BACHELOR

Comparison of formaldehyde columns observed with aircraft and the SCIAMACHY and OMI satellite instruments during the 2006 AMMA campaign

de Wild, M.

Award date:
2012

Comparison of formaldehyde columns observed with aircraft and the SCIAMACHY and OMI satellite instruments during the 2006 AMMA campaign

M. de Wild, August 2012, R-1808-A

Supervisors: M.H.A. van Geel (TU Eindhoven)
K.F. Boersma (TU Eindhoven, KNMI)

Abstract

The goal of this study is to compare local formaldehyde (HCHO) mixing ratios obtained during the 2006 AMMA campaign with atmospheric column densities retrieved by the OMI and SCIAMACHY satellites. Therefore we converted the measurements of the DLR-F20 and BAe-146 aircraft to atmospheric columns. Since the aircraft did not perform measurements from the ground to the tropopause, we used multiple extrapolation methods. We also investigated the error in the computed HCHO column density.

When comparing the two aircraft column densities with each other, we see that the downward extrapolation of the DLR-F20 aircraft (21%) is greater than the downward extrapolation of the BAe-146 aircraft (11%). The upward extrapolation of the BAe-146 aircraft (38%) is greater than the upward extrapolation of the DLR-F20 aircraft (16%). This can be directly related to the altitude at which the aircraft has performed measurements; the BAe-146 aircraft covered a lower part of the troposphere than the DLR-F20 aircraft. In total, the DLR-F20 column densities seem more reliable due to the lower total extrapolated fraction.

When we compare the aircraft data with SCIAMACHY, we only find two matches due to SCIAMACHY's poor daily coverage, a global coverage in 3-6 days. The two matches suggest that the column density retrieved by SCIAMACHY is higher than the column density computed using the average of the different extrapolation methods.

When we compare the aircraft data with OMI, more matches are found due to the daily global coverage of the satellite. Plotting the satellite column density versus all of the aircraft column densities, the retrieved slope from the linear regression is 0.59, the intercept lies at $3.78 \cdot 10^{15}$ molec./cm^2 and it has a Pearson coefficient of 0.27.

Table of contents

Introduction

Formaldehyde (HCHO) is a toxic and corrosive gas in the atmosphere [1]. It is mainly produced by photochemical oxidation of methane and other hydrocarbons. Isoprene, the most dominant source of HCHO, is a gas emitted by vegetation. Since isoprene emissions increase with sunlight, temperature and soil moisture, the concentrations of isoprene, and therefore HCHO, peak over tropical rainforests [2]. HCHO is also a primary product of biomass burning. HCHO has a short lifetime of approximately one to two hours, implying that it cannot diffuse much. Since the 1990s, formaldehyde can be measured by satellite instruments which provide global coverage and long-term data records. This provides a tool to infer trends in HCHO concentrations, and improve estimates of highly uncertain isoprene emissions. Before the satellite measurements can be relied upon, they need to be validated with independent measurements. Data gathered by aircraft provide such an independent measurement set.

In June, July and August 2006 a multi-aircraft campaign, called the African Monsoon Multidisciplinary Analysis (AMMA) campaign, took place over West Africa [3]. The monsoon is a part of the year in which the wind blows in one direction, causing monotonic weather. If the wind direction changes, usually a 180 degree turn, the weather will also change. Depending on the time of the year, the monsoon either causes draught or heavy rainfall.

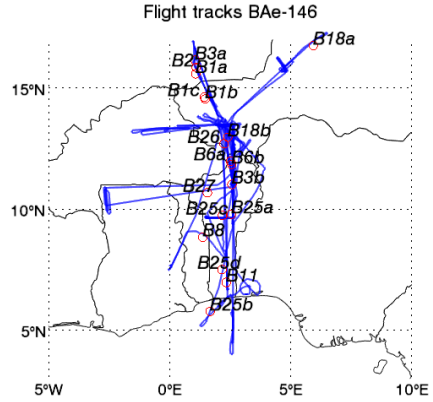

Figure 1: The flight tracks of the British aircraft plotted over the map of West Africa. The circles and the corresponding label represent the average location of the measurements resulting in an HCHO profile.

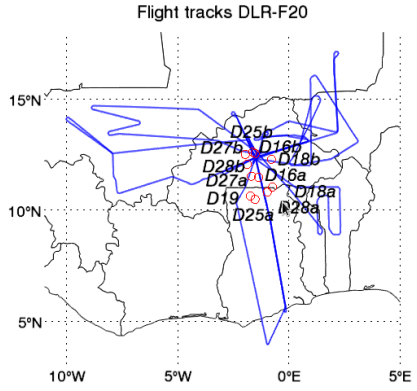

Figure 2: The flight tracks of the German airc plotted over the map of West Africa. The cir and the corresponding label represent average location of the measurements resul in an HCHO profile.

During the AMMA campaign the composition of the atmosphere was measured to obtain a better understanding of atmospheric conditions [4]. HCHO measurements were performed aboard the British BAe-146 (Figure 1) and German DLR-F20 (Figure 2) aircraft. The British aircraft made its HCHO measurements between July 17 and August 17 and the German aircraft between August 1 and August 18, of which four measurement flights are transfer flights between Europe and Africa. Profiles measured over the ocean during these transfer flights can be used to determine the background HCHO concentration as the measurements take place far from sources of isoprene emission.

The red open circles in Figure 1 and Figure 2 are the average coordinates of the vertical profiles we could obtain from the measurements. The markings are labeled with a three part code. The first letter B or D indicates which aircraft performed the measurements: B stands for BAe-146 and D is short for DLR-F20. The number tells us what day the measurement is performed, setting the first day at which the BAe-146 aircraft performed a measurement that could be converted into a column (July 20) to 1, the day after that (July 21) to 2, and so on. The last and optional symbol indicates which profile is meant. If there is no a, b, c or d, only one profile was obtained that day.

Using the HCHO mixing ratios measured from the aircraft, an atmospheric column can be determined. This will be compared to satellite data of SCIAMACHY (**SC**anning **I**maging **A**bsorption spectro**M**eter for **A**tmospheric **CH**artograph**Y**) [5] and OMI (**O**zone **M**onitoring **I**nstrument) [6], which both provide a calculated vertical HCHO column density. As of 2007, a third satellite, GOME-2, has been measuring the atmospheric composition [7]. Since GOME-2 was not active during the 2006 AMMA campaign, we cannot use the satellite. SCIAMACHY has a swath width of 60 km by 30 km [5] and OMI has pixels ranging from 13 km by 24 km at nadir to 28 km by 135 km at the edge of the measurement area [8]. The overpass times for SCIAMACHY and OMI are approximately 10.00 and 13.40 local time, respectively [8] [9].
Since we would like to compare the satellite measurements with the aircraft measurements, we do not want profiles where the longitudinal and latitudinal displacement was very large. When the displacement of the aircraft is large, many satellite pixels have to be taken into account, which is not preferable for this comparison study. To be able to compare the satellite measurements with the aircraft measurements, both the satellite and the aircraft should have measured approximately the same volume of air. This means, in addition to the small displacement, the time of the aircraft measurements should also be as close as possible to the satellite overpass time. For the longitude and latitude we have only taken into account flight tracks that have less than 1.5 degree longitudinal or latitudinal displacement.
To be able to compare the aircraft HCHO with the HCHO retrieved from the satellites, the aircraft HCHO profiles need to be converted into an atmospheric column density, which is the vertically integrated amount of particles per square centimeter. In this report four approximation methods regarding the conversion of the aircraft data into a column are discussed. An average will be used to determine the column value which is compared with the column values retrieved by the satellites. The aircraft data is used to test the accuracy of the satellite data.
The satellites use the fingerprints in an absorption spectrum to recognize the different gasses in the atmosphere. Using various assumptions they convert it into a vertical column. This report will not give an in-depth explanation of the calculation used by the satellites. Instead, it will try to make a quantitative approach to the accuracy of the satellite measurements.

Depending on the flight time, HCHO columns from either SCIAMACHY or OMI are compared with the aircraft HCHO column. A timespan of two hours around the approximate OMI overpass time and a timespan of two and a half hours around the SCIAMACHY overpass time are used to determine whether or not the measurements

can be compared. This is large compared to the lifespan of HCHO, one to two hours, but do to the continuous isoprene emissions during daytime, we assume the HCHO concentration does not fluctuate a lot during a two hour timespan.

The first theory subsection of this report focuses on how to compute the column densities from the aircraft measurements with the extrapolation methods we used. The second subsection of the theory focusses on an estimate for the uncertainty in the column densities computed with the aircraft measurements as well as the column densities retrieved by the satellites. In the results section, firstly the column densities established from the measurements of the two aircraft will be compared. Secondly, the aircraft column densities will be compared to the column densities retrieved by the satellites. By using a regression technique satellite HCHO column retrieval will be compared against the aircraft HCHO columns. Metrics such as the Pearson correlation coefficient, and the fit will be used as quantitative markers of the quality of the satellite retrievals.

Theory

1. Determining the column from aircraft measurements

The HCHO atmospheric column density is the amount of formaldehyde molecules in a vertical volume with a base of 1 cm^2 [10]

$$\Omega_{\mathrm{HCHO}} = \int_{z_{ground}}^{z_{trop}} \frac{A_v\, p(z)\, C_{\mathrm{HCHO}}(z)}{R\, T(z)}\, dz, \tag{1}$$

where Ω_{HCHO} represents the column density [molec. cm^{-2}] in the slice of the atmosphere from the ground (z_{ground} [m]) to the tropopause (z_{trop} [m]), z is the height [m], A_v is Avogadro's number, $6.022 \cdot 10^{23}$ mol^{-1} [1], p is the pressure [hPa or 10^{-2} N cm^{-2}], $C_{\mathrm{HCHO}}$is the HCHO mixing ratio [mol mol^{-1}], R is the universal gas constant [J mol^{-1} K^{-1} or N m mol^{-1} K^{-1}] and T is the temperature of the atmosphere [K]. The relation between altitude and pressure is [11]

$$\frac{dz}{dp} = \frac{1}{\frac{dp}{dz}} = -\frac{R\, T(z)}{M_{air}\, g\, p(z)}, \tag{2}$$

where M_{air} is the average molecular weight of air ($28.97\ \mathrm{g\,mol^{-1}}$) and g the gravitational constant. g has a slight height dependence, which will be ignored in this study. Equation (1) can now be written as

$$\Omega_{\mathrm{HCHO}} = \int_{p_{trop}}^{p_{ground}} \frac{A_v\, C_{\mathrm{HCHO}}(p)}{M_{air}\, g}\, dp, \tag{3}$$

p_{ground} and p_{trop} respectively the ground pressure and tropopause pressure obtained by the chemical transport model TM5.

Since the aircraft measurements only cover a part of the vertical distance between the earth's surface and the tropopause height, the column is split up into three parts, namely

$$\begin{aligned}\Omega_{\mathrm{HCHO}} = &\int_{p_{zmin}}^{p_{ground}} \frac{A_v\, C_{\downarrow}(p)}{M_{air}\, g}\, dp + \int_{p_{zmax}}^{p_{zmin}} \frac{A_v\, C_{\mathrm{HCHO}}(p)}{M_{air}\, g}\, dp \\ &+ \int_{p_{trop}}^{p_{zmax}} \frac{A_v\, C_{\uparrow}(p)}{M_{air}\, g}\, dp,\end{aligned} \tag{4}$$

hereafter referred to as *a*, *b* and *c*, respectively, where p_{zmin} and p_{zmax} are the pressures corresponding to respectively the lowest and highest altitude the measurements have taken place. Terms *a* and *c* do not contain data from measurements. The extrapolated HCHO mixing ratios $C_{\downarrow}$ and $C_{\uparrow}$ represent downward or upward extrapolation respectively. For the extrapolated values the available measurements are used. In Figure 3 four possible extrapolation methods are shown. Methods (1) and (3) assume the mixing ratio remains constant from the average of the last three measurements to the desired pressure. Methods (2) and (4) represent a different extrapolation, in which the HCHO increases linearly with the pressure, and decreases linearly as pressure decreases. In method (2) the mixing ratio goes to zero at the tropopause pressure. In the last method, method (4), a linear fit along the curve is used. The linear fit procedure returns a value and approximate error for the intercept (a_0) at zero pressure and the slope (a_1), as in

$$C_{\mathrm{HCHO}}^{fit}(p) = a_0 + a_1\, p. \tag{5}$$

The final column amount will be the average of those expansions, as represented by the red lines. We have chosen not to compare all five possible outcomes with the satellite, as that would not be meaningful in a comparison study, however we have based our result on the average of all of them.

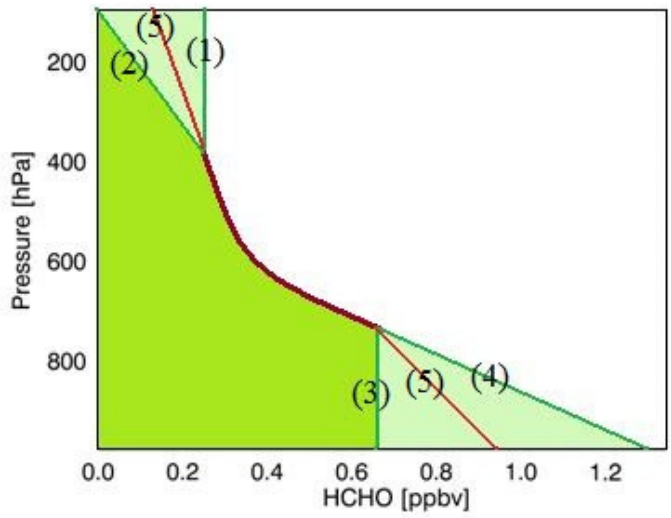

Figure 3: An example HCHO profile where all four expansions are visualized. The bold dark red line represents the aircraft measurements (schematic). (1) and (3) *indicate the vertical extrapolation,* (2) *represents the diagonal expansion where the HCHO mixing ratio going to zero at tropopause pressure and (4) indicates a linear fit along the curve. The two red lines* (5) *are the average of the two extrapolation methods and these will be used to compare the aircraft data with satellite measurements.*

The most important part of the column, *b*, includes the aircraft measurements. Using the trapezoidal rule, the integral will be numerically evaluated using the HCHO measurements as follows

$$\int_{p_{zmax}}^{p_{zmin}} \frac{A_v\, C_{\mathrm{HCHO}}(p)}{M_{air}\, g}\, dp \approx \frac{A_v}{M_{air}\, g} \sum_i (p_{i-1} - p_i)\, \frac{1}{2}\left(C_{\mathrm{HCHO},i} + C_{\mathrm{HCHO},i-1}\right), \qquad (6)$$

where the subscript indicates the i'th measuring point. This method is used because it gives a good approximation for the numerical integral without making the calculation complicated.

In order to compare the aircraft HCHO to the satellite measurements taken at specific locations and overpass times, we need to attribute a time and location to the aircraft measurement. Therefore we compute weighted averages of time, longitude and latitude. HCHO has its most important source near the ground. Therefore, measurements made at lower altitudes are more important for determining the average location rather than the measurements at higher altitude. Therefore, a weighting factor which is inversely proportional to the height, is used. As the pressure decreases with the altitude, the pressure is chosen to be the weigh factor, so that the average of a quantity *x* becomes

$$\langle x \rangle_p = \frac{\sum_i p_i x_i}{\sum_i p_i}. \quad (7)$$

Using this method, the location and time where the highest HCHO concentrations are present, which is near the ground, are given more importance.
The same method is used for averaging the satellite pixels. The satellite pixel containing an aircraft measurement is weighted with the pressure of that specific measurement. In this way, the satellite pixels have the same weight as the measurements. Therefore, satellite pixels corresponding to aircraft measurements near the ground are more important than pixels corresponding to measurements near the tropopause.

2. Error calculation.

The total error in the aircraft column depends on, firstly, the direct measurement errors of the aircraft and, secondly, the method of extrapolation used. The measurement errors consist of the errors in the HCHO measurements and the errors in the ground and tropopause pressure calculated by TM5. It is assumed that the error in the measured local pressure is negligible.
The estimated fraction of the column calculated by extrapolation is given by

$$f = \frac{A_v}{M_{air}\, g} \frac{(p_{ground} - p_{zmin}) C_{\downarrow}^* + (p_{zmax} - p_{trop}) C_{\uparrow}^*}{\Omega_{\mathrm{HCHO}}}, \quad (8)$$

where $C_{\downarrow}^*$ and $C_{\uparrow}^*$ are the effective mixing ratios for extrapolating downward and upward respectively. The value of the effective mixing ratio depends on the extrapolation method used: for method (1) $C_{\uparrow}^*$ equals the average of the last measured values, but for method (2) the effective mixing ratio is only half of the average of the last measured values. A psychological barrier is reached when the fraction exceeds 0.5, at which the larger share of the HCHO column is from extrapolation rather than calculated from measurements. The fraction can be split up into an upper and lower part, to distinguish between the extrapolated lower and top part. The two partial fractions are given by

$$f^- = \frac{A_v}{M_{air}\, g} \frac{(p_{ground} - p_{zmin}) C_{\downarrow}^*}{\Omega_{\mathrm{HCHO}}}$$
$$f^+ = \frac{A_v}{M_{air}\, g} \frac{(p_{zmax} - p_{trop}) C_{\uparrow}^*}{\Omega_{\mathrm{HCHO}}}, \quad (9)$$

where f^- is the fraction of downward extrapolation and f^+ is the fraction of upward extrapolation.

Since formula (4) and (5) are linear, the error calculation is relatively simple and given by

$$\sigma_{\Omega_{\mathrm{HCHO}}}^2 = \sum_i \left(\frac{\partial \Omega_{\mathrm{HCHO}}}{\partial x_i} \sigma_{x_i} \right)^2, \quad (10)$$

where σ_{x_i} is the uncertainty in a measured quantity x_i, and $\sigma_{\Omega_{\mathrm{HCHO}}}$ is the uncertainty in the whole column Ω_{HCHO}. For the aircraft measurements, the uncertainty in the whole column contains contributions from the three labeled parts, *a*, *b* and *c*,

$$\sigma_{\Omega_{\mathrm{HCHO}}}^2 = \sigma_a^2 + \sigma_b^2 + \sigma_c^2, \quad (11)$$

where σ_a, σ_b and σ_c represent the errors in their respective column parts. We assumed that the errors in *a, b* and *c* have no correlation. We also assumed that the height dependence of the gravitational constant is negligible. It is also assumed that the pressure measured by the aircraft is correct. The errors in the extrapolated parts *a* and *c* are approximately 100% of the extrapolation, meaning $\sigma_a \approx a$ and $\sigma_c \approx c$, as can be seen in Figure 3. The error in the measured fraction of the column density will be calculated using error propagation. We need the partial derivatives in *b,* which are

$$\left|\frac{\partial b}{\partial C_{HCHO,i}}\right| = \left|\frac{\partial b}{\partial C_{HCHO,i-1}}\right| = \frac{1}{2}\frac{A_v}{M_{air}g}(p_{i-1} - p_i). \tag{12}$$

The total error in *b* is approximately

$$\sigma_b^2 \approx \frac{1}{4}\frac{A_v^2}{M_{air}^2 g^2}\sum_i (p_{i-1} - p_i)^2\left(\sigma_{C_{HCHO,i}}^2 + \sigma_{C_{HCHO,i-1}}^2\right), \tag{13}$$

with $\sigma_{C_{\mathrm{HCHO},i}}^2$ the uncertainty in the *i*'th measuring point.

The total error in the aircraft column density is approximately given by

$$\sigma_{\Omega_{HCHO}}^2 \approx a^2 + \frac{1}{4}\frac{A_v^2}{M_{air}^2 g^2}\sum_i (p_{i-1} - p_i)^2\left(\sigma_{C_{HCHO,i}}^2 + \sigma_{C_{HCHO,i-1}}^2\right) + c^2. \tag{14}$$

The weighted average column of the satellite is given by

$$\langle \Omega_{HCHO} \rangle_p = \frac{\sum_i p_i \Omega_i}{\sum_j p_j}, \tag{15}$$

with Ω_i the column corresponding to the location of the *i*'th measuring point. This means that the error in the weighted average becomes

$$\sigma_{\langle \Omega_{HCHO} \rangle_p}^2 = \frac{\sum_i \left(p_i \sigma_{\Omega_i}\right)^2}{\left(\sum_j p_j\right)^2}, \tag{16}$$

where σ_{Ω_i} is the uncertainty in the satellite measurements computed by the satellite.

Results and discussion

We started by qualifying which parts of the flights could be used to obtain HCHO columns. A perfect flight track would be one which is solely vertical. Realizing that is not possible, it is preferred that the horizontal displacement is as small as possible. In order to limit the amount of satellite pixels needed, we decided that the displacement in longitudinal en latitudinal displacement should be less than 1.5˚. Assuming the earth is perfectly spherical with a radius of $6.371 \cdot 10^6$ m [1], that would equal to 165 km. For SCIAMACHY, with pixels of 30 km by 60 km, this would mean three to six pixels and for OMI, with pixels of at least 13 km by 24 km, seven to thirteen pixels, but it can go as low as a single pixel.

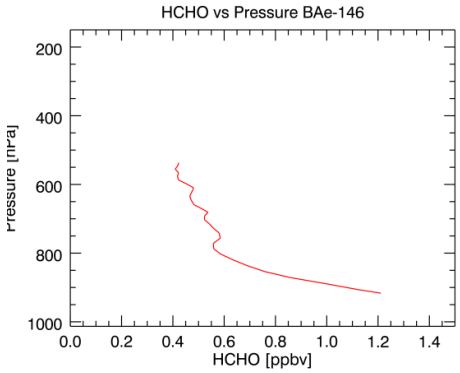

Figure 4: A single column determined using measurements from the BAe-146 aircraft. The low aircraft ceiling and low aircraft bottom represent the BAe-146 flights. This specific column is the second column on the flight of July 22 (B3b).

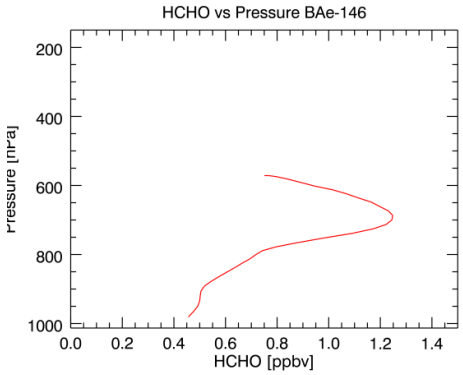

Figure 5: The measurement points plotted against the altitude of the BAe-146 aircraft. It is the second profile on August 13 (B25b). It looks as if the profile is upside-down, which might be the result of biomass burning.

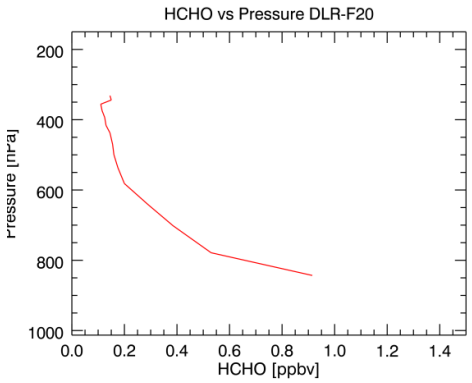

Figure 6: A randomly picked visualization of measurements of the DLR-F20 aircraft; it is the first column of the flight of August 13 (D25a). The high aircraft bottom and high aircraft ceiling are common among the DLR-F20 flights.

Figure 4 and Figure 6 show example profiles of the BAe-146 and DLR-F20 aircraft respectively. The DLR-F20 aircraft performed its measurements at a higher altitude

than the BAe-146 aircraft, covering slightly more of the pressure range in the troposphere. Out of all profiles, particular profile stands out and is shown in Figure 5. The second profile on August 13 (B25b in Figure 1) looks like it is upside-down. This is the result of biomass burning [12], in which locally a relatively large HCHO mixing ratio is found aloft, presumably in the smoke plume downstream of the fires.

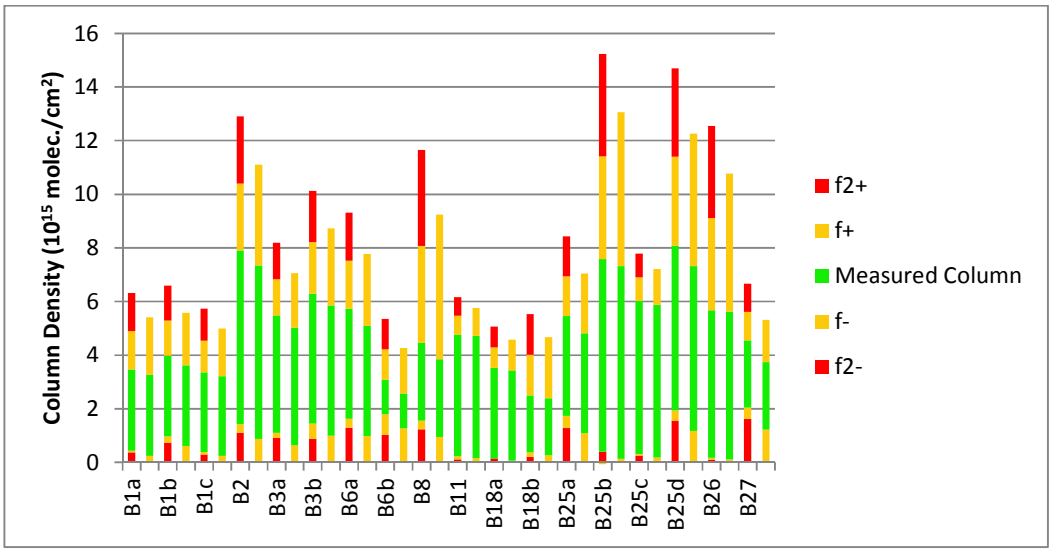

Figure 7: Column densities of the BAe-146 aircraft plotted in 10^{15} molec./cm^2. Each profile has two stacked columns associated with it. The left stacked column, consisting of five parts, is the maximum column, with the fractions of each of the four expansions visualised. The right column, consisting of three parts, is the average of both expansions and only has the average of the downward and upward extrapolation fraction visualised.

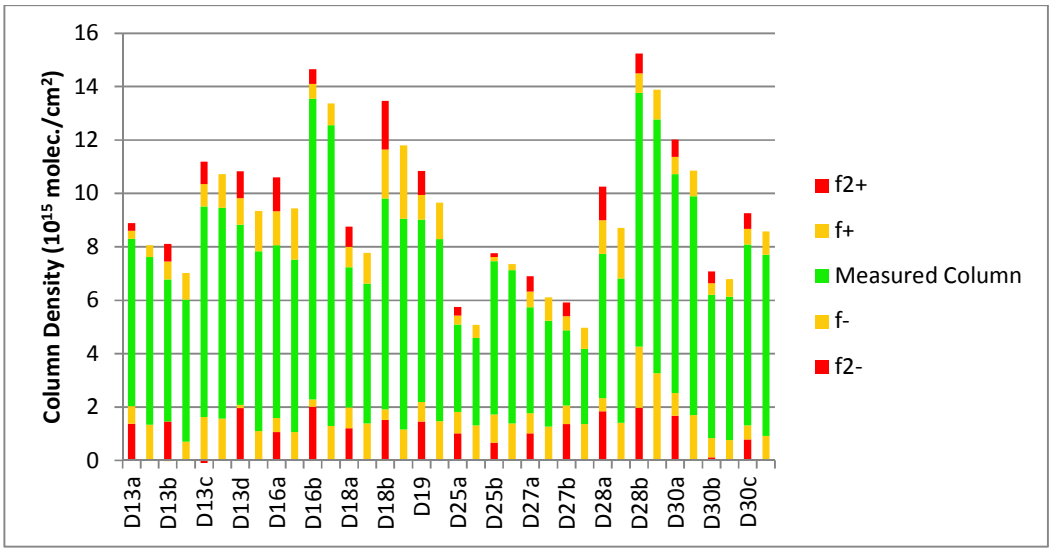

Figure 8: Column densities of the DLR-F20 aircraft plotted in 10^{15} molec./cm^2. Each profile has two stacked columns associated with it. The left stacked column, consisting of five parts, is the maximum column, with the fractions of each of the four expansions visualised. The right column, consisting of three parts, is the average of both expansions and only has the average of the downward and upward extrapolation fraction visualised.

Figure 7 shows the column densities in 10^{15} molec./cm^2 of the BAe-146 aircraft and Figure 8 shows the column densities in 10^{15} molec./cm^2 of the DLR-F20 aircraft. The

red fractions in both figures, labeled f2- and f2+ in the graphs, are the areas which are marked with a greyish color in Figure 3, labeled with a minus sign for downward and marked with a plus sign for upward extrapolation. When the figures of the two aircraft are compared, it can be seen that the downward extrapolation of the DLR-F20 aircraft is greater than the downward extrapolation of the BAe-146 aircraft, which can be related directly to the lower bottom altitude of the BAe-146 and higher flight levels of the DLR-F20 aircraft as illustrated in Figure 4 and Figure 6. The average downward extrapolation for the DLR-F20 aircraft is 21% of the total column and the average downward extrapolation of the BAe-146 aircraft is 11%. The difference in the altitude of the measurements also explains why the upward extrapolation of the BAe-146 aircraft is greater than the upward extrapolation of the DLR-F20 aircraft. The average upward extrapolation of the BAe-146 aircraft is 38%, while the average upward extrapolation of the DLR-F20 aircraft is only 16%. The data from Figure 7 and Figure 8 can also be found in the Appendix, along with the corresponding average UTC time and the average longitude and latitude of the aircraft measurements.

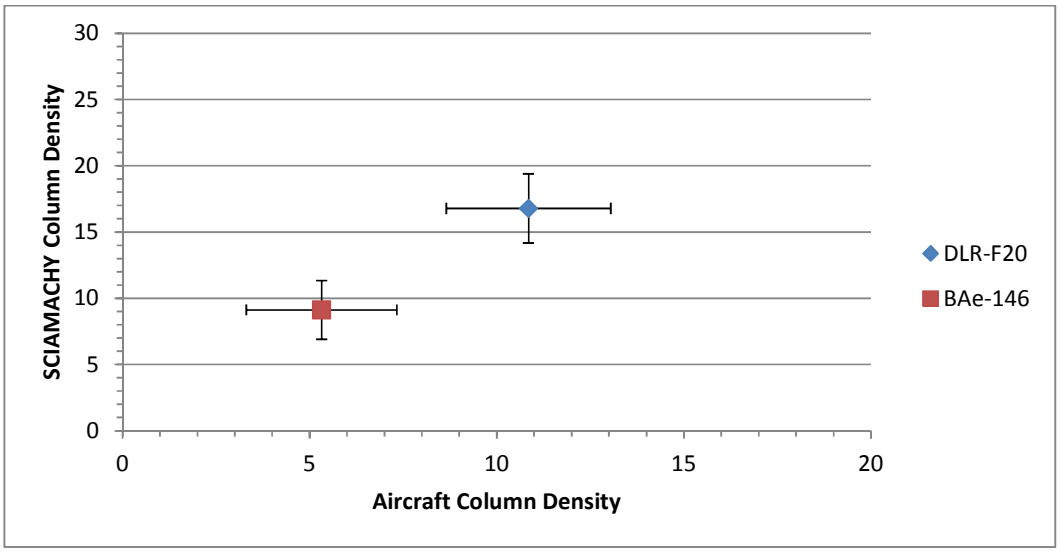

Figure 9: The computed aircraft HCHO column density versus the SCIAMACHY HCHO column density (in 10^{15} molec/cm^2). The blue diamond is the DLR-F20 column density and the red square is theBAe-146 column density. Due to the poor daily coverage of SCIAMACHY, only two profiles found a match.

Figure 9 compares the aircraft HCHO column density with the SCIAMACHY HCHO column density in a scatter plot. Due to SCIAMACHY's poor daily coverage – it needs 3-6 days for global coverage [13] – and the time restrictions – the aircraft measurements should be within two and a half hours from the overpass time of SCIAMACHY – only two matches were found. The two points are located above the diagonal $y = x$ line, which means the column values retrieved by SCIAMACHY are higher than the measured values. The two matches are also found in Table 1, along with the corresponding date, flight label, longitude, latitude and time in UTC. The aircraft measurements have been performed well before the overpass time. As HCHO is mainly produced during daytime, the time difference provides a good explanation for the difference between the SCIAMACHY and the aircraft column densities.

Table 1: Detailed information about the column densities of the two flights which found a match with the SCIAMACHY satellite. From left to right, the columns represent the date (day/month), the profile label, the aircraft column density and the computed uncertainty, SCIAMACHY's column density and its uncertainty, the average longitude and latitude, the time in UTC of the aircraft and the time in UTC of the SCIAMACHY measurement.

Date	Profile	Aircraft Column	Aircraft Uncer-tainty	Satellite Column	Sat. Uncer-tainty	Lon	Lat	Aircr. UTC	Sat. UTC
18/8	D30a	10.85	2.20	16.78	2.61	-1.03	12.43	07:43	09:55
15/8	B27	5.32	2.01	9.12	2.21	1.56	10.68	07:51	09:50

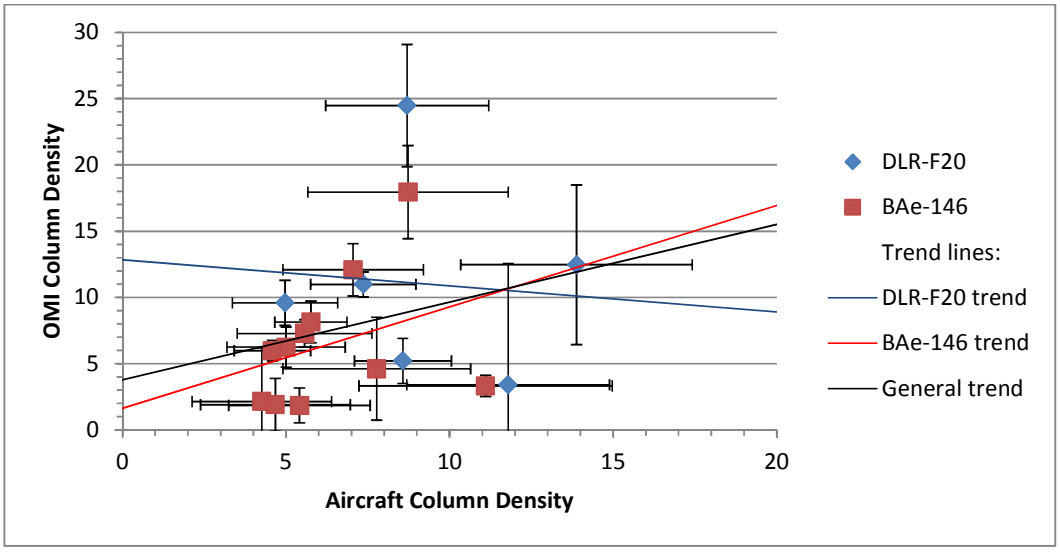

Figure 10: The computed aircraft HCHO column density versus the OMI HCHO column density (in 10^{15} molec/cm^2). The blue diamonds are the DLR-F20 column densities and the red squares are the BAe-146 column densities. The blue and red lines are the trend lines of the satellite when compared to the DLR-F20 and BAe-146 aircraft respectively. The black line is the trend line of the whole data set.

Figure 10 shows the scatter plot of the computed HCHO column density based on the measurements of the DLR-F20 and BAe-146 aircraft versus the HCHO column density retrieved by OMI. These measurements can also be found in Table 2, along with the corresponding date, flight label, longitude, latitude and time in UTC. In the graph, two trend lines are plotted; the red one represents the trend in the comparison of OMI measurements with the BAe-146 aircraft measurements and the blue one represents the trend in the comparison with the DLR-F20 aircraft. The black line is the trend line of the whole data set. The trend lines have been calculated by fitting the data sets using linear regression without error weighting.

Table 2: Detailed information about the column densities of the flights which found a match with the OMI satellite. From left to right, the columns represent the date (day/month), the profile label, the aircraft column density and the computed uncertainty, OMI's column density and its uncertainty, the average longitude and latitude and the time in UTC.

Date	Label	Aircraft Column	Aircraft Uncertainty	Satellite Column	Satellite Uncertainty	Lon	Lat	UTC
6/8	D18b	11.79	3.10	3.40	9.17	-0.79	12.28	12:10
13/8	D25b	7.36	1.61	10.98	0.94	-1.63	12.61	13:29
15/8	D27b	4.97	1.61	9.59	1.71	-1.97	12.50	11:43
16/8	D28a	8.70	2.49	24.48	4.61	-0.98	10.81	12:28
16/8	D28b	13.88	3.54	12.46	6.02	-1.88	12.06	16:30
18/8	D30c	8.57	1.48	5.21	1.70	8.10	32.23	12:06
20/7	B1a	5.41	2.16	1.85	1.32	1.08	15.59	14:02
20/7	B1b	5.57	2.06	7.28	1.03	1.47	14.55	14:51
20/7	B1c	5.00	1.81	6.25	1.52	1.43	14.63	15:25
21/7	B2	11.10	3.87	3.32	0.79	1.10	15:86	13:58
22/7	B3a	7.05	2.15	12.08	1.98	1.05	16.15	15:09
22/7	B3b	8.73	3.06	17.94	3.52	2.57	11.02	16:58
25/7	B6a	7.77	2.87	4.61	3.88	2.51	11.98	12:42
25/7	B6b	4.26	2.13	2.15	4.04	2.56	11.85	17:42
30/7	B11	5.76	1.10	8.14	1.58	2.34	6.97	12:37
6/8	B18a	4.58	1.17	5.98	0.78	5.93	16.76	14:27
6/8	B18b	4.67	2.29	1.90	2.00	2.40	12.97	12:18

The blue DLR-F20 trend line shows that the expected satellite column density decreases as the aircraft column density increases. This is counter intuitive. One would expect that the satellite column density increases as the aircraft column density increases. The fit has a slope of -0.20 and an intercept of 12.83 (10^{15} molec./cm^2). The Pearson correlation coefficient for this measurement set is -0.08, meaning the data sets have almost no correlation. Because there are only six matches between the DLR-F20 aircraft and the OMI satellite, we can safely assume this coefficient has a huge uncertainty. Leaving out the match with profile D18b, the correlation coefficient changed from -0.08 to $+0.14$, which underlines the big margin in the coefficient. This was done to test the robustness of the regression.

The red BAe-146 trend line shows that the expected satellite column density increases as the aircraft column density increases with a slope of 0.77. The intercept of the fit is 1.64 (10^{15} molec./cm^2). The slope is close to 1 and the intercept is close to 0, meaning the BAe-146 and the OMI column densities match very well. The Pearson coefficient of the set has a value of $+0.33$, which means the correlation between OMI measurements and the BAe-146 aircraft is bigger than the correlation between OMI and the DLR-F20 aircraft. This could be the result of having more matches, meaning random errors have less influence. It could also be the result of less difference in the time of the measurements and of a smaller distance between the center of the satellite pixel and the average location of the aircraft measurements.

When all data is combined, the Pearson coefficient is 0.27. The black trend line has a slope of 0.59 and an intercept of 3.78 (10^{15} molec./cm^2).

Conclusion and outlook

All profiles established from the aircraft measurements show that the HCHO mixing ratio decreases as the altitude increases. One profile, B25b, is different as it shows a much larger HCHO mixing ratio at a pressure of 700 hPa than at 1000 hPa. This is of biomass burning.
When comparing the DLR-F20 and BAe-146 column densities quantitatively, we see that the downward extrapolation of the DLR-F20 aircraft is greater than the downward extrapolation of the BAe-146 aircraft. The upward extrapolation of the BAe-146 aircraft is greater than the upward extrapolation of the DLR-F20 aircraft. This can be related directly to the altitude at which the aircraft have performed measurements. In total, the DLR-F20 column densities seem more reliable due to the lower total extrapolated fraction.
If we compare the aircraft data with SCIAMACHY, we only find two matches. The two matches suggest that the column density retrieved by SCIAMACHY is higher than the column density computed using the average of the introduced extrapolation methods. If we compare the aircraft data with OMI, more matches are found. When compared to OMI, the six DLR-F20 column densities return a slope of -0.20, an intercept at 12.83 (10^{15} molec./cm^2) and a Pearson correlation coefficient of -0.08, which tells us that the spread in comparison points is big and it has a negative slope. This might be caused by having only six reference points. Removing the first row in Table 1 results in a positive Pearson coefficient, which indicates that having more comparison points might give a better result.
When compared to OMI, the eleven BAe-146 column densities return a slope of 0.77, an intercept at 1.64 (10^{15} molec./cm^2) and a Pearson correlation coefficient of 0.33. Not only does this have a positive slope close to 1 and an intercept close to 0, it is also more linear than the comparison of OMI with the DLR-F20. This could be the result of having more matches, meaning random errors have less influence.
When comparing all aircraft data with OMI, the fit returns a slope of 0.59, an intercept at 3.78 (10^{15} molec./cm^2) and a Pearson correlation coefficient of 0.27. Although it has an upward slope, the data points still do not show a strong correlation.

The matches have all been weighted equally. However, the distance, both in space and time, is different for each match and each pixel. The accuracy of this study can be further improved by taking into account the distance from the pixel in spacetime to the aircraft column densities. The difference in displacement of the aircraft has also been ignored. Giving less weight to matches with a large displacement profile could help to improve the trend lines.
The accuracy of the aircraft column densities suffer mostly from the extrapolation. Higher accuracy in the aircraft column densities can be achieved by using data of an aircraft which flies as low and as high as possible.

Bibliography

[1] G. Verkerk, J. Broens, R. Bouwens, P. de Groot, W. Kranendonk, M. Volgelezang, J. Westra and I. Wevers-Prijs, BINAS, 5 ed., Groningen: Wolters-Noordhoff, 2004, p. 6A.

[2] A. Guenther, C. N. Hewitt, D. Erickson, R. Fall, C. Geron, T. Graedel, P. Harley, L. Klinger, M. Lerdau, W. McKay, T. Pierce, B. Scholes, R. Steinbrecher, R. Tallamraju, J. Taylor and P. Zimmerman, "A global model of natural volatile organic compound emissions," *Journal of Geophysical Research,* pp. 8873-8892, 1995.

[3] J. Ferreira, C. E. Reeves, J. G. Murphy, L. Garcia-Carreras, D. J. Parker and D. E. Oram, "Isoprene emissions modelling for West Africa: MEGAN model evaluation and sensitivity analysis," *Atmospheric Chemistry and Physics,* vol. 10, pp. 8453-8467, 2010.

[4] M. Saunois, C. Reeves, C. Mari, J. Murphy, D. Stewart, G. Mills, D. Oram and R. Purvis, "Factors controlling the distribution of ozone in the West African lower troposphere during the AMMA (African Monsoon Multidisciplinary Analysis) wet season campaign," *Atmospheric Chemnistry and Physics,* no. 9, pp. 6135-6155, 2009.

[5] I. De Smedt, J.-F. Müller, T. Stavrakou, R. van der A, H. Eskes and M. Van Roozendael, "Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors," *Atmospheric Chemistry and Physics,* vol. 8, pp. 4947-4963, 2008.

[6] N. L. Boeke, J. D. Marshall, S. Alvarez, K. V. Chance, A. Fried, T. P. Kurosu, B. Rappenglück, D. Richter, J. Walega, P. Weibring and D. B. Millet, "Formaldehyde columns from the Ozone Monitioring Instrument: Urban versus background levens and evaluation using aircraft data and a global model," *Journal of Geophysical Research,* vol. 116, 2011.

[7] TEMIS, "TEMIS - Tropospheric N20," KNMI, 07 May 2012. [Online]. Available: http://www.temis.nl/airpollution/no2.html. [Accessed 05 June 2012].

[8] K. F. Boersma, H. J. Eskes, J. P. Veefkind, E. J. Brinksma, R. J. van der A, M. Sneep, G. H. J. van den Oord, P. F. Levelt, P. Stammes, J. F. Gleason and E. J. Bucsela, "Near-real time retrieval of tropospheric NO2 from OMI," *Atmospheric Chemnistry and Physics,* no. 7, pp. 2103-2118, 2007.

[9] K. F. Boersma, D. J. Jacob, H. J. Eskes, R. W. Pinder, J. Wang and R. J. van der A, "Intercomparison of SCIAMACHY and OMI tropospheric NO2 columns: Observing the diurnal evolution of chemnistry and emissions from space," *Journal of Geophysical Research,* no. 113, 2008.

[10] D. J. Jacob, Introduction to atmospheric Chemnistry, Princeton, 1999, pp. 3-11.

[11] J. H. Seinfeld and S. N. Pandis, Atmospheric Chemistry and Physics, Wiley, 2006.

[12] J. Murphy, D. Oram and C. E. Reeves, "Measurements of volatile organic compounds over West Africa," *Atmospheric Chemnistry and Physics,* no. 10, pp. 5281-5294, 2010.

[13] "SCIAMACHY Measurement Modes," IUP Bremen, [Online]. Available: http://www.iup.uni-bremen.de/sciamachy/instrument/modes/index.html. [Accessed 5 August 2012].

Appendix

Detailed tables on the aircraft measurements

Table 3: Detailed information on the BAe-146 column density as visualised in Figure 8. From left to right, the columns represent the date (day/month), the profile label, the column density in 10^{15} molec./cm^2, the four fractions as explained in the Results and Discussion section, the average longitude and latitude and the time in UTC time.

Date	Profile	Column	f2-	f-	f+	f2+	Lon	Lat	UTC
20/7	B1a	6,31	0,38	0,05	1,42	1,42	1.08	15.59	14:02
	B1a	5,41	0	0,24	2,13	0			
20/7	B1b	6,59	0,74	0,24	1,30	1,30	1.47	14.55	14:51
	B1b	5,57	0	0,60	1,95	0			
20/7	B1c	5,74	0,29	0,10	1,19	1,19	1.43	14.63	15:25
	B1c	5,00	0	0,24	1,79	0			
21/7	B2	12,90	1,10	0,33	2,50	2,50	1.10	15.86	13:58
	B2	11,10	0	0,88	3,75	0			
22/7	B3a	8,18	0,92	0,19	1,35	1,35	1.05	16.15	15:09
	B3a	7,05	0	0,65	2,03	0			
22/7	B3b	10,12	0,88	0,57	1,91	1,91	2.57	11.02	16:58
	B3b	8,73	0	1,01	2,87	0,00			
25/7	B6a	9,31	1,29	0,34	1,79	1,79	2.51	11.98	12:42
	B6a	7,77	0	0,98	2,68	0			
25/7	B6b	5,35	1,04	0,76	1,13	1,13	2.56	11.85	17:42
	B6b	4,26	0	1,27	1,70	0			
27/7	B8	11,66	1,23	0,33	3,60	3,60	1.37	8.84	11:23
	B8	9,25	0	0,95	5,39	0			
30/7	B11	6,16	0,12	0,11	0,70	0,70	2.34	6.97	12:37
	B11	5,76	0	0,17	1,04	0			
6/8	B18a	5,04	0,15	-0,02	0,77	0,77	5.93	16.76	14:27
	B18a	4,58	0	0,06	1,15	0			
6/8	B18b	5,53	0,21	0,16	1,51	1,51	2.40	12.97	12:18
	B18b	4,67	00	0,27	2,27	0			
13/8	B25a	8,42	1,28	0,45	1,48	1,48	2.55	9.79	07:38
	B25a	7,04	0	1,09	2,23	0			
13/8	B25b	15,17	0,40	-0,06	3,82	3,82	1.67	5.77	08:46
	B25b	13,06	0	0,13	5,74	0			
13/8	B25c	7,78	0,26	0,06	0,89	0,89	2.18	9.74	10:16
	B25c	7,21	0	0,19	1,33	0			
13/8	B25d	14,69	1,56	0,39	3,30	3,30	2.16	7.50	08:01
	B25d	12,26	0	1,17	4,95	0			
14/8	B26	12,54	0,11	0,06	3,43	3,43	2.25	12.70	04:18
	B26	10,77	0	0,11	5,15	0			
15/8	B27	6,67	1,63	0,41	1,06	1,06	1.56	10.68	07:51
	B27	5,32	0	1,23	1,59	0			

Table 4: Detailed information on the DLR-F20 column density as visualised in Figure 8. From left to right, the columns represent the date (day/month), the profile label, the column density in 10^{15} molec./cm^2, the four fractions as explained in the Results and Discussion section, the average longitude and latitude and the time in UTC time. If the longitude and latitude are colored, it means the displacement was more than one degree longitude or latitude.

Date	Profile	Column	f2-	f-	f+	f2+	Lon	Lat	UTC
1/8	D13a	8,89	0,18	0,27	0,04	0,04	9.41	42.92	03:03
	D13a	8,05	0	0,17	0,05	0			
1/8	D13b	8,08	0,22	0,21	0,10	0,10	-7.38	31.34	06:24
	D13b	7,02	0	0,10	0,14	0			
1/8	D13c	11,10	-0,009	0,14	0,08	0,08	-7.75	28.27	07:53
	D13c	10,72	0	0,15	0,12	0			
1/8	D13d	10,83	0,22	0,23	0,11	0,11	-1.85	12.84	10:58
	D13d	9,34	0	0,12	0,16	0			
4/8	D16a	10,61	0,11	0,17	0,13	0,13	-1.37	11.47	07:58
	D16a	9,44	0	0,11	0,20	0			
4/8	D16b	14,65	0,16	0,18	0,04	0,04	-1.53	12.54	11:51
	D16b	13,37	0	0,10	0,06	0			
6/8	D18a	8,76	0,16	0,26	0,10	0,10	-0.74	11.03	08:37
	D18a	7,77	0	0,18	0,15	0			
6/8	D18b	13,46	0,13	0,16	0,15	0,15	-0.79	12.28	12:10
	D18b	11,79	0	0,10	0,23	0			
7/8	D19	10,84	0,16	0,23	0,10	0,10	-1.69	11.52	11:11
	D19	9,66	0	0,15	0,14	0			
13/8	D25a	5,75	0,21	0,38	0,07	0,07	-1.53	10.48	09:37
	D25a	5,08	0	0,26	0,10	0			
13/8	D25b	7,77	0,09	0,24	0,02	0,02	-1.63	12.61	13:29
	D25b	7,36	0	0,19	0,03	0			
15/8	D27a	6,90	0,17	0,30	0,10	0,10	-1.73	10.63	08:14
	D27a	6,10	0	0,21	0,14	0			
15/8	D27b	5,92	0,30	0,45	0,12	0,12	-1.97	12.50	11:43
	D27b	4,97	0	0,27	0,16	0			
16/8	D28a	10,25	0,22	0,28	0,15	0,15	-0.98	10.81	12:28
	D28a	8,70	0	0,16	0,22	0			
16/8	D28b	15,24	-0,13	0,15	0,05	0,05	-1.88	12.06	16:30
	D28b	13,88	0	0,24	0,08	0			
18/8	D30a	12,01	0,16	0,24	0,06	0,06	-1.03	12.43	07:43
	D30a	10,85	0	0,16	0,09	0			
18/8	D30b	7,07	0,02	0,12	0,06	0,06	7.14	32.41	11:09
	D30b	6,79	0	0,11	0,10	0			
18/8	D30c	9,26	0,09	0,15	0,07	0,07	8.10	32.23	12:06
	D30c	8,57	0	0,11	0,10	0			