#Referee 1

Thanks the reviewer for the valuable comments. We have revised the manuscript accordingly, with a point-by-point reply to the comments, and a marked-up manuscript version showing the changes made (text in red).

- 1. The transferability of LUR models across areas or cities varies greatly, given the large regional variability of predictive performances of LUR models. The authors are encouraged to strengthen the motivation why they developed one model to predict the air quality in northern Taiwan, or highlight/discuss the strengths of their LUR model, compared with previously established model.
  Response: In our original manuscript, we did include the motivation part (Line 57-68 and Line 15-17).

  P2, Line 57-68: "**In addition, most previous Taiwan LUR studies used data from purpose-designed monitoring networks or combined purpose-designed and routine monitoring networks (Ho et al., 2015; Lee et al., 2014; Lee et al., 2015). …As a result, a general limitation of LUR models upon purpose-designed monitoring networks is that the established models may only reflect the situation the measurement period (Hoek et al., 2008; Shi et al., 2020). Therefore, the development of long-term average LUR models for specific air pollutants using only routine monitoring networks should be explored, which is especially critical for epidemiological studies.**"
  P1, Line 15-17: "**To provide long-term air pollutant exposure estimates for epidemiological studies, it is essential to test the feasibility of developing land-use regression (LUR) models using only routine air quality measurement data and to evaluate the transferability of LUR models between nearby cities.**"

- 2. L33-34: The health effect of aerosol is not adequately cited since air pollution has been well recognized to adversely affect cardiovascular diseases. The authors are suggested to consider citing Sun et al. 2011 (doi: 10.1161/CIRCULATIONAHA.109.893461); Yin et al. 2020 (doi: 10.1021/acs.estlett.9b00735)
  Response: We added these two articles as references (Line 34-35, Line 507-508, and Line 553-555).

  P2, Line 34-35: "**…such as lung function, and respiratory and cardiovascular diseases (Çapraz et al., 2017; Sun et al., 2010; Yin et al., 2020; Zhou et al., 2020).**"

  P16, Line 507-508: **Sun, Q., Hong, X. and Wold, L.E.: Cardiovascular effects of ambient particulate air pollution exposure. Circulation 121(25), 2755-2765, 2010.**

  P17, Line 553-555: **Yin, P., Guo, J., Wang, L., Fan, W., Lu, F., Guo, M., Moreno, S.B., Wang, Y., Wang, H., Zhou, M. and Dong, Z.: Higher risk of cardiovascular disease associated with smaller size-fractioned particulate matter. Environ. Sci. Technol. Lett. 7(2), 95-101, 2020.**

- 3. L39: "estimating" -> "estimate"
  Response: Revised as suggested.

  P2, Line 41: "**… estimate population exposure…**"

- 4. L213: "Traffic emission is a major source of air pollution in urban areas of the TKMA (Lee et al., 2014; Wu et al., 2017)." Please be more specific regarding the contribution of traffic emission to air pollution in TKMA, e.g., what is the percentage?
  Response: We did include the contribution percentage of traffic emission to air pollution ($PM_{2.5}$ in the cited study) in TKMA (Line 216-218).

  P7, Line 216-218: "**For instance, it was reported that gasoline and diesel vehicle emissions contributed approximately half of $PM_{2.5}$ concentrations in Taipei City based on source apportionment analysis (Ho et al., 2018).**"

- 5. L317-320: I am confused with the logic that the weak correlation between air pollution LUR model derived results and nearby-station measurements (Figure 6), makes the author believe in the

notion that thereby air pollution LUR models may provide more accurate exposure estimates than nearby-station measurements. Please clarify it.

Response: We have corrected our text to make the meaning clearer (Line 323-328).

P11, Line 323-328: "**A possible explanation is that LUR-model-based exposure estimates generally accounted for neighbourhood-scale variations of air pollutant concentrations, while the nearby-station measurements usually only revealed the urban-scale variability of air pollution (e.g., urban area versus suburban area versus rural area) (Marshall et al., 2008). The LUR-model-based exposure estimates and nearby-station measurements should be further validated if the air quality measurement data at residential locations of cohort participants (if not all, at least some of the participants) are available.**"