Response of the authors to comments by Paul Laris on the manuscript: "Dynamic savanna burning emission factors based on satellite data using a machine learning approach"

Roland Vernooij (corresponding author) on behalf of the authors:

We sincerely thank Paul Laris for taking the time and effort to read and comment on our manuscript, and the detailed and constructive comments both on this platform and in earlier conversations, which helped to improve the quality of this paper. Please find below our point-to-point response to the review. The revised text and updated figures are included in the updated manuscript. A separate 'track-changes' document is included to highlight the changes we made to the manuscript.

Detailed comments	**Author's response, reasoning and comments**
Can you clarify that these fires were all **"head" fires** as opposed to backfires? And, if so, can you comment on why the following dimensions are adequate? We question wether 10m is wide enough for head fires to fully develop. This width is fine for backfires. Also, if only head fires were examined, can you justify given that many fires are purposefully set as backfires in Africa. Headfires have long been used in research on African fires, yet research finds more backfires are set.	While it is indeed correct that most of the measurements have been taken during 'headfires' and 'sideward propagating' fires we also measured backfires. We tried to obtain measurements proportionately to the area burned by the different types within these fires. However, we did also conduct measurements where we tried to distinguish the different fire propagation directions, which in a changing wind regime was much more challenging than we previously anticipated. Individual 35-second bag samples more often than not contained smoke from multiple changing wind directions. In agreement with the findings by Wooster et al. (2011) and Laris et al. (2021) we found "back" fire samples to have slightly higher combustion completeness compared to "head" fire samples. A possible explanation being that slower lofting smoke from the residual smouldering does not mix with the flaming emissions in these measurements, like it does in head fires. There was no significant difference between head and sideward propagating fires. The early dry season fires were all lit by land managers under the guidance of prescribed burning experts. This meant that head fires were only used if the conditions allowed them (which was most often the case), to prevent runaway fires. Although these experts deemed the measured EDS fires representative of prescribed fires, you

	are correct that pure backfires may burn more efficiently. To clarify this in the text, we have added the following in (P3 L32): “Fires were lit with the aim of being representative of early dry season (EDS, often prescribed) fires and late dry season (LDS) non-prescribed fires. Although some backfires were sampled during the initial phase of the fires, the majority of samples were obtained from the faster ‘head’ fires, which consumed most of the biomass. Fire sizes generally ranged between 2 to 10 hectares based on UAS drone imagery described by Eames et al. (2021), with exceptions of some fires that would not light and conversely, some fires that burned several hundred hectares. In the EDS, fire size was primarily limited by environmental conditions and fires ceased burning as humidity increased overnight whereas in the LDS, fire size was confined by low-fuel areas like burn scars, roads and prepared fire breaks. Particularly in the LDS, this means a limited fire size does not necessarily indicate limited fire intensity. Emissions were sampled at altitudes between 5−50 m depending on flame height for a duration of 35 seconds, resulting in 0.7 litres per gas sample. On average, we took 35 samples per fire. The sampling methodology involved taking samples from a fire passing a certain point −while correcting for wind direction and severity− until no more visual smoke passed the drone anymore. From earlier work (Vernooij et al., 2022a), where we compared the average of these measurements to results using continuous measurements taken at a mast, we have some confidence in the fidelity of this approach.” Also, in the discussion (P14 L19) we added: “The samples were predominantly collected over “head” fires, which in the measured fires typically represented most of the burned area. A common approach for prescribed fires is burning against the wind

	(backburning), to minimise both the impact on vegetation and risk of spread. In accordance with Wooster et al. (2011) and Laris et al. (2021), we found higher MCE in samples from backfires, which indicates these types of fires may emit less CH_4 and CO. Another possible explanation is that slower lofting RSC smoke does not mix with the flaming combustion emissions in these measurements, like it does in "head" fires."
You do not appear to have published local or ground data on weather conditions. While T and H can be collected from regional weather stations, **wind speed is critical to determining fire intensity and will influence MCE as well.** Do you have wind data, it would seem critical for accurate fire intensity and MCE results.	Unfortunately, we have only started to log windspeed (from a Kestrel 5500FW Fire Weather Meter) in the very last campaigns. We agree that windspeed is most likely a more significant predictor that the models suggest based on the ERA5-land data. Note that although WS is not often seen as a major predictor, the FWI which contains WS is. While it would be very interesting to verify the windspeed from ERA5-land with the on-site windspeed, more accurate on-site windspeed measurements could not be used for the spatiotemporal extrapolation, and therefore would not improve the model.
I wonder about this comment: " The grasslands with the highest EFs (found in high-rainfall savanna Dambos) were **"uncharacteristically green for the time of the season"** given that many fires are set to "green" grasses in African savannas, especially the perennials (See Le Page who documented this back in 2010 as well as many West African case studies).	The vegetation we refer to with this comment was highly limited in its spatial extent to relatively deep and clayey Dambos with widths often smaller than a 500-meter MODIS pixel. Since the water availability and grass curing state in these areas is highly dependent on soil type and geomorphology, these characteristics are poorly captured by the much coarser seasonal features (e.g. soil moisture and VPD). The Dambos where we measured the highest EFs (also in the LDS) had just fallen dry and were still very green, whereas other Dambos close by had already fully cured and showed very low EFs. By this statement we mean that because of the dominant role of soil type and geomorphology, the EFs measured in those Dambos were a poor indicator of the seasonal cycle in other grasslands.

	We added the following text to the discussion (P14 L36): "Although burning grasslands under green conditions releases more CH_4 and CO, there are valid reasons to do so. For example to remove moribund grass that remains after the dry season with minimal damage to the grass sward (Nieman et al., 2021; Le Page et al., 2010). In its current form, the model may not always pick up on those landscape features."
I think Laris et al found very similar results to: "The strongest predictors for the MCE and the CO and CH4 EF were the **tree density in the plots, the grass to litter ratio, the combustion completeness** and the **moisture** content of the consumed fuel. It might be useful to compare and to consider the hypothesis that burning of green leaves on shrubs and trees vs. dried leaves on the ground may explain why EF CH4 is not linerally related to MCE. This reasoning may also explain the following finding, "For CO and CH4, the dominant effect is a spatial redistribution with higher CO and CH4 EFs in mesic, high-tree cover savannas and lower EFs in xeric savannas compared to previous estimates. The Higher CH4 EF in mesic may well be a function of leaf burning. This is logical given the findings from Senegal research by Barker finding burning trees emitted smoke with the highest methane EF. This needs further explanation: "Although CO and CH4 followed the same spatial pattern, we found that MCE affected the CH4 which resulted in lower CH4 to MCE ratios in open **(lower tree density?) savannas…. Do you mean higher CH4/MCE in wooded savannas as compared to grass-dominated ones? What is "open"? Clarify.** Again, see works in Mali and Senegal which agree with this finding.	We indeed find higher CH_4/MCE ratios in tree-dominated savannas compared to grass-dominated fires. In our previous work on isotopes, we found CH_4 EFs to be more ^{13}C depleted compared to CO emissions when burning wooden logs. This may indicate CH_4 is more RSC driven than CO and possibly stronger dominated by the pyrolysis of lignin rather than cellulose and hemicellulose. In P12 L37 we added the following text: "In accordance with previous studies (e.g. Korontzi et al., 2003b; van Leeuwen and van der Werf, 2011; Barker et al., 2020), we found steeper CH_4 EF to MCE regression slopes in woodlands compared to grasslands. Our data indicated a positive correlation of the CH_4 EF to MCE slope with the FTC based on MOD44Bv006. The MCE is a simplified form of the combustion efficiency and only calculated using CO and CO_2 emissions. Being less oxidized than CO (which is still common in flaming combustion), CH_4 emissions have a stronger dependency on the actual combustion efficiency (CO_2 divided by all carbon emissions). While most studies describe the relationship between the CH_4 EF and the MCE as being linear (Korontzi et al., 2003; van Leeuwen and van der Werf, 2011; Selimovic et al., 2018; Yokelson et al., 2003), we found that for individual bag samples it was better described using a nonlinear function (Fig. 9), in line with findings by Meyer et al. (2012) for Australian savanna measurements. Figure 9 represents

	individual bag measurements rather than fire averages (for which the spread in MCE is much lower). Laboratory experiments described by Selimovic et al. (2018) showed that the CH_4 to CO ratio is strongly dependent on flaming or smouldering phases if the fire. Individual bag samples –which often hold emission from a single phase– therefore show much more variation compared to fire averages. Stable carbon isotopes also point to CH_4 emissions being more depleted in heavy carbon (^{13}C) compared to CO in both mixed (C3 and C4) and single-fuel-type experiments, indicating a stronger dominance of RSC and the pyrolysis of lignin in its total emissions (Vernooij et al. 2022b). This explains both why studies that are skewed towards either smouldering or flaming phase emissions find different CH_4 EF to MCE slopes using linear regressions and why this slope varies with FTC." With "Open savannas" we indeed meant lower tree density. To avoid this confusion, we changed the text to: 'savannas with lower tree density'
Not sure I agree with this logic: **"**Contrary to previous research which indicated that dryer conditions in the LDS would lead to higher-MCE fires late-LDS conditions (Fig. 3). In part, this may be because our measurement campaigns missed the peak-season fires when the fires may be hotter..." Winds are the critical factor here. When do they peak in areas studied. High winds (especially if fires studied are head fires) result in higher intensity regardless of fuel moisture. Laris also found lower MCE in LDS due to leaf litter in the fuel load and lighter winds with much higher winds in MDS for the region studied. Note that these factors are key reasons why binary (LDS/EDS) is problematic for determining emissions.	While we did not include windspeed in the field measurements and therefore in the intermediate explanatory field drivers. However, we agree that it is a very influential driver of fire behavior. In the future we will include windspeed measurements on the ground. Although this means we currently cannot correlate reliable measurements of the actual windspeed during the fire with satellite derived proxies, we do include windspeed proxies in the model. We added the following text (P12 L24): "Contrary to previous research which indicated that dryer conditions in the LDS would lead to higher-MCE fires in both grasslands and savanna woodlands (Korontzi, 2005), we found lower MCE in these regions under late-LDS conditions (Fig. 3). This may be because our

	measurement campaigns missed the peak-season fires when the fires may be hotter due to stronger winds (Laris et al., 2021; N'Dri et al., 2018)." We acknowledge that the binary (LDS/EDS) classification is in many ways flawed, as you rightfully point out in your earlier work. With this study, we hope to work towards getting rid of the EDS and LDS classes altogether when it comes to savanna EFs. In the introduction (P2 L37) we state: "EFs used for the accreditation of such projects currently assume a dichotomy of early- and late dry season averages, determined by a cut-off date. However, as discussed by Laris (2021), the fuel and meteorological conditions thought to drive EFs vary more gradually over the season and are subjected to substantial inter-annual and spatial variability. Incorporating spatiotemporal variability in inventories makes emission inventories more dynamic and better equipped for assessing seasonal fluctuations."
Again, see research in Mali and Senegal which support this finding: In accordance with previous studies (e.g. Korontzi et al., 2003b; van Leeuwen and van der Werf, 2011), we found steeper CH4 EF to MCE regression slopes in woodlands compared to grasslands. Comments Figure 3. What is **"typical savanna"** there is no such thing. Also, use more specific terminology, what is "open"?	These classes serve to indicate that the prevalence of trees was a useful feature for clustering the EFs. In Figure 3, we removed the classes and replaced those with rough FTC bands (0-2%, 2-10% and 10-50%)
This and other data rely on 500 x 500 MODIS is this relevant given efforts to burn patchy EDS fires which operate at a hectare level scale? Can you justify using 500m data for the following? For fire severity proxies we used the differential	That is indeed an issue and could be one of the main reasons why the models did not pick out any of these features as strong indicators of the fire. Although not mentioned in the list of features, we also used Landsat retrievals for the

Normalized Burn Ratio (dNBR) and 5 the differential Normalized Difference Vegetation Index (dNDVI) retrieved before and after the fire. These were based on the MODIS surface spectral reflectance, corrected for atmospheric conditions (MOD09GAV6; Vermote)	abovementioned spectral indices. While the spatial resolution is better, it goes at the cost longer intervals between cloud-free scenes and just as with MODIS data our model did not find these features were important. In their current form these models were developed with the application of global modelling in mind. This means that using high resolution (e.g. Landsat and Sentinel) data becomes computationally heavy. Although it could be possible to retrieve the training data at higher resolution and subsequently use courser products (e.g. MODIS) for the spatiotemporal extrapolation, using different data for training and final usage is risky as tree-based models use absolute split values. Therefore, the consistency of these datasets would have to be proven for the entire savanna biome first. We added the following text in the discussion (P15 L4): "Fire intensity proxies (dNDVI and dNBR from MODIS) were poor predictors for the EFs. A potential explanation is that these features were at times heavily diluted, as many of the measured fires only affected part of the pixel. Similar misrepresentation errors can be expected for the NDVI before the fire, FPAR and the Pgreen. Particularly in the LDS, we were often limited to areas that were enclosed by recent fire scars (0-2 years) or other non-flammable boundaries. Although these areas were sizable (several hectares) many of the retrievals in these pixels may poorly represent the burned vegetation. Along with inconsistent retrievals related to cloud cover, this may be an important reason why these features were deemed poor predictors by the models while seen as strong predictors in previous research (Korontzi et al., 2004). Higher resolution features may increase the representativeness of the pixels for the actual burned vegetation."