We thank the editor for the comments. We feel that they have helped improve the manuscript. Our response is color-coded in blue.

Can the random training/test split produce samples in each subset that are close in space and time so that the corresponding footprints are very similar? Since such correlations between training and test samples would lead to an overestimation of test performance and hence generalisability, can this be ruled out?

We thank the editor for this comment. In the first version of FootNet, we attempted to address this issue by using footprints from two regions with different meteorology. As such, this version of FootNet is generalizable to both simple (Barnett Shale) and complex (SF Bay Area) regions. You are correct that the performance may be optimistic in this work. The focus of this manuscript is the development of a proof of concept: FootNet v1. We have a companion paper that uses this model within a flux inversion system to test the viability of quantifying emissions using a machine learning surrogate model (Dadheech et al., ACP under review). We also have preliminary results from ongoing work generalizing the FootNet model to any region. The preliminary results indicate comparable performance to what is shown here, but that analysis is very much ongoing.

In terms of peer-review, we submitted this manuscript as the proof of concept that we can effectively emulate the source-receptor relationship under simple and complex cases, a second paper showing that the emulator can be used within a flux inversion framework, and aim to have a third manuscript generalizing this framework.

Line 197: "Due to the computational cost required by the generation of high-resolution footprints, we only included footprints generated from previous studies for the two locations in training version 1.0 of FootNet. We are actively generating new footprints at 1 km from a broader region to further improve the emulator's performance, especially in regions with different meteorological conditions from the two locations used in this study (Dadheech et al., under review). Generalizing this source-receptor emulator to other regions is being tackled in the next version of FootNet."

Code and data availability section: Please add the archives
https://zenodo.org/records/12752655
https://zenodo.org/records/12803617
https://zenodo.org/records/12803736
https://zenodo.org/records/12803855
Thank you. The archives are now added in the revised manuscript.

Line 78: In the units of H the per area seems to be missing. Are the units pbb / (nmol m-2 s-1) per grid cell (i.e. per km2)? In general, I think that the Footprint H, being the central quantity of this study, deserves to be introduced with an equation.

Thank you for the comment. The footprints can be thought of as the mapping between emissions and concentrations. The concentrations are the product of the spatial map of emissions (in units of nmol/m^2/s) multiplied by the footprint. When a spatial map of gridded emissions is multiplied by the spatial footprints, the resulting quantity is a concentration in units of ppb. We have added a paragraph introducing footprint and its unit in the paper.

Line 64: "The output of FootNet is a source-receptor relationship (i.e., footprint, **H**), which represents the sensitivity of atmospheric concentrations at a receptor site to emissions upwind of the receptor. This relationship between the measured concentrations and the emissions in the upwind area can be formulated as

$$\mathbf{y} = \mathbf{Hx} + \mathbf{b}$$

where **y** represents the measured concentration, **x** is the emission fluxes in a domain around the measurement location, and b is the background concentration upwind of the domain. The units of **y** and **x** can be expressed as dry air mixing ratio (ppb) and flux rates (nmol m^{-2} s^{-1}), respectively (Lin et al., 2003). The source-receptor relationships, $\mathbf{H} = \partial \mathbf{y}/\partial \mathbf{x}$, therefore have units of ppb / (nmol m^{-2} s^{-1})."

Line 89: Is the target also scaled?
The target is log-transformed to reduce the skewness of the distribution of footprint values. This information is added in the paragraph.

Line 78: "The output of FootNet is measurement footprints and is transformed by the natural logarithm function to reduce the skewness of the distribution of footprint values."

Line 113 and Figures: Does "log" mean "ln" or "log_10"? Could you please clarify the units in the log(H) plots, what are the units of H before applying the log?
"Log" means "natural logarithm". This information is corrected throughout the paper. The unit of H is clarified in all the log(H) plots now.

Line 120, Eq. (1): Please explicitly introduce the |.| notation.
The |.| notation is now introduced in line 111.

Line 111: "The absolute value bars ($| \cdot |$) here refer to the area of a region. Specifically, $|Y \cap \hat{Y}|$ represents the area of the region where both the truth and FootNet predictions show non-zero footprints. Similarly, $|Y \cup \hat{Y}|$ calculates the area of the region where either the truth or FootNet predictions show non-zero footprints."