SEAC4RS Data-Analysis Techniques Map True Emission Factors from Western Forest Fires, Revealing Why Some Produce Their Own Ozone

Robert B. Chatfield

NASA Ames Research Center, USA

THE SEAC4RS and DISCOVER-AQ Science Teams

Especially the Anderson, Blake, Diskin, Fried, Weinheimer, and Wisthaler Groups

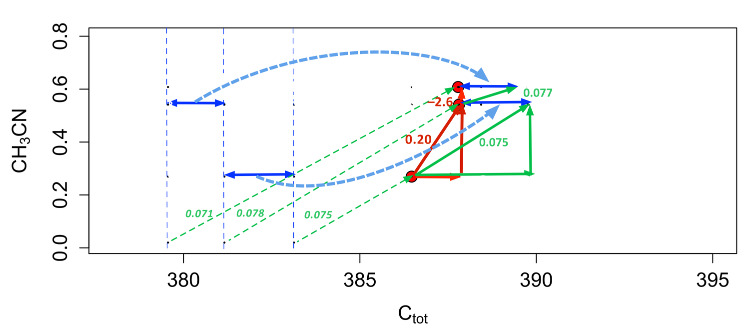
Abstract:

We describe new analysis techniques for quantifying emissions factors for 446 samples of forest fire plumes in the Western US during NASA's SEAC4RS and the ARCTAS airborne campaigns. An unmixing analysis suggested clues about which fires produce ozone without needing addition of urban nitrogen oxides. Preliminary emission factors and associated relationships for key fire-type indicators species CO, HCHO, CH_4, CH3CN/NOy, C_4H_{10}, SO_2, NOz, and several other species are described.

Regarding technique for emissions, we describe a method to compare and categorize individual forest-fire plume samples from the two missions in a manner that minimizes effects of intake-CO_2 and mixing of air-masses. This is known to affect previous estimates of emission factors. Sequential samples are not required. Mixed-effects regression (ME) methods allow us to estimate pre-fire values of the C_{tot} =(CO_2 + CO) (and plume ΔC_{tot}) simultaneously with emissions factors for individual species i, $\Delta x_i / \Delta C_{tot}$. Another simple but approximate methodology is evaluated for errors. In alternation with mixed-effects regression, another method, we employ another method, non-negative matrix factorization (NMF) is employed. NMF allows classification when only mixtures are found in the observations (i.e., "unmixing"). (Observations are expected to describe contributions from differing fire chemistries at a small scale.) Particle emission characteristics (b_{scat}, b_{abs}, SSA, AAE) are closely and individually related to gas-phase composition. Sample scrubbing for major non-fire influences on "signal" species with other large proximate sources, e.g. CO and CH_4, was critical to a clear analysis.

Regarding ozone production in plumes, we distinguish fire types as high-NOy high-VOC, or both. The fire-types with optimal ratios, higher NOy but not too much reactive VOC associate with the highest additional ozone. Urban NOy plays a small role for our selected samples, but fuel nitrogen apparently does.

Why Sequence Analysis ➔ Emission Factor ("NERM", "ERM") Often Fails



Comparison of difference methods for biomass burning emissions assuming knowledge of pre-fire background concentrations of C_{tot} = CO_2 + CO.

Actually observed C_{tot} the fire tracer CH_3CN are plotted as the red dots.

The example uses actual (rather typical) data from the ARCTAS flight April 22, 2008.

Normalized Emission Ratio Method. The difference method estimates as emission factors the slopes deduced form the red triangle (upper triangle too small to depict), i.e., 0.2 and –2.6.

Enhancement Ratio Method. Regression methods would average this variability but behave similarly, in this case suggesting high emission factors.

Estimates of background C_{tot} are shown as dashed, light blue vertical lines, and the differences of C_{tot} between samples as dark blue arrows.

The same difference method corrected for the change in background C_{tot} provides two much more consistent estimates (green triangle estimates). However it is possible to estimate emissions factors directly for each point using the slopes of the green dashed lines (light green numbers).

Additionally, *consecutive* plume observations are *not* required.

Several tracers *are* required.

This behavior of difference-based estimates is true of essentially all the 60-s data we analyzed. Some more consistent estimates are likely with 10-s data, but biases do remain [Yokelson et al, 2013]. Our early estimation using differences with the ARCTAS 10-s data gave many estimates of negative emission factors.

Three Interlinked Types of Variables to Solve for Simultaneously:

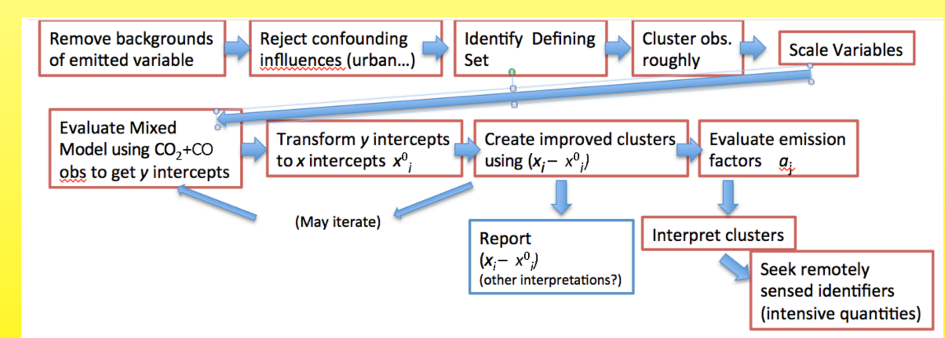
- "Background C" and so $C_{burn} = C_{tot} - C_{back}$ at instance (time) i
- The emission factor a_j of species j as defined by fire type (think of CO(flaming), CO(smoldering),… and
- The fire type at instance (time) i same for all species j

Why it's so complicated:

The first two are solved using mixed-effects (ME) modeling and the fire type is solved for using non-negative matrix factorization(NMF). The ME step solves *most easily* if there are rough estimates of the fire types: rough C_{burn} and rough fire types made by NMF. Once the ME step is complete, more informative NMF assignments and classes are possible.

Experience shows that the ME estimates of C_{back} are surprisingly robust. Fire types depend on good C_{back} and, to some extent, what distinctions are most important to the scientist.

We chose an NMF that distinguished the traditional classification by "modified combustion efficiency" (essentially the ratio Δ CO / ΔC_{burn}).



Flowchart of the algorithm used to estimate biomass burning emissions. Some steps may not be needed in simpler datasets.

Mixed-Effects Regression

$$C_{burn} = x_i - \hat{x}_i^0$$

Take *two or more* chemical species or integrated aerosol properties.

These sets yield an array y_{ij} of species j measured at instance i.

We remove any general background value y_j^0 for these y_{ij}, easier to estimate and of less numeric importance than x_i^0. For most VOCs, $y_j^0 \approx 0$.

C_{burn} estimated from known C_{tot} and varying background C_{backg}

$$y_{ij} = a_j\left(x_i - x_i^0\right) + e_{ij}$$

X_i is (CO+CO_2):

Two sources of variation to be estimated, a_j for the species and x_i^0 for the instances.

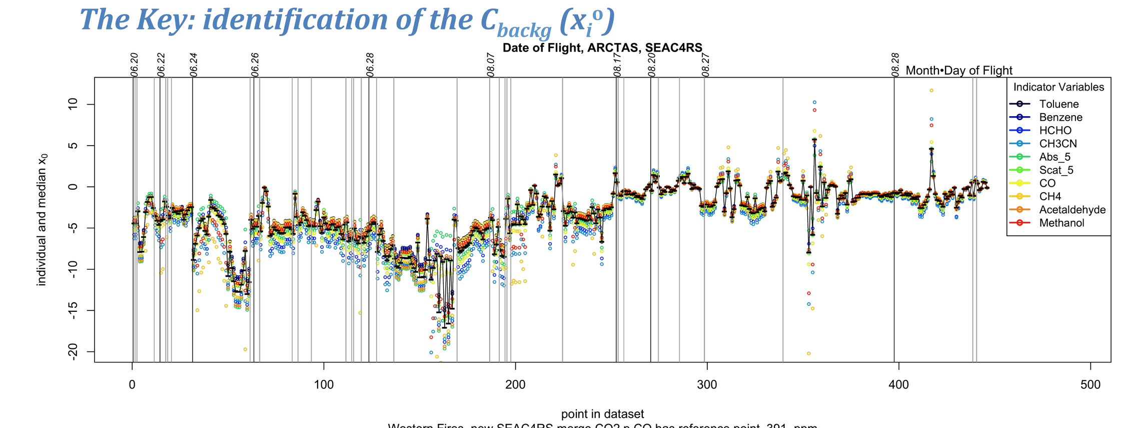
"Pointer compounds" with nearly complete measurements, varied chemical character (in fires and atmosphere) and available at 60 sec

Concentration / property	Abbreviation	Technique	Group
Toluene	$C_6H_5CH_3$	PTRMS	Wisthaler
Benzene	C_6H_6	PTRMS	Wisthaler
Formaldehyde	HCHO	LAS	Fried
Acetonitrile	CH_3CN	PTRMS	Wisthaler
Absorption Coefficient Dry, Total, 532 nm	b_{abs}, Abs_5	Nephelometry	Beyersdorff
Scattering Coefficient, Dry, Submicron 550 nm	b_{acat}, Scat_5	Nephelometry	Beyersdorff
Carbon monoxide	CO	LAS, GC, *mice()*	Diskin, Blake
Methane	CH_4	LAS, GC, *mice()*	Diskin, Blake
Acetaldehyde	CH_3CHO	PTRMS	Wisthaler
Methanol	CH_3OH	PTRMS	Wisthaler

Tests with ARCTAS 10-sec vs 60-sec measurements suggested essential variability captured.

Separating Effects:

The Key: identification of the C_{backg} (x_i^o)



The C_{backg} is shown by flight date and individual plume sample.

All estimates are made independently of their neighbors; the consistency is reassuring.

The use of medians to summarize evidence from differing indicators makes the estimates robust against difficulties of individual species. Methane, (yellow-orange) though it had troubled measurements in SEAC4RS, and could be confuse with other sources in ARCTAS, does not show difficulties. Similarly with CO (yellow).

Examining FireTypes:

Predominant Class Keyed by Color

HCHO-like or b_{Abs}-like

HCHO is likely formed both from VOC oxidation as well as emitted. The fires with highest C_{burn} were in this class, including the very high-burn Rim Fire of Yosemite sampled in SEAC4RS. Regression lines are shown for each class. Values of C_{burn} ~6 ppm tended to control the slope of these lines in many cases. Note the similarities in pattern of HCHO and b_{Abs}.

NO_y-rich or CH_3CN-like

This class was found only in California, and not in the Sierra samples. The NOy seems to come from high-nitrogen fuels, since CH_3CN is the highest in these fires. HCHO (VOC) is lower than the red HCHO-like class. The locations tend to be in Southern California, suggesting possibly chapparal or other dry woodland fuel. There is no correlation with CH_3Cl, an urban emission. We attempted to remove urban-dominated emissions by various techniques, though some samples were not far from urban areas. C_{burn} ranged to 10ppm

CH_4-like

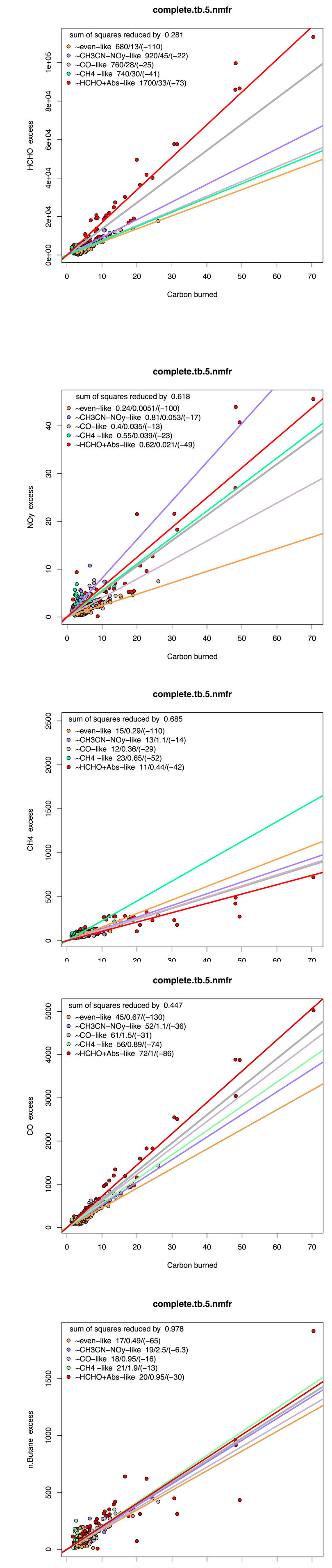
This type was seen only in California and Nevada, and found in both missions. It also had high NOy but even lower HCHO (VOC). C_{burn} ranged to 12ppm. CH_4 in SEAC4RS had to be estimated from can samples, and samples with CH_4 from other large sources had to be removed (otherwise: negative emission factors). Both acetone and methanol showed similar behavior.

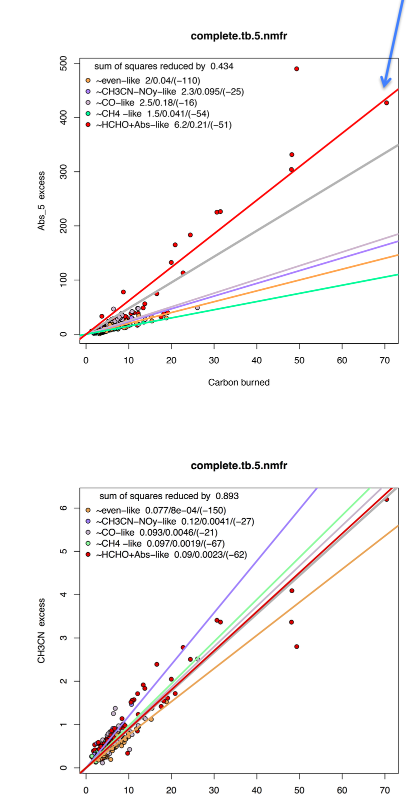
CO-like

CO had the most similar emission factors among fire types. HCHO-like fires emitted slightly higher amounts of CO, but this fire-type also had high CO emissions, hence the name. The fire type was common in both missions, and distinguished plume origins in the Montana multi-smoke-plume sampling.

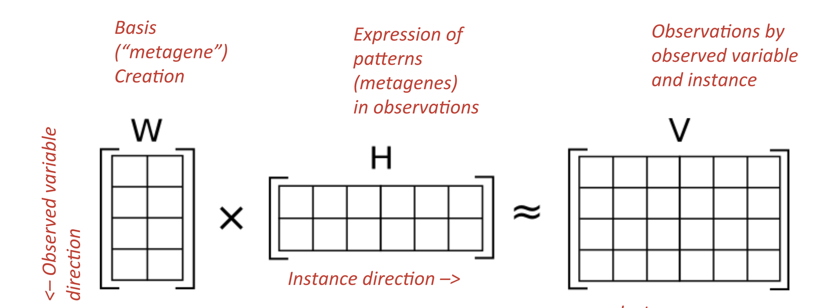
"Even"-like

These samples including more aged plumes: note the progression in the Rim Fire plume moving north into Nevada. "Even"-like samples may have other origins. n-Butane is shown: the similarity of all classes in the graph suggests that dilution effects are not distorting our results very much.





Non-Negative Matrix Factorization



Advantages of NMF description:

While instances may be assigned classes, variables (b_{scat}, CH_3CN) at each instance are easily seen as mixtures of more basic fire types.

Composition attributes ("metagenes") following different basic chemistries of combustion may mix, even at the few-meter scale, and yet still be recognizable. This general description of samples is appealing. The pursuit of "pure flaming combustion" in aircraft or ground samples is not required. Basic chemistries may still be identified. ("Chemistries" is shorthand for fuel composition *and* burn chemistry, i.e. effects on emissions such as fuel type and moisture, stage and intensity of combustion, etc.)

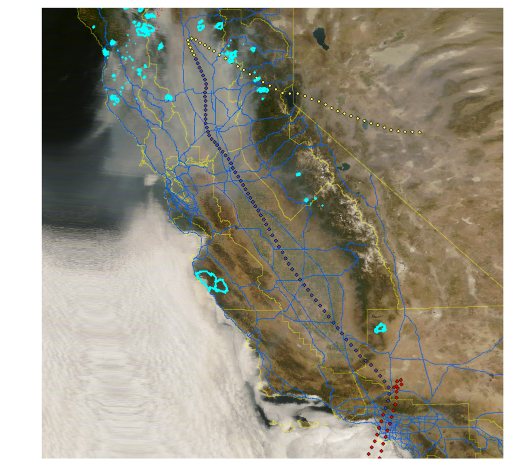
Classes may be assigned based on predominant metagene. The NMF technique used here, "Brunet" NMF, has been shown to be equivalent to multinomial principle components analysis. NMF techniques relate also to popular *K*-means clustering techniques. Additionally they describe the closeness to the "pure" predominant class and measures of contributions of other class. Clearly, the factorization and the resulting classifications depend on the indicator species chosen, and selection of a number of classes (metagenes, here 5), is influenced by the scientists' interests. Here we chose a factorization emphasizing distinctions in CO, "modified combustion efficiency.")

Character of the NMF Classes

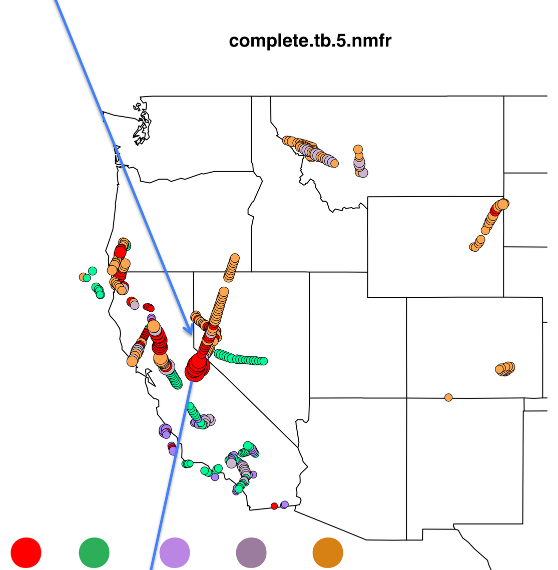
We have resisted the urge to give names like "flaming" or "smoldering" or "distillation" although some likely correspondences may appeal to you. Instead we give names based on species which have high emission ratios in this class. Connection to histories and fuels burned is left for future work.

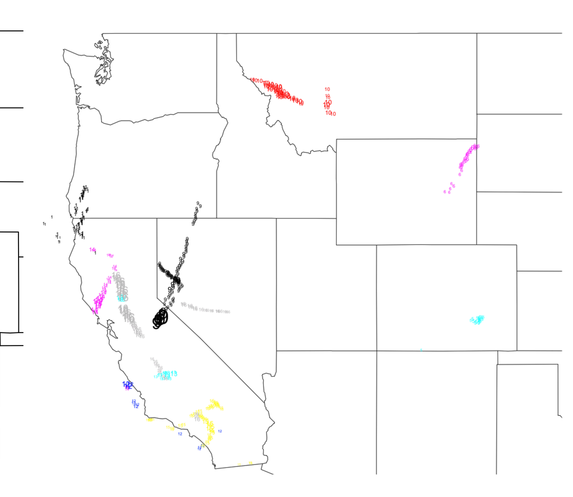
Fire Regimes Vary Widely

Thanks, Amber Soja !



The Rim Fire





Map of all samples classified by fire types, (colors). Area of dots indicate fire size, C_{burn}

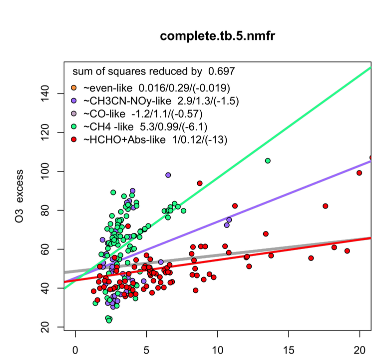
Map of flight numbers in SEACR4S and ARCTAS Size of numerals indicates C_{burn}

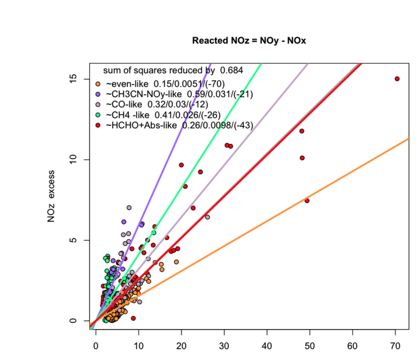
Ozone Levels in Fires Which Are Not Greatly Affected by Urban Emissions

Ozone levels affected by many processes besides fire. A "background" concentration is not easy to estimate due to many ozone sources and dilution effects.

For the ozone graph below, we allowed any intercept for ozone at C_{burn} = 0. To our surprise, the very distinct classes HCHO-like, CH_3CN-like and CH_4-like (red, green, violet) showed rather different linear relationships to C_{burn}. The CH_4-like class was associated with the highest ozone, and the CH_3CN class was second. Very high levels of C_{burn} produced notable ozone, e.g., in the Rim Fire plume, even if not efficiently.

NOz = NOy – NOx showed corresponding behavior, reversing green and violet. NOz is associated with ozone production, since both originate from radicals + NOx.





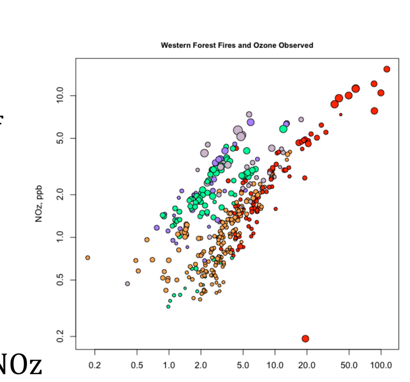
VOC-NOz Relationships to Ozone Are Very Strong

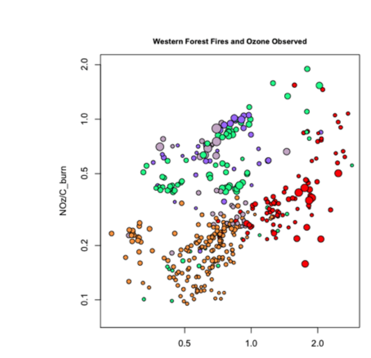
The bottom two figures show the relationship of ozone (size of dots) to HCHO and NOz.

(Left) High ozone is achieved with more emissions,

… but optimal ratios of NOz to HCHO allow high ozone at lower total pollution (green dots vs red dots).

This follows the theory of photochemical smog, but the results are unexpectedly clear.

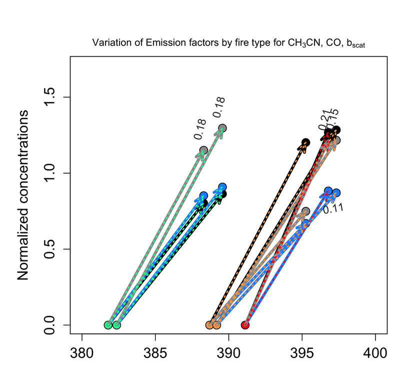




Why the Strong Bifurcation?

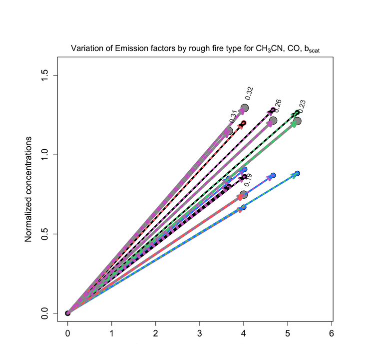
The right figure echoes the left but in terms of emission ratios HCHO/ C_{burn} and NOz/ C_{burn}. The bifurcation of the dots and classes is even more clear, High NOz & Low HCHO vs Low NOz & High HCHO. Is this due to photochemistry?

A Messy Part: Emission Factors Vary … by Fire Type (Affects All Methods)



The same species may have different emission factors (shown as slopes) for different fires, as illustrated using 5 of the 446 plume observations. Plot shows values of CO (gray), b_{scat} (black), and CH_3CN (green), all normalized by their mean at the observed values of $x_i = C_{tot}$. With knowledge of x_i^0, shown as values along x axis, emission factors are defined, and they vary among types of fires. Types are indicated by the colors of the dots showing x_i^0. Scaled emission factors, $1/\left(x_i - x_i^0\right)$, ppm⁻¹, tend to have different values according to fire type, as lab observations suggest. Solid colors of the arrows indicate species, over-plotted dashed-line colors indicate fire type eventually identified.

Details: Bootstrapping Past The Messy Part



$$\check{y}_{ij} = y_{ij} \Big/ \underset{i}{\text{mean}}\ y_{ij}, \quad \forall i$$

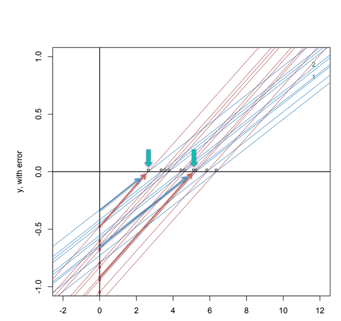
and then form a consensus mean

$$\hat{X}_i = \check{y}_{ij} \Big/ \underset{j}{\text{mean}}\ \check{y}_{ij}$$

Here the symbol reinforces the idea that $\hat{X}_i$ is just an estimate and will have a different scale than than the actual C_{burn}. A regression of $\hat{X}$ against C_{burn} can estimate the scale, but an associated value of a desired quantity like X_i^0 , or indicated zero-point without fire emissions is elusive.

Before estimation of x_i^0, emission factors a_j, and fire types, approximations $\hat{X}_i$ of $\left(x_i - x_i^0\right)$ can be made using a consensus estimate based on the scaled quantities. $\hat{X}_i$ need only be measured from 0 (see text). Three rough fire types (classes) are shown using a different color set. This process produces aids the convergence and assessed fit of the subsequent mixed-effect model estimation step.

Details: But We Want x-Intercepts Not y-Intercepts to Get C Backgrounds!



$$y_{ij} = a_j x_i + c_i^0 + e_{ij}$$

where the y-intercepts c_i^0 are estimated for each instance. While this misses the true model, it is worthwhile to use geometry to find approximations for the x_i^0. Given regression estimates for a_j and c_i^0, $\hat{a}_j$ and $\hat{c}_i^0$, we may make *many* estimates of x_i^0, i.e., $\hat{x}_{ij}^0$,

$$\hat{x}_{ij}^0 = -\frac{\hat{c}_i^0}{\hat{a}_j}$$

$$y_{ij} = \sim \hat{a}_j x_i - \hat{a}_j \hat{c}_i^0/(-\hat{a}_j)$$

See the figure for a geometrical interpretation of using the slope estimates to move from y- to x-intercept. We may then take the median over j of these $\hat{x}_{ij}^0$ for a reasonable estimate of $\hat{x}_i^0$ We may also understand different behavior among the indicator variables. With the background $\hat{x}_i^0$ and the carbon burned, $C_{burn} = x_i - \hat{x}_i^0$, we may procede with ordinary regression estimates. This allows us a *linear* mixed-effects model, easier to compute. For this solution we used `lmer()` in the **R** package `lme4`.

Conclusions

As the abstract suggests, the combination of mixed-effects modeling and non-negative matrix factorization has solved the threefold problem of C_{backg}, emission factors a_i, and variation among fire types ("chemistries") However, this first attempt to classify fire types has led to new questions about fire plume chemistry. Statistics has done its work; other approaches may now be called for.