G. S., Huey, G., Oltmans, S. J., Thouret, V., Weinheimer, A., and

962   Wisthaler, A.: Characterizing summertime chemical boundary conditions for air

963   masses entering the US West Coast, Atmos. Chem. Phys., 11, 1769–1790,

964   doi:10.5194/acp-11-1769-2011, 2011.

965   Pierce, T., C. Geron, L. Bender, R. Dennis, G. Tonnesen, and A. Guenther: Influence of

966   increased isoprene emissions on regional ozone modeling, J. Geophys. Res.,

967   103(D19), 25611–25629, doi:10.1029/98JD01804, 1998.

968   Poisson, N., M. Kanakidou, P. J. Crutzen: Impact of Non-Methane Hydrocarbons on

969   Tropospheric Chemistry and the Oxidizing Power of the Global Troposphere: 3-

970   Dimensional Modelling Results, J. Atmos. Chem., 36, 157-230, 2000.

971   Rasmussen, R. A.: What do the hydrocarbons from trees contribute to air pollution? J. Air

972   Poll. Con. Asso., 22, 7, 537-543, 1972.

973   Ryerson, T. B., Andrews, A. E., Angevine, W. M., Bates, T. S., Brock, C. A., Cairns, B.,

974   Cohen, R. C., Cooper, O. R., de Gouw, J. A., Fehsenfeld, R. C., Ferrare, R. A.,

975   Fischer, M. L., Flagan, R. C., Goldstein, A. H., Hair, J. W., Hardesty, R. M.,

976   Hostetler, C. A., Jimenez, J. L., Langford, A. O., McCauley, E., McKeen, S. A.,

977   Molina, L. T., Nenes, A., Oltmans, S. J., Parrish, D. D., Pederson, J. R., Pierce, R. B.,

978   Prather, K., Quinn, P. K., Seinfeld, J. H., Senff, C. J., Sorooshian, A., Stutz, J.,

979   Surratt, J. D., Trainer, M., Volkamer, R., Williams, E. J., and Wofsy, S. C.: The 2010

980   California Research at the Nexus of Air Quality and Climate Change (CalNex) field

981   study, J. Geophys. Res., 118, 5830-5866, doi:10.1002/jgrd.50331, 2013.

982   Sakulyanontvittaya, T., Duhl, T., Wiedinmyer, C., Helmig, D., Mat- sunaga, S.,

983   Potosnak, M., Milford, J., and Guenther, A.: Monoter- pene and sesquiterpene

984   emission estimates for the United States, Environ. Sci. Technol., 42, 1623–1629,

985   2008.

986   Sanderson, M.G., Jones, C.D., Collins, W.J., Johnson, C.E. and Derwent, R.G.: Effect of

987   Climate Change on Isoprene Emissions and Surface Ozone Levels. Geophysical

988   Research Letters 30: doi: 10.1029/2003GL017642. issn: 0094-8276, 2003.

989   Schurgers, G., A. Arneth, R. Holzinger, and A. Goldstein: Processbased modelling of

990   biogenic monoterpene emissions combining production and release from storage,

991   Atmos. Chem. Phys., 9, 3409–3423, 2009.

992   Shilling, J. E., R. A. Zaveri, J. D. Fast, L. I. Kleinman, M. L. Alexander, M. R.

993   Canagaratna, E. Fortner, J. M. Hubbe, J. T. Jayne, A. Sedlacek, A. Setyan, S.

994   Springston, D. R. Worsnop, and Q. Zhang: Enhanced SOA formation from mixed

995   anthropogenic and biogenic emissions during the CARES campaign. Atmospheric

996   Chemistry and Physics, 13(4):2091-2113, doi:10.5194/acp-13-2091-2013, 2013.

997   Shrivastava MKB, JD Fast, RC Easter, Jr, WI Gustafson, Jr, RA Zaveri, JL Jimenez, P

998   Saide, and A Hodzic: Modeling Organic Aerosols in a Megacity: Comparison of

999   Simple and Complex Representations of the Volatility Basis Set Approach.

1000   Atmospheric Chemistry and Physics 11(13):6639-6662. doi:10.5194/acp-11-6639-

1001   2011.

1002   Shrivastava MKB, A Zelenyuk, D Imre, RC Easter, Jr, J Beranek, RA Zaveri, and JD

1003   Fast: Implications of Low Volatility SOA and Gas-Phase Fragmentation Reactions on

1004   SOA Loadings and their Spatial and Temporal Evolution in the Atmosphere. Journal

1005   of Geophysical Research. D. (Atmospheres), 118(8), 3328-3342,

1006   doi:10.1002/jgrd.50160, 2013.

1007   Stauffer, D. R. and Seaman, N. L.: Use of four-dimensional data assimilation in a limited-

1008   area mesoscale model, Part I: Experiments with synoptic-scale data, Mon. Weather

1009   Rev., 118, 1250–1277, 1990.

1010   Still, C. J., J. A. Berry, G. J. Collatz, and R. S. DeFries, Global distribution of C3 and C4

1011   vegetation: Carbon cycle implications, Global Biogeochem. Cycles, 17(1), 1006,

1012   doi:10.1029/2001GB001807, 2003.

1013   Went, F. W.: blue hazes in the atmosphere, Nature, Vol. 187, 4738, 641-643, 1960.

1014   Wild, O., X. Zhu, and Prather, M. J.: Fast-J: Accurate simulation of in- and below-cloud

1015   photolysis in tropospheric chemical models. Journal of Atmospheric Chemistry,

1016   37(3), 245 - 282. doi: 10.1023/A:1006415919030, 2000.

1017   Willmott, C. J. and K. Matsuura: Terrestrial Air Temperature and Precipitation: Monthly

1018   and   Annual   Time   Series   (1950   -   1999),   2011

1019   (http://climate.geog.udel.edu/~climate/html_pages/README.ghcn_ts2.html).

1020   Zaveri, R. A., Shaw, W. J., Cziczo, D. J., Schmid, B., Ferrare, R. A., Alexander, M. L.,

1021   Alexandrov, M., Alvarez, R. J., Arnott, W. P., Atkinson, D. B., Baidar, S., Banta, R.

1022   M., Barnard, J. C., Beranek, J., Berg, L. K., Brechtel, F., Brewer, W. A., Cahill, J. F.,

1023   Cairns, B., Cappa, C. D., Chand, D., China, S., Comstock, J. M., Dubey, M. K.,

1024   Easter, R. C., Erickson, M. H., Fast, J. D., Floerchinger, C., Flowers, B. A., Fortner,

1025   E., Gaffney, J. S., Gilles, M. K., Gorkowski, K., Gustafson, W. I., Gyawali, M., Hair,

1026   J., Hardesty, R. M., Harworth, J. W., Herndon, S., Hiranuma, N., Hostetler, C.,

1027   Hubbe, J. M., Jayne, J. T., Jeong, H., Jobson, B. T., Kassianov, E. I., Kleinman, L. I.,

1028   Kluzek, C., Knighton, B., Kolesar, K. R., Kuang, C., Kubátová, A., Langford, A. O.,

1029   Laskin, A., Laulainen, N., Marchbanks, R. D., Mazzoleni, C., Mei, F., Moffet, R. C.,

1030   Nelson, D., Obland, M. D., Oetjen, H., Onasch, T. B., Ortega, I., Ottaviani, M.,

1031   Pekour, M., Prather, K. A., Radney, J. G., Rogers, R. R., Sandberg, S. P., Sedlacek,

1032   A., Senff, C. J., Senum, G., Setyan, A., Shilling, J. E., Shrivastava, M., Song, C.,

1033   Springston, S. R., Subramanian, R., Suski, K., Tomlinson, J., Volkamer, R., Wallace,

1034   H. W., Wang, J., Weickmann, A. M., Worsnop, D. R., Yu, X.-Y., Zelenyuk, A., and

1035   Zhang, Q.: Overview of the 2010 Carbonaceous Aerosols and Radiative Effects Study

1036   (CARES), Atmos. Chem. Phys., 12, 7647-7687, doi:10.5194/acp-12-7647-2012,

1037   2012.

1038 Zeng, X., M. Shaikh, Y. Dai, R. E. Dickinson, and R. Myneni (2002), Cou-pling of the

1039   common land model to the NCAR community climate model,J. Clim., 15, 1832–

1040   1854, doi:10.1175/1520-0442(2002)015<1832:COTCLM>2.0.CO;2.

1041 Zhao, C., Hu, Z., Qian, Y., Ruby Leung, L., Huang, J., Huang, M., Jin, J., Flanner, M. G.,

1042   Zhang, R., Wang, H., Yan, H., Lu, Z., and Streets, D. G.: Simulating black carbon

1043   and dust and their radiative forcing in seasonal snow: a case study over North China

1044   with field campaign measurements, Atmos. Chem. Phys., 14, 11475-11491,

1045   doi:10.5194/acp-14-11475-2014, 2014.

1046 Zimmerman, P. R., R. B. Chatfield, J. Fishman, P. J. Crutzen, and P. L. Hanst, Estimates

1047   the production of CO and H2 from the oxidation of hydrocarbon emissions from

1048   vegetation, Geophys. Res. Lett., 5, 679-682, 1978.

1049 Zimmerman, P.: Testing of hydrocarbon emissions from vegetation, leaf litter and aquatic

1050   surfaces and development of a method for compiling biogenic emission inventories,

1051   Tech. Rep. EPA-450-4-70-004, U.S. Environmental Protection Agency, Research

1052   Triangle Park, California, USA, 1979.

1053

1054

1055
1056
1057
1058
1059   **Table 1** Average percentage of PFTs over the simulation domain

PFT # and description		USGS	VEG1	VEG2	VEG3
0	Bare soil	26.0	7.6	38.1	41.6
1	Needleleaf evergreen tree – temperate	13.0	12.5	9.1	10.7
2	Needleleaf evergreen tree - boreal	0.0	0.1	0.0	4.9
3	Needleleaf deciduous tree – boreal	0.1	0.0	0.0	0.0
4	Broadleaf evergreen tree – tropical	0.0	0.0	0.0	0.0
5	Broadleaf evergreen tree – temperate	0.0	0.4	1.9	0.0
6	Broadleaf deciduous tree – tropical	2.9	0.0	0.0	0.0
7	Broadleaf deciduous tree – temperate	1.5	0.4	1.8	1.5
8	Broadleaf deciduous tree – boreal	0.0	0.0	0.0	0.3
9	Broadleaf evergreen shrub - temperate	21.1	5.3	0.0	0.3
10	Broadleaf deciduous shrub – temperate	20.0	37.5	27.4	10.8
11	Broadleaf deciduous shrub – boreal	0.9	0.2	0.0	1.0
12	C_3 arctic grass	0.0	0.0	1.2	2.2
13	C_3 grass	1.0	28.0	14.9	18.9
14	C_4 grass	10.4	0.0	0.0	0.0
15	Crop	3.2	6.5	4.1	6.3

1060
1061   [1]USGS is the 16-PFT dataset converted from the default 24 USGS land cover dataset based on a lookup
1062   table derived from Bonan et al. [1996];
1063   [2]VEG1 is from the PFT fractional cover product by Ke et al. [2012];
1064   [3]VEG2 is obtained from the NCAR CESM data repository [Oleson et al., 2010];
1065   [4]VEG3 is derived from a dataset over the U.S. with 16 PFT classifications by combining the National Land
1066   Cover Dataset (NLCD, Homer et al., 2004) and the Cropland Data Layer (see
1067   http://nassgeodata.gmu.edu/CropScape/).
1068
1069
1070
1071   **Table 2** Experiments of WRF-Chem

	Surface scheme	BVOC scheme	Plant Function Type Dataset				
			USGS/VEG-M	USGS	VEG1	VEG2	VEG3
WRF-Chem	CLM4.0	MEGANv2.0 MEGANv2.1	Mv20CLM -	- Mv21USGS	- Mv21V1	- Mv21V2	- Mv21V3
	Noah	MEGANv2.0	Mv20Noah	-	-	-	-

1072
1073
1074
1075

1076

1077

1078

1079

1080

1081

1082

1083

1084

1085
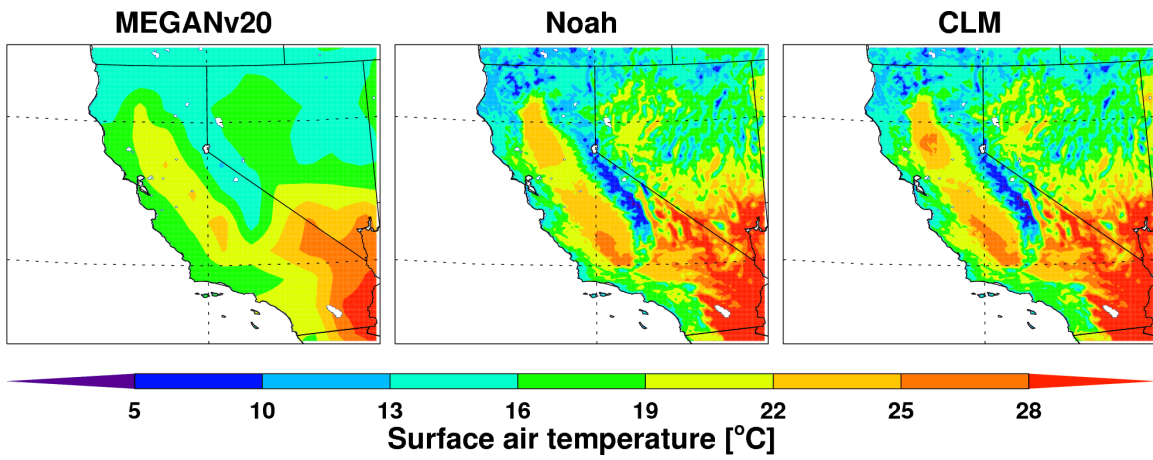
Figure 1 Spatial distributions of monthly mean surface air temperature in June 2010 from the MEGAN v2.0 climatology dataset (MEANv20, prescribed) and the WRF-Chem simulations with the Noah (Noah, simulated) and CLM4 (CLM, simulated) land surface schemes.

1087
1088
1089
1090
1091

1092

1093

1094

1095

1096

1097

1098

1099

1100

1101

1102
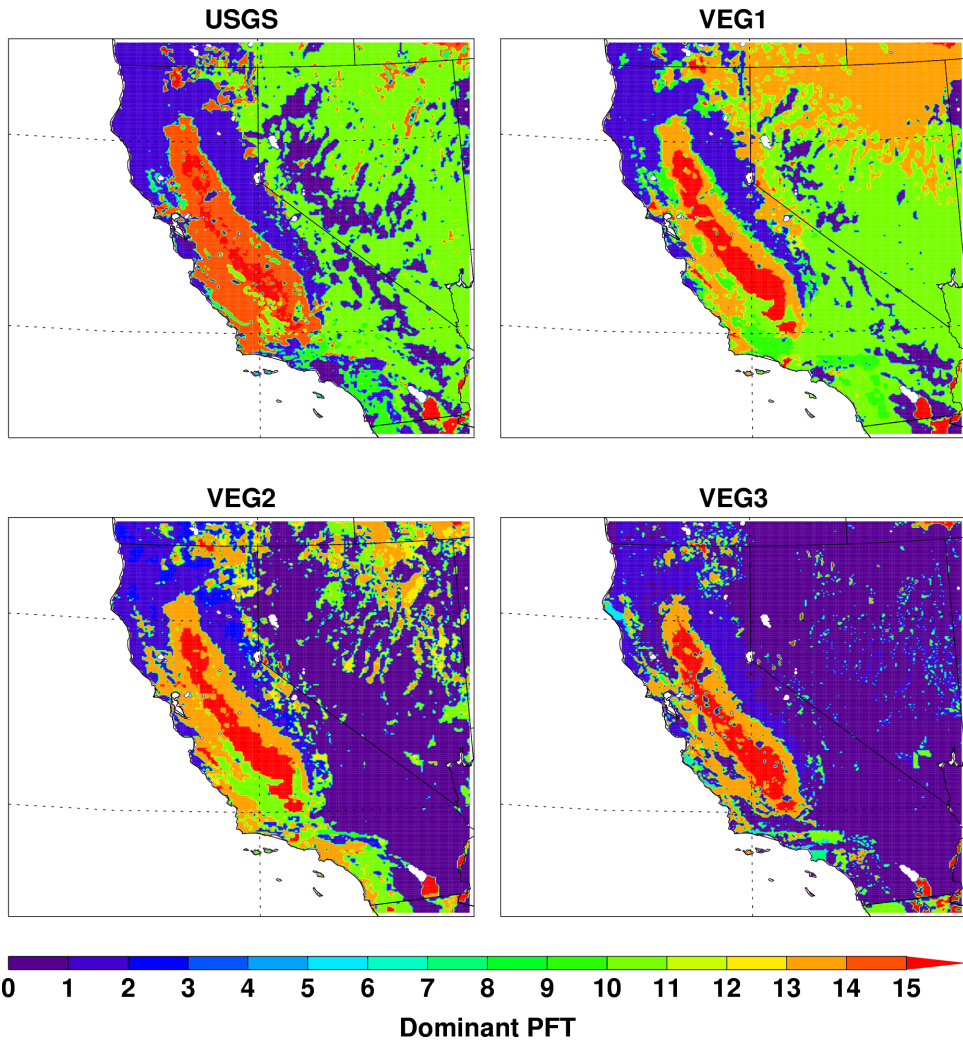
Figure 2. Spatial distribution of dominant PFTs over the simulation domain from the four datasets: USGS, VEG1, VEG2, and VEG3. The PFT number is referred to the list in Table 1.

1103

1104
1105
1106

1107

1108

1109

1110

1111

1112

1113

Figure 3. Spatial distribution of percentage of the four PFTs from the VEG-M used by MEGAN v2.0 over the simulation domain.
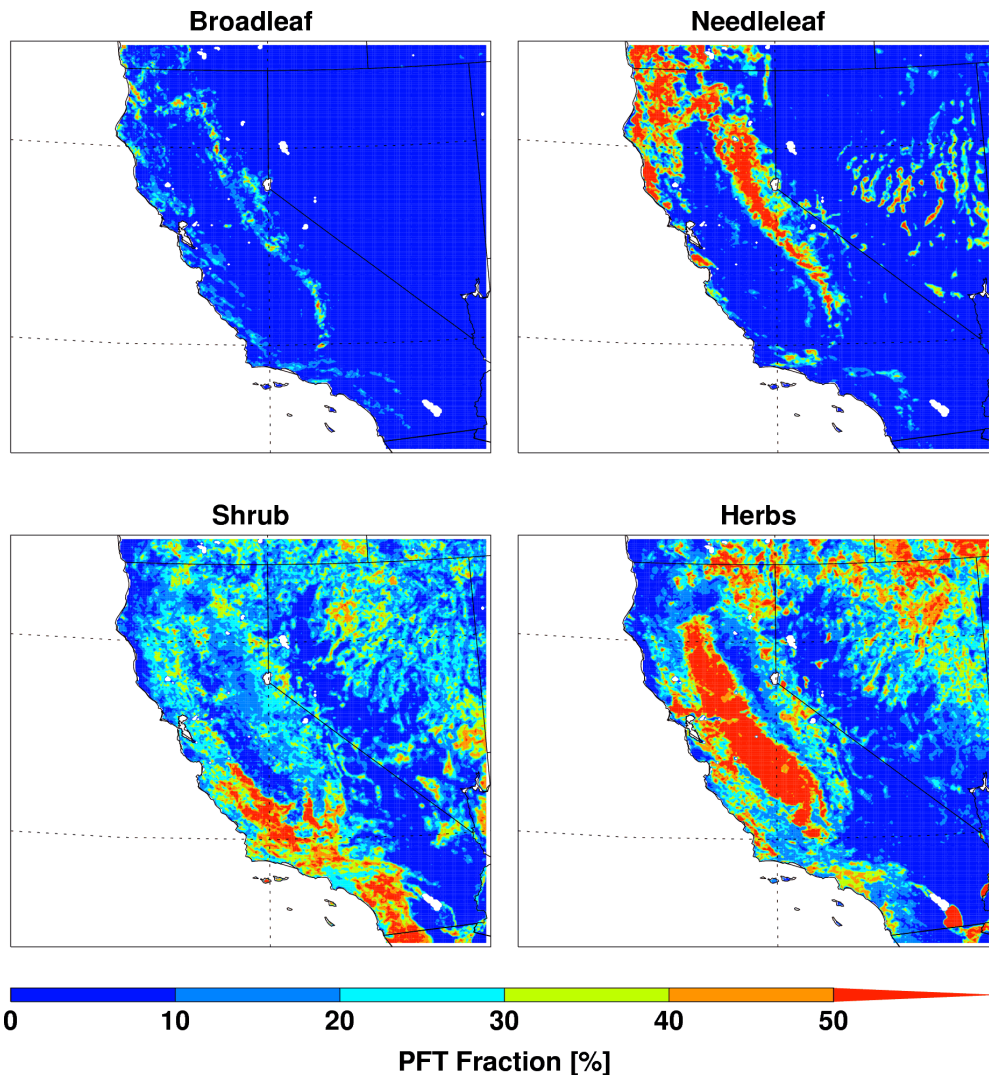
1114
1115
1116
1117

1118

1119

1120

1121

51

1122

1123

1124

1125
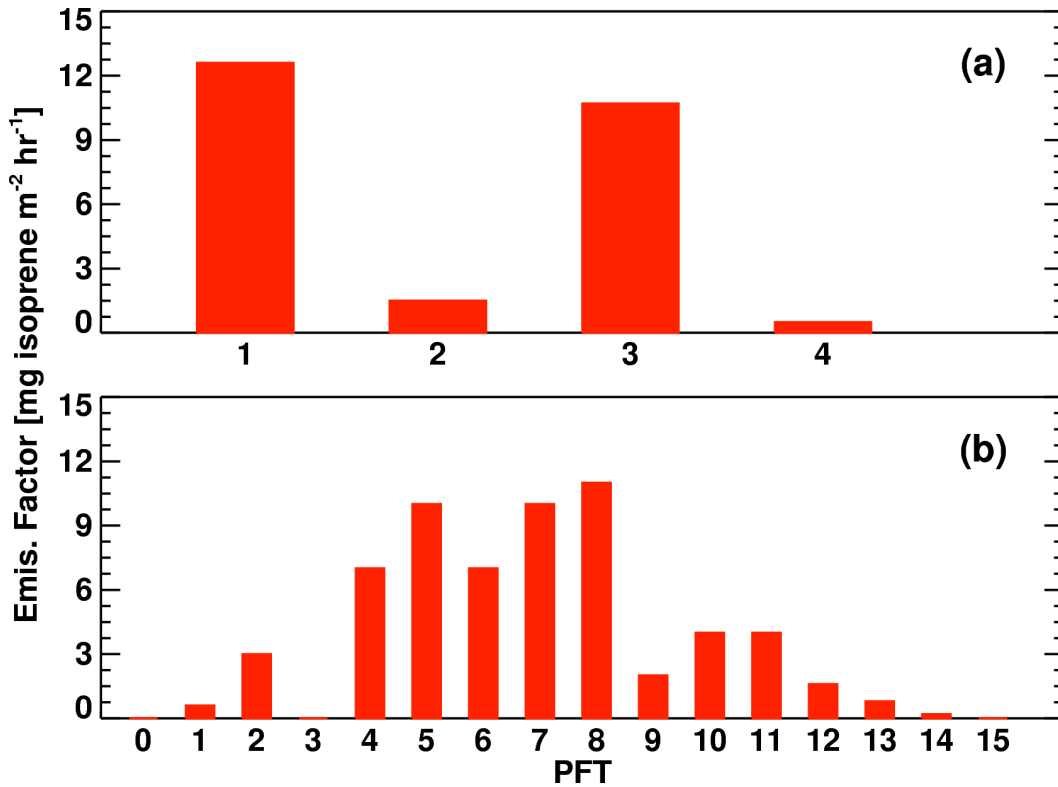
Figure 4. Biogenic isoprene emission factor for each PFT in (a) MEGAN v2.0, the PFT number 1-4 is referred to Broadleaf, Needleleaf, Shrub, and Herbs, respectively; (2) MEGAN v2.1, the PFT number 0-15 is referred to the list in Table 1.

1130

1131

1132

1133

1134

1135

1136

1137

1138
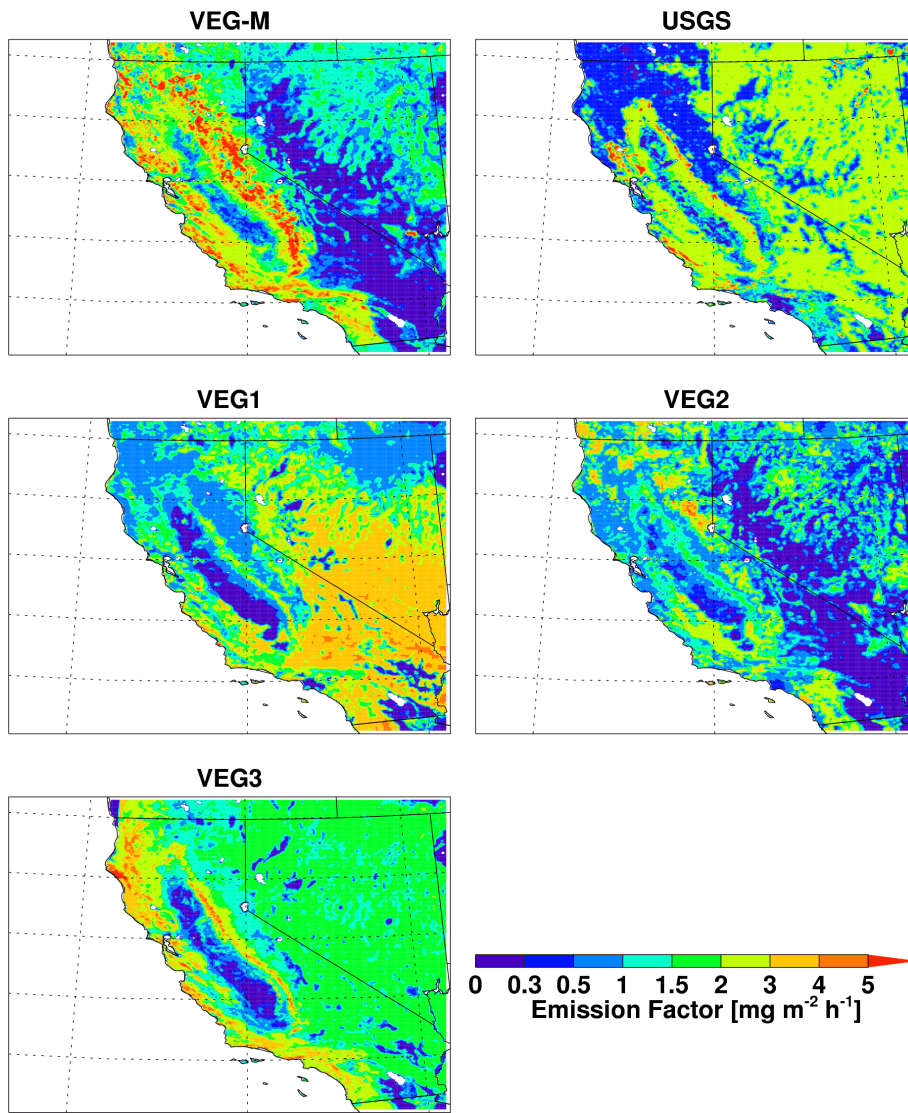
Figure 5. Spatial distribution of PFT-weighted mean biogenic isoprene emission factor derived with the VEG-M in MEGAN v2.0 and the USGS, VEG1, VEG2, and VEG3 in MEGAN v2.1.

1140
1141
1142
1143

1144

1145

1146

1147

1148

1149
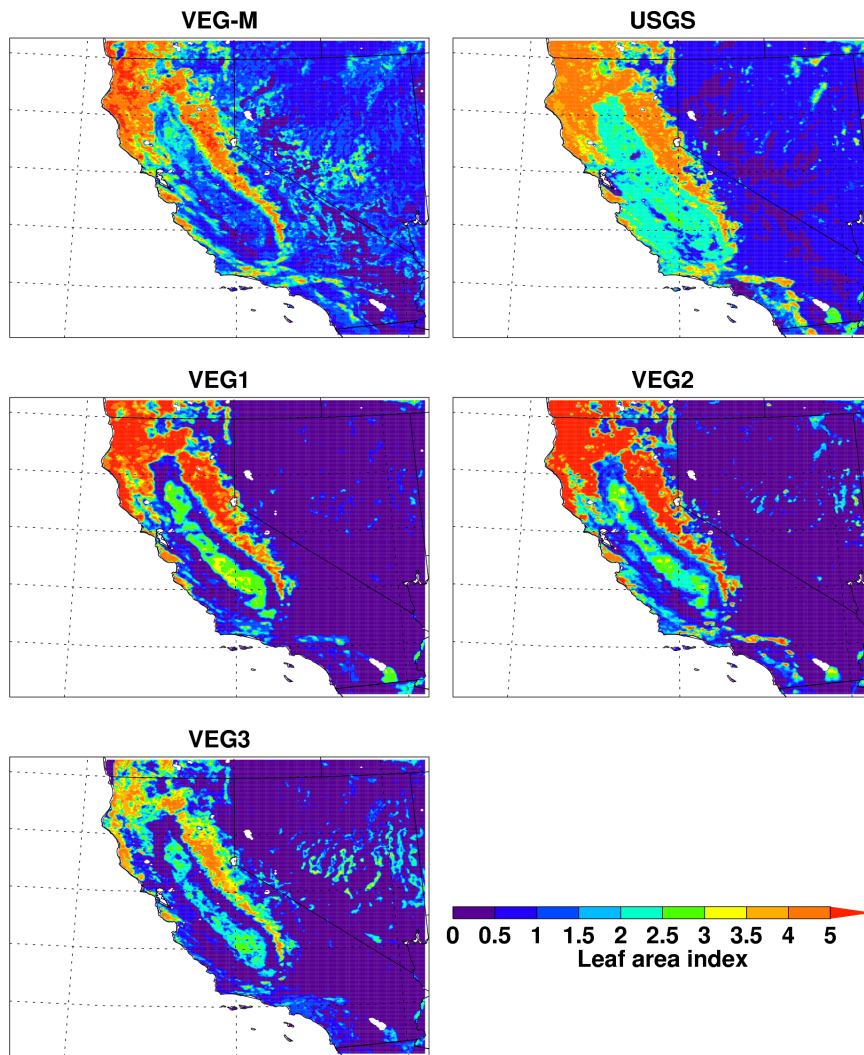
Figure 6. Spatial distribution of leaf area index (LAI) from the VEG-M in MEGAN v2.0
and from the USGS, VEG1, VEG2, and VEG3 in MEGAN v2.1.

1150

1151

1152

1153

1154

1155

1156

1157

1158

1159

1160

1161
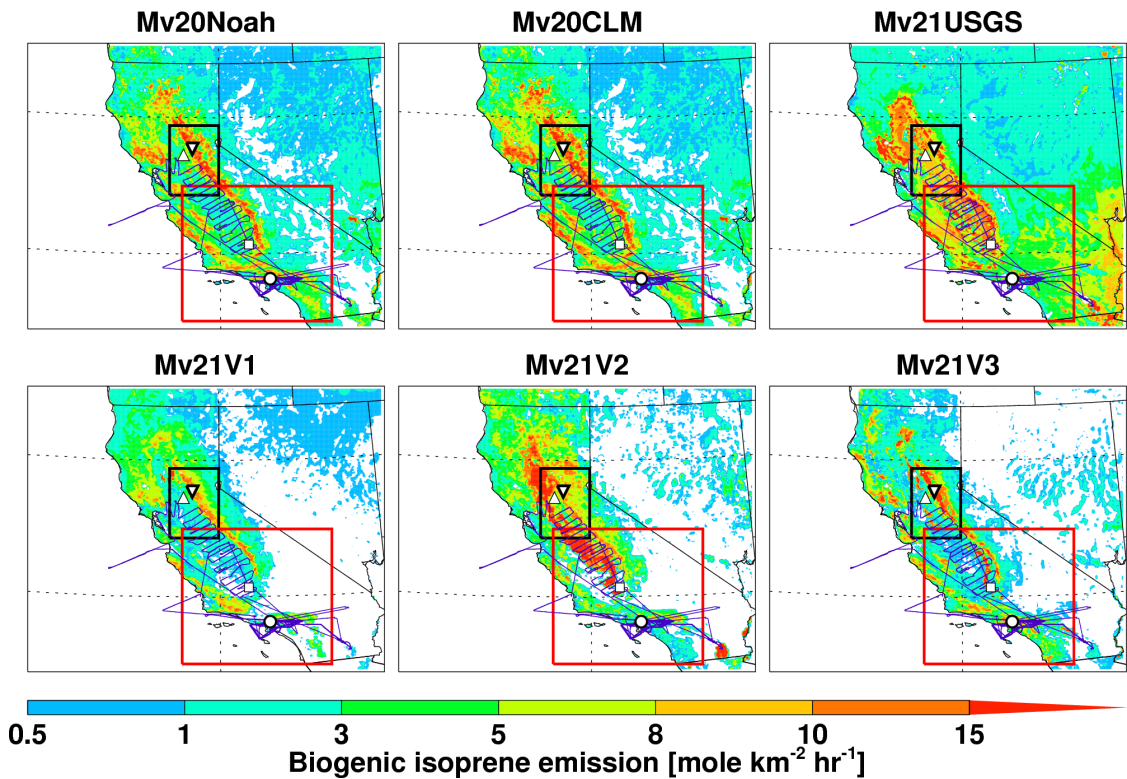
Figure 7. Spatial distributions of biogenic isoprene emissions averaged in June estimated in the six simulations as listed in Table 2. The four observation sites are shown as T0 (white upward triangle), T1 (white downward triangle), Bakersfield (white square), and Pasadena (white circle). The CalNex WP-3D flight tracks below 1 km (blue line) during June 2010 are also shown. The black and red boxes denote the predominant CARES and CalNex regions, respectively.

1163
1164
1165
1166
1167
1168
1169

1170

1171

1172

1173

1174

1175

55

1176

1177

(a) Biogenic isoprene emission [mole km^{-2} hr^{-1}]

(b) Biogenic isoprene emission [mole km^{-2} hr^{-1}]
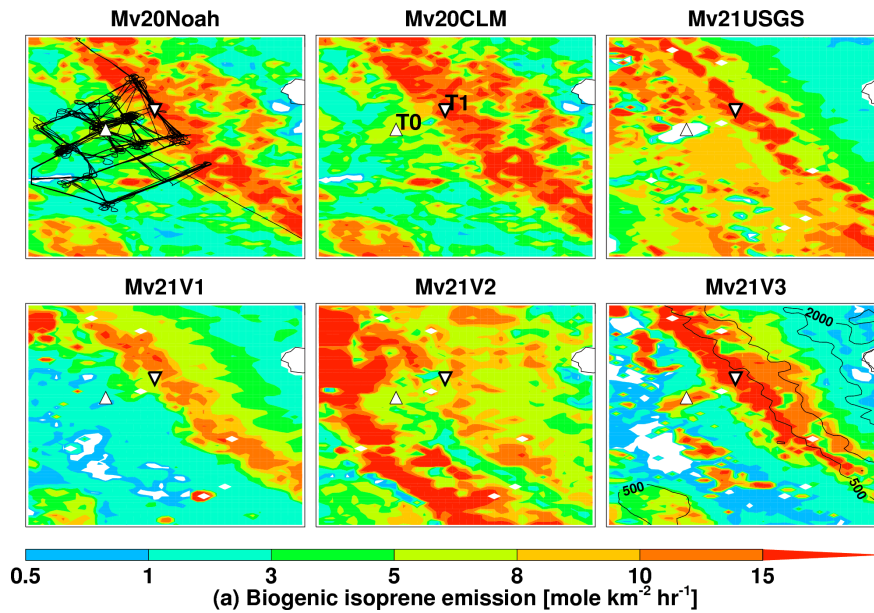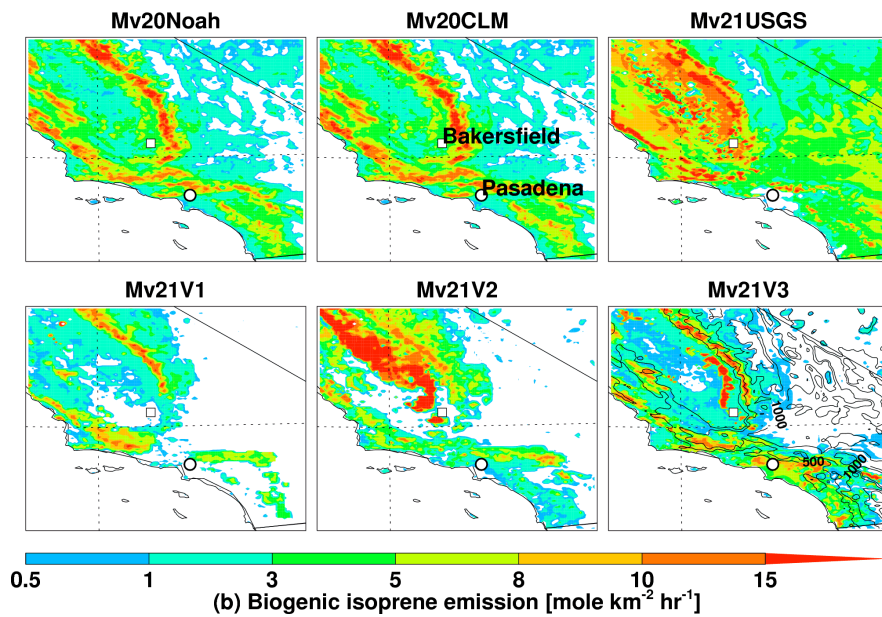
1178

1179 **Figure 8.** a) Spatial distributions of biogenic isoprene emissions around the CARES
1180 observational sites T0 and T1 (the black box shown in Fig. 7) estimated in the six
1181 simulations as listed in Table 1. The CARES G-1 flight tracks below 1 km (black line)
1182 during June 2010 are also shown with the Mv20Noah result; the terrain height is also
1183 shown as the black contour lines with the Mv21V3 result. b) Same as a) except around
1184 the CalNex observational sites Bakersfield and Pasadena (the red box shown in Fig. 7).

1185

1186

56

1187

(a) Biogenic monoterpene emission [mole km^{-2} hr^{-1}]

1188

(b) Biogenic monoterpene emission [mole km^{-2} hr^{-1}]
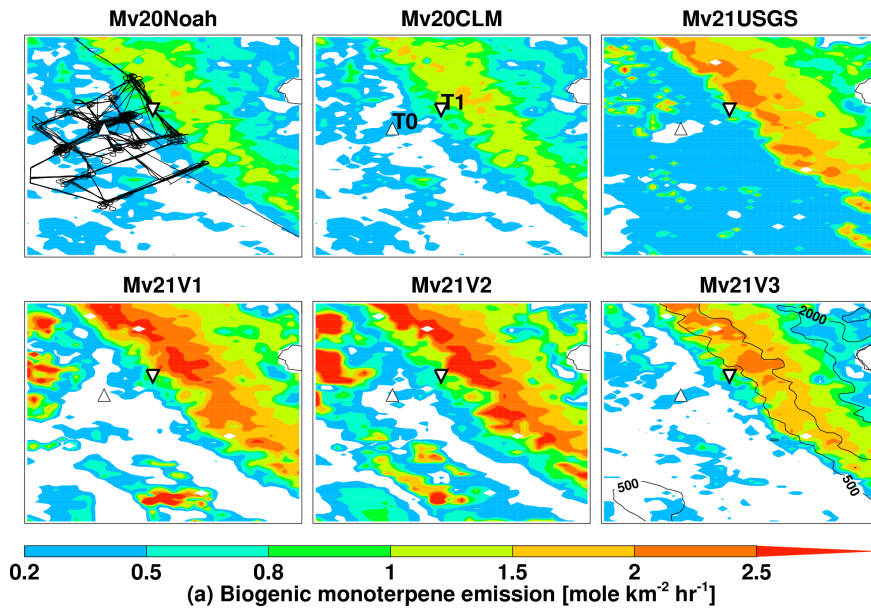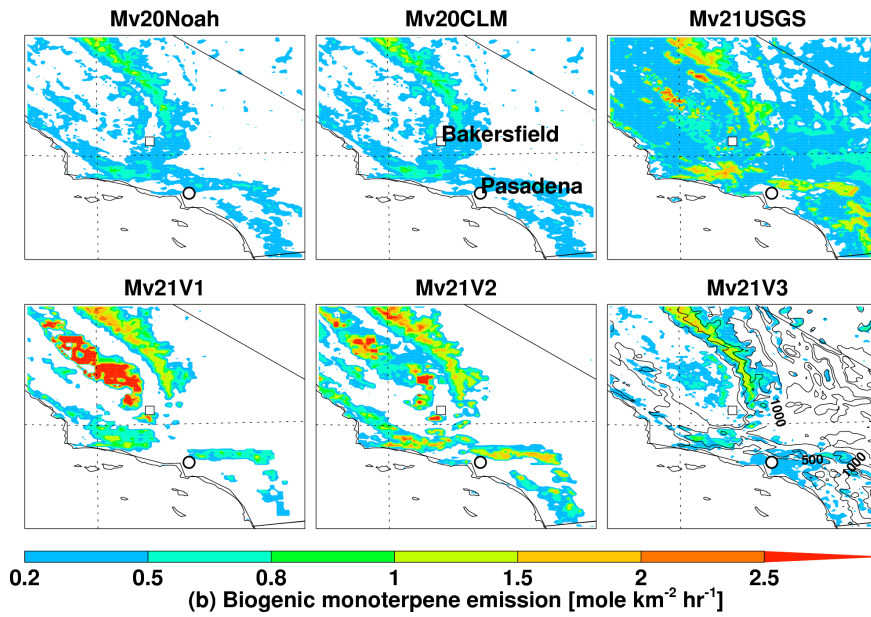
1189

1190   **Figure 9.** Same as Fig. 8, except for biogenic monoterpene emissions.

1191

1192

1193

1194

1195

1196
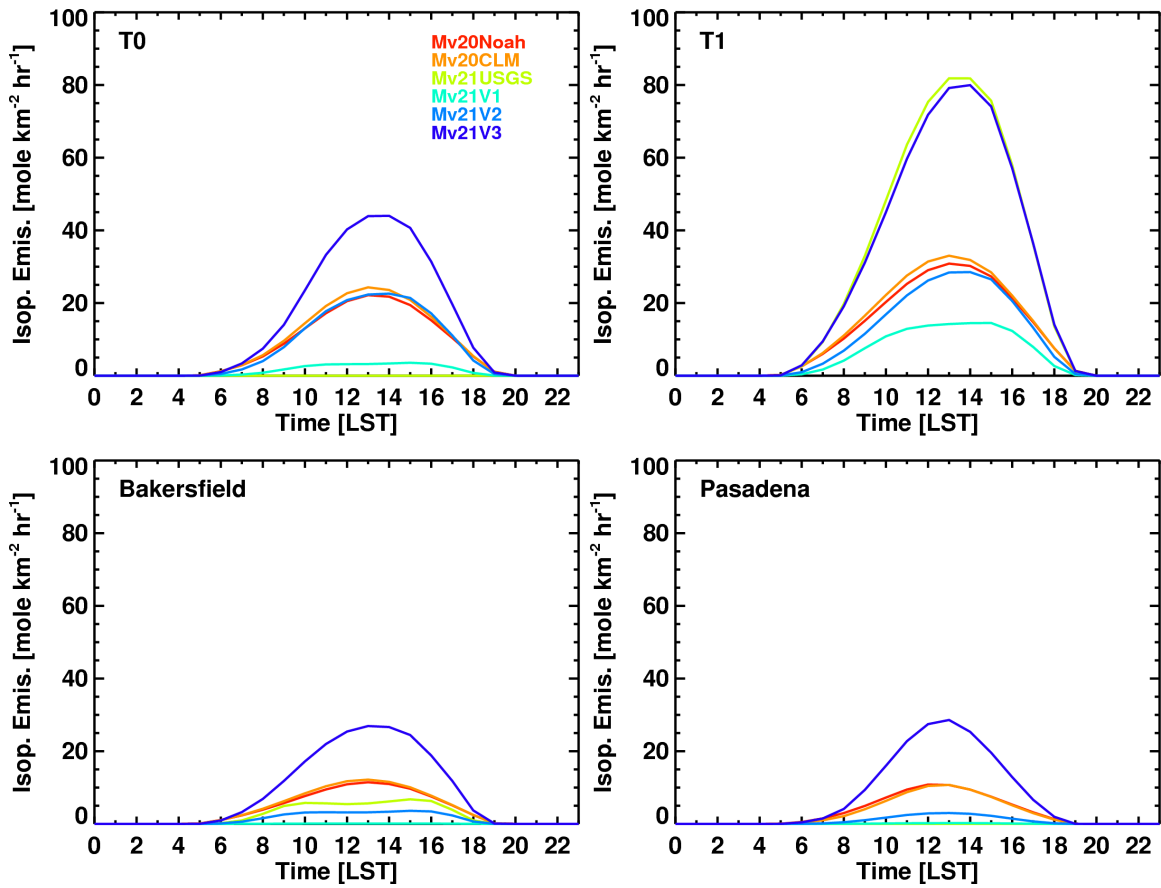
1197

Figure 10. Average diurnal variation of biogenic isoprene emissions at the four

observation sites from the six simulations listed in Table 1.

1200

1201

1202

1203

1204

1205

1206

1207

1208

1209
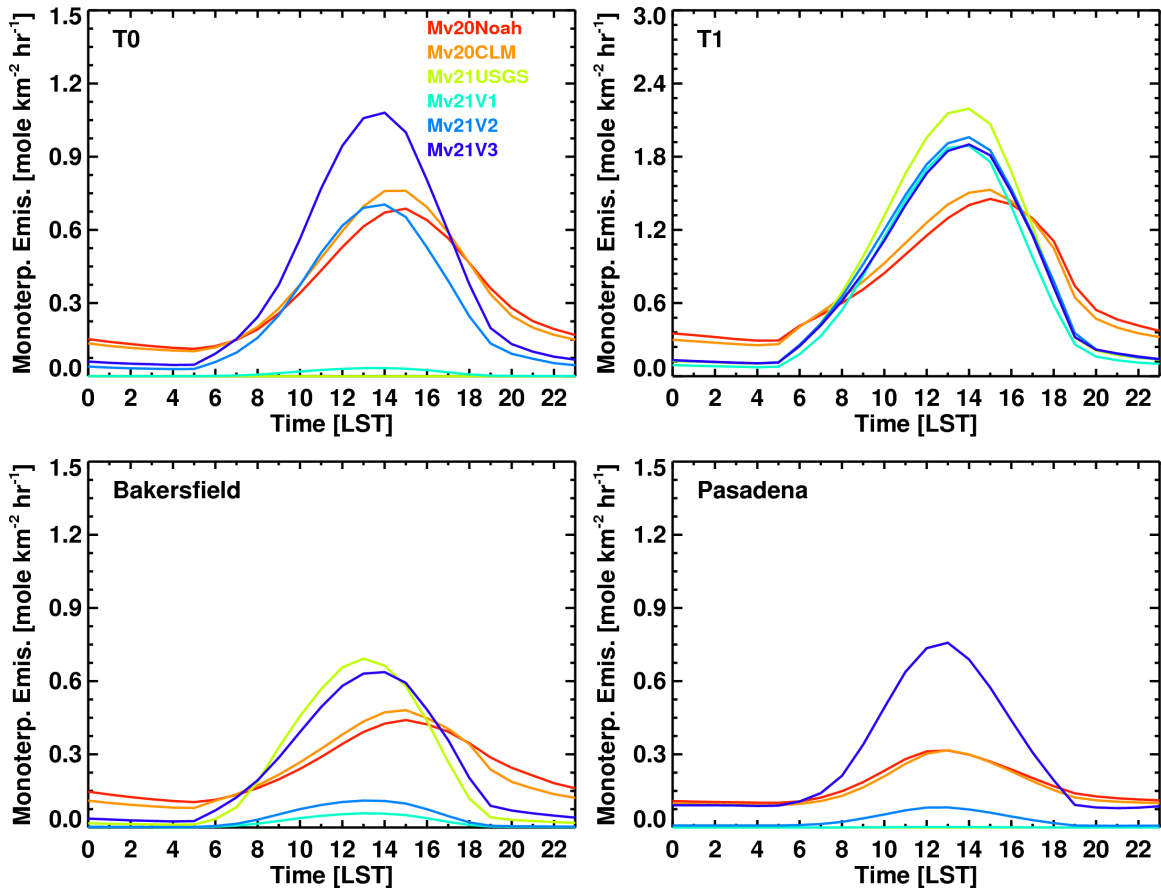
Figure 11. Same as Fig. 10, except for biogenic monoterpene emissions.

1211

1212

1213

1214

1215

1216

1217

1218

1219

1220

1221

a) Isoprene+MVK+MACR [ppbV]   b) Isoprene+MVK+MACR [ppbV]
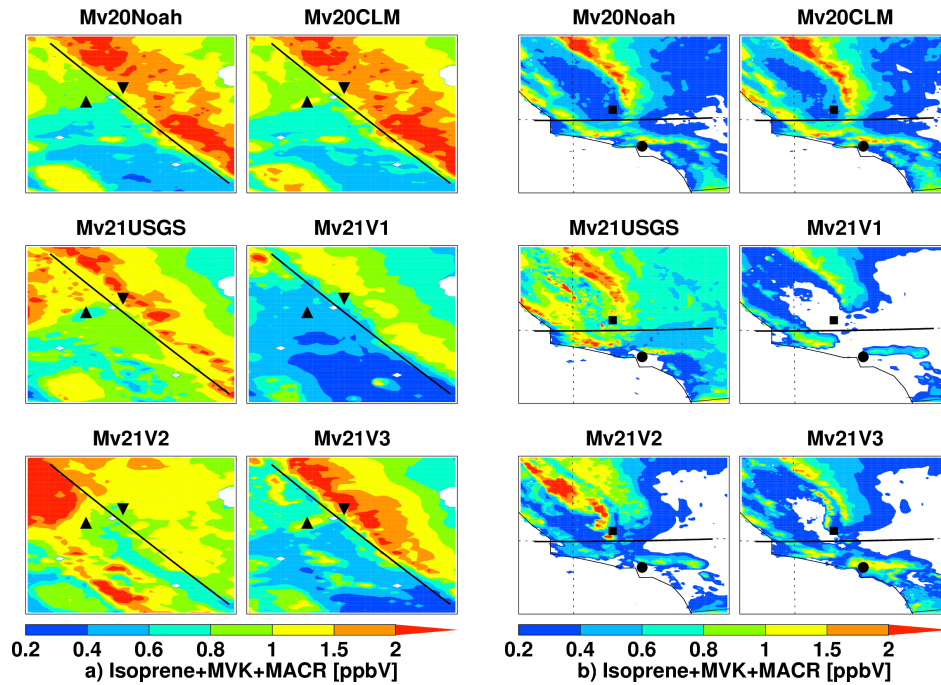
1222

Figure 12. a) Spatial distributions of monthly averaged surface isoprene mixing ratios around the CARES T0 and T1 observational sites from the six simulations as listed in Table 1. The black lines parallel to the Sierra Nevada divide the region to the Southwest and the Northeast for comparison with CARES G-1 aircraft measurements shown in Fig. 16 and 17. b) Same as a) except around the CalNex observational sites Bakersfield and Pasadena. The black lines divide the region to southern California and the Central Valley for comparison with CalNex WP-3D aircraft measurements shown in Fig. 16 and 17.

1230

1231

1232

1233

1234

1235

1236

1237

1238

1239
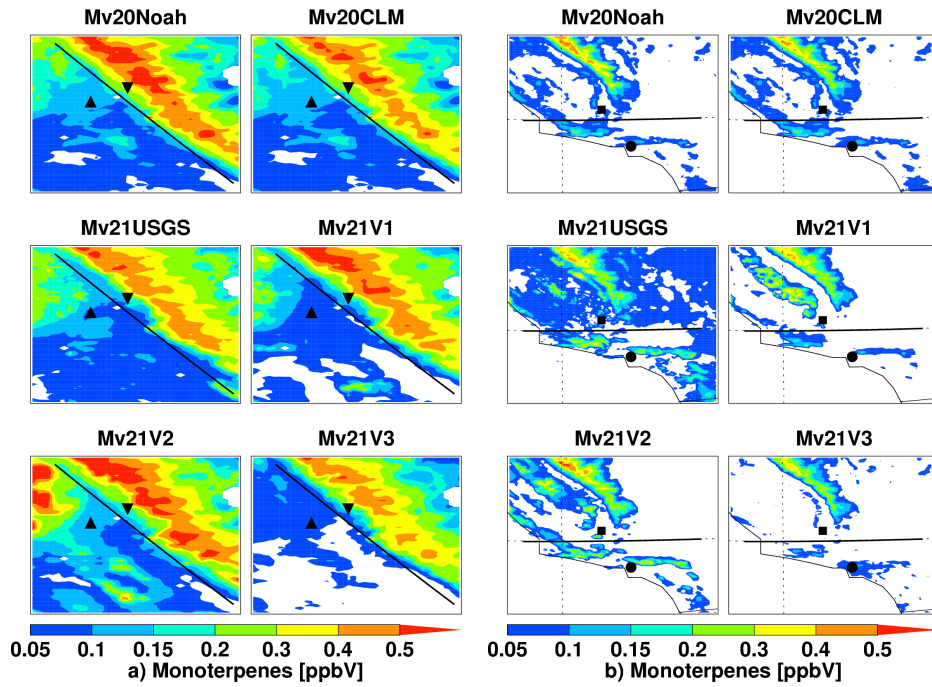
Figure 13. Same as Fig. 12, except for monoterpene.

1242

1243

1244

1245

1246

1247

1248

1249

1250

1251

1252

1253

1254
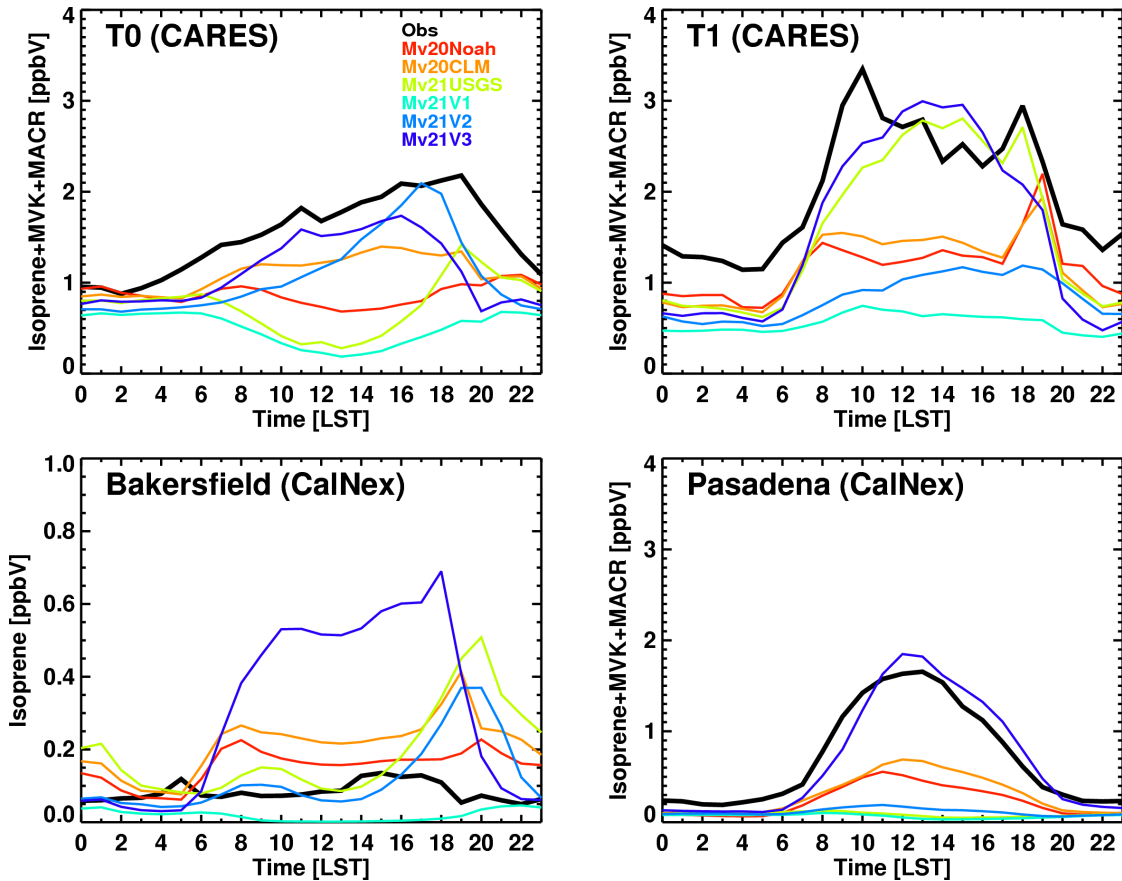
Figure 14. Monthly averaged diurnal variation of surface isoprene+MVK+MACR

mixing ratios at the three observation sites and isoprene mixing ratios at the Bakersfield

site from the observations and six simulations listed in Table 2. The simulated values for

the Bakersfield and Pasadena sites are averaged for the first two weeks of June to be

consistent with the observations.

1261

1262

1263

1264

1265

1266

1267

1268
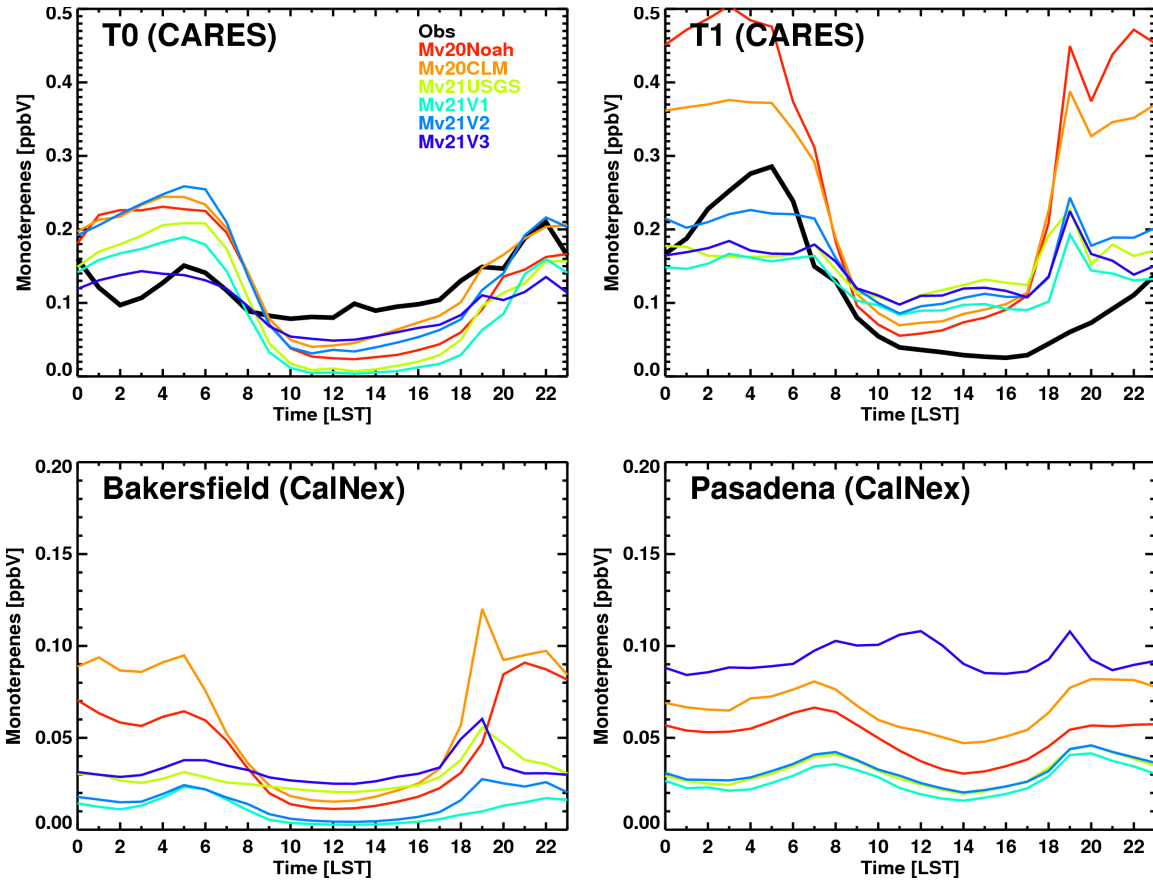
1269

Figure 15. Monthly averaged diurnal variation of surface monoterpene mixing ratios at

the four observation sites from the observations and six simulations as listed in Table 2.

There are no observations for the Bakersfield and Pasadena sites in June.
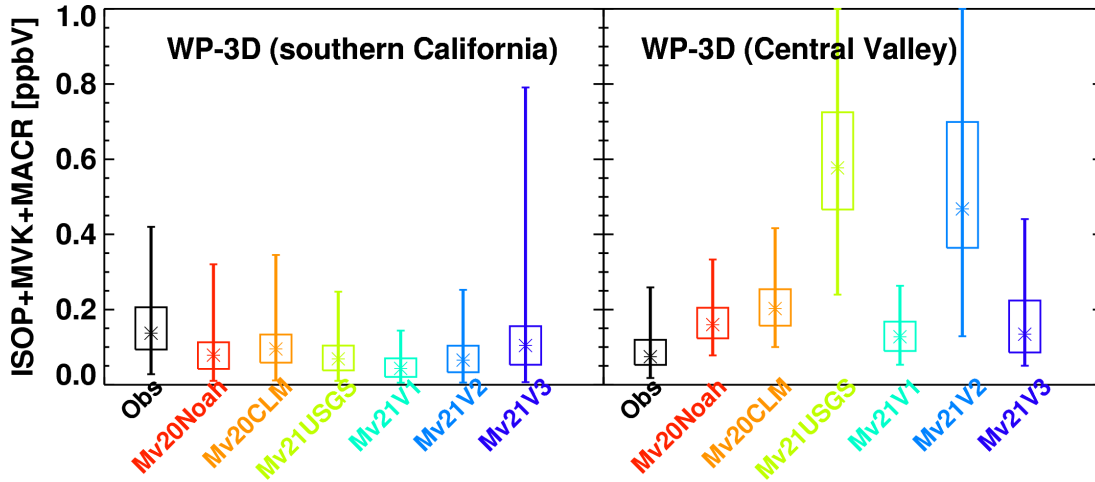
1273

1274

1275

1276

1277

1278

1279
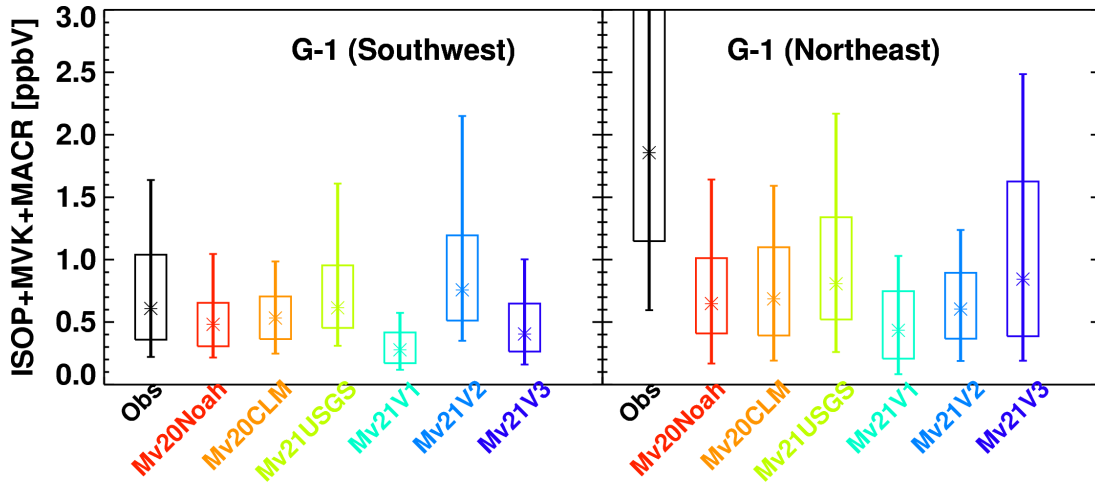
1280

1281

Figure 16. Comparison of isoprene+MVK+MACR mixing ratios averaged below 1 km
from the observations by G-1 flights over the Southwest and Northeast regions (as
marked in Fig. 12a) and WP-3D flights over southern California and the Central Valley
(as marked in Fig. 12b) and the corresponding simulations. Asterisk denotes the 50th
percentiles. Vertical lines denote 10th and 90th percentiles, and the boxes denote the 25th
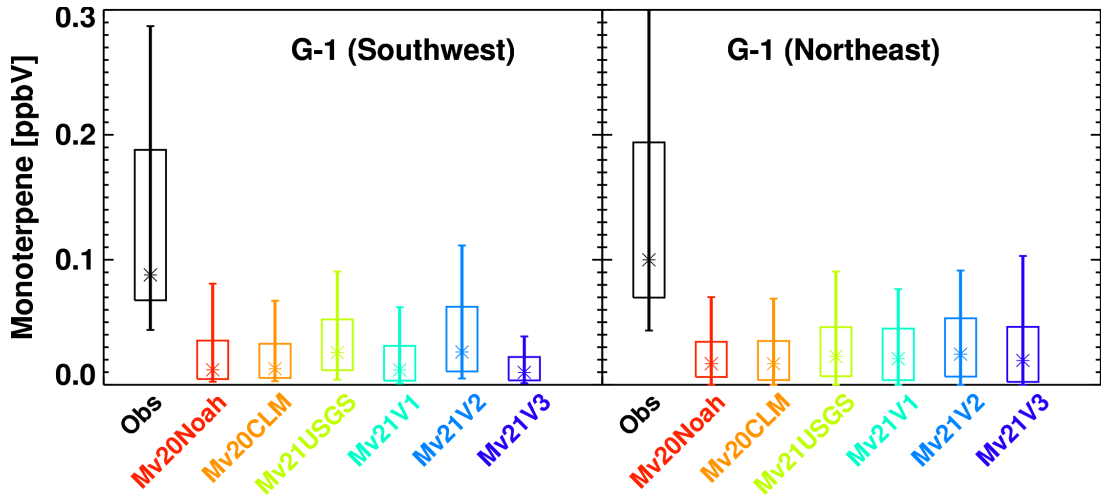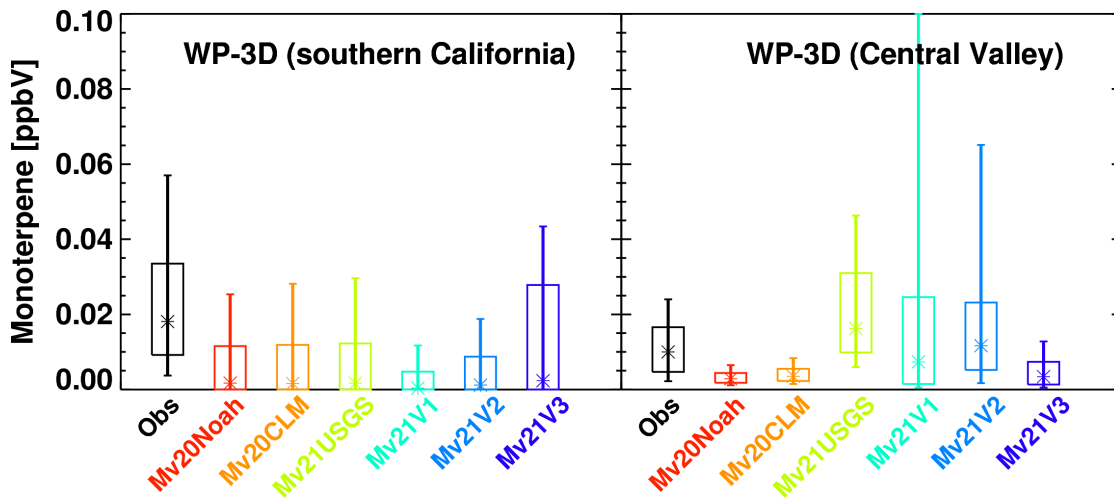and 75th percentiles.

1288

1289

1290

1291

1292

1293

1294

1295

1296

1297

Figure 17. Same as Fig. 16 except for monoterpene mixing ratios.

1299