Responses to reviewers:

We thank the two anonymous reviewers for their helpful comments and suggestions. We have revised the manuscript following the comments as described below. Reviewer comments are shown in blue. Our responses are shown in black. The revised texts are shown in italics.

Reviewer #1

(1) Treatment of uncertainty: aside from several brief mentions of NMB or similar metrics, there is little discussion of model uncertainties. For example, uncertainties in emission inventories surely propagate into derived O₃ partitioning. In particular, it would be helpful to estimate the confidence in the values shown in Fig. 9, which is the key result of the paper.

Thanks for the comment. We added discussions on uncertainties of (1) emission inventory, (2) chemical mechanism, and (3) biogenic VOC measurements

(1) Emission inventory

Anthropogenic emissions were obtained from the Multi-resolution Emission Inventory for China (Li et al., 2017). Emission estimates from bottom-up inventories are uncertain due to lack of complete knowledge of human activities and emission from different sources. The overall uncertainties of the bottom-up SO₂, NOₓ, NH₃, VOC, and PM₂.₅ emission estimates are ±31%, ±37%, ±153%, ±78% and ±133%, respectively, with 95% confidence intervals (CI) (Zhang et al., 2009; Lei et al., 2011; Zhao et al., 2011; Lu et al., 2011; Kurokawa et al., 2013). We updated Table S1 and added text in the Section 2.3 of the revised manuscript to describe this issue.

The emission estimates and uncertainties of VOCs, SO₂, NOₓ, NH₃, and PM₂.₅ in the domain during the simulation period are summarized in Table S1.

(2) Chemical mechanism

The uncertainty of chemical mechanism is stated in the comment (4).

(3) Biogenic VOCs

The uncertainties of biogenic VOC observation and simulation are stated in the comment (7). In addition, we added discussion about sub-month variability of the simulated biogenic VOCs in Section 3.2 in the revised manuscript.

We also analyzed the temporal variation of simulated biogenic VOC during the whole simulation period and found the sub-month variability was relatively small (the standard deviation < 25%).

(2) NOₓ measurements: the “NOₓ” measurement uses a hot molybdenum converter, which is known to convert a lot more than NO₂ (probably most of NO₃). This point may or may not impact the model measurement comparison shown in Fig.
3 depending on the NO\textsubscript{y} partitioning, but it should be acknowledged and addressed.

Thank you for pointing this out. The observed NO\textsubscript{x} (NO+NO\textsubscript{2}) concentration in the center of Xi’an was measured by a gas-phase chemiluminescence detection analyzer. The principle of this method is converting ambient NO\textsubscript{2} to NO on the hot surface of a molybdenum oxide (MoO) catalyst followed by a chemiluminescence detection of NO. The conversion of the thermal catalyst is not specific to NO\textsubscript{2}, but also other nitrogen-containing components (NO\textsubscript{z}: NO\textsubscript{z}=NO\textsubscript{y}-NO\textsubscript{x}, including PANs, HNO\textsubscript{3}, HO\textsubscript{2}NO\textsubscript{2}, HONO, RO\textsubscript{2}NO\textsubscript{2}, and organic nitrates), resulting in a higher NO\textsubscript{2} concentration. However, despite its drawbacks, it remains a widely used method in air-quality monitoring networks and research, owing in part to its low cost, sensitive detection, and ease of operation.

Xu et al. (2013) evaluated this potential uncertainty of the MoO converter at four different sites in China, and found that the MoO converter worked well at the urban site (overestimation less than 10%), but overestimated NO\textsubscript{2} more than 30% at suburban and background sites. In the Guanzhong basin, we used the simulated NO\textsubscript{y} to evaluate the uncertainty of the NO\textsubscript{x} measurements, due to the lack of referenced NO\textsubscript{y} or NO\textsubscript{x} observation. The calculated results indicated that the NO\textsubscript{z} accounted for 11% of the total NO\textsubscript{y} at the urban Xi’an during the no-raining period of Fig. 3. We noted the uncertainty in the NO\textsubscript{x} measurements, but thought this would not crucially impact the model-measurement comparison.

We modified the text in Section 2.1.2 and 3.3 in the revised manuscript to describe this issue.

Section 2.1.2

NO\textsubscript{x} was measured by a gas-phase chemiluminescence detection analyzer EC9841, coupled with a hot molybdenum converter.

Section 3.3

It is worth noting that the observed NO\textsubscript{x} were detected by a chemiluminescence analyzer coupled with a hot molybdenum converter (MoO), and this method was recognized to cause higher NO\textsubscript{2} concentration due to the positive interference of other nitrogen-containing components (NO\textsubscript{z}, e.g. PAN, HNO\textsubscript{3} and HONO). Xu et al. (2013) found that the uncertainty caused by the MoO converter was much smaller at urban sites (less than 10%) than that at suburban and background sites (more than 30%). In the GZ basin, to evaluate the uncertainty, we estimated the ratio of NO\textsubscript{z}/(NO\textsubscript{x}+NO\textsubscript{z}) from the model. The calculated results indicated that the NO\textsubscript{z} accounted for 11% of the total NO\textsubscript{x}+NO\textsubscript{z} at urban Xi’an during the no-raining period. We noted the uncertainty in our NO\textsubscript{x} measurements, but thought this would not crucially impact the model-measurement comparison.

(3) Sect. 2.2.1: It should be made clear here that only 6 samples total were collected.

Thanks for the suggestion. We modified the text in Section 2.1.1 in the revised manuscript as follows.

We selected six field sites in the Qinling Mountains (Fig. 1b, the triangles) and collected one ambient air sample at each site on 6\textsuperscript{th} – 7\textsuperscript{th} August 2011 under sunny
weather conditions (details are presented in Table 1).

(4) Page 7, Line 5: This mechanism seems outdated given the recent leaps in understanding of isoprene chemistry. Of particular concern is treatment of alkyl nitrates (which can be temporary or permanent NO\textsubscript{x} sinks, depending on the mechanism) and the general assumed fate of Isoprene RO\textsubscript{2} radicals. The authors should consider either 1) justifying why this does not impact their results, or 2) assessing potential uncertainties arising from use of an outdated mechanism.

Thank you for pointing this out. Isoprene is the most important NMVOC (Guenther et al., 2006), affecting tropospheric O\textsubscript{3}, OH and aerosols in complex ways. Being highly reactive, isoprene has a lifetime of only ~3h at OH concentration of 1 × 10\textsuperscript{6} mol cm\textsuperscript{-3}. In the daytime, isoprene reacts with OH radical to form hydroperoxy radicals (ISOPO\textsubscript{2}). Subsequently, in the presence of NO\textsubscript{x}, ISOPO\textsubscript{2} reacts with NO leading to the production of hydroxynitrates (ISOPN) by a minor branch, which sequester NO\textsubscript{x} and therefore regulate O\textsubscript{3} formation locally. With the presence of a double bond, ISOPNs are highly reactive with respect to OH and O\textsubscript{3}. The products from ISOPN oxidation might either release NO\textsubscript{x} or generate secondary organic nitrates. Some of the secondary organic nitrates products are found to be considerably longer-lived than ISOPN and can serve as temporary reservoirs for NO\textsubscript{x} (Paulot et al., 2009a, b; 2012). Besides daytime oxidation, nighttime oxidation of isoprene by NO\textsubscript{3} contributes significantly to the budget of organic nitrates (Horowitz et al., 2007; von Kuhlmann et al., 2004; Xie et al., 2013).

Many laboratory, filed observations and simulation studies (e.g. von Kuhlmann et al., 2004; Perring et al., 2009; Paulot et al., 2009a) agreed that tropospheric O\textsubscript{3} production was highly sensitive to the ISOPN yield (4.4% to 15%) and NO\textsubscript{x} recycling efficiency (3% to 50%). However, large uncertainties still remain in estimating these parameters. Horowitz et al. (2007) found a 4% ISOPN yield best captured the alkyl and multifunctional nitrates measured by aircraft, and unreasonably high ISOPN yield (18%) would let ISOPN to be a terminal sink for NO\textsubscript{x} (Hudman et al., 2009).

Some recent modeling studies evaluated the sensitivity of tropospheric O\textsubscript{3} to isoprene oxidation chemistry. Xie et al. (2013) incorporated recent advances in isoprene oxidation chemistry in CMAQ (including a more explicit representation of isoprene nitrate formation from OH/NO and NO\textsubscript{3} pathways as well as modification to the isoprene chemistry under low-NO\textsubscript{x} conditions) and found the model can capture observations by assuming the ISOPN yield of 6% to 12%. O\textsubscript{3} increased by 2 ppbv in the eastern U.S as a result of these updates. Also focusing on the eastern U.S., Mao et al. (2013) implemented an updated isoprene chemistry mechanism in GEOS-Chem and found O\textsubscript{3} increased by 3-5 ppbv.

We added the discussion above in Section 2.2 to state the uncertainty.

We noted some advances in isoprene nitrate chemistry in recent year. A number of laboratory, filed observation and simulation studies (e.g. Paulot et al., 2009a, b, 2012; Horowitz et al., 2007; Hudman et al., 2009) highlighted the importance of the yield of isoprene nitrate (4-15%), the NO\textsubscript{x} recycling efficiency (3-50%) and the representation of secondary nitrate photochemistry for simulation of tropical O\textsubscript{3}. They all agreed there were still large uncertainties in isoprene nitrate chemistry. Some recent modeling studies evaluated the sensitivity of tropospheric O\textsubscript{3} to isoprene
oxidation chemistry. Xie et al. (2013) incorporated recent advances in isoprene oxidation chemistry in CMAQ and found the model can capture observations by assuming the isoprene nitrate yield of 6% to 12%. Simulated $O_3$ increased by 2 ppbv in the eastern U.S. as a result of these updates. Also focusing on the eastern U.S., Mao et al. (2013) implemented an updated isoprene chemistry mechanism in GEOS-Chem and found $O_3$ increased by 3-5 ppbv.

(5) Page 9, Line 5: “Brute force” does not, to the reviewer’s knowledge, refer a specific method of source characterization. Please refine or define.

Thank you for the comment. We defined the “Brute force” method in Section 2.4 in the revised manuscript as follows.

The approach referred to as the “brute-force” method (sensitivity analysis used to measure the model output response to emission changes) is traditionally used in air quality model to identify source contributions from specific non-reactive species in a linear process....

(6) Page 10, Line 22: Does the model use assimilated meteorology? If so and the model is being nudged with observations, this agreement may not be especially remarkable.

We appreciate the reviewer’s concern. We did not include assimilated meteorology in this study.

(7) Sect. 3.2: There are only 6 observations here, and looking at Table 1 the model seems to have little skill in capturing the variability of those. It would be worthwhile to point this out and justify why that’s not a big deal for the present analysis.

We agree that our model didn’t well capture the variability of the observed BVOCs. We think it is mainly because that, in our 3x3 km grid, different terpene emitters are not homogeneously distributed in the grid and the point measurements are influenced by the microenvironment and meteorology (Zare et al., 2012; Kota et al., 2015).

However, in this study, our goal is estimating the biogenic effect on $O_3$ formation in an urban area 50 km away from the foothills. We pay more attention to the magnitude of BVOC concentration at the regional scale, instead of capturing the hourly- and microenvironment-scale variability in either the observation or the simulation. Therefore, we think it is reasonable to compare the averaged simulations of 6 grids with the measurements.

We modified the text in Section 3.2 of the revised manuscript to state the uncertainty and to describe the above rationale.

In general, different terpene emitters are not homogeneously distributed in a kilometer-scale grid and the point measurements are influenced by the microenvironment and meteorology (Zare et al., 2012; Kota et al., 2015). However, in this study, our goal is to estimate the biogenic effects on urban $O_3$ 50 km away from
the foothills, which requests more concern on the regional scale VOC level, rather than the microenvironment-scale variability in either the observation or the simulation. Thus, we compared the average of VOC measurements with model simulations to validate whether the calculated results were reasonable. The isoprene mean concentration simulated in the six grids (corresponding to the time of observations) was 1.4 ppb, which is close to the observed average value of 1.3 ppb at the six sampling sites. Monoterpenes performed quite similarly, simulated 0.22 ppb comparing with observed 0.21 ppb. We also analyzed the temporal variation of simulated biogenic VOC during the whole simulation period and found the sub-month variability was relatively small (the standard deviation < 25%). The evaluation indicates that biogenic VOCs simulations reasonably agreed with the observations in the Qinling Mountains, on average, which provides a basis for us to further evaluate biogenic effects on O3.

(8) Page 11, Line 30: “agree well” seems optimistic. Looking at Fig. 4, 5 of the 11 non-rain days show model measurement disagreement by more than a factor of 2, and it is not evident that the model captures observed day-to-day variability.

Thank you for pointing this out. We modified text in Section 3.3 in the revised manuscript to make our statements more precise.

The model predicted PM$_{2.5}$ concentration to be 94.6±28.2 μg m$^{-3}$, slightly lower (NMB=-12%) than measured 107 μg m$^{-3}$ averaged for the no-raining period, but didn’t perform well in capturing temporal correlation (r=0.17). The simulated PM$_{2.5}$ showed the similar compositions to the observation (Fig. S1b and S1c)....

(9) Page 13, Line 7: Is this JNO$_2$ calculation based on model output? Please clarify.

Yes, the J(NO$_2$) is calculated by model output “photor_no2”. We added the text in Section 4.1 in the revised manuscript as follows.

Figure 6 shows the changes of J(NO$_2$) (calculated by model track output photor_no2) with the participation of PM averaged for urban Xi’an.

(10) Page 13, Line 15: presumably, aerosol lifetime is also longer than NO$_x$.

Thank you for the suggestion. We added this in Section 4.1 in the revised manuscript.

The spatial distribution of high-values of PM$_{2.5}$ was similar to that of NO$_x$, but covered a wider area mostly in the downwind region of urban Xi’an, which is expected due to longer lifetime of aerosols compared with NO$_x$ and the time required for secondary aerosol formation, thus further dispersion.

(11) Page 15, Line 18: Calling this the “original state” is confusing. If it is the anthropogenic simulation, just refer to it as such.

Thank you for the comment. We modified the text in Section 4.3 in the revised manuscript.
manuscript to make our statements clearer.

To make it more specific, we started the discussion from the result of ANTH simulation without the biogenic sources (Table 2). Figure 10b shows the O₃ production regime in the ANTH simulation.

(12) Page 16, Line 10: It is not clear that the discussion of the “actual” contributions adds much to the message of the paper. Indeed, it is a bit confusing because “pure” and “actual” have similar connotations, and because it convolutes several of the separated contributions.

Thank you for the comment. We deleted the discussion of the actual contributions to avoid ambiguity.

(13) Sect. 4.3: The lack of a “synergistic” contribution to PM is noteworthy and may deserve a few more sentences of discussion, especially given its importance for air quality. In other regions (e.g. the SE US), there seem to be relatively strong links between anthro/bio emissions and PM.

Thank you for the suggestion. We added the text in Section 4.3 in the revised manuscript to discuss this issue.

It is worth noting that the biogenic contribution to PM₂.₅ is not obvious (less than 3%) in GZ basin, which might be different from some other regions (e.g. Fu et al., 2012; Li et al., 2013). The main reasons are that 1) organic matter, the most important biogenic PM₂.₅ component, only accounted for 14-16% of PM₂.₅ in GZ basin in August; 2) Undeniably, uncertainties still exist in organic matter simulations in the model.

(14) Table 2: Not sure this is necessary for the story. Could be moved to supplement. Your call.

Thanks for the suggestion. We moved this table to the supplement of the revised manuscript.

(15) Page 12, Line 14: discuss
Thanks. Corrected

(16) Page 16, Line 20: delete “and discussion”
Thanks. Corrected

(17) Fig. 12: suggest modifying colorbar to better separate NOx vs VOC controlled regions (e.g. blue-whitered gradient) and mentioning the limits for each regime in the caption.
Thank you for the suggestion. We modified the colorbar of Fig. 10 in the revised manuscript.

**Figure 10.** The monthly mean ratio of H$_2$O$_2$/HNO$_3$ during the daily O$_3$ peak time (14:00-18:00 local time) in August 2011 in the (a) Base simulation and (b) the simulation without biogenic sources.

**Reviewer #2**

1. My major concern with the work comes from the model uncertainty analysis, particularly in the simulation of VOCs and in emission inputs. In my opinion, six off-line VOC measurement points in space and time are not sufficient to validate the model accuracy for the following ozone source apportionment analysis. In addition, an agreement in Aug. 6 & 7 would not necessary indicate an agreement in the rest of month August. The current state of Section 3.2 is insufficient. More material is needed to show and quantify the model uncertainty in VOC estimation for the area under study. The uncertainty in emission input is needed as it is one of the main sources of model uncertainties.

   Thanks for the comment. We added discussions about uncertainties of emission inventory and biogenic VOC simulations.

1. Biogenic VOC

   We agree that only six off-line measurement points are not sufficient to describe the spatial and temporal variability of BVOCs. However, in this study, our goal is estimating the biogenic effect on O$_3$ formation in urban Xi’an. We pay more attention to the magnitude of BVOC concentration at the regional scale, instead of capturing the hourly- and microenvironment-scale variability in either the observation or the simulation.

   To verify the representativeness of BVOC observation (in Aug. 6 & 7), we analyzed the magnitude and temporal variability of simulated BVOC in the Qinling Mountains for the whole simulation period. The results indicated that the mean BVOC (isoprene + monoterpenes) concentration is 1.73 ppb, close to the observed value (1.51 ppb) during Aug 6 & 7, and the sub-month variability (characterized by
the standard deviation of daily mean simulation (averaged for 08:00-16:00 local time) were <25%, indicating that the sub-month variability was relatively small.

We modified the text in the Section 3.2 of the revised manuscript to state the uncertainty.

In general, different terpene emitters are not homogeneously distributed in a kilometer-scale grid and the point measurements are influenced by the microenvironment and meteorology (Zare et al., 2012; Kota et al., 2015). However, in this study, our goal is to estimate the biogenic effects on urban O3 50 km away from the foothills, which requests more concern on the regional scale VOC level, rather than the microenvironment-scale variability in either the observation or the simulation. Thus, we compared the average of VOC measurements with model simulations to validate whether the calculated results were reasonable. The isoprene mean concentration simulated in the six grids (corresponding to the time of observations) was 1.4 ppb, which is close to the observed average value of 1.3 ppb at the six sampling sites. Monoterpenes performed quite similarly, simulated 0.22 ppb comparing with observed 0.21 ppb. We also analyzed the temporal variation of simulated biogenic VOC during the whole simulation period and found the sub-month variability was relatively small (the standard deviation < 25%). The evaluation indicates that biogenic VOCs simulations reasonably agreed with the observations in the Qinling Mountains, on average, which provides a basis for us to further evaluate biogenic effects on O3.

(2) Emission inventory

Anthropogenic emissions were obtained from the Multi-resolution Emission Inventory for China (Li et al., 2017). Emission estimates from bottom-up inventories are uncertain due to lack of complete knowledge of human activities and emission from different sources. The overall uncertainties of the bottom-up SO2, NOx, NH3, VOC, and PM2.5 emission estimates are ±31%, ±37%, ±153%, ±78% and ±133%, respectively, with 95% confidence intervals (CI) (Zhang et al., 2009; Lei et al., 2011; Zhao et al., 2011; Lu et al., 2011; Kurokawa et al., 2013). We updated Table S1 and added text in Section 2.3 of the revised manuscript to describe this issue.

The emission estimates and uncertainties of VOCs, SO2, NOx, NH3, and PM2.5 in the domain during the simulation period are summarized in Table S1.

(2) Another minor comment is with the application of brute-force comparison method in assessing the contributions of different emission sources to ozone concentration. The brute-force method has inherent disadvantages when applied to secondary species such as ozone and PM due to the non-linearity in responses. A critical question I would expect the authors to discuss in the manuscript is about the difference between actual and pure contribution from anthropogenic and biogenic.

The sensitivity analysis (brute-force method) used to measure the model output response to emission changes, is traditionally used in air quality model to identify source contributions from non-reactive species in a linear process. However, as the reviewer mentioned, it is not proper to secondary species such as O3 and PM, because secondary species are generated nonlinearly, and interactions between different
sources cannot be ignored. The difference between the actual and pure contribution from anthropogenic or biogenic source is the synergistic effect between the two sources. In this work, we combined the brute-force method and Factor Separation Technique (FST) (section 2.4) to discuss the importance of synergistic effect on O$_3$ formation. We modified the text in the Section 2.4 in the revised manuscript to make our statement clearer.

O$_3$ is formed by complicated nonlinear reactions of anthropogenic and biogenic precursors (NO$_x$ and VOCs) in the presence of sunlight. The approach referred to as the “brute-force” method (sensitivity analysis used to measure the model output response to emission changes) is traditionally used in air quality model to identify source contributions from specific non-reactive species in a linear process, but it cannot straightforward apply to secondary species due to the non-linearity in responses. In practice, the actual impact of one factor in a nonlinear process in the presence of others can be separated into 1) pure impact from the factor and 2) interactional impacts from all those factors. In this study, we adopted the factor separation approach (FSA) (Stein and Alpert, 1993) to decompose the pure contribution of a factor from its interaction with other factors.

(3) In Fig. 7, is the black line based on theoretical calculation or model simulation? It seems to me that the smooth black line is based on the theoretical calculation for clear sky condition. If it was true, then I doubt if the decrease (red line) includes both the cloud and aerosol effects.

The black line is based on model simulated output “photor_no2”. We modified the text in Section 4.1 in the revised manuscript to make this issue clear.

Figure 6 shows the changes of J(NO$_2$) (calculated by model track output photor_no2) under the participation of PM averaged for urban Xi’an.

(4) In Section 3.2, is the simulated isoprene/monoterpene mean concentration the mean over the one-month simulation periods or Aug. 6-7th or the mean of six data points corresponding to the time and location of the six measurements?

The simulated isoprene and monoterpene concentrations in Section 3.2 were the mean of six data points corresponding to the time and location of the six measurements. We modified the text in Section 3.2 in the revised manuscript to make the statement clearer.

The isoprene mean concentration simulated in the six grids (corresponding to the time of observations) was 1.4 ppb, which is close to the observed average value of 1.3 ppb at the six sampling sites.

(5) Why modeled wind speed is discontinuous in Fig. 3c?

We appreciate the reviewer's concern. Figure 3c is presented as a “feather picture”, in which arrow direction indicates wind direction and length of arrow indicates wind speed. We give a line graph as follow to show wind speed more clearly.
Reference


