The authors used their own deep learning model, Res2-Unet+, to generate the first high-resolution 1 (1-2m) dataset of storage tanks from over 4000 images and assign the year of each storage tank. The dataset is then used for CH4 emission analysis. The dataset is useful and the analysis is reasonable. It is suitable for publication in ESSD. However, I have the following comments:

Q1. Abstract Line 30 "based on high spatial resolution images (1-2m) …". I suggest the authors add the number of images.
"based on over 4000 high spatial resolution images (1-2m)… "
Answer: Thanks very much for your suggestion. We have revised the corresponding sentence to be 'we generated a storage tank dataset (STD) by implementing a deep learning model with manual refinement based on 4,403 high spatial resolution images (1-2m)' in **Line 28-30** of the revised manuscript.

Q2. 3.3 Land use land cover product. Line 263-264. "historical high spatial resolution images, high spatial resolution images collected, and field survey from Google Earth", please refine the sentence. from my understanding, only historical high spatial resolution images are from Google Earth.
Answer: Thank you for your suggestion. We have revised the sentence correspondingly.

Q3. Section 4.2.1
Line 309 "Res2-Unet+ by Yu et al. (Yu et al., 2021)". "by Yu et al." should be removed. The same issue is in line 352.
Line 329 "Our proposed Res2-UnetA", does it mean Res2-Unet+?
equation (1)-(7), what do "f, m, n, h, etc" stand for?
Answer: Thanks very much for your suggestions. We have removed in **Line 312** of the revised manuscript. Res2-UnetA is newly proposed in this work, stemming from Res2-Unet+, as illustrated in **Line 311-313**.

In terms of Equations (1)-(7), *f* indicates feature map, *m* and *n* are the size of feature map *f*, *h* is the channel number of feature map *f*. Sorry for the missing introduction. We have revised the corresponding sentences in **Line 337-348** as follows:

'Detailed calculation of channel-wise and spatial attention modules can be referred to Equations (1)-(7). Spatial average pooling (sa) and spatial maximum pooling (sm) operations are calculated as the average value and maximum value of input feature map *f* with size of $m \times n$, as described in Equations (1)-(2). Correspondingly, the channel-wise average (ca) and maximum pooling (cm) operations are the average feature values of all the *h* channels and the maximum feature values of all the channels in Equations (3)-(4). The output feature map of the spatial attention module (SA) and channel attention module (CA) are calculated according to Equations (5)-(6), respectively, and the synthesis of the feature maps from the channel and spatial attention modules is organized by multiplication, as illustrated in Equation (7).'

Q4. Section 4.3 The STD dataset covers 2000-2021. How do you define the year if the storage tank was built before 2000, and how many such cases are there?

Answer: Due to the limited accessibility of high spatial resolution images before 2000 from Google Earth, the storage tanks with the record of first year image with storage tank, but without that of last year image without storage tank in our proposed dataset STD are possibly constructed before the year of 2000. Upon checking the storage tanks in our dataset, there are 188 storage tanks lacking the record of last year image without storage tanks due to the lacking of continuous historical high spatial resolution images from Google Earth, which are possibly constructed in year before 2000.
To make the manuscript easier to understand, we have added the corresponding sentence in Line 3396-399 as 'For the storage tanks built before 2000, they are recorded with the first year image with storage tank in the shapefile, but lacking the last year image without storage tank in our proposed dataset STD due to the limited accessibility of high spatial resolution images before 2000 from Google Earth'

Q5. Section 5.1 Line 404-405. "It may be seen that 404 storage tanks of 500-1000 m2 are more than those of larger sizes". Then, what are the criteria for threshold 500m^2?
Answer: The threshold 500m^2 was selected by experience, given that large storage tanks may emit significant levels of CH_4.

Q6. Section 5.2 What is the year for density calculation in Fig 9?
Answer: The storage density calculation in response to different CH_4 emissions in the atmosphere is explored in year of 2020. We have added the corresponding illustration in **Line 444-445** as 'we explored the spatial consistency between estimated CH4 from energy emission products in year of 2020 and the density of storage tanks in our proposed dataset STD over the study area.'

Q7. Section 6.1
NEPU-OWOD dataset, an oil storage tank dataset from platform Kaggle, and the STD dataset.
They should be separated into different paragraphs.
Answer: Thanks very much for your suggestion. We have separated the illustration of our proposed STD dataset over the published works of NEPU-OWOD V1.0 dataset, the Oil and Gas Tank Dataset, and the oil storage tank dataset from platform Kaggle into two paragraphs in the revised manuscript.

Q8. References. The references need to revised in uniform format. Some references are missing. Some references have doi.
Answer: Sorry for the uniform format of references. We have checked and revised all the references to maintain consistency.