HEI Research Report 240

Predictive, Source-Oriented Modeling and Measurements to Evaluate Community Exposures to Air Pollutants and Noise from Unconventional Oil and Gas Development

Lea Hildebrandt Ruiz et al.

Additional Materials F: Chapter 9

Correspondence may be addressed to Dr. Lea Hildebrandt Ruiz, The University of Texas at Austin, 200 E. Dean Keeton St., Austin, TX 78712; email: lhr@che.utexas.edu.

Although this report was produced with partial funding by the United States Environmental Protection Agency under Contract No. 68HERC19D0010 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and may not reflect the views of the Agency; thus, no official endorsement by the Agency should be inferred. This report also has not been reviewed by private party institutions, including those that support HEI Energy, and may not reflect the views or policies of these parties; thus, no endorsement by them should be inferred.

Additional Materials F: Chapter 9

Emissions Sources Data

Table F-1. Numbers of Emissions Sources and Locations Included in the Exposure Assessment

Location Name	TB	CWP	CWK
No. of Locations	402	314	975
No. of Sources	1089	936	4130

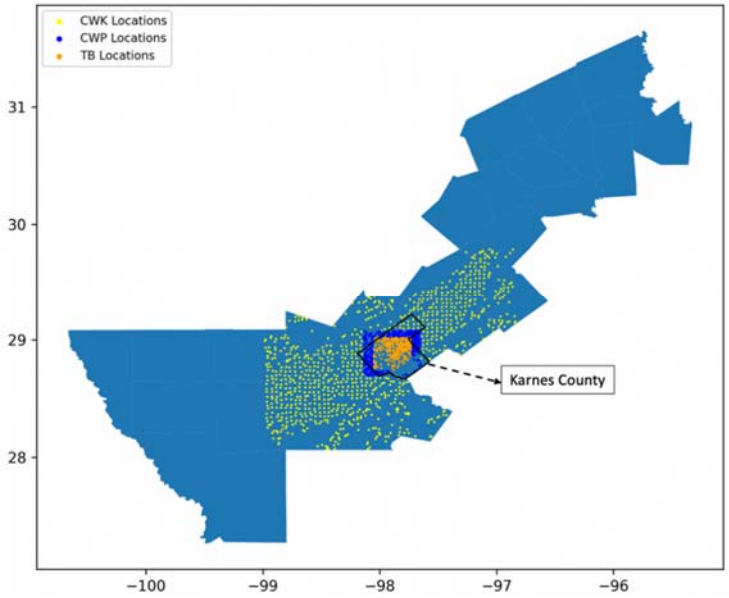


Figure F-1. Emissions sources and weighted locations included in the exposure assessment.

We modeled hourly emissions of methane, ethane, propane, and n-hexane using a modified version of the MEET tool.[3.17] For each site, we quantified emissions from leaks, ground leaks, pneumatic controllers, tank flash, and unloading, each of which has its own emissions factors and temporal emissions profiles. We aggregated hourly emissions to either 24-hour or 12-hour totals to estimate daily exposure or daytime and nighttime exposure.

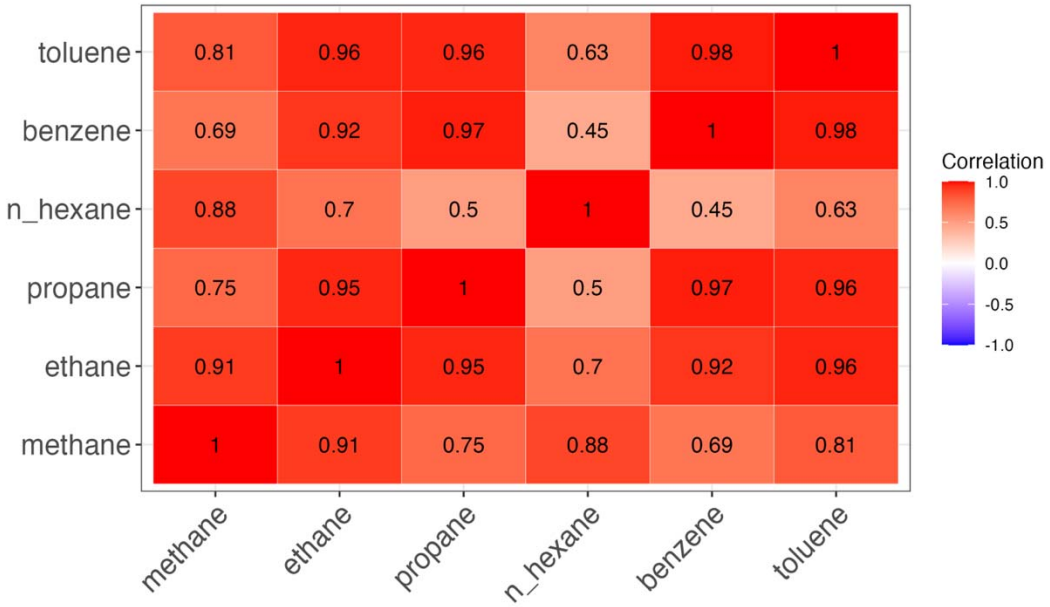


Figure F-2. Correlation matrix of hourly emissions for different pollutants from oil and gas extraction sources at EFS.

Pollutant concentration timeseries at the Karnes city measurement site

For methane, all models, except IDW (no meteorology), show good agreement with observational data, with correlation values near r = 0.8. This reflects performance by AERMOD, CALPUFF, Gaussian, and IDW-met in capturing methane variability at the Karnes City site.

The evaluation of propane concentrations reveals that both IDW-met and IDW-no-met achieve the highest correlation with observations (r = 0.72) However, IDW-no-met generally underestimates observed peaks and produces overly smoothed profiles, suggesting that while its overall trends align, it lacks sensitivity to short-term variability. AERMOD follows with a solid performance (r = 0.69), capturing daily trends reasonably well but slightly underestimating high-concentration events. CALPUFF, on the other hand, demonstrates large deviations and sharp peaks (r=0.58), especially in late March and mid-April, indicating possible over-responsiveness to meteorological shifts or localized emissions. The Gaussian model shows the weakest correlation (r = 0.57), underestimating both variability and magnitude.

For n-hexane, IDW-met achieves the strongest correlation with observations (r = 0.76), followed closely by Gaussian (r = 0.61) and AERMOD (r = 0.64), indicating relatively consistent temporal agreement among these models. The time series reveals that AERMOD captures the timing of most concentration peaks—particularly around late March and early April—but tends to overestimate baseline levels in later periods. CALPUFF performs moderately (r = 0.63), tracking general trends but underestimating key episodic peaks. While Gaussian produces smoothed outputs, it aligns well with observed fluctuations.

For benzene, AERMOD shows the good agreement with observations (r = 0.71), followed by CALPUFF (r = 0.62), indicating their better capability to capture temporal variability and peak events. The time series confirms that AERMOD tracks daily fluctuations closely and reasonably captures elevated concentrations during high-activity periods in early April and mid-April. CALPUFF, while occasionally aligning with observation peaks, tends to produce exaggerated variability, including multiple sharp overestimations, particularly around April 17.

For toluene, IDW-met emerges as the best-performing model with the highest correlation to observations (r = 0.88), indicating strong temporal agreement and an ability to reflect day-to-day variability. AERMOD and CALPUFF follow with more modest performance (r = 0.75 and r = 0.54, respectively), both capturing general trends but often overestimating or misaligning with episodic peaks. The time series shows that CALPUFF consistently overpredicts concentrations—especially in mid-April—producing large and unrealistic spikes, while AERMOD better tracks the shape and magnitude of observed fluctuations.

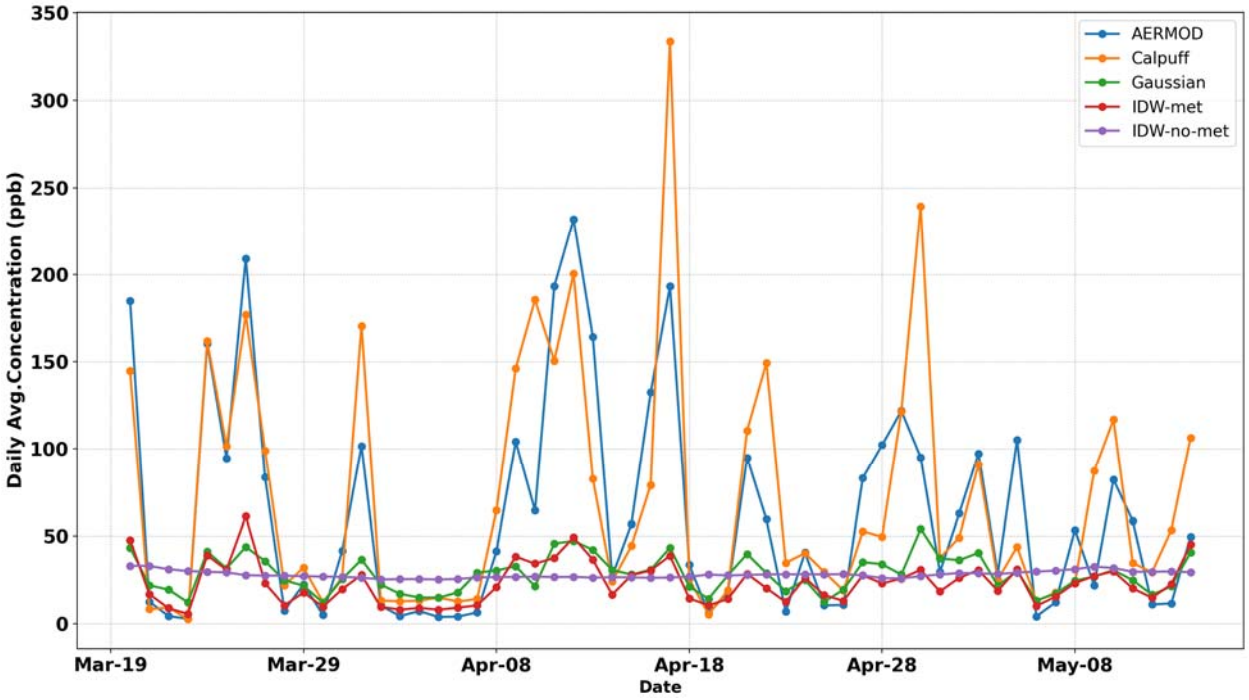


Figure F-3. Daily average methane concentrations from each of the models at the Karnes City measurement site.

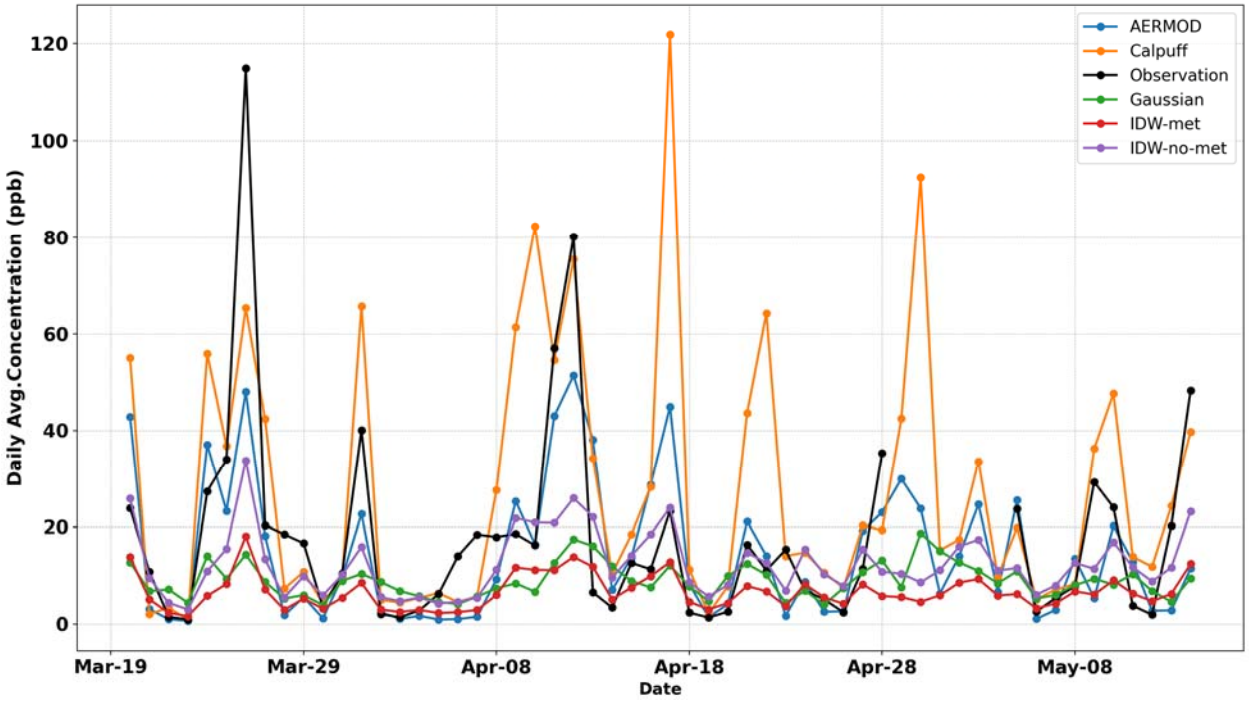


Figure F-4. Daily average propane concentrations from each of the models at the Karnes City measurement site.

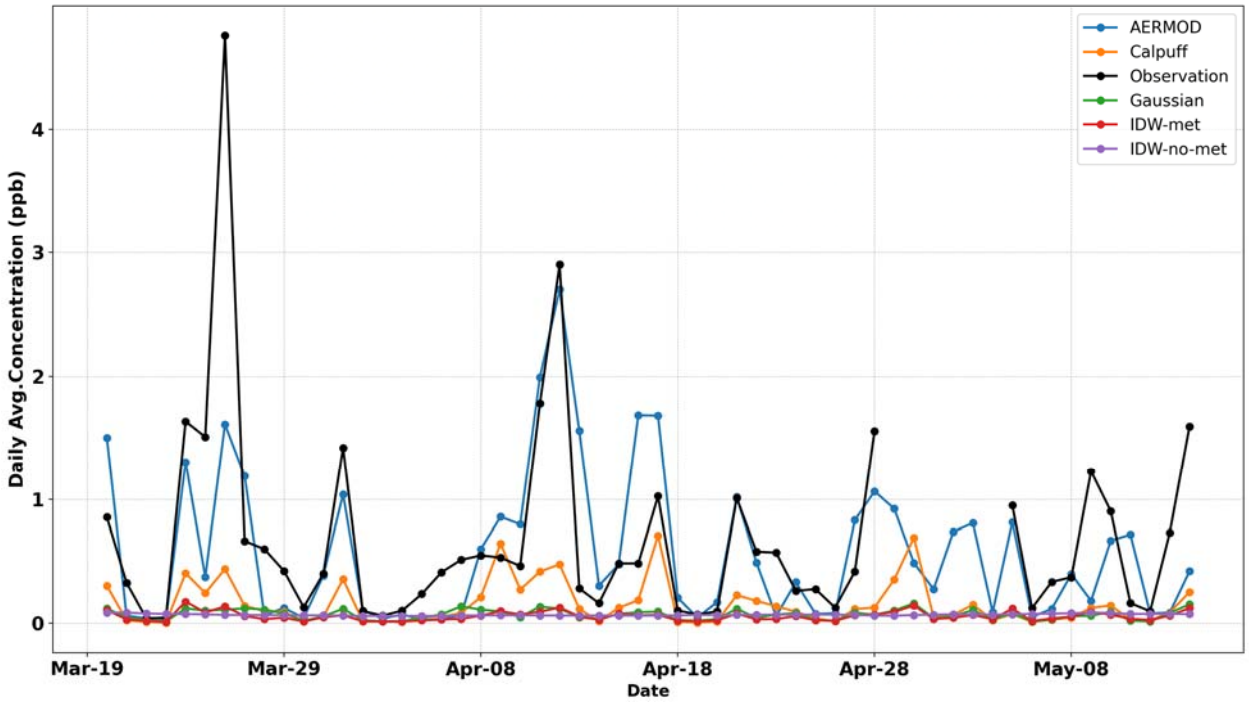


Figure F-5. Daily average n-hexane concentrations from each of the models and observations at the Karnes City measurement site.

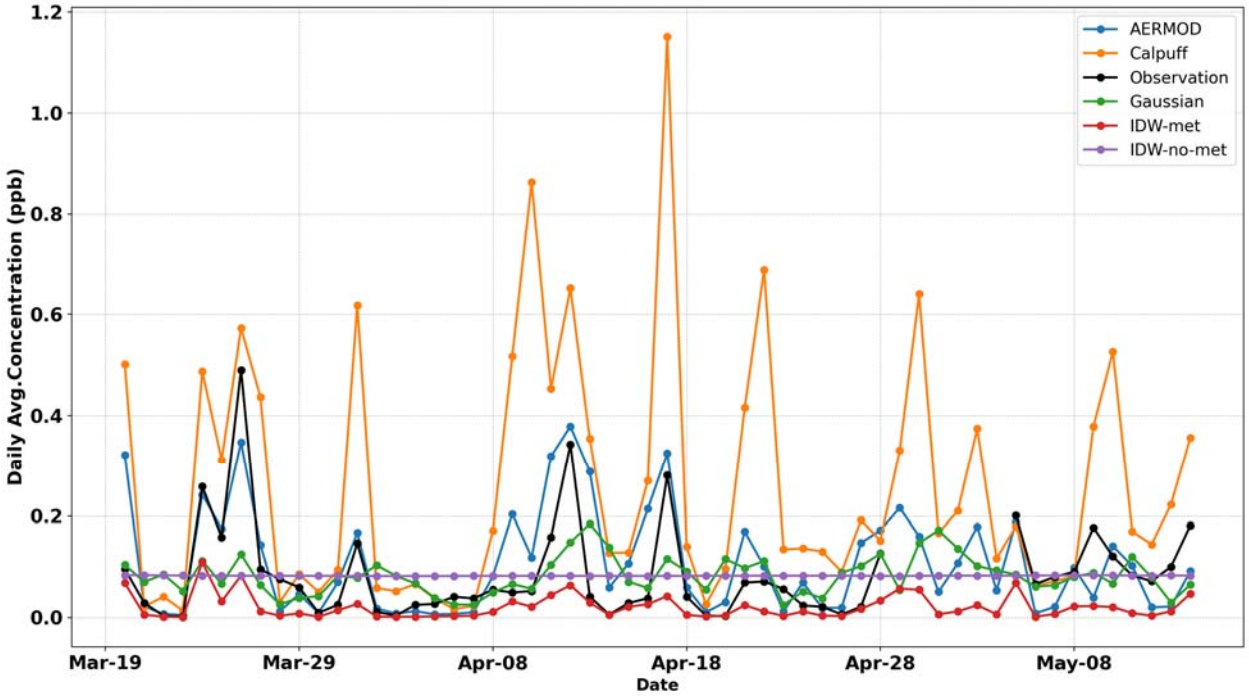


Figure F-6. Daily average benzene concentrations from each of the models and observations at the Karnes City measurement site.

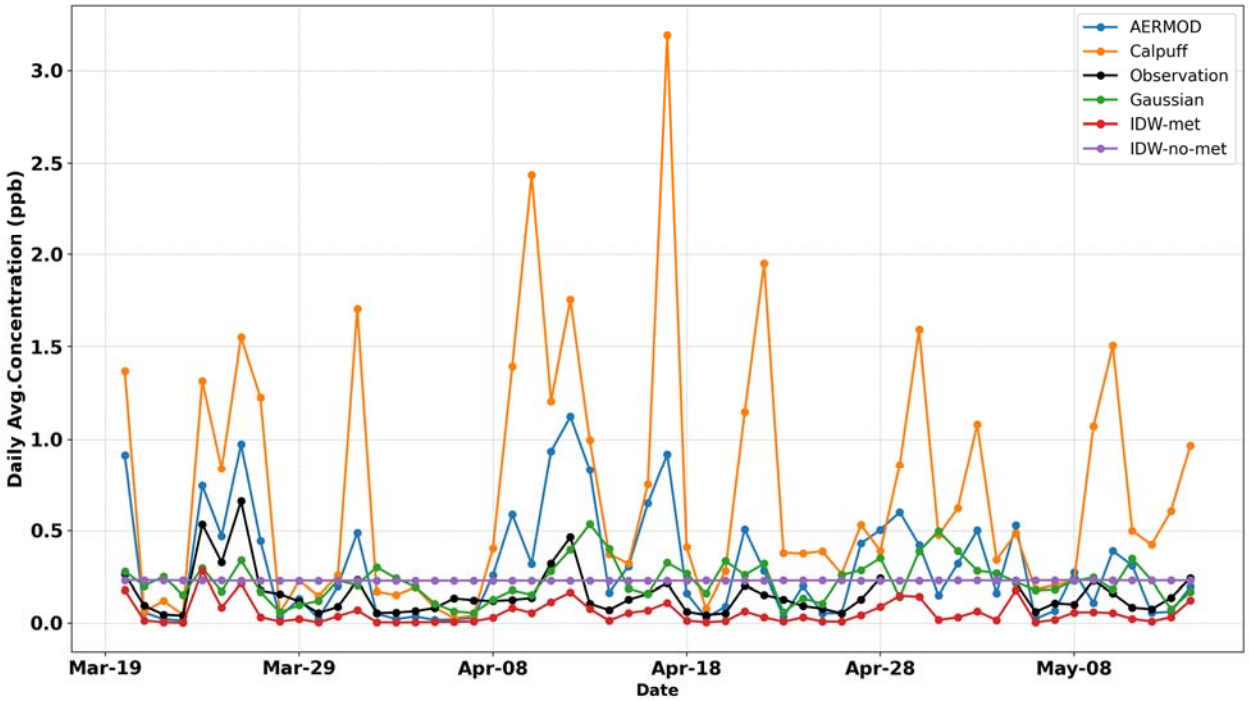


Figure F-7. Daily average toluene concentrations from each of the models and observations at the Karnes City measurement site.

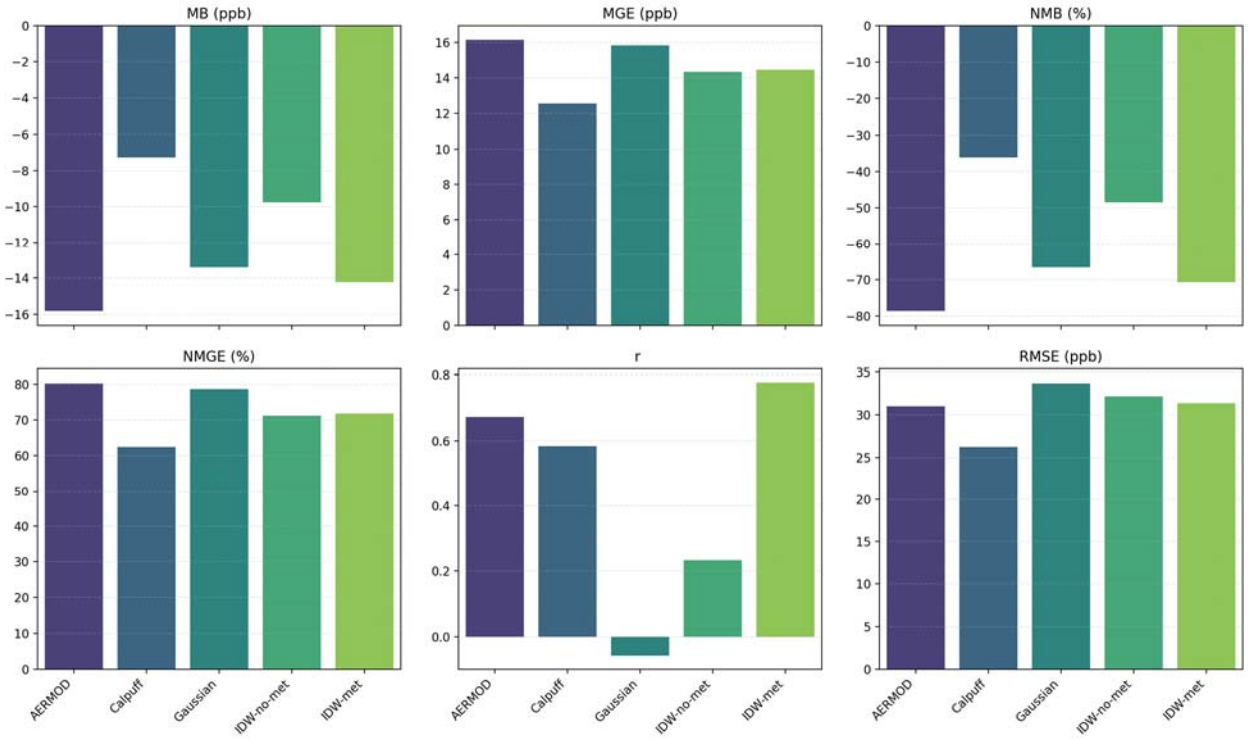


Figure F-8. Evaluation statistics for the models for daytime average ethane concentrations observed at Karnes City from March 20 – May 14, 2023.

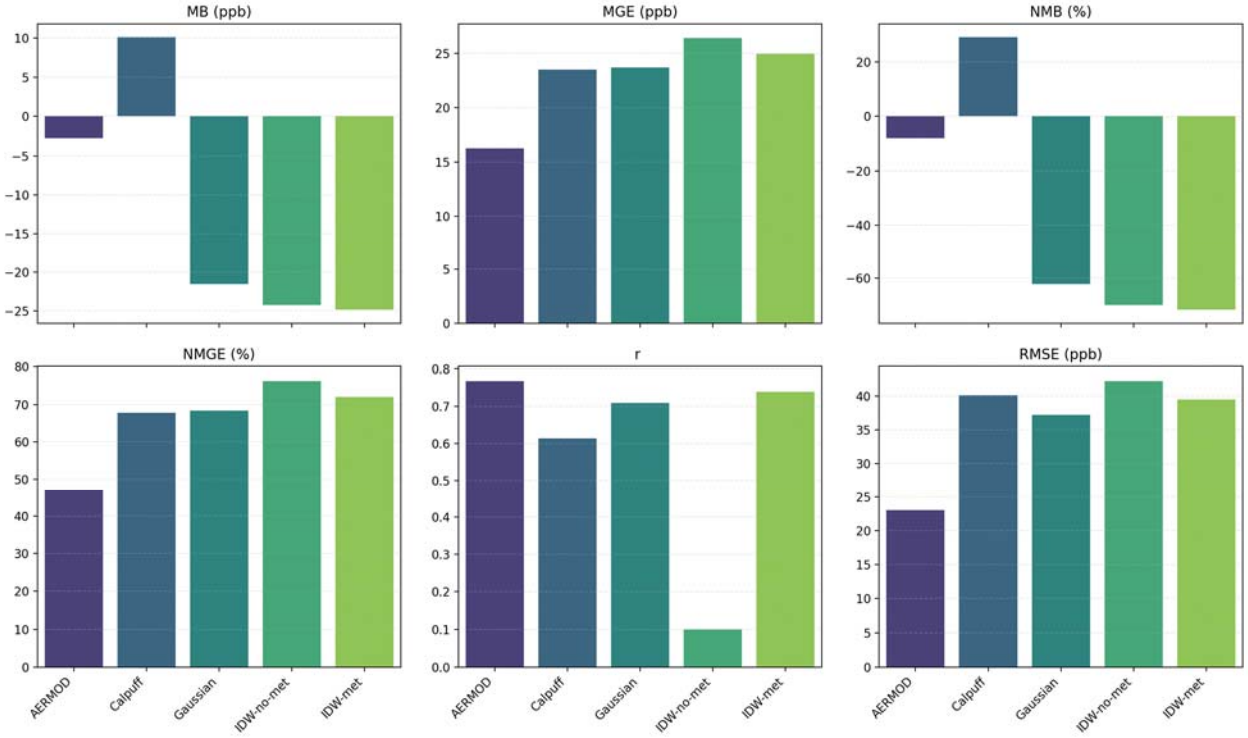


Figure F-9. Evaluation statistics for the models for night time average ethane concentrations observed at Karnes City from March 20 – May 14, 2023.

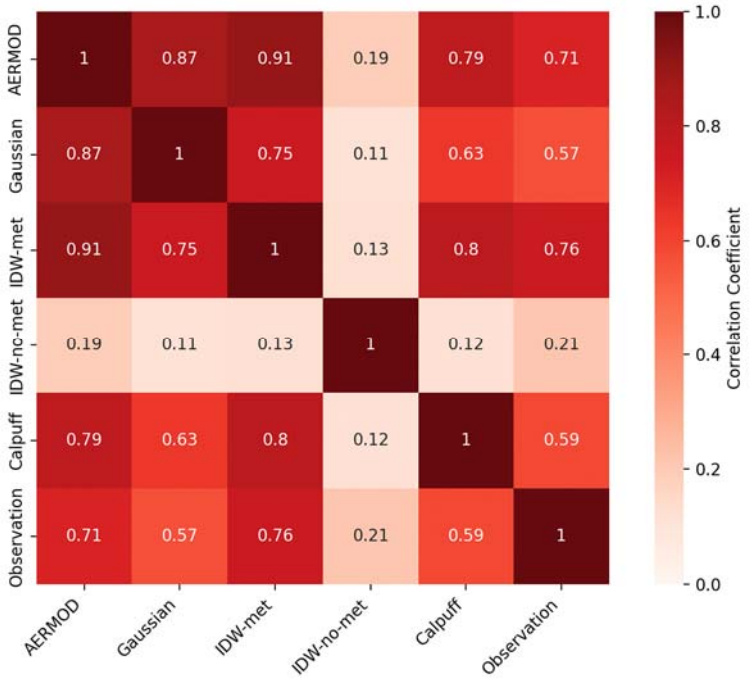


Figure F-10. Correlation matrix between exposure models for daily average concentration estimates of ethane at the Karnes City monitor.

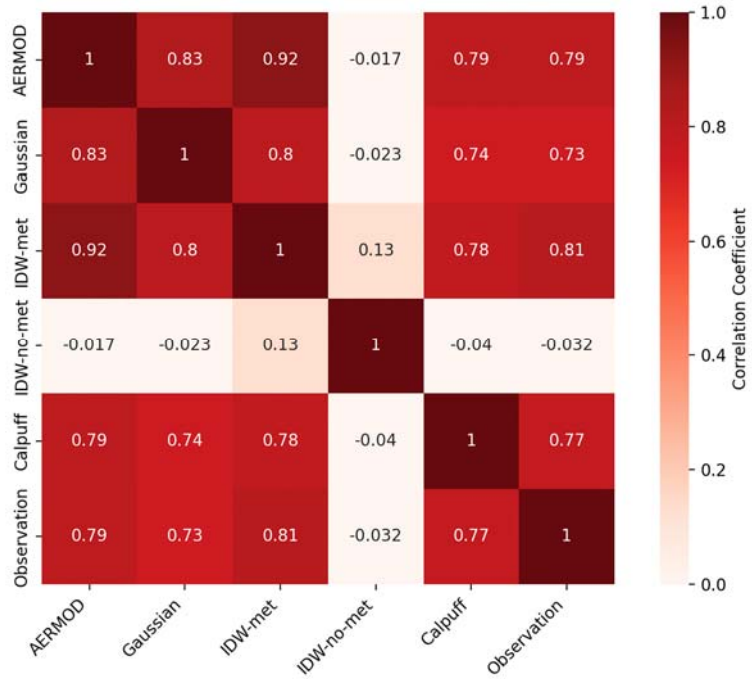


Figure F-11. Correlation matrix between exposure models for daily average concentration estimates of methane at the Karnes City monitor.

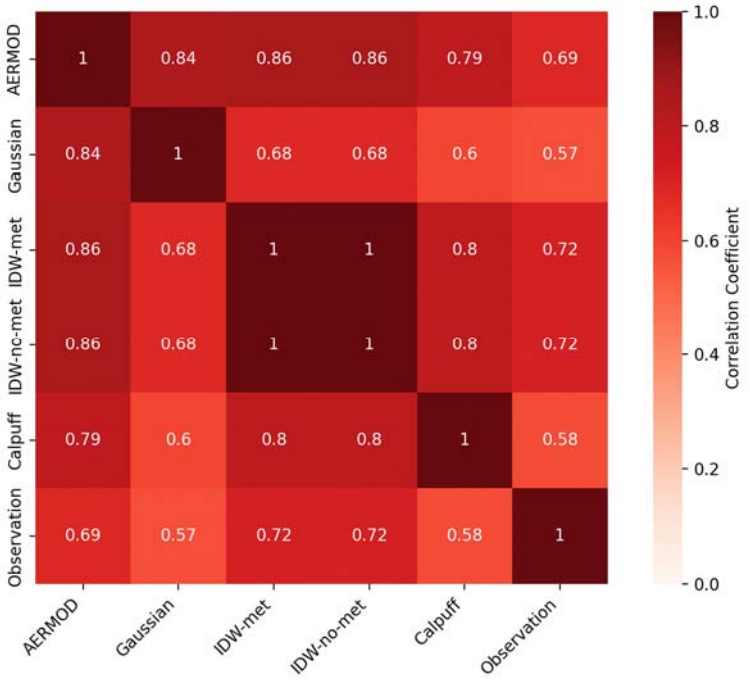


Figure F-12. Correlation matrix between exposure models for daily average concentration estimates of propane at the Karnes City monitor.

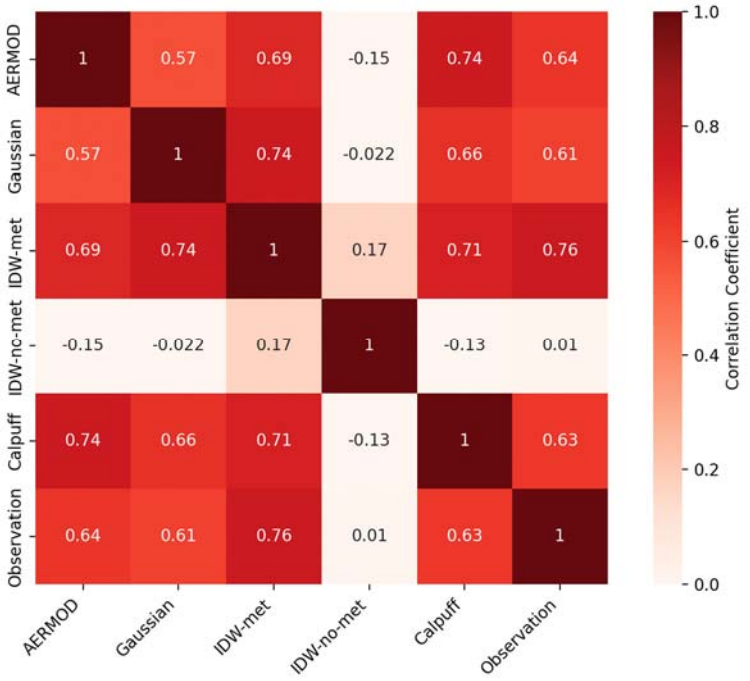


Figure F-13. Correlation matrix between exposure models for daily average concentration estimates of n-hexane at the Karnes City monitor.

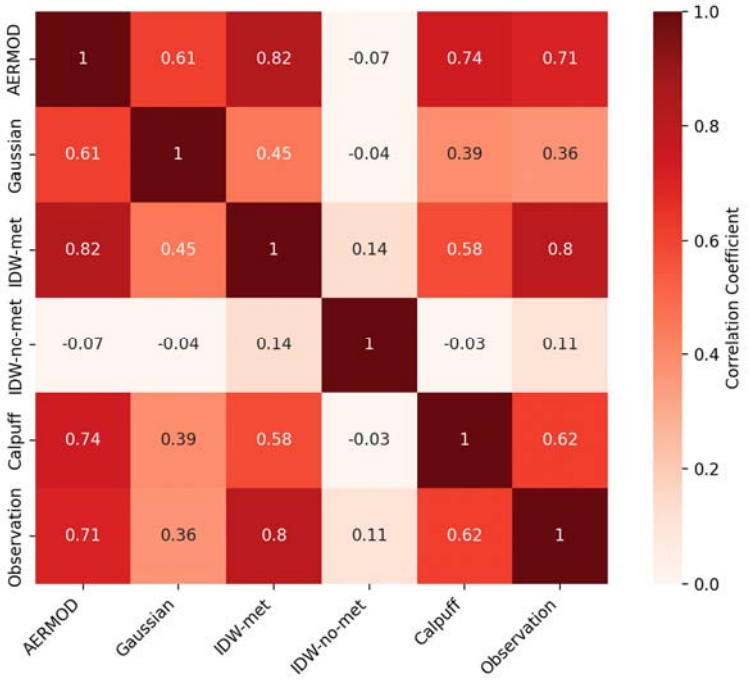


Figure F-14. Correlation matrix between exposure models for daily average concentration estimates of benzene at the Karnes City monitor.

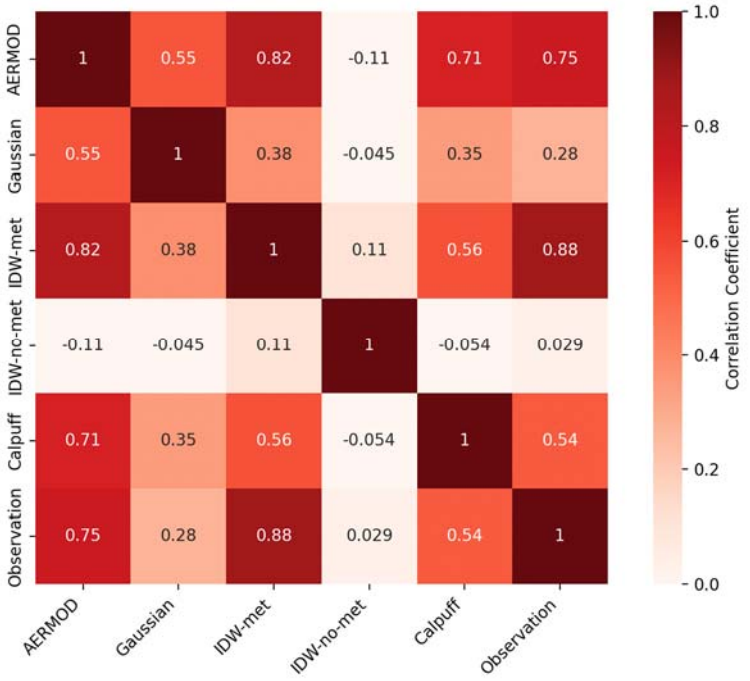


Figure F-15. Correlation matrix between exposure models for daily average concentration estimates of toluene at the Karnes City monitor.

Factors leading to concentration variability

Relative influences of meteorological and emission variability

The two factors most impactful on modeled concentrations at the measurement site are source emissions and meteorology (primarily wind speed and direction, which additionally influence stability class). To quantify their influences on measurement site concentrations separately, we modeled two scenarios: 1) holding ethane emissions constant at their first day's values, and 2) holding wind speed and direction constant at their first day's values. We compare the variability in modeled concentrations from all sources at the measurement site by dividing each scenario by its first day's concentrations. Meteorological variations result in greater variation in observed concentrations than emissions variability.

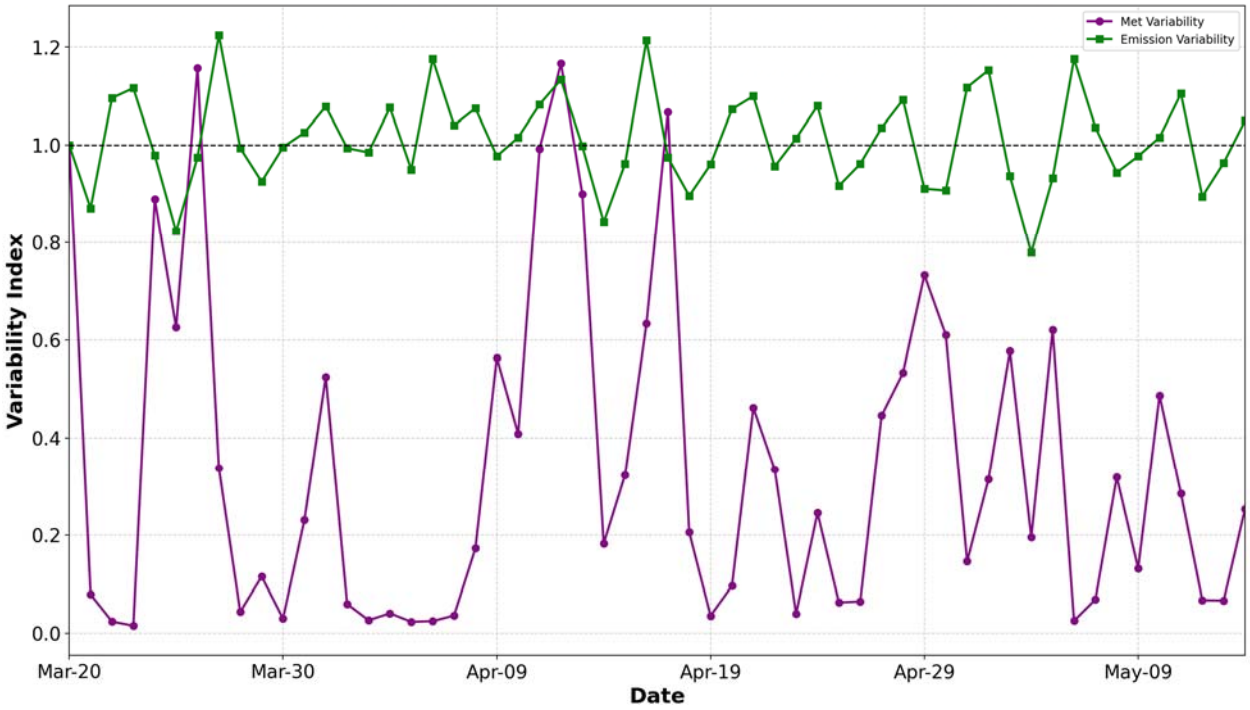


Figure F-16. Comparison of meteorological and emission variability relative to the first day of simulation (March 20) for ethane.

Meteorology leads to greater variation in observed concentrations than emissions variability (Figure E-16). Daily average concentrations from the AERMOD run with constant emissions shows a moderate autocorrelation with at a 1-day lag (Figure E-17), with positive correlations observed at the first three lags before declining toward zero or negative values. This pattern suggests some systemic memory for up to three days due to meteorology. In contrast, for constant meteorology, the ACF shows weaker and less consistent autocorrelation across all lags, with most values fluctuating narrowly around zero

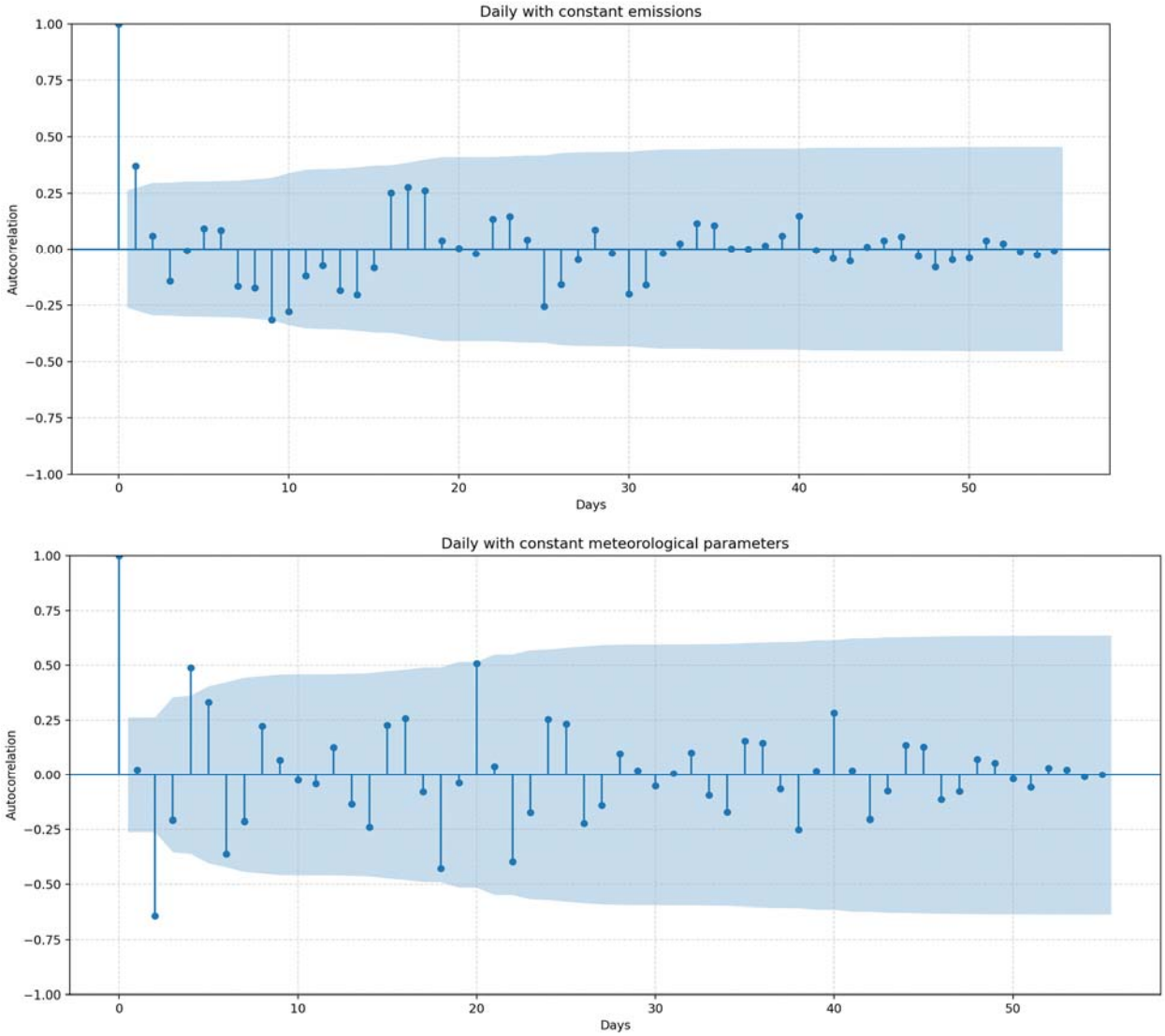


Figure F-17. Autocorrelations of daily average modeled concentrations under constant meteorological and emission conditions from AERMOD.

Influence of wind speed and direction on modeled concentrations

Wind speed is associated with increased transport and dilution of atmospheric pollutants in the atmosphere. Higher wind speeds typically reduce pollutant concentrations by dispersing them over a larger area, while lower wind speeds allow pollutants to accumulate near their sources. We found a large decline in ethane concentration with increasing wind speed (Figure E-18). However, the models predict much flatter trends, suggesting they underestimate the sensitivity of ethane concentration to wind speed compared to real-world observations. This highlights potential limitations in the models' ability to capture the full impact of wind speed on pollutant dispersion.

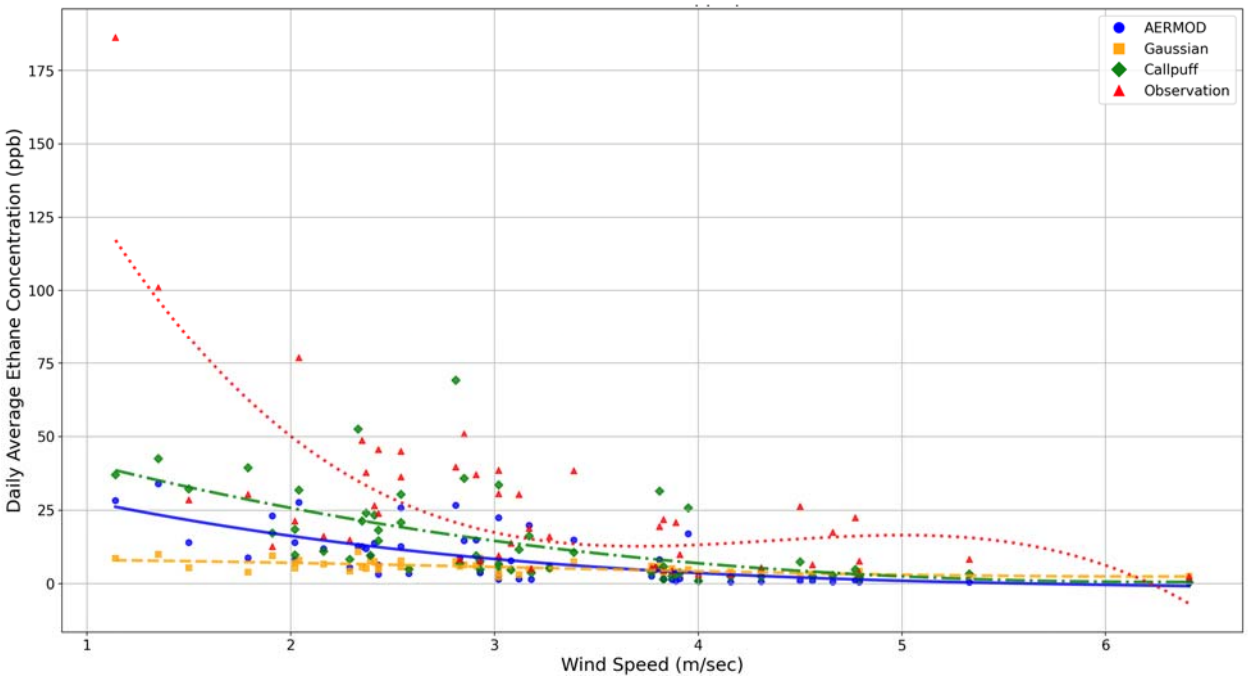


Figure F-18. Relationship between wind speed and daily average ethane concentrations, third-degree polynomial regression was applied to the data for each model.

Effect of elevation on the average concentration

We assessed the influence of AERMOD's consideration of terrain elevation on the average daily ethane concentrations from March 20 to April 20 at the measurement site (Figure E-19). The purpose is to determine whether the elevation adjustment impacts concentration levels and whether it is necessary for accurate analysis. The two series show nearly identical trends, with only slight reductions in concentration when elevation is considered.

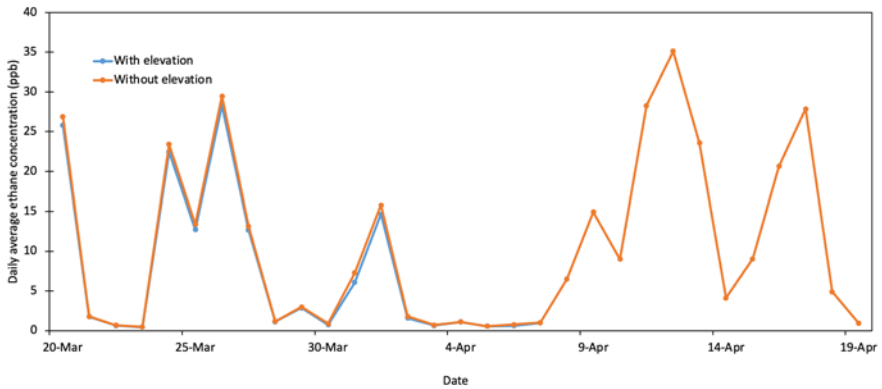
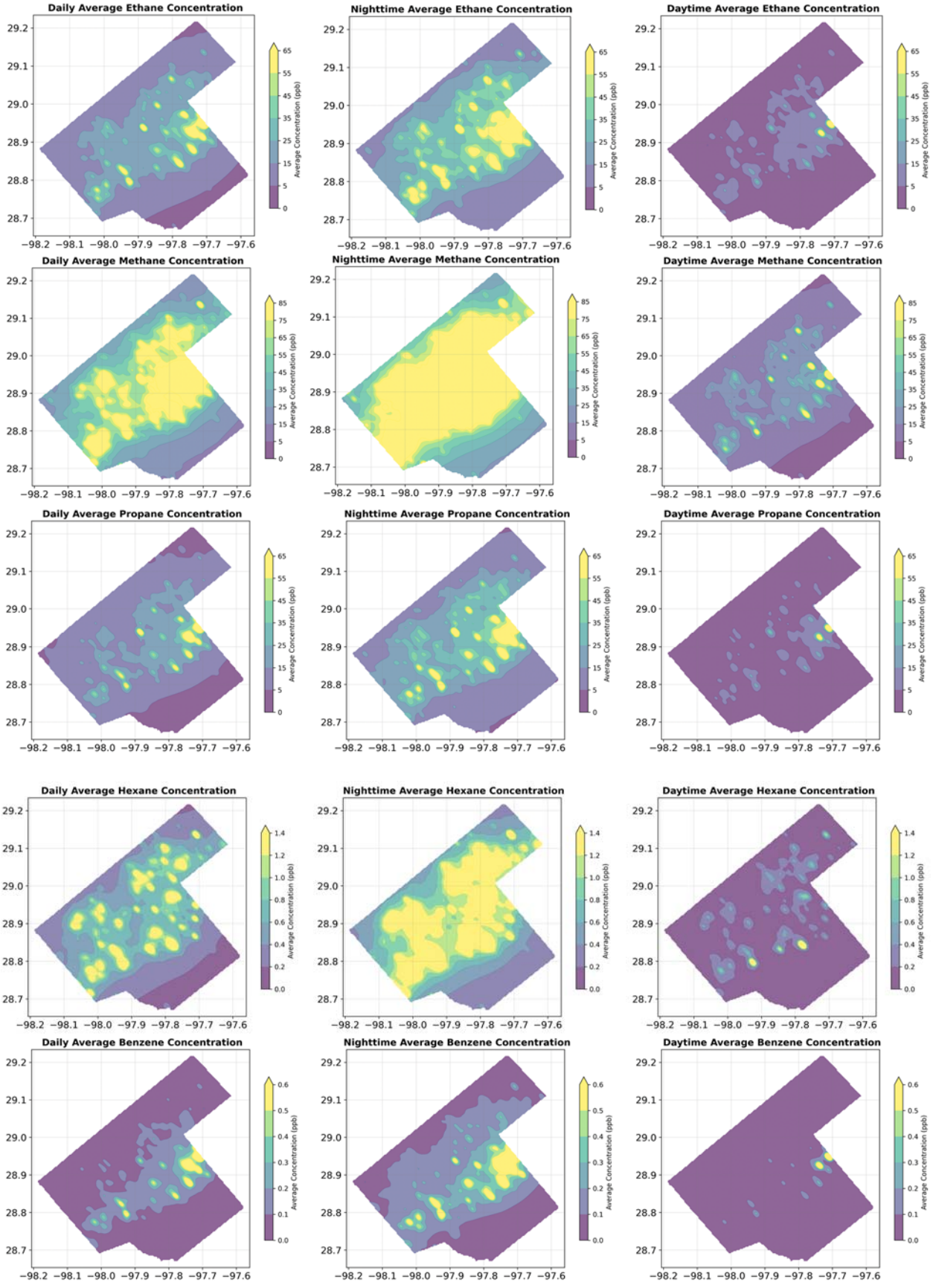


Figure F-19. Daily variability in ethane concentrations at the Karnes City monitor from AERMOD modeled with and without adjusting for variable elevation in the modeling domain.

Spatial concentration variability assessed with AERMOD

Spatial concentration contributions from all UOGD operations follow similar patterns for all pollutants (Figure E-20). Concentrations are elevated in the central part of the county in a band running from southwest to northeast. The highest concentrations are observed in the central region of the county, following a southwest-to-northeast band, which aligns with areas of intensive UOGD activity and emissions. The southeastern part of the county consistently shows lower concentration levels, likely due

to reduced UOGD activity and the prevailing southeasterly wind. The day-night concentration variations largely reflect meteorological influences, rather than changes in emissions. For all pollutants, nighttime concentrations are consistently higher than daytime levels.


Daily Average Ethane Concentration
Nighttime Average Ethane Concentration
Daytime Average Ethane Concentration
Daily Average Methane Concentration
Nighttime Average Methane Concentration
Daytime Average Methane Concentration
Daily Average Propane Concentration
Nighttime Average Propane Concentration
Daytime Average Propane Concentration
Daily Average Hexane Concentration
Nighttime Average Hexane Concentration
Daytime Average Hexane Concentration
Daily Average Benzene Concentration
Nighttime Average Benzene Concentration
Daytime Average Benzene Concentration
Average Concentration (ppb)

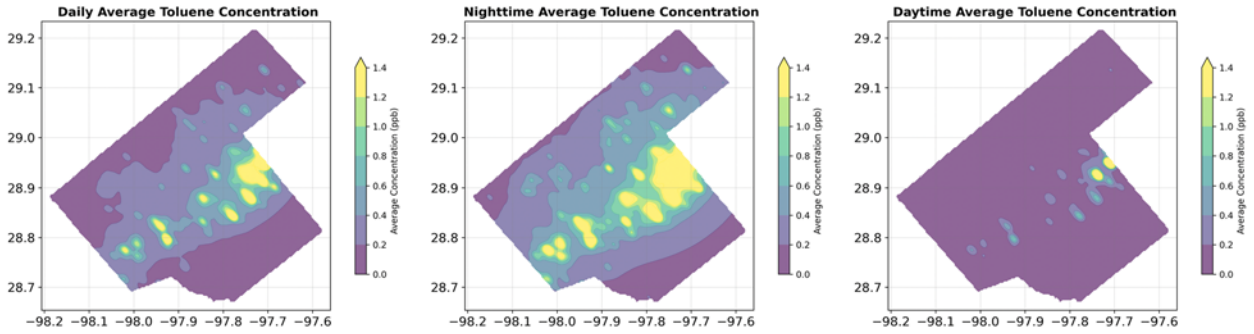


Figure F-20. AERMOD concentration maps from UOGD sites in Karnes County for ethane, methane, propane, n-hexane, benzene, and toluene.

UOGD concentration contributions by location and process type:

Spatially (Figure E-21), CWK sources influence only the northernmost portion of the county. Tank Flash, Pneumatics, and Unloading operations are the major contributors to ethane concentrations. Tank Flash emissions (998 kg/hr) alone account for 44% of the modeled concentration, followed by Pneumatics (792 kg/hr, 33%) and Unloading (1026 kg/hr, 20%). Spatially, these sources enhance concentrations in unique spatial patterns in Karnes County. panel displays high-intensity emissions with significant spatial coverage. In contrast, other sources such as Leaks, Ground Leaks, and Flares have minimal emissions (<50 kg/hr) and contribute less than 1% to observed concentrations, with negligible presence in the spatial distribution maps (Figure E-22).

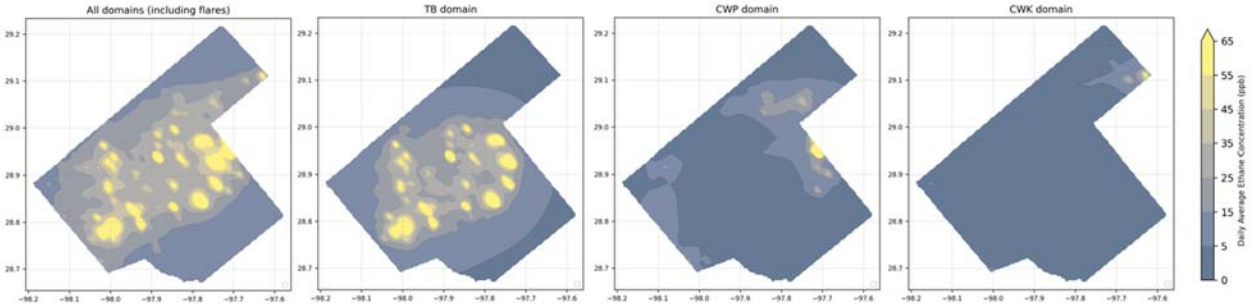


Figure F-21. Spatial concentration contributions to ethane from UOGD sources separated by spatial domain on March 20.

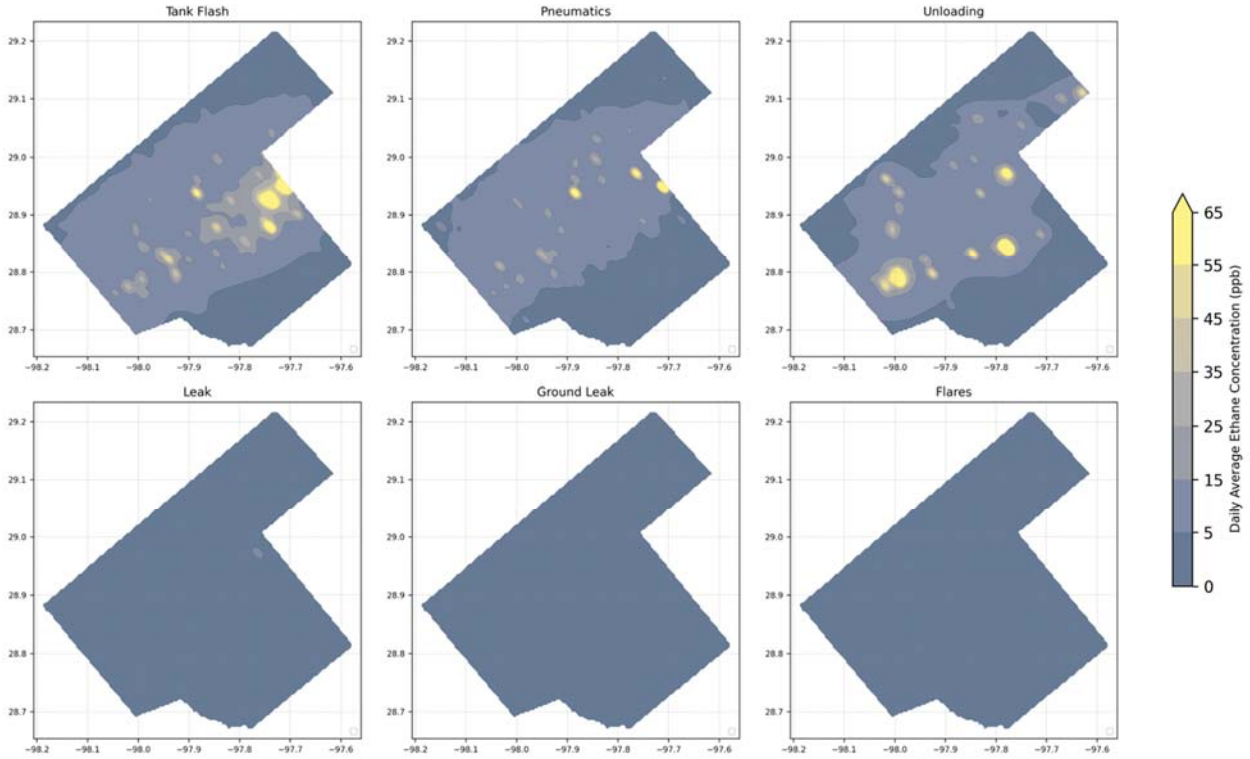


Figure F-22. Spatial concentration contributions to ethane from UOGD sources separated by process type.

Summary statistics of Census geographies in EFS

Here, we investigate the spatial modeling resolution needed to represent exposure variability across geophysical boundaries in EFS. The density and percentile plots for block groups, tracts, and counties provide insight into the distribution of land areas within these Census geographies in EFS, TX. For block groups, the distribution is bimodal, with density around 1 km^2 and 100 km^2 suggesting that block groups tend to fall into these two size categories. The percentile plot reveals that more than 75% of block groups are smaller than 1000 km^2, indicating a concentration of relatively smaller land areas. Overall, block groups show a wide range of sizes, with a majority skewed toward the smaller end of the spectrum.

Moving to tracts, the density plot indicates a broader range of land areas compared to block groups, with peaks at 10 $km^2$and 1000 km^2. This bimodal distribution suggests that tracts, like block groups, are split between smaller and larger size categories, though the areas are generally larger. The percentile plot supports this, showing that approximately 75% of tracts are smaller than 2000 km^2. The median tract size appears to fall around 100 km^2, highlighting a wider but still moderate land area distribution.

For counties, the data reveals a more centralized distribution around a peak near 3000 km^2, with most counties having much larger land areas compared to block groups and tracts. The percentile plot shows that 75% of counties are under 4000 km^2, indicating a somewhat compressed range of county sizes. Unlike the other geographies, the distribution for counties is more unimodal, suggesting less variability in land area, with the bulk of the data concentrated around the median of 3000 km^2. This highlights that counties tend to occupy much larger and more uniform areas in EFS, TX.

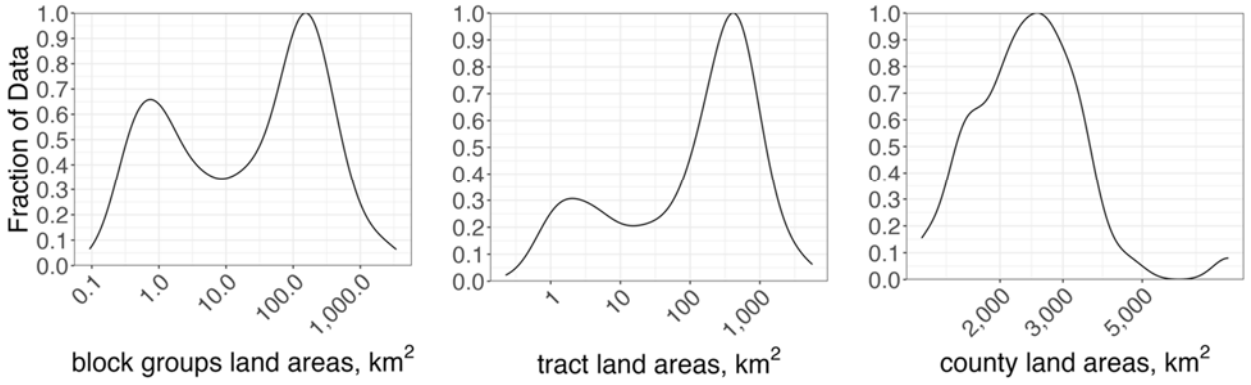


Figure F-23. Density of block group land areas in EFS region for different census geographies.

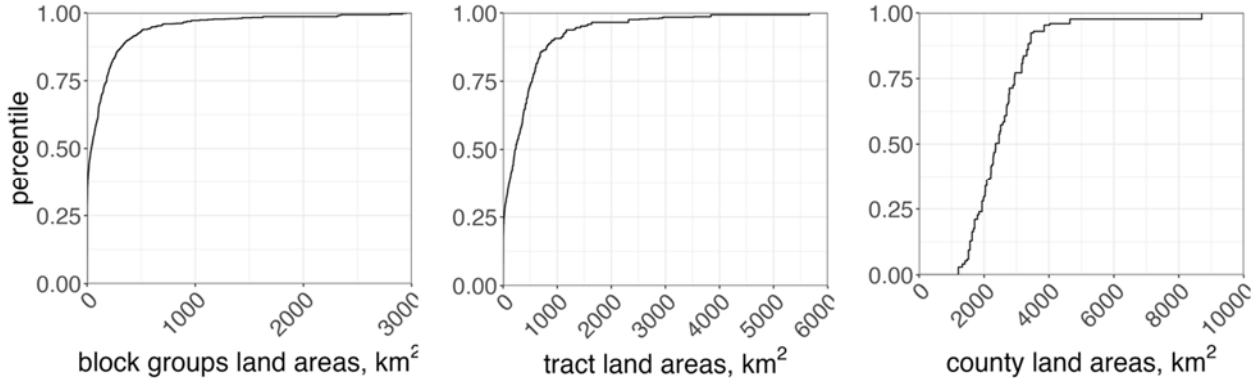


Figure F-24. Cumulative percentiles of land areas in EFS region for different census geographies.

The 1.33 km×1.33 km grid size is a balanced choice in terms of spatial resolution and computational efficiency. From the density plot, we see that finer grids such as 500 m and 1,000 m (1 km) resolutions have a higher density of grid cells in block groups, which means significantly more grid cells to process. However, these finer grids may not provide substantially more information for regional studies, making the higher computational cost unnecessary. The selected resolution of 1.33 km×1.33 km offers an appropriate level of detail while being computationally manageable.

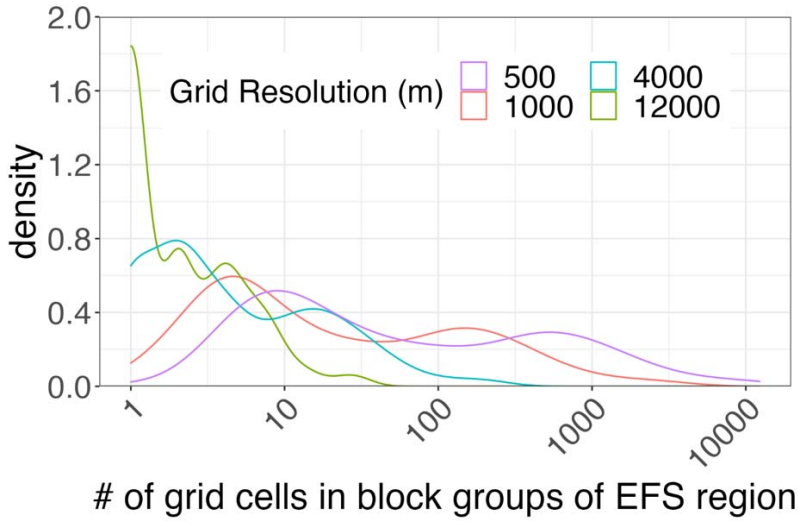


Figure F-25. Distribution of grid cell counts per block group for various grid resolutions in EFS region.

Population demographics and income in EFS

The Eagle Ford Shale (EFS) region is composed of multiple counties, each with diverse demographic distributions. The population is categorized into several key groups: White, Black, Native American, Asian, Native Hawaiian and Pacific Islander (HIPI), and Hispanic. Across the entire EFS region, the White population forms the largest demographic group, with block group populations frequently exceeding 1,500 individuals in certain areas. The Hispanic population is also significant, with several block groups showing more than 2,000 individuals. In contrast, the Black population is more modest, with most block groups having fewer than 750 individuals. Native American and HIPI populations are generally small across the EFS area, with most block groups reporting fewer than 20 individuals. The Asian population also remains relatively low, with fewer than 100 individuals in most block groups.

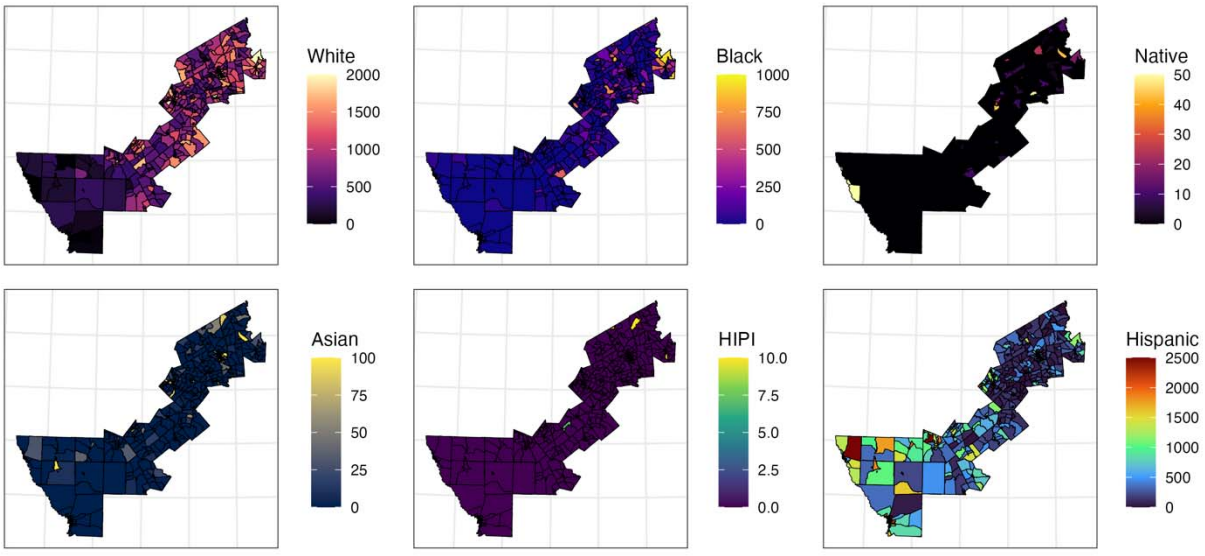


Figure F-26. Demographic distribution across the EFS region by census block groups (2020).

In Karnes County, the population demographics reflect a smaller, but still diverse, subset of the overall EFS region. The White population in Karnes County is concentrated in specific block groups, with estimates reaching over 1,500 individuals in the most populous areas. The Hispanic population plays a significant role here as well, often exceeding 2,000 individuals in certain block groups, making it one of the dominant demographics in the region. The Black population, while present, is generally lower, with most block groups having fewer than 500 individuals. The Native American**,** Asian**,** and HIPI populations remain minimal, with most areas reporting fewer than 50 individuals in these groups.

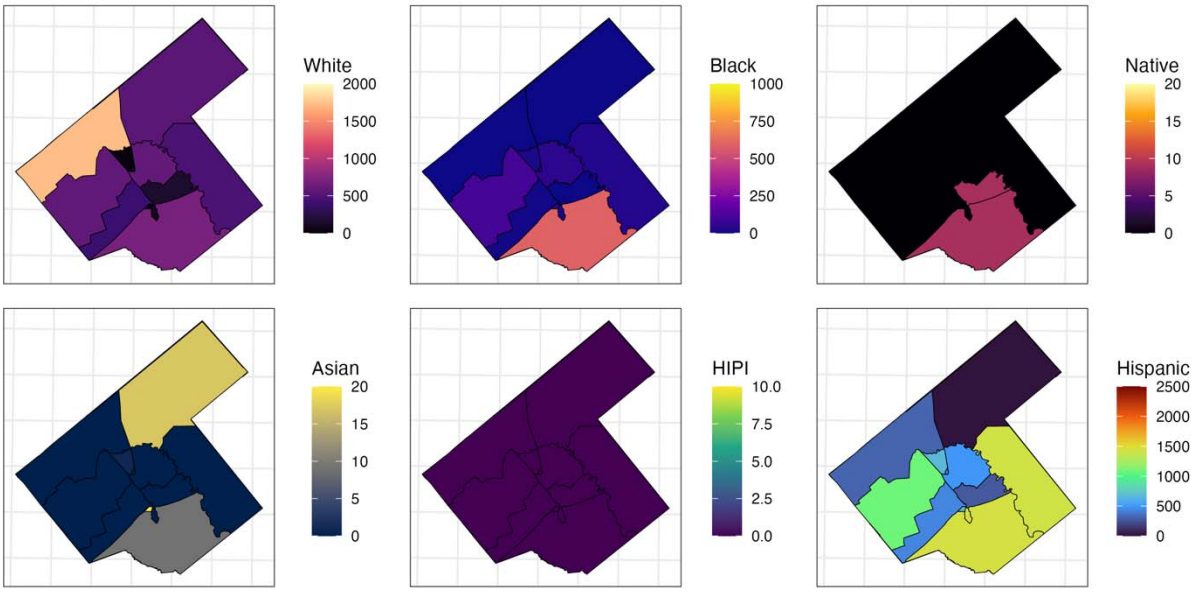


Figure F-27. Demographic distribution in Karnes County by Census Block Groups (2020).

EFS income estimates have a much wider range, from $3,901 to a significant $185,515. The median income is also $59,048, aligning with Karnes County, but the mean income for EFS is $59,048, lower than Karnes County's $62,163. This wider spread suggests both wealthier and lower-income areas within the EFS region.

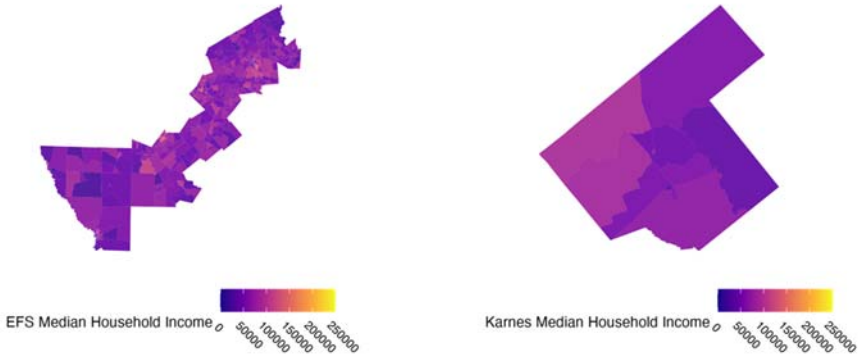


Figure F-28. Median household income ($59,048 yr^{-1}) in EFS and Karnes County Census blocks.

The income estimates for Karnes County (2020) range from $26,698 to $90,625, with a median household income of $59,048 and a mean income of $62,163. This indicates that while the median income is somewhat near the mean, there are likely some higher income earners skewing the mean upwards.