

*Investigations on the Smile-Effect for Airborne
Spectroscopy of Methane and Carbon Dioxide*

Bachelor Thesis

*Institute of Environmental Physics (IUP)
University of Bremen, Germany*

Submitted by: Hannes Schnisa
Email: hschnisa@iup.physik.uni-bremen.de

Supervisor: Dr. Michael Buchwitz

Examiners: Prof. John P. Burrows
Prof. Justus Notholt

Submitted on: 30.06.2009

Danksagung

Im Folgenden möchte ich einigen Personen danken, die zum Gelingen dieser Arbeit beigetragen haben.

Vielen Dank an alle Menschen, die in meinem Umfeld gearbeitet haben, denn sie haben es geschafft eine angenehme Arbeitsatmosphäre zu schaffen, so dass ich die Zeit der Bachelor-Arbeit mit viel Spaß und Freude erlebt habe.

Insbesondere möchte ich mich bei Dr. Oliver Schneising und Ralf Bauer bedanken. Einen Monat lang teilten wir ein Büro und während dieser Zeit waren sie stets bereit mir Fragen zu IDL und Linux zu beantworten. Außerdem möchte ich mich bei Konstantin Gerilowski und Thomas Krings bedanken, die diese Aufgabe für das MAMap Projekt übernommen haben. Ohne sie alle hätte ich nicht so effektiv arbeiten können.

Ganz besonderer Dank gilt meinem Betreuer Dr. Michael Buchwitz, der mir nicht nur während der Auswertungsphase mit Rat und Tat zur Seite stand, sondern mich auch bei der Strukturierung und Ausarbeitung der Thesis unterstützt hat.

Abschließend einen Dank an meine Familie und alle Menschen, die mich in meinem Leben begleitet haben. Ohne sie würde ich nicht an diesem wichtigen Punkt meines Lebens stehen: dem Bachelor-Abschluss in Physik.

Content

<i>1. Abstract</i>	4
<i>2. Introduction and motivation</i>	5
<i>3. Foundations and theory</i>	7
3.1 Remote sensing based on spectroscopic measurements	7
3.2 Spectrometers and the Smile-Effect	9
3.3 Atmospheric radiative transfer	10
3.4 Inverse methods.....	11
4.1 The MAMap Project	17
4.2 The MAMap instrument.....	18
4.2 Previous work on MAMap data analysis	23
4.3 MAMap and the Smile-Effect	26
<i>5. Analysis of MAMap data</i>	30
5.1 Objectives of this thesis and selected approach	30
5.2 Discussion of CCD signals of the oxygen (O ₂ -A-band) spectrometer	31
5.3 Definition of Inhomogeneity Indices (IH _i).....	33
5.4 Correlation analysis of Inhomogeneity Indices with GHG retrievals	36
5.5 Discussion of the Results	44
<i>6. Conclusion and outlook</i>	48
<i>A. References</i>	50

1. Abstract

Carbon dioxide (CO_2) and methane (CH_4) are the two most important anthropogenic greenhouse gases (GHG) contributing to global warming. Despite their importance, our knowledge about their surface sources and sinks has significant gaps, which limits the reliability of climate predictions. Most of our knowledge about their surface fluxes stems from accurate but sparse surface observations. Only recently global satellite data have become available, albeit only at coarse horizontal resolution (~ 50 km). In order to fill the gap between the in-situ point measurements and the low spatial resolution satellite observations, the Methane Airborne Mapper (MAMap) instrument has been built at IUP, University of Bremen, in cooperation with the GFZ Potsdam. MAMap has been specified to retrieve CH_4 and CO_2 at a spatial resolution of ~ 100 m. First test flights have been performed and the in-flight data have been analyzed using an initial retrieval algorithm, which permits the retrieval of total columns of CH_4 and CO_2 with sufficient accuracy to assess the instrument performance. It has been found that the instrument is very stable and that the noise of the spectral measurements is as specified. It has however also been found that the standard deviation of the retrieved CH_4 and CO_2 columns and column ratios are often a factor of 3 larger than expected ($\sim 1\%$ for an exposure time of 58 ms). A possible explanation for this could be retrieval errors caused by, for example, unaccounted variations of the instrument line-shape (ILS) function due to scene inhomogeneity caused by optical effects such as the “Smile Effect”. As the SWIR spectrometer (~ 1600 nm) used for the GHG retrievals is only equipped with a linear detector array, it is not possible to obtain across flight track scene information. MAMap is however also equipped with a similarly designed second spectrometer covering the 760 nm spectral region (O_2 A-band) with a 2D CCD array which provides this information. Therefore, an analysis of the O_2 A-band spectra can in principle be used to assess to what extent the GHG retrievals are affected by scene inhomogeneity. This has been investigated in this thesis. “Inhomogeneity Indices” (IH_i) have been defined and derived from the O_2 -A-band spectra, which have been correlated with the independent GHG retrievals. It was expected that because of the 3 times larger variability up to about 89% of the retrieved GHG variability could be explained by scene inhomogeneity. This would imply a maximum correlation coefficient of $R = 0.94$ between the IH_i and the GHG retrievals (as $R^2 = 0.89$ can be interpreted as “explained variance”). The MAMap data have been analyzed using cross-correlation techniques applying different quality filter to the retrieved GHG columns and column ratios. The correlation coefficient was less than 0.3 in most cases, i.e., the expected significant correlation has not been found. Possible reasons for this finding are discussed.

2. Introduction and motivation

Our climate system is driven by radiation of the sun, properties of the earth's surface and the atmosphere. The radiation budget in the atmosphere highly depends on its chemical composition. At this juncture the main components oxygen and nitrate are having less relevance although they represent about 99% of the mass of the atmosphere. Gases with much more relevance are water vapor (H_2O), carbon dioxide (CO_2) and methane (CH_4). These gases absorb and emit the sun's radiation within the thermal infrared range and therefore are the fundamental cause of the greenhouse effect. That is why they are called greenhouse gases (GHG). The relative impact of each of these gases is different, with CH_4 20 times more effective than CO_2 (Wallace and Hobbs, 2006). If methane accounts for 20% only of the overall greenhouse radiative forcing this is primarily because of the different atmospheric mixing ratios: 380ppm on average for CO_2 and 1.75ppm on average for CH_4 . Due to human activities, the global mixing ratios of CO_2 and CH_4 have risen to levels never recorded in the last 650,000 years (Forster et al., 2007).

Therefore, to predict future climate changes caused by the enhanced greenhouse effect and to evaluate the degree of anthropogenic impact it is necessary to characterize the global distribution of greenhouse gases' sources and sinks.

This bachelor thesis is based on the master thesis "Mapping of Tropospheric Greenhouse Gases using Airborne Near-Infrared/Shortwave-Infrared Spectroscopy" from Pietro P. Bertagnolio (Bertagnolio, 2008). His work gives an introduction to the Methane Airborne Mapper (MAMap) Project, which is capable of providing information on the mixing ratios of methane (CH_4) and carbon dioxide (CO_2) on a scale between ground-based local measurements and satellite-based datasets.

Bertagnolio (2008) analyzed three different flights where the MAMap instrument measured CH_4 and CO_2 at an anthropogenic carbon dioxide source (Schwarze Pumpe coal power plant), a natural methane source (Zarnekow wetlands) and an artificial methane source (Ketzin experiment). He found that the standard deviation of the retrieved CH_4 and CO_2 columns and column ratios are often ~3-4% (for an exposure time of 58 ms), which is about a factor of 3 larger than expected, based on simulated retrievals. He also concluded that the "intensity of the incoming radiance (...) has the largest impact in determining the precision of the fit results" and that this "implies that the illumination of the target scene is a crucial factor for the planning and the measurement analysis of the flight missions" (Bertagnolio, 2008:112).

The conclusions of Bertagnolio (2008) concerning the importance of the level of the incoming radiance are consistent with simulated retrievals: the lower the radiance, the lower the signal-to-noise ratio, and the lower the GHG column retrieval precision. The radiance level is mainly determined by the surface reflectivity and the solar zenith angle. Apart from higher noise due to lower signal, radiance variability due to variations of the surface reflectivity is not a major problem for the GHG retrievals, as the WFM-DOAS retrieval algorithm considers radiance fluctuations due to surface reflectivity by including a low order spectral polynomial in the spectral fitting procedure. WFM-DOAS is based on fitting a linearized radiative transfer model to the measured spectra. The algorithm assumes that the instrument line-shape (ILS) function of MAMAP is known and constant. If the ILS would vary from measurement to measurement this would result in high-frequency variations of the retrieved GHG columns in addition to random errors due to instrument noise. Significant variability of the ILS could be caused by inhomogeneously reflecting scenes, where the surface reflectivity varies especially in the across-track direction. In this case the spectrometer entrance slit will not be homogeneously illuminated. Depending on how the slit will be illuminated, the ILS is supposed to change its shape, its width, and its spectral position. These ILS variations are not considered for current MAMAP GHG retrievals. It cannot be ruled out that the observed GHG variability of 3-5%, which is about a factor 3 higher than the expected retrieval precision of 1-2% estimated from simulated retrievals, is to a large extent caused by unaccounted ILS variations due to scene inhomogeneity via the well known optical Smile- and Keystone-distortion effects.

The focus of this thesis is to investigate if the observed higher than expected variability of the retrieved GHG columns and column ratios is due to scene inhomogeneity. Unfortunately this cannot be directly be investigated using MAMap's GHG channel, as this channel consists of a 1D linear detector array which does not permit do obtain sub-scene across-track information. Fortunately, MAMap is also equipped with a second (nearly) identically designed spectrometer covering the O₂-A-band spectral region at 760 nm with a 2D CCD detector, with higher spatial resolution and sampling in the across-track direction. These 2D CCD measurements enable to investigate the Smile and Keystone-distortion effects and to derive "Inhomogeneity Indices" (IH_i) for each measurement. In this thesis the O₂-A-band CCD readouts have been analyzed by deriving Inhomogeneity Indices and by correlating them with the GHG retrievals in order to find out if the expected correlation can be observed and if it is possible to determine to what extent the observed unexplained GHG excess variability can be explained by scene inhomogeneity.

3. Foundations and theory

The Methane Airborne Mapper (MAMap) is a spectrometer designed to measure the concentration of atmospheric CH_4 and CO_2 via their spectral absorption features. In order to understand the information, which is included in the MAMap spectral measurements, it is necessary to understand how remote sensing spectrometry works (presented in Chapter 3.1, based on Bertagnolio, 2008) and how radiative transfer in the atmosphere influences the measurements (presented in Chapter 3.3). The relation between spectrometers and the Smile-Effect is presented in Chapter 3.2 and inversion methods, used in remote spectroscopy, are discussed in Chapter 3.4, which is based on Rodgers (2000), Schlitzer (Introduction to Inverse Methods and Mathematical Data Analysis, personal communication, 2006), Richter (Introduction to Measurement Techniques in Atmospheric Physics: Differential Optical Absorption Spectroscopy (DOAS), personal communication, 2007) and Buchwitz et al. (2000).

3.1 Remote sensing based on spectroscopic measurements

Remote sensing is a method to gain information about the earth surface or other objects, which are not directly accessible. Measuring natural radiation that is emitted or reflected by the object or the surrounding area being observed does this. Electromagnetic radiation serves as the information carrier.

A distinction is drawn between passive and active sensors, both capable of analyzing a huge area of the electromagnetic spectra. Passive sensors measure the sun radiation reflected or emitted from the earth surface and/or the atmosphere. Active systems, on the other side, emit micro- or laser radiation and collect the reflected parts.

The advantage of remote sensing spectrometry is the possibility to measure in areas otherwise difficult to access. Automated instruments, used in remote sensing processes, are able to collect large time series and covering large areas and so replace costly and slow data collecting ground instruments. It also provides information on several different physical quantities.

The disadvantage of remote sensing is the fact that only indirect measurements can be gathered, so that several processes and factors cannot be excluded. This makes it very difficult

to interpret the measurements accurately without the help of different models and assumptions. All in all remote sensing is a very powerful tool, which is used in a variety of scientific areas like earth-, weather-, atmospheric- and ocean observation.

The practical way to measure the radiation transmitted through the atmosphere at different wavelengths is by separating the different spectral components using a dispersing element. Due to its elevated resolving power and its stability, the diffraction grating is the most common dispersing element used in atmospheric spectroscopy, together with Fourier Transform spectrometers. A grating is a reflecting or transmitting surface carved with several (n) grooves called rules, each of them independently diffracting light, at a distance g from each other. The different light waves coming from the different grooves then interfere with each other, and their interference maxima occur at angles α so that

$$\sin \alpha = \frac{m\lambda}{g},$$

where m is an integer called the diffraction order. At different dispersion angles α then different wavelengths λ will have their interference maxima, and can then be dispersed on a surface.

Several detectors can be used then, to measure the dispersed radiation at different wavelengths, for example the photodiode array and the couple-charged device. A photodiode array is a linear detector formed by a series of adjacent photodiodes (semiconductor diodes) with a surface exposed to the incoming photons. There are two kinds of photodiode arrays. In the first one, when the detector absorbs photons, they create an electron-hole pair that drifts to the next p-n junction and discharges it. After a time interval, called exposure time, the detector is read out by sequentially charging the capacitors corresponding to each pixel, and the current needed is proportional to the number of photons absorbed in the exposure time. The second kind instead works by having the junctions depleted of charge, and then measuring the current produced when the electrons created by photon absorption are read out. There are photodiode arrays with 256 to 2048 pixels, which allow measuring several wavelengths at the same time.

Another kind of radiation detector that can measure several wavelengths is the charge-coupled device or CCD. It is a two-dimensional array of detector pixels (with size ranging from 256 to 4096 per each side), also exploiting the properties of semiconductors. When photons hit one of the pixels, the resulting electrons are collected in the corresponding

uncharged depletion zone. The readout then consists in shifting the charges sequentially from row to row. The lowest row is then readout and digitized. The main advantages of this kind of detectors are the high sensitivity to radiation and its 2D shape that allows measuring several wavelengths, on one axis, coming from different viewing angles, on the other axis, effectively creating a multi-spectral imaging sensor. A drawback is that the capacity of the single pixels is lower than that of other sensors, so shorter exposure times are needed not to fill the holes with electrons (saturate). Moreover, a long time is needed to read out sequentially each different pixel that can last up to several seconds.

3.2 Spectrometers and the Smile-Effect

The MAMap is equipped with a spectrometer covering the 760nm spectral region (O₂ A-band) with a 2D CCD array. Those rectangular photo detectors have pixels arranged in rows (parallel to spatial axis defined by a straight slit) and columns (parallel to the spectral axis). The light, which is collected by the photo detector, enters the spectrometer through a slit. The point or pixel along one spatial axis corresponds to a point or pixel along the slit in the scene, which is observed. Thus each point or pixel on a line that crosses the scene gives readout of the spectrum for each column of pixels. Figure 1a shows a spectrometer, which is similar to the MAMap O₂-A-Band spectrometer. Those types of spectrometers are called “Czerny-Turner”-spectrometers.

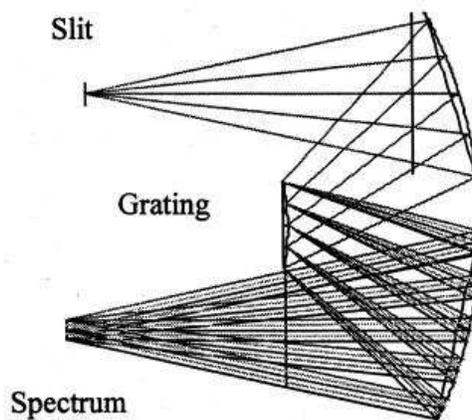


Figure 1a: Schematic view of a "Czerny-Turner"-spectrometer. (Fischer et al. 1998)

It is important to have an exact configuration for extracting detailed quantitative information from the readouts. To achieve a constant length of monochromatic slit image and a constant length of the spectrum remain it is necessary to adjust the monochromatic image of the slit straight and parallel to the rows of the photo detector array at any wavelength and to adjust the spectrum of any point along to the slit as straight and parallel as possible.

When such conditions are achieved, then, deviations are donated as spectral and spatial distortion, which are called Smile- and Keystone-Distortion or Smile- and Keystone-Effect.

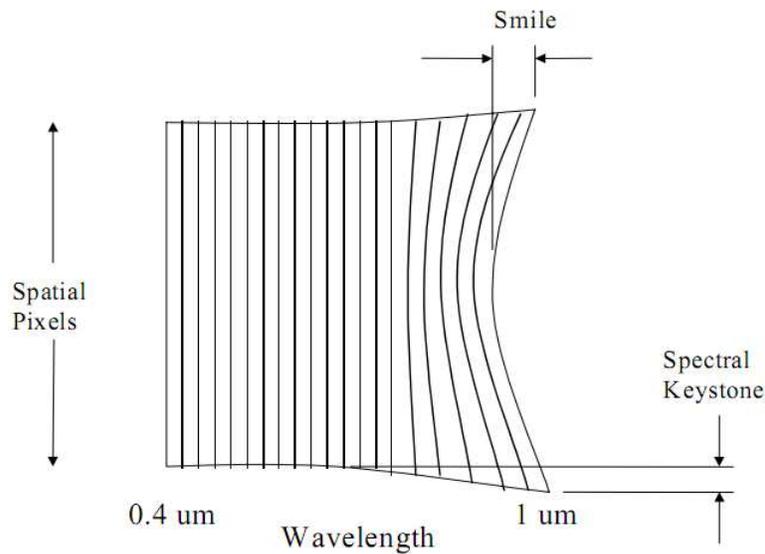


Figure 1b: Definition of Smile and Keystone Distortion. (Fischer et al. 1998)

Smile is a change in dispersion with field position, and spectral keystone is a change in magnification with wavelength. Spectral keystone involves mixing of spectra from adjacent field positions. (Fischer et al. 1998)

3.3 Atmospheric radiative transfer

Electromagnetic radiation is used in remote sensing because the characteristics of interaction with matter are well known. While the radiation moves through the atmosphere the photons can be absorbed by atoms and molecules, there is elastic and inelastic scattering within the interaction between radiation and molecules or aerosols. There is stimulated and thermal emission and geometric reflection by macroscopic surfaces. All these processes are dependent

on the radiation wavelength and therefore it is possible to differ between the effects and identify their components (Bertagnolio, 2008:16).

An example is the optical spectrum of the sun. Fraunhofer found dark lines in that spectrum, which are the absorption lines of specific elements. Each element absorbs electromagnetic radiation at a specific wavelength so that every element has a unique spectrum. This makes it possible to get detailed information about the composition of an object one is looking at, e.g. the sun or trace gases in the earth's atmosphere.

3.4 Inverse methods

Airborne spectroscopic measurements of the atmospheric composition, like all remote observations, are inherently indirect. The parameter of interest (here the composition of the atmosphere) is not directly measured, must be inferred from the measured spectrum. The measured spectrum must depend of the parameters of interest in a known way (here via forward radiative transfer simulation). This inference, which is the inversion of the radiative transfer simulation, is often complex, and belongs to the class of problems called inverse problems. Rodgers (2000) defines them as the question of finding the best representation of the required parameter given the measurements made, together with any appropriate prior information that may be available about the system and the measuring device.

The general inverse problem can be regarded as a question of setting up and solving a set of simultaneous linear or non-linear equations, in the presence of experimental error of some of the parameters, the measurements, and quite possibly in the presence of approximations in the formulations of the equations.

The quantities to be retrieved can be represented by a state vector, \bar{x} with n elements x_1, x_2, \dots, x_n . It could represent a profile of some quantity given at a finite number of levels, or any set of relevant variables, such as coefficients for another representation, or decomposition of the profile itself, or again, as in the case of MAMap, it may include a range of different types of parameters, like the vertical profile scaling parameter of different gases, a temperature profile shift, and polynomial coefficients for the low-frequency spectrum.

The quantities actually measured in order to retrieve, \bar{x} can be represented by a measurement vector \bar{y} , with m elements y_1, y_2, \dots, y_m . This vector should include all the quantities measured that are functions of the state vector. Measurements are made to a finite accuracy. Random error or *measurement noise* will be denoted by the vector $\bar{\epsilon}$.

For each state vector there is a corresponding ideal measurement vector \vec{y}_{mod} , determined by the physics of the measurement. The physical details are approximated by a *forward model* $\vec{F}(\vec{x})$, so that

$$\vec{y} = \vec{F}(\vec{x}) + \vec{\epsilon}$$

to construct a forward model we must of course understand how the quantity measured, that is the absorbed solar infrared radiation, is related to the quantity that is really wanted, in this case the vertical column of CO_2 or CH_4 .

The quantities to be retrieved in most inverse problems are continuous functions, while the measurements are always of discrete quantities. Thus most inverse problems are formally ill-posed or under constrained in this trivial sense. This is simply dealt with by replacing the truly continuous state function, corresponding to an infinite number of variables, with a representation in terms of a finite number of parameters. After discretisation the problem may or may not be under constrained, depending on the information content of the measurement.

Linear least-squares method

The simplest way to address a problem is to start with its linear form. A linearization of the forward model about a reference state \vec{x}_0 is adequate provided that $\vec{F}(\vec{x})$ is linear within the error bounds of the retrieval. When we write

$$\vec{y} - \vec{F}(\vec{x}) = \left. \frac{\partial \vec{F}(\vec{x})}{\partial \vec{x}} \right|_{\vec{x}_0} (\vec{x} - \vec{x}_0) + \vec{\epsilon} = K(\vec{x} - \vec{x}_0) + \vec{\epsilon}$$

we define a $m \times n$ weighting function matrix $K = \partial \vec{F}(\vec{x}) / \partial \vec{x}$ not necessarily square, in which each element is the partial derivative of a forward model element with respect to a state vector element. If $m < n$ the equations are described as under constrained (or ill-posed or under-determined) because there are fewer measurements than unknowns. Similarly if $m > n$ the equations are often described as over constrained or over-determined, as long as all the equations are linearly independent and carry enough independent information.

The term *weighting function* is peculiar to the atmospheric remote sensing literature, and it arose because in the early applications of nadir sounding for temperature the forward model takes the form of a weighted mean of the vertical profile of the Planck function.

For a fundamentally over constrained problem, like that of the MAMap total column retrieval (which assumes sufficiently good knowledge or little impact of the vertical profile shape),

where the measurement vector has considerably more elements than the state vector, and the algebraic form of the model is known from sound physical reasoning, an appropriate approach is the *least-squares method*.

In the case where there are more measurements than unknowns, an exact solution is not possible in general. Therefore we look for a solution that minimizes the sum of the squares of the differences between the actual measurements and those calculated from the forward model using the solution. That is, we minimize:

$$[\bar{y} - \bar{F}(\bar{x})]^T [\bar{y} - \bar{F}(\bar{x})] \quad \text{or} \quad (\bar{y} - K\bar{x})^T (\bar{y} - K\bar{x})$$

In the linear case a derivative with respect to \bar{x} leads immediately to the normal equations:

$$\hat{x} = G\bar{y} = (K^T K)^{-1} K^T \bar{y}$$

where G is the approximate inverse of K .

If the measurement error $\bar{\epsilon}$ is known, it must be taken into account by weighting both the model and the measurement vectors by it, so that the resulting system of linear equation is balanced:

$$y_i = \sum_j K_{ij} x_j + \sigma_{yi} \Rightarrow \frac{y_i}{\sigma_{yi}} = \frac{1}{\sigma_{yi}} \sum_j K_{ij} x_j + 1$$

where $i \in [1, m], j \in [1, n]$ and the σ_{yi} are the errors on the single measurements, the elements of $\bar{\epsilon}$. Then a covariance matrix $cov(\bar{y})$ could be defined, that if the measured quantities are completely independent of each other will have the form $cov(\bar{y}) = I\bar{\epsilon}$ otherwise in general

$$cov(\bar{y}) = \begin{pmatrix} \sigma_{y1}^2 & \sigma_{y2,y1}^2 & \cdots & \sigma_{ym,y1}^2 \\ \sigma_{y1,y2}^2 & \sigma_{y2}^2 & & \vdots \\ \vdots & & \ddots & \\ \sigma_{y1,ym}^2 & \cdots & & \sigma_{ym}^2 \end{pmatrix}$$

where $\sigma_{y_i}^2$ is the variance on the i -th measurement and σ_{y_i, y_j}^2 the covariance of the i -th and the j -th measurement. The errors of the retrieved variables σ_{x_i} then can be easily calculated by the covariance matrix of the unknown $\text{cov}(\vec{x})$, that is

$$\text{cov}(\vec{x}) = G \cdot \text{cov}(\vec{y}) \cdot G^T$$

In the case instead that there is no information on the uncertainties of the single measurements σ_{y_i} , the uncertainties on the parameters σ_{x_i} can still be estimated following Press et al. (1992). The covariance matrix of the unknown $\text{cov}(\vec{x})$ expresses the dependence of each parameter x_j from the forward model, and it can be shown that it is equal to

$$\text{cov}(\vec{x}) = (K^T K)^{-1}$$

The relative weight of the error on any single parameter x_j is then proportional to the j -th element of the diagonal of the covariance matrix. To take into account instead the overall uncertainty on the parameters coming from the approximation of the least-square fit, they can be weighted by the fit residuum

$$\sigma_{x_j}^2 = \text{cov}(\vec{x})_{jj} \frac{\|RES\|^2}{m - n}$$

divided by $m - n$ that is the number of degrees of freedom. The fit residuum is defined as:

$$\|RES\|^2 = \sum_i \left(y_i - \sum_j K_{ij} x_j \right)^2$$

This is the method used in the WFM-DOAS fit and in the preliminary MAMap retrieval algorithm, where the uncertainty on the measurements from each single spectral pixel is not estimated with sufficient accuracy.

One special method for the remote sensing of atmospheric trace gases in the atmosphere is the so-called *differential optical absorption spectroscopy* or DOAS. Its main difference from other measurement methods based on absorption spectroscopy in the infrared wavelength range is that, to avoid problems with extinction by scattering or changes in the

instrument throughput, only signals that vary rapidly with wavelength are analyzed (thus the differential in DOAS). The other, broadband structures - that are e.g. caused by a changing cloud cover and that are difficult to quantify - are approximated by a polynomial. Therefore, only those species can be observed that show significant and structured absorption in the near infrared wavelength range, such as CO_2 , CH_4 , H_2O , and CO . Unlike other techniques such as ozone sondes, LIDAR measurements or microwave radiometry, DOAS measurements provide little information on the vertical profile of the absorbers but rather the integrated column in the atmosphere. This disadvantage is compensated by the simplicity of the experiment and the relatively large number of species that can be measured simultaneously.

The Earth-reflected spectrum I_λ is first divided by the solar spectrum measured at the top of the atmosphere $I_{\lambda 0}$, to eliminate the structures of the solar spectrum itself, like the Planck blackbody shape and the Fraunhofer lines. The second step is to fit a low-order polynomial $P(\lambda)$ to the logarithm of the resulting spectrum, and divide the logarithm by the polynomial. The resulting differential absorption spectrum I_λ^d then contains only variation on the same order of the absorption by the trace gases of interest. They can be separated as a linear combination of the absorption spectra of the different species i :

$$\ln I_\lambda^d = \frac{1}{P(\lambda)} \ln \left(\frac{I_\lambda}{I_{\lambda 0}} \right) = \sum_i \alpha_{i\lambda} c_i$$

where $\alpha_{i\lambda}$ is the absorption cross-section of the gas i and c_i a linear coefficient proportional to the quantity of absorbers along the light path. Figure 4a is an example of a differential spectrum from a MAMap CH_4 measurement. The quantity of the different absorbers (in this example methane, carbon dioxide and water) is estimated with a least-squares fit, and their sum results in the differential spectrum, plus a difference called *residuum*.

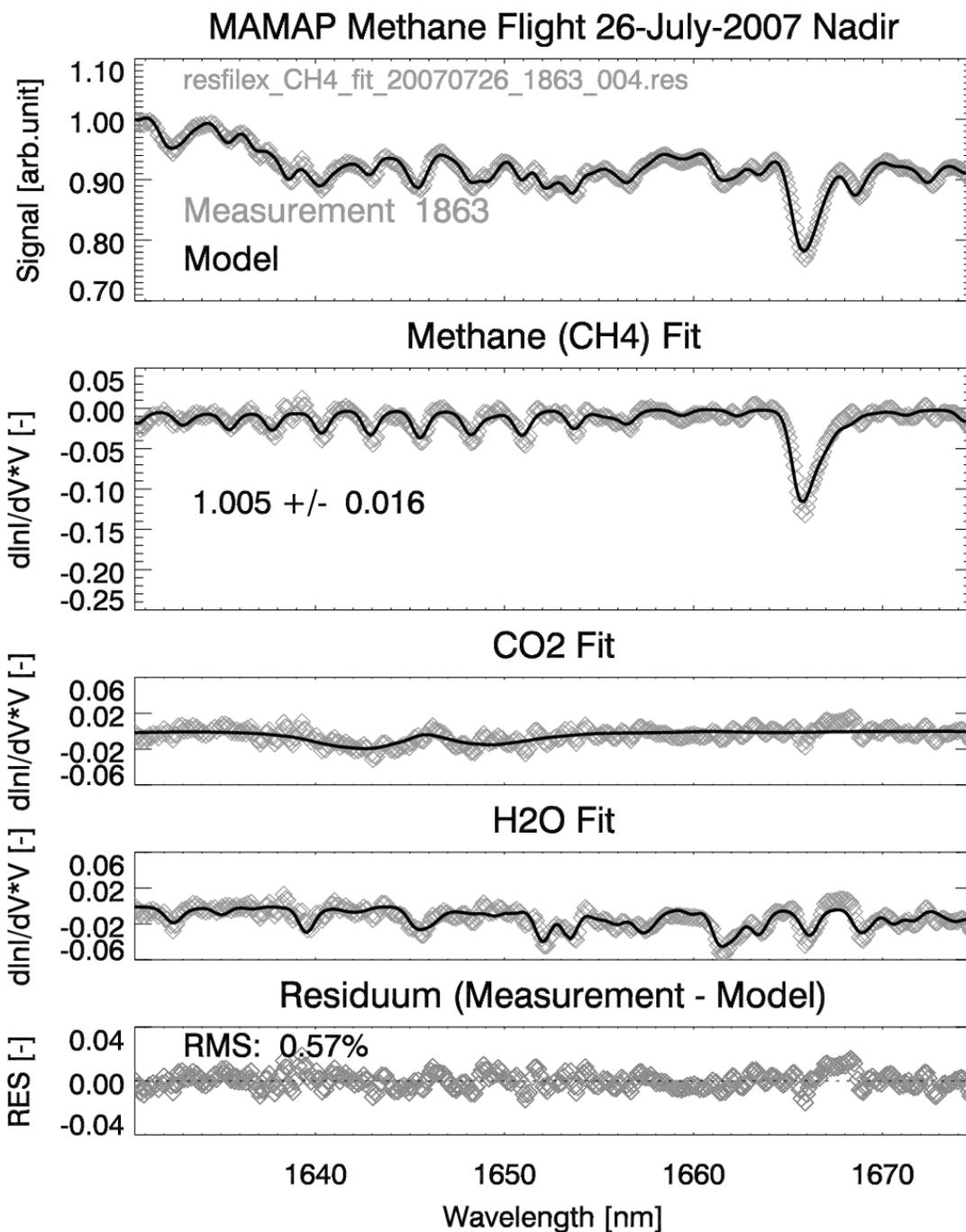


Figure 4a: Structure of a MAMap WFM-DOAS fit in the spectral region used for the retrieval of the methane profile-scaling factor: the measured spectrum (grey symbols in top panel) is modeled as the sum of the contributions of the different absorbers. Courtesy of Dr. Michael Buchwitz (IUP, University of Bremen)

4. The Methane Airborne Mapper (MAMap)

The following chapters describe the MAMap project and instrument (4.1 and 4.2 are based on Bertagnolio (2008)). Then a short summary of the previous work on MAMap data analyses by Bertagnolio (4.3) is given and the relation of Smile-Effect and MAMap (4.4) is explained.

4.1 The MAMap Project

The Methane Airborne Mapper (MAMap) project has been initiated to provide information on the mixing ratios of CH_4 and CO_2 on an intermediate scale between ground-based local measurements and satellite-based global datasets. These greenhouse gases have physical properties that make them particularly suited to be detectable by remote sensing from an airborne passive spectrometer. First, the sources and sinks for these gases are located at the Earth surface, and the greatest spatial variations take place in the atmospheric boundary layer, that is in the lowest 1000m. Instead from the free troposphere above the vertical profile is almost constant in space and time. This means that an instrument carried by a small plane can observe most of the variability.

Second, CH_4 and CO_2 molecules absorb strongly the sunlight in the short wave infrared (SWIR) spectral region around 1600 nm. For this reason, the sunlight reflected by the Earth surface can be used as a source for a passive spectrometer. Moreover, by measuring alternately in zenith- and nadir-looking geometry it is theoretically possible to distinguish the contribution to the absorption due to the gas amount present directly under the plane from the absorption along the light path through the entire atmosphere. This light path could have a complex geometry due to multiple scattering by air molecules and the effect of clouds, aerosols and different surface albedos. In order to minimize the influence of those factors, the mixing ratio of oxygen (O_2) is also measured and used as a constant reference, since it is well mixed at every location.

The ratio between the column number density of absorbing gas and the column number density of dry air (referred to as dry column mixing ratio) can be then retrieved from the detected spectra using a differential optical absorption spectroscopy (DOAS) inversion algorithm.

4.2 The MAMap instrument

The MAMap instrument is a two-channel airborne NIR/SWIR grating spectrometer. Both channels have separate optics for zenith- and nadir looking, and a fold-mirror allows changing between the two operating modes. The instrument can be installed in an airplane with a down-looking window for the nadir telescopes, and light can be fed to the zenith telescopes with diffuser plates and optical fibers.

The main channel (SWIR) is designed to detect the absorption lines of CO_2 (1590-1620 nm) and CH_4 (1630-1750 nm). The detector is an InGaAs linear photodiode array, with 1024 pixels, 25.6 mm long. Due to the disperse properties of the optics this channel has an overall effective spectral resolution of 0.82 nm. The detector is cooled with liquid nitrogen to an operating temperature of -120°C , to strongly reduce the detector dark current. The signal-to-noise ratio varies then between 1000 (over land, albedo 0.18) and 350 (over water, albedo 0.01), when 10 single spectra are co added for an overall exposure time of 1 s. The instantaneous field of view (IFOV) of the SWIR channel is 1.34° cross-track (CT) and 0.02° along-track (LT). For nominal flight parameters (altitude 700 m, speed 200 km/h) and detection times (800 ms, albedo 0.18) the ground-projected IFOV is 20 m (CT) x 30 m (LT).

The second channel (NIR) is adapted to the absorption line of O_2 in the near infrared spectral range at 760 nm. To make sure that both channels observe exactly the same ground scene, this channel is equipped with a two-dimensional detector, that produces a one-dimensional cross-track image of the ground (pushbroom imaging spectrometer). Since the NIR IFOV is 5.85° CT and 0.072° LT, it is more than 4 times wider, in CT direction, than the SWIR IFOV. Optimal co-alignment between the two channels is then possible by choosing, during the data analysis process, the right portion of the NIR window. An accurate matching of the ground-projected fields-of-view has not yet been implemented.

The 2D-detector is a frame transfer (FT) CCD, with 85 (binned) pixels on the imaging axis, and 256 (binned) pixels on the spectral axis. Due to the grating optical parameters, the spectral window detected is approximately 17 nm, with a spectral resolution of 0.46 nm. The signal-to-noise ratio obtained during test measurements is greater than 850 for a single imaging pixel, and greater than 4000 after 21 imaging pixels are averaged for the alignment with the SWIR channel.

<i>MAMap instrument specifications</i>			
	<i>CH₄/CO₂</i>	<i>O₂-A</i>	
Focal Length (optic)	300 mm	80 mm	
Detector Type:	Focal Plane Array (FPA) detector	Charged Coupled Device (CCD)	
Cooling system	Liquid nitrogen	Thermoelectric	
Number of pixels	1024	512x512 (256x85 binned)	
Spectral Range	1590-1690nm	755-785nm	
Spectral Resolution	0.82nm FWHM	0.46nm FWHM	
Spectral Sampling	8 pixels/FWHM	6 pixels/FWHM	
Signal-to-Noise Ratio	1000	4000	For exposure time 0.8s and surface albedo A=0.18
Instantaneous Field-of-View	1.34° (CT) x 0.02° (LT)	5.85° (CT, divided into 85 pixel) x 0.072° (LT)	
Ground Projected Co-Added Pixel Size	50-80m (LT) over land (albedo A=0.18) 4-5km (LT) over water (A=0.01)	50-80m (LT) over land (albedo A=0.18) 4-5km (LT) over water (A=0.01)	4km flight height, ground speed 300 km/h (DLR Do 228 aircraft)

Table 1: MAMap instrument technical specifications and nominal performance.

The spectrometer system is integrated by a device system called “Observer”, that provides information about the position of the aircraft and the ground scene during the measurement: a GPS positioning system provides accurate latitude, longitude and altitude measurements; a triaxial gyro sensor records the orientation of the aircraft; a digital imaging interlined CCD camera coaxial to the spectrometer captures visible images of the ground scene, which can be used to extract information about surface spectral reflectivity, vegetation biotype, cloud cover.

Both the spectrometer systems and the Observer are activated by the same trigger signal to obtain optimal synchronicity. Three autonomous processing units manage the data readout and storage for the main subsystems (SWIR channel, NIR channel, Observer).

The main goal of the instrument is to measure the concentration of greenhouse gases below the aircraft. Measuring the radiation emitted by the Sun and reflected by the surface of the Earth performs this. The radiation, in fact, carries the information on the total number of molecules of each absorber it met along its path (called *light path*) from the top of the atmosphere to the instrument.

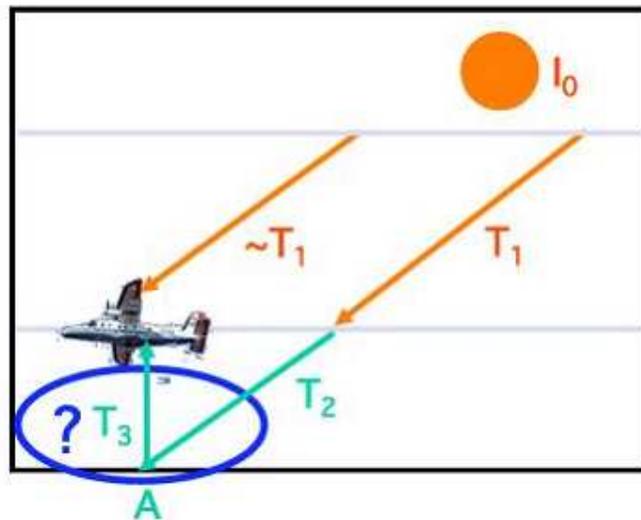


Figure 4b: Zenith and nadir light paths and operation modes. I_o denotes the solar irradiance, A the surface albedo, and the T_i are transmittances in different parts of the atmosphere along different light paths.

The problem is then how to separate the different contributions to the absorption, one by the molecules located under the aircraft and the other from all the rest of the atmosphere.

This is done by measuring alternately with two different geometries: *nadir-looking* and *zenith-looking*.

In nadir mode, the instrument is pointed directly towards the surface. The majority of the photons coming directly from down below will have been emitted by the Sun (intensity I_0), gone through the atmosphere above the aircraft (with a transmissivity T_1), then through the atmosphere below the aircraft (transmissivity T_2), reflected at the surface (angular reflectivity or albedo A) in the direction of the instrument, again through the atmosphere (T_3) and then detected. The rest of the photons will have undergone multiple scattering and, as such, have a different light path, but in this case they are only a small fraction of the whole incoming radiation and can be neglected. The signal detected will be then $S_{down} = I_0 T_1 T_2 A T_3$. But only T_2 and T_3 are relevant for this research.

In zenith mode, instead, the instrument is looking directly above the aircraft with a diffuser, so that the sunlight is measured directly. Their light path then will result in $S_{up} = I_0 T_1$. We can assume that the transmissivity through the upper atmosphere T_1 is the same for nadir and zenith photons because solar rays run parallel, when the solar zenith angle is small.

By taking the ratio

$$\frac{S_{down}}{S_{up}} = \frac{I_0 T_1 T_2 A T_3}{I_0 T_1} = A T_2 T_3$$

then the transmissivity below the aircraft $T_2 T_3$ is isolated. The factor A can be assumed as a variation with a low spectral frequency, and as such not relevant for the DOAS retrieval.

The different components of the instrument have different operation times, high- and low resolution. Currently, in a measurement cycle of 2.00 seconds, the CH₄/CO₂ detector measures 10 spectra, the O₂A-band detector measures 2 spectra (because of the longer readout times of the CCD sensor), the Observer system records once the GPS and gyro information and a digital picture. The rest of the time is used to store the data on the on-board hard disks.

This structure has been chosen to have repeated, closely spaced measurements and have a better signal-to-noise ratio than would be possible with only one measurement. The 10 SWIR spectra, in fact, can be averaged together, and the resulting spectrum be used for the retrieval of the gas columns (low-resolution retrieval), or from each single spectrum a column value can be retrieved (high-resolution retrieval).

Both modes have advantages and disadvantages. In both cases, the data points (the spectra or the columns) have to be filtered for outliers and bad data, and then averaged, to reduce the background noise. Since the current method for the exclusion of unreliable data is based on the fit residuum that is how different a measured spectrum is from the model, only the column factors can be filtered, and not the spectra before the fit. The high-resolution mode then allows filtering out single measurements, and the others can still be averaged together, instead of throwing away a full set of 10 spectra.

The exposure time can be changed to match the illumination (solar zenith angle) and the surface albedo: till now 58 ms has been chosen for the summer flights, over land, 78 ms for the autumn flights, over land, and 1.998 s for measurements over open water.

4.2 Previous work on MAMap data analysis

Bertagnolio has performed several analyses in his master thesis “Mapping of Tropospheric Greenhouse Gases using Airborne Near Infrared/Shortwave-Infrared Spectroscopy” to assess the capabilities of the MAMap instrument, which is based on WFM-DOAS v.HR-002 high resolution retrieval algorithm.

Also a theoretical estimate of the signal-to-noise levels and the corresponding GHG column retrieval precisions for the fast single detector readouts (exposure time $t_{exp}=58$ ms) has been carried out by Bertagnolio (2008: 112). The results are shown in Table 2.

	ρ Surface Reflectance [-]	\bar{R} Mean Radiance [photons/ s nm cm ² ster]	\bar{S} Mean Signal [counts]	\bar{S}/N Mean SNR [-]	σ_x/x Relative Column Error [%]
$x = CH_4$	0.01	3.75E+11	821.08	24.93	10.24
	0.02	7.50E+11	1342.15	49.43	5.16
	0.05	1.87E+12	2905.31	120.47	2.11
	0.10	3.75E+12	5510.77	231.57	1.10
	0.15	5.62E+12	8116.08	334.81	0.76
	0.20	7.50E+12	10721.53	431.39	0.58
	0.40	1.50E+13	21143.07	767.14	0.33
$x = CO_2$	0.01	3.84E+11	833.63	25.53	14.75
	0.02	7.68E+11	1367.27	50.60	7.44
	0.05	1.92E+12	2968.15	123.25	3.05
	0.10	3.84E+12	5636.29	236.71	1.59
	0.15	5.76E+12	8304.58	342.01	1.10
	0.20	7.68E+12	10972.72	440.38	0.85
	0.40	1.54E+13	21644.89	781.68	0.48

Table 2: Signal-to-noise estimates based on instrument simulation valid for $t_{exp}=58$ ms.
(Bertagnolio, 2008)

From Table 2 it can be concluded that the theoretical CH_4 / CO_2 column retrieval precision due to instrument noise for an exposure time of 58ms (single, not co added detector readout) and a surface albedo of 0.20, typical of a vegetated land ground scene, is on the order of 1.03% for (in Tables 2 precision for the vertical columns are listed: CH_4 : 0.58%, CO_2 : 0.85%). With 10 co-added detector readouts (typical operation mode of MAMap resulting in an total co-adding time of 0.58 sec = 0.058*10 sec) this would imply a precision of 0.33%

(1.03% / $\sqrt{10}$). This is what the MAMap would achieve if the instrument detector noise were the only relevant noise contribution.

Bertagnolio analyzed three different flights where the MAMap instrument measured CH_4 and CO_2 at an anthropogenic carbon dioxide source (Schwarze Pumpe coal power plant), a natural methane source (Zarnekow wetlands) and an artificial methane source (Ketzin experiment). Here the 10 SWIR spectra were not co-added to retrieve a single column value. By analyzing these data he found out that the standard deviation of the retrieved CH_4 and CO_2 columns of the single detector readouts and column ratios are often around 3-4% (for 58 ms retrievals). With 10 co-added SWIR spectra this would result in a precision of $\sim 1.1\%$ ($= 3.5\% / \sqrt{10}$), which is about a factor of 3 larger than the expected precision of 0.33 %. This is illustrated in Figure 4c (source: Bertagnolio (2008)), where it is shown that for 58ms the retrieval precision is often $\sim 3\%$ instead of the expected $\sim 1\%$, i.e., a factor of 3 larger than expected from detector signal-to-noise considerations.

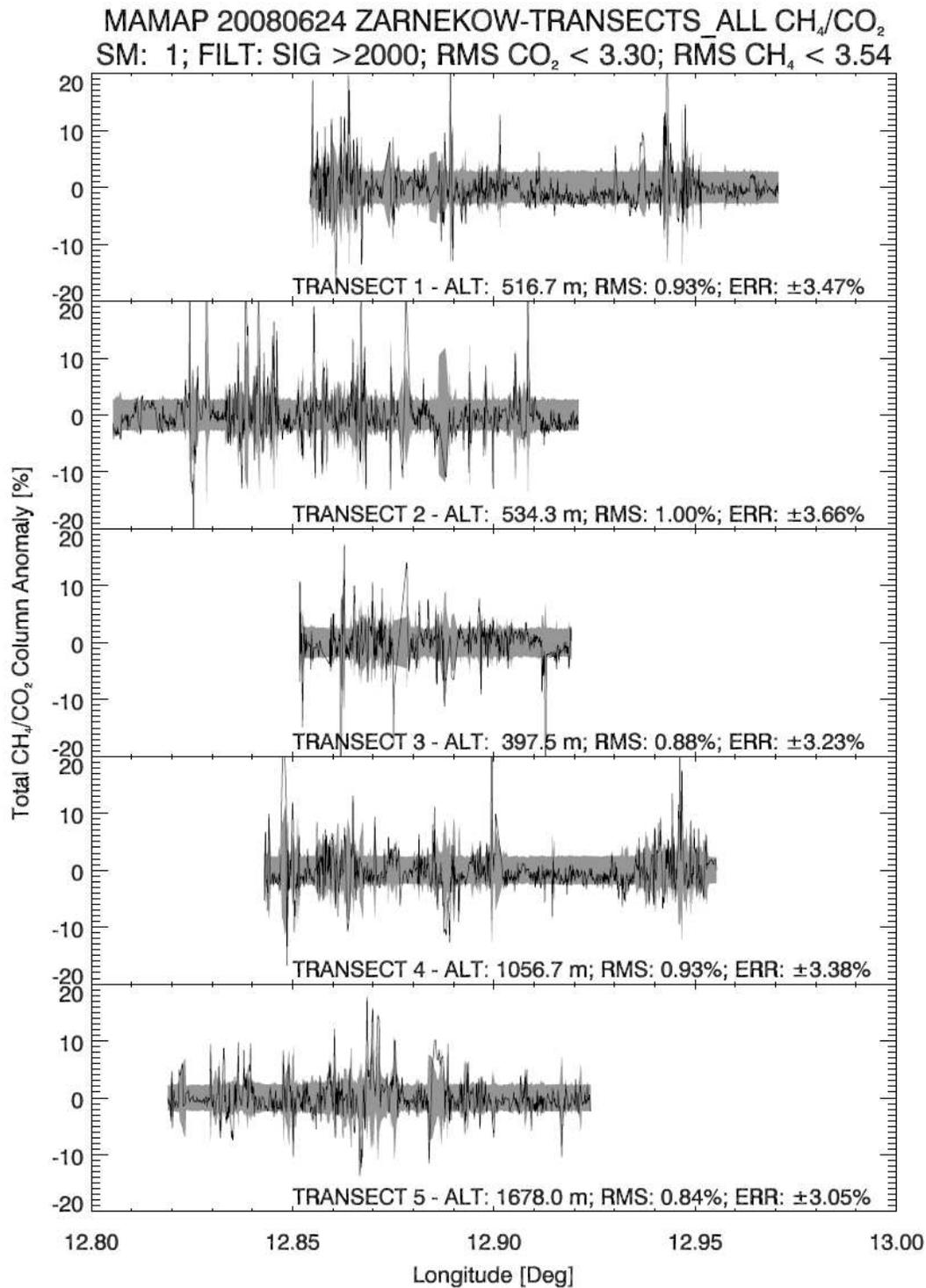


Figure 4c: MAMap results from the flight over the Zarnekow wetlands on 24-06-2008. Shown is the relative anomaly of the retrieved CH₄/CO₂ column ratio. No data averaging has been performed, and all valid data points are displayed. The error range from the average value is displayed in grey. The scatter (standard deviation) of the values in the different transects ranges between 3.3% and 4.7% for single, not co added, detector readout Bertagnolio (2008).

4.3 MAMap and the Smile-Effect

The algorithm of Buchwitz et al (2000), which has been applied to MAMap in a modified form, assumes that the instrument line-shape (ILS) function of MAMAP is known and constant. If the ILS would vary from measurement to measurement this would result in high-frequency variations of the retrieved GHG columns in addition to random errors due to instrument noise (main signal + dark current shot noise of the detector). Significant variability of the ILS could be caused by inhomogeneously reflecting scenes and therefore due the inhomogeneously illuminated slit of the spectrometer system. For inhomogeneous scenes, i.e. scenes where the surface reflectivity varies especially in the cross-track direction, the spectrometer entrance slit will not be homogeneously illuminated.

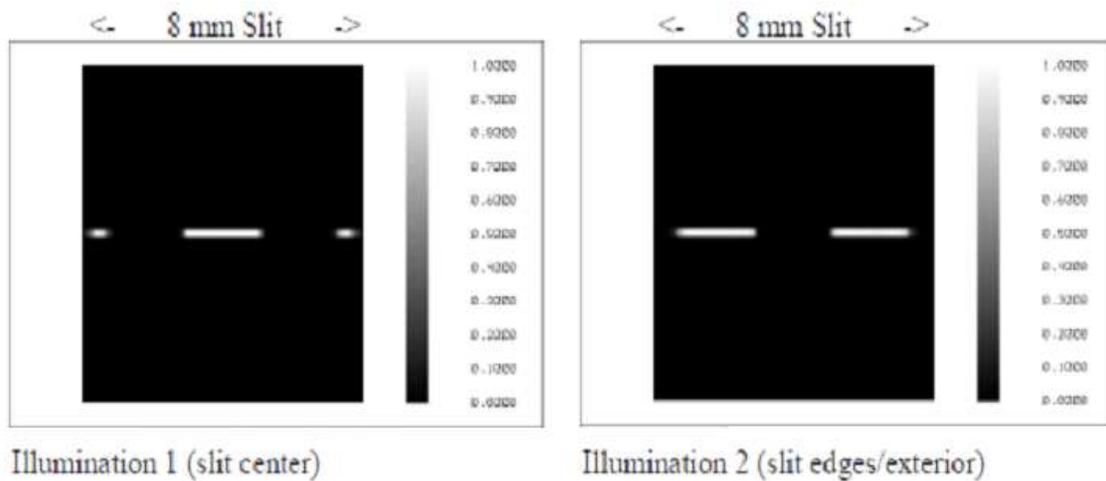


Figure 5a: ZEMAX Simulation of the MAMap slit function under 2 different illumination conditions: illumination 1 (slit center, left panel) and illumination 2 (slit edges/exterior, right panel). Parameters: 8 mm slit height, 0.1 mm slit width, simulation on 128x128 pixel grid.

(Gerilowski, 2009)

Depending on how the slit will be illuminated, the ILS is supposed to change its shape, its width, and its spectral position. These ILS variations are not considered for current MAMAP GHG retrievals. It cannot be excluded that the observed GHG variability of 3-5% (single 58 ms measurements not co-added detector readout), which is about a factor 3 higher than the expected retrieval precision of 1-2% estimated from simulated retrievals, is to a large extent

caused by unaccounted ILS variations due to scene inhomogeneity induced by the optical Smile- and Keystone-distortion effects of the Spectrometer. Figure 5a and 5b show a ZEMAX simulation of the MAMap slit function under the illumination conditions shown in Fig 5a.

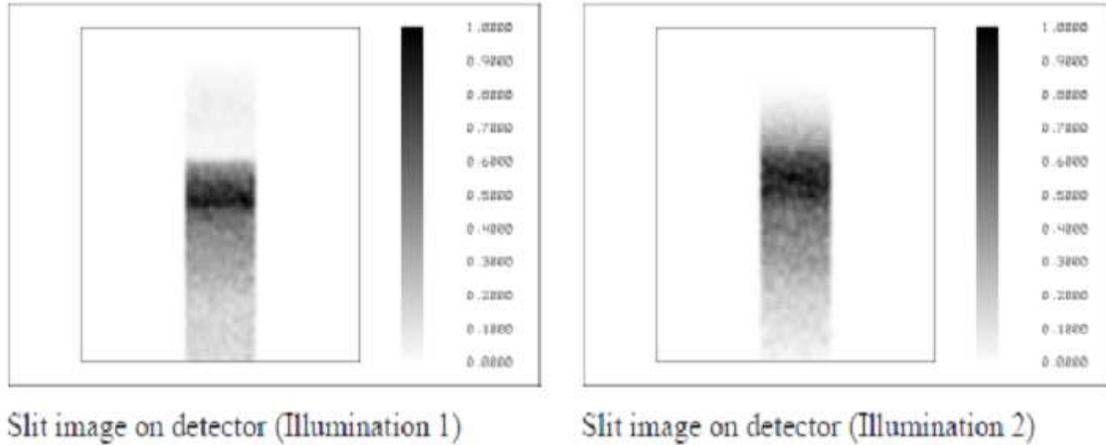


Figure 5b: ZEMAX ray tracing analysis of the MAMap slit image on the detector for illumination 1 (slit center, left panel) and illumination 2 (slit edges/exterior, right panel). Parameters: Detector height = 225 μ , simulated box size 1x1 mm² on 128x128 pixels grid. (Gerilowski, 2009)

The two tested illuminations where the slit center and small outer areas are lighted and one where the center is dark but the outer areas are lighted. These different illuminations result in two different slit functions, which are presented in Figure 5c.

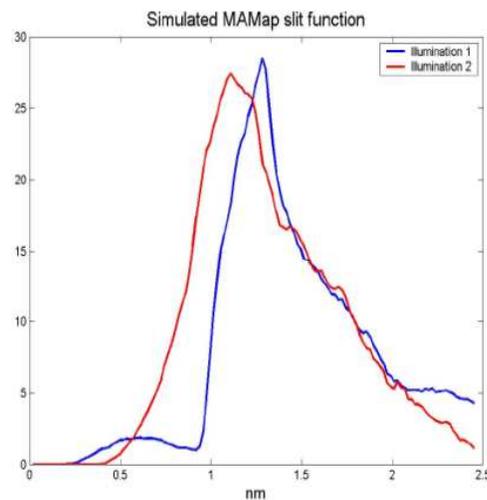


Figure 5c: Simulated slit function for two different illumination conditions. Integral of the ray tracing simulation showed in Fig. 5b over the detector high of 225 μ . (Gerilowski, 2009)

Figure 5c shows that, depending on illumination; the slit function changes its shape, its width and its spectral position. The shift changes with field position across the slit (spatial direction), due to the illumination of a scene. This caused by the Smile-Effect and it can be observed on contour plots of the O_2 -A-Band on the MAMap CCD, shown in Figure 5d.

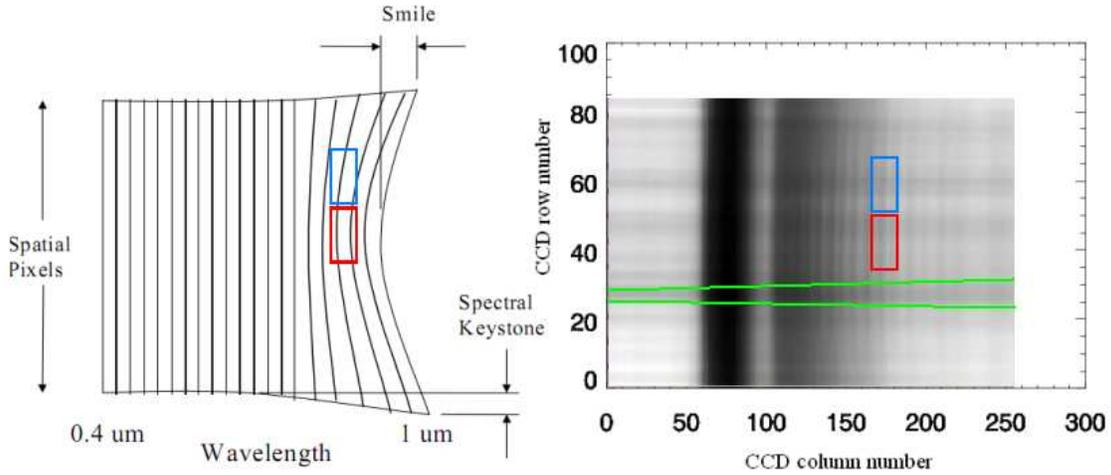


Figure 5d: Comparison between distortion effects (left, from: Fischer et al. 1998) and O_2 -A-Band (right) with marked areas to indicate smile- and keystone effect.

Figure 5d not only shows a comparison between smile- and keystone-error, two areas (red and blue) symbolize that two integrals over those areas, would result in different values. The bending of the slit function is the reason for this, which is a result of the Smile-Effect. Also the keystone-effect is slightly observable and is indicated with green lines. The detection and quantification of the influence the smile-effect has on the GHG retrieval is the focus of this thesis.

As stated before it was expected that because of the 3 times larger variability up to about 89% of the retrieved GHG variability could be explained by scene inhomogeneity. This would imply a correlation coefficient of $R = 0.94$ (as $R^2 = 0.89$ can be interpreted as “explained variance”). The expected 89% variability originates from the assumption that the square of the observed standard deviation, i.e. the observed variance, given by the sum of the variances due to instrument noise (with standard deviation) $\sigma_{Noise} = 1\%$ and the inhomogeneity σ_{IH} .

The inhomogeneity variance can be determined from this relation, which assumes uncorrelated noise terms:

$$\sigma_{Obs}^2 = \sigma_{Noise}^2 + \sigma_{IH}^2. \quad (1)$$

From this equation it follows that the variance of the inhomogeneity is $\sigma_{IH}^2 = 8 (x10^{-4})$. Comparing this variance with the total observed variance of $\sigma_{Obs}^2 = 9 (x10^{-4})$, i.e. 3%x3% quantifies the maximum value of the fraction of variance due to inhomogeneity to the total observed variance, i.e., the maximum value of the variance, which is “explained” by scene inhomogeneity. As R^2 can be interpreted as “explained variance”, the following relation holds:

$$R^2 = \frac{\sigma_{IH}^2}{\sigma_{Obs}^2} = \frac{8}{9} = 0.89. \quad (2)$$

The “explained variance” declares how much percentage deviation of a dependent variable is ascribable to an independent variable. The root of the “explained variance” R^2 of (2) delivers the correlation $R=0.94$, which is “needed” to explain the factor of 3 to high observed retrieval precision of ~3% instead of ~1%.

5. Analysis of MAMap data

5.1 Objectives of this thesis and selected approach

The flights that were analyzed by Bertagnolio showed that there is a standard deviation of 3.3% to 4.7%, for single (not co-added) detector readouts with exposure time of 0.058 s, from which the theoretical estimate of the signal-to-noise levels can only describe 1%. A possible explanation can be scenes of inhomogeneity, which affect the measurements. The main objective of this thesis is to find out if the observed higher than expected variability of the retrieved GHG columns and column ratios is due to scene inhomogeneity.

Unfortunately this cannot be directly be investigated using MAMAPs GHG channel, as this channel consists of a 1D linear detector array which does not permit do obtain sub-scene across-track information.

Fortunately, MAMAP is also equipped with a second (nearly) identically designed spectrometer covering the O₂-A-band spectral region at 760 nm with a 2D CCD detector, with higher spatial resolution and sampling in the across-track direction. These 2D CCD measurements enable to investigate the Smile and Keystone-distortion effects and to derive “Inhomogeneity Indices” (IH_i) for each measurement.

In this thesis the O₂-A-band CCD readouts have been analyzed by deriving Inhomogeneity Indices and by correlating them with the GHG retrievals in order to find out if the expected correlation can be observed and if it is possible to determine to what extent the observed unexplained GHG excess variability can be explained by scene inhomogeneity.

This work analyzes the flight from 24.06.2008, which was also analyzed by Bertagnolio, to be able to compare the results.

5.2 Discussion of CCD signals of the oxygen (O_2 -A-band) spectrometer

As stated before, the 2D CCD O_2 -A-Band is utilized to get information about the inhomogeneity of a scene. Two examples for the O_2 -A-Band spectrometer follow, one with a homogeneous scene (Figure 6a) and one with a very inhomogeneous scene (Figure 6b). What is shown in the two figures are contour plots of the O_2 -A-Band CCD readouts.

The curved stripes in the right half of Figure 6a are indicating the Smile-Effect, which is due to a shift and bending of the slit functions. The coefficient of inhomogeneity for Figure 6a (its precise definition is given in the next chapter) is close to zero, as it should be for a homogeneous scene.

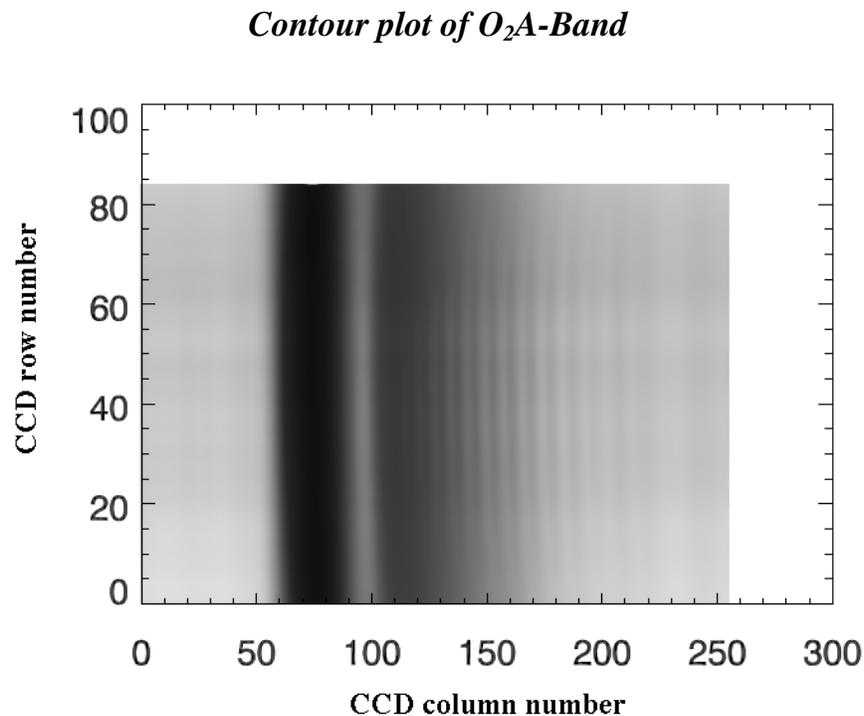


Figure 6a: O_2 -A-Band. Zarnekow wetlands, from 24.06.2008 at 16:26:35.8 o'clock. A very homogeneous scene with $IH_1 = 0.02$, i.e., a "Inhomogeneous Index" close to zero.

Figure 6b shows the CCD signals over an inhomogeneous scene. As it should be the Inhomogeneity Index significantly differs from zero ($IH_1 = -0.44$) for this scene. Responsible for such sceneries is a change of albedo inside the scene.

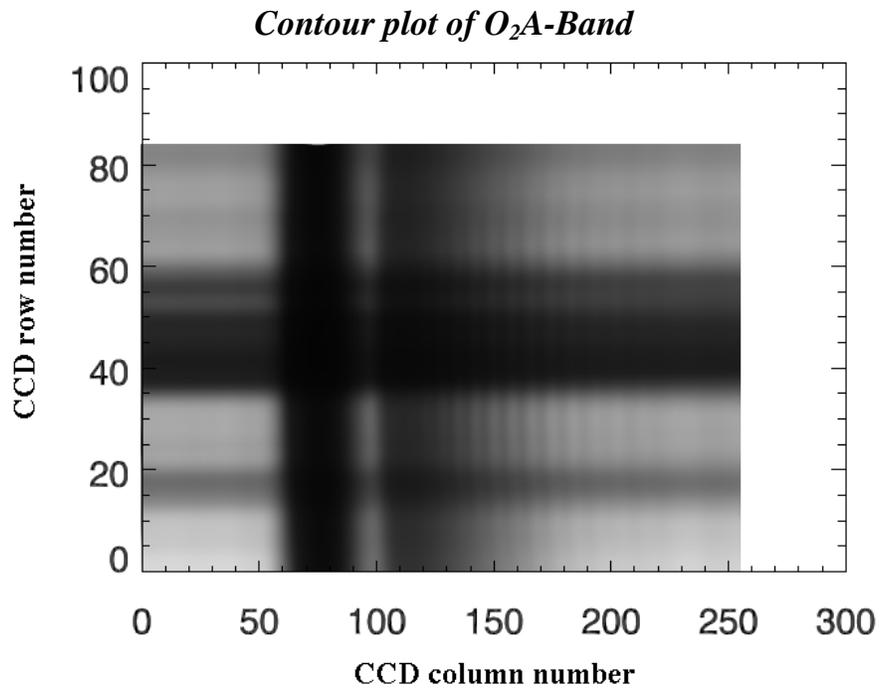


Figure 6b: O₂A-Band. Zarnekow wetlands, from 24.06.2008 at 17:01:57.0 o'clock. A very inhomogeneous scene with $IH_1 = -0.44$.

5.3 Definition of Inhomogeneity Indices (IH_i)

Figure 7 shows the 2D CCD (256x85) O_2 -A-Band spectrometer. The areas, which are marked red and blue, were used to calculate the different inhomogeneity indices.

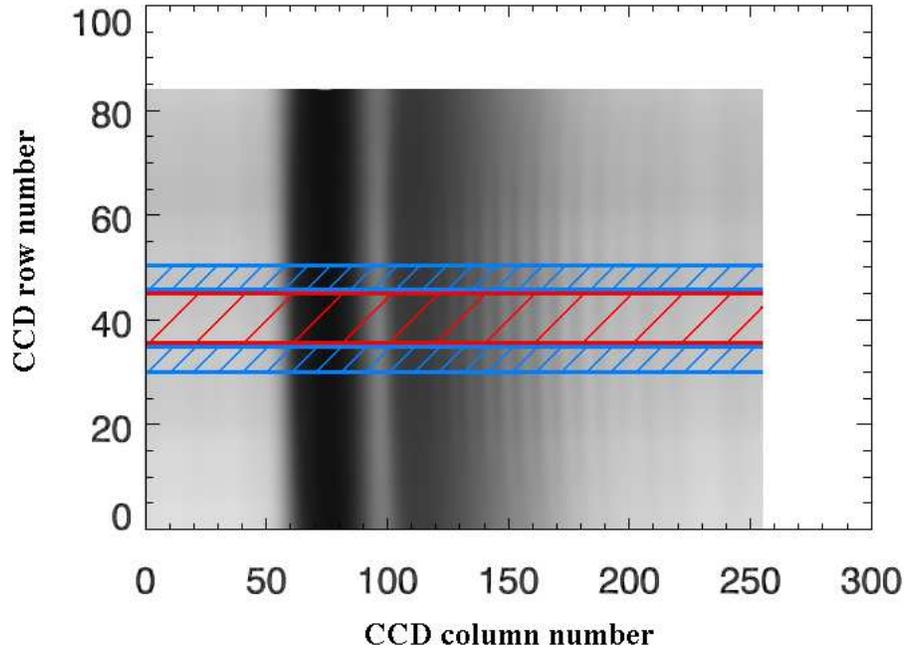


Figure 7: View of CCD array. Red (spectra 36-45) and blue areas (spectra 30-34 and 47-51) were used to determine the inhomogeneity indices (IH_i).

The specific areas were summed up and an arithmetic mean was built to have three different factors, which represent the three areas (A_{red} , $A_{blue(30-34)}$, $A_{blue(47-51)}$). These areas were chosen because they represent the (intensity weights) area of a scene, which is mapped on the CH_4/CO_2 detector and accord to the simulation of the slit function, shown in Figure 5c.

The two inhomogeneity indices IH_1 and IH_2 used in this study are defined as follows:

$$IH_1 = \frac{|A_{red}| - (|A_{blue(30-34)}| + |A_{blue(47-51)}|)}{|A_{red}| + (|A_{blue(30-34)}| + |A_{blue(47-51)}|)} \quad (3)$$

$$IH_2 = |IH_1| \quad (4)$$

As equations (3) and (4) show IH_1 is standardized to -1 and +1 and IH_2 is in the range 0 and 1. Figure 6a and 6b show two examples where IH_1 was used.

To be able to compare the calculated Inhomogeneity Indices of each scene it is important to synchronize the scenes of the O_2 -A spectrometer with the correct GHG retrievals. To achieve that the GPS time, which both systems have saved, were analyzed to be able to synchronize inhomogeneity coefficients with the measurements. To reach synchronization the first GPS time of the GHG retrieval was compared with the equivalent GPS time of the O_2 -A spectrometer. The resulting column number was taken as the start, so that the following columns are always synchronized.

After the inhomogeneity indices were determined and the synchronization was done, a comparison between them and the GHG retrievals was plotted. The results are shown in Figure 8a and 8b.

Comparison between GHG retrievals and inhomogeneity index IH_1

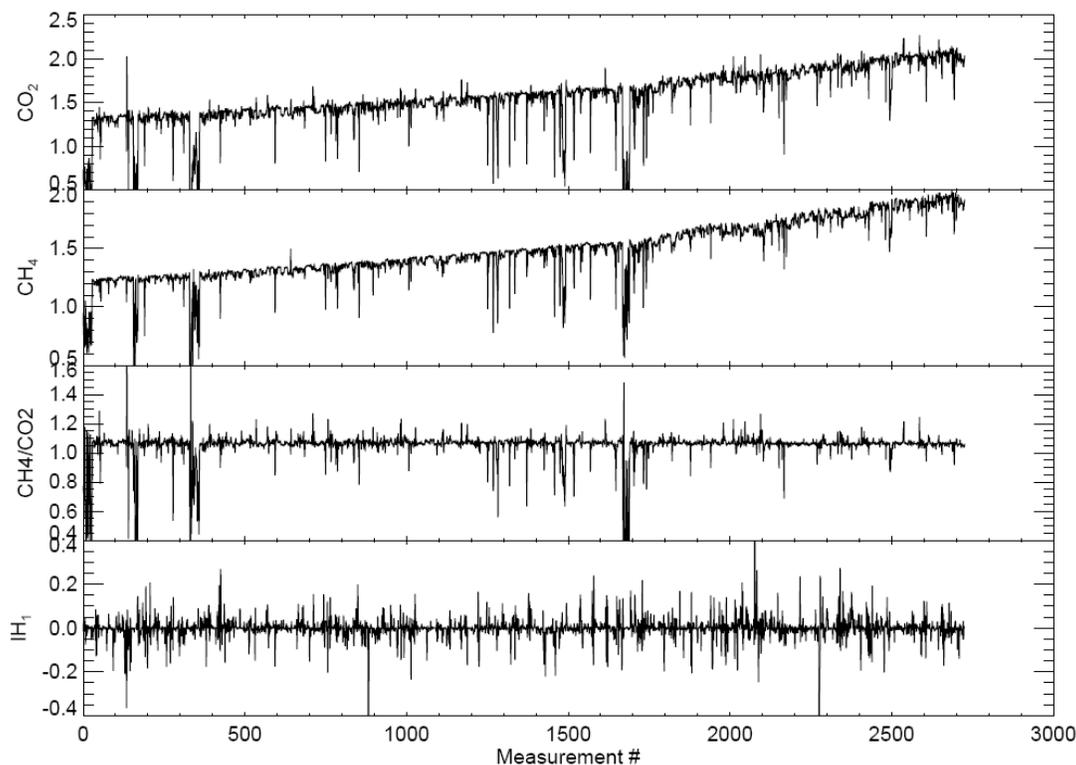
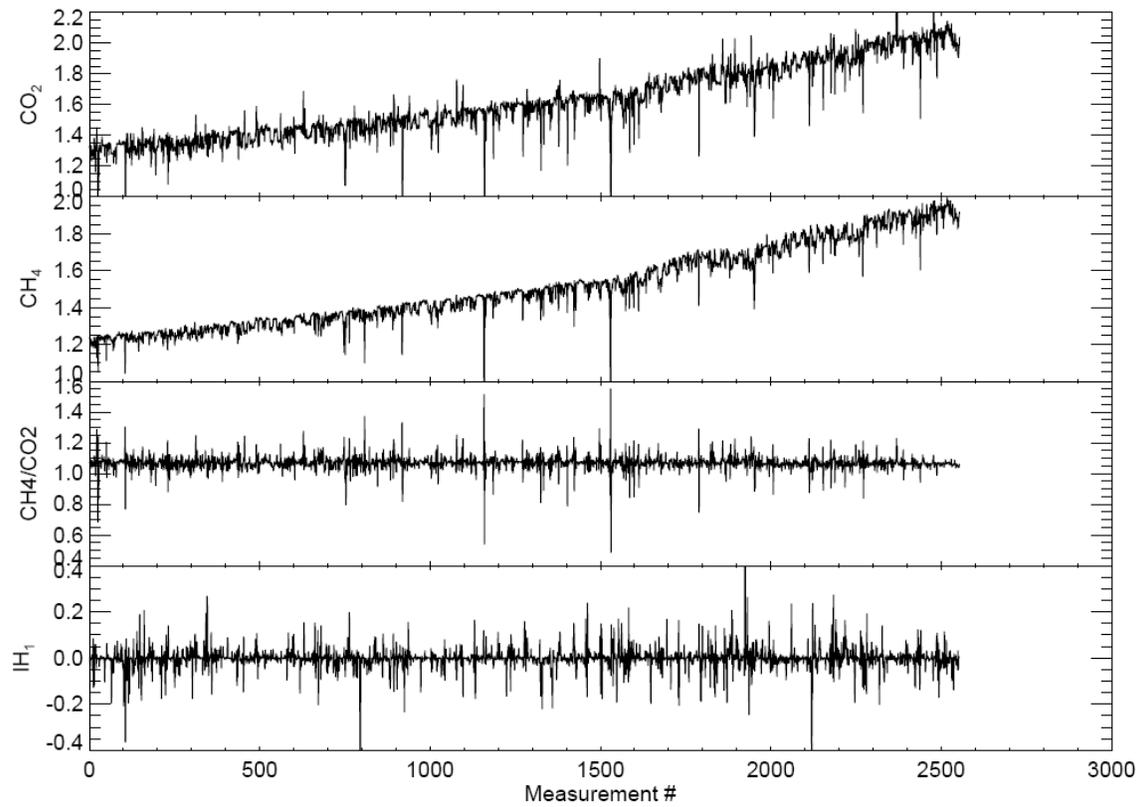


Figure 8a: Comparison between GHG retrieval and Inhomogeneity index IH_1 . The entire flight from 26.04.2008 over the Zarnekow wetlands is shown. The comparison does not show any obvious correlation. No signal filters were used.

Comparison between GHG retrieval and inhomogeneity IH_1 with Filters



*Figure 8b: As Figure 8a but with Filters: $SIG > 2000$; $RMS CO_2 < 3.30\%$; $RMS CH_4 3.54\%$.
Again, no obvious correlation is observable.*

5.4 Correlation analysis of Inhomogeneity Indices with GHG retrievals

To quantify the results a correlation analysis was done. The correlation coefficient, R , indicates the strength and direction of a linear relationship between two random variables. In this case a relationship between the GHG retrievals and IH_i is checked. R is a scalar quantity in the interval $[-1, 1]$ and is defined as the ratio of the covariance of the sample populations to the product of their standard deviations:

$$R = \frac{\frac{1}{N-1} \sum_{i=0}^{N-1} \left(x_i - \left[\sum_{k=0}^{N-1} \frac{x_k}{N} \right] \right) \left(y_i - \left[\sum_{k=0}^{N-1} \frac{y_k}{N} \right] \right)}{\sqrt{\frac{1}{N-1} \sum_{i=0}^{N-1} \left(x_i - \left[\sum_{k=0}^{N-1} \frac{x_k}{N} \right] \right)^2} \sqrt{\frac{1}{N-1} \sum_{i=0}^{N-1} \left(y_i - \left[\sum_{k=0}^{N-1} \frac{y_k}{N} \right] \right)^2}}$$

A value of $R = +1$ or $R = -1$ indicates a perfect fit to a positive or negative linear model, respectively. A value of R close to $+1$ or -1 indicates a high degree of correlation and a good fit to a linear model. A value of R close to 0 indicates a poor fit to a linear model. (Bronstein, et al., 2008)

For accurate analyzes it is important that the inhomogeneity indices and the GHG retrievals are perfectly synchronized. To review this, an additional cross correlation was performed. This gives the opportunity to shift both variables against each other so that a possible better configuration is made visible. An example was done for CH_4 and CO_2 retrievals, which is shown in Figure 9. The shift is indicated by the lag, which goes from -100 to $+100$ in this example. The Figure shows that CH_4 and CO_2 are well correlated. The correlation gets smaller as the shift (lag) gets larger. The cross correlation is defined as:

$$R_{cross}(L) = \begin{cases} \frac{\sum_{k=0}^{N-|L|-1} (x_{k+|L|} - \bar{x})(y_k - \bar{y})}{\sqrt{\left[\sum_{k=0}^{N-1} (x_k - \bar{x})^2 \right] \left[\sum_{k=0}^{N-1} (y_k - \bar{y})^2 \right]}}, & Lag < 0 \\ \frac{\sum_{k=0}^{N-L-1} (x_k - \bar{x})(y_{k+L} - \bar{y})}{\sqrt{\left[\sum_{k=0}^{N-1} (x_k - \bar{x})^2 \right] \left[\sum_{k=0}^{N-1} (y_k - \bar{y})^2 \right]}}, & Lag \geq 0 \end{cases}$$

The correlation analyzes was done for different areas of the GHG measurements. Although the comparison shown in Figure 8a and 8b does not show an observable correlation there are specific areas where a bigger correlation probably exists. Those areas were analyzed with cross correlation (see Figure 10 and 11).

Cross correlation between CH₄ and CO₂

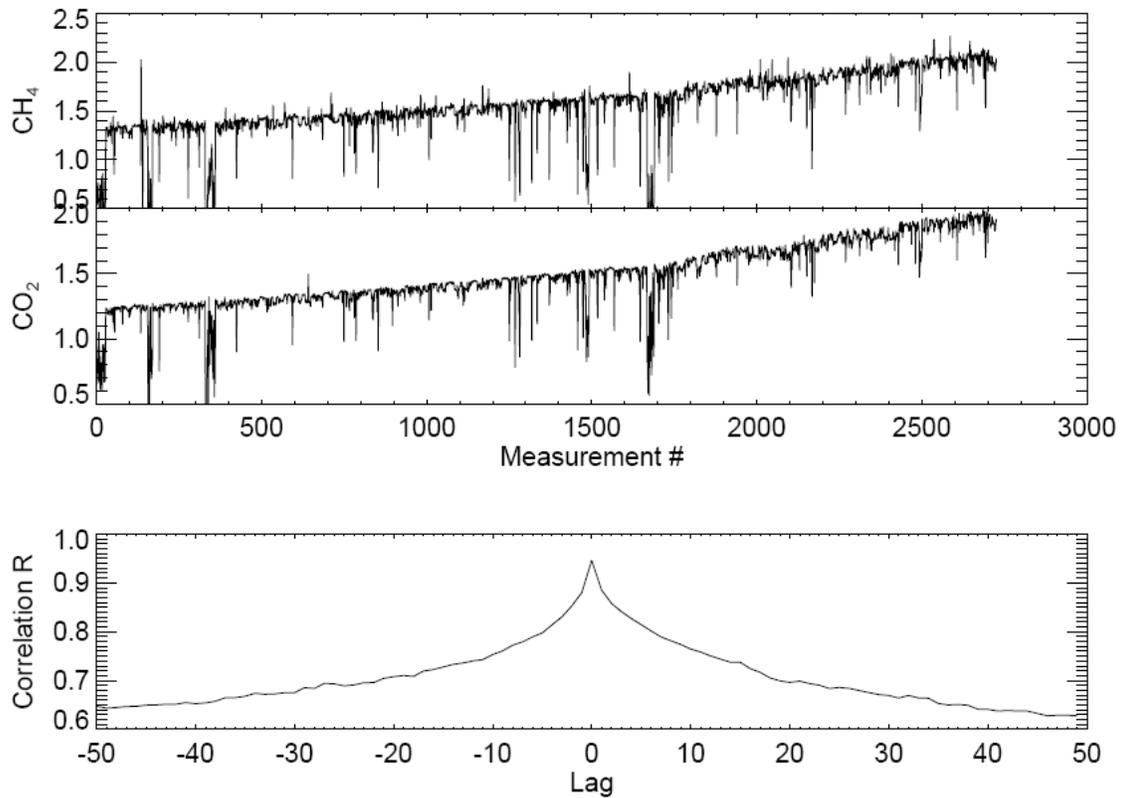


Figure 9: Cross correlation between CH₄ and CO₂ for entire flight from 24.06.2008 over the Zarnekow wetlands. The cross correlation uses a lag from -50 to 50 and shows a perfect correlation at lag = 0 with R=0.95.

***Cross correlation between CH_4 and inhomogeneity
at smaller measurement range***

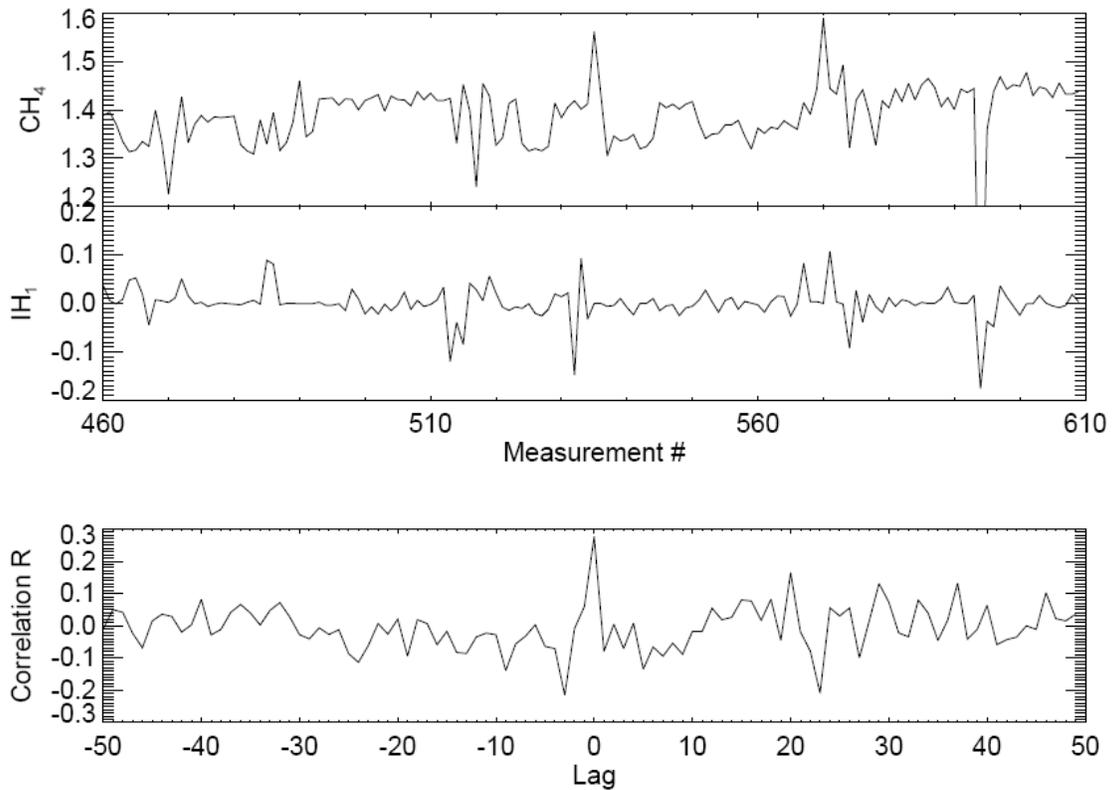


Figure 10a: Cross correlation between CH_4 and inhomogeneity coefficient IH_1 , with lag from -50 to 50, performed for the flight from 26.04.2008 over Zarnekow wetlands at specific measurements (from number 460 to 610). Best correlation at lag 0. A Correlation is observable but the correlation factor only is at $R = 0.28$ with IH_1 .

Cross correlation between CO₂ and inhomogeneity at smaller measurement range

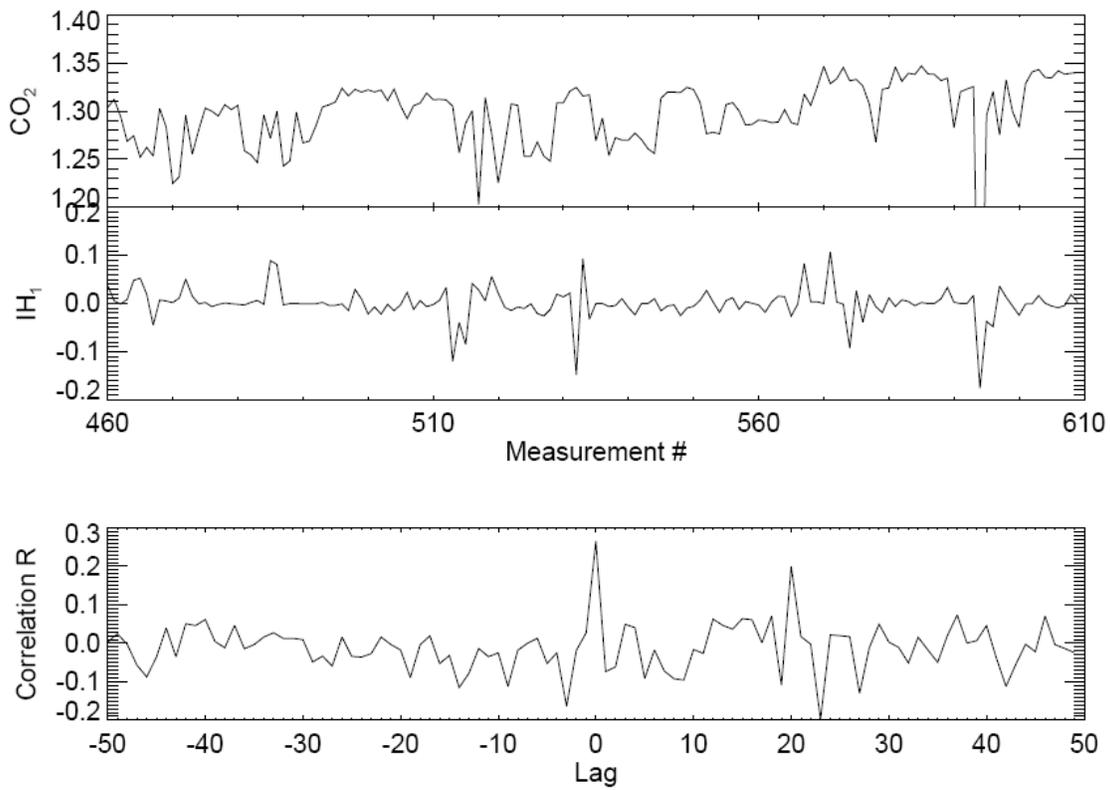


Figure 10b: As 10a but for CO₂ and IH₁. Correlation is at R = 0.27 with IH₁.

Cross correlation between CH_4 / CO_2 ratio and inhomogeneity at smaller measurement range

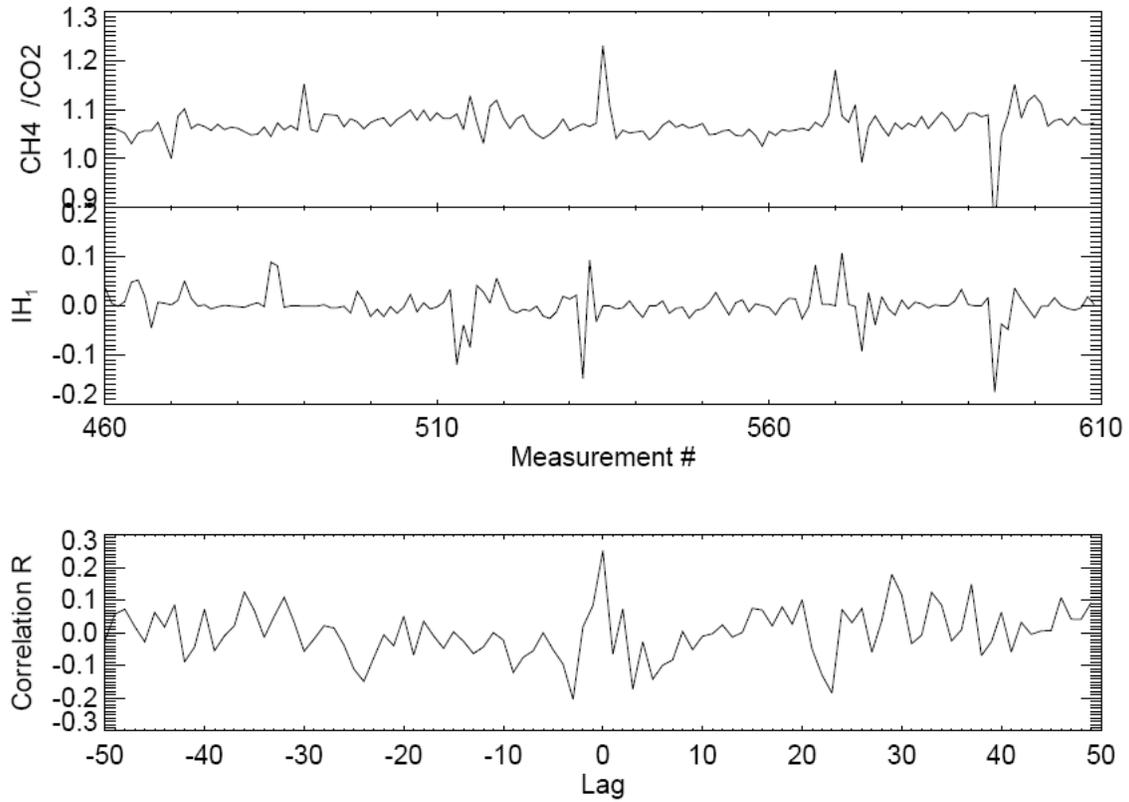


Figure 10c: As 10b but for the CH_4 / CO_2 ratio. Correlation is at $R = 0.25$ with IH_1 .

Cross correlation between CH₄ ratio and inhomogeneity at smaller measurement range

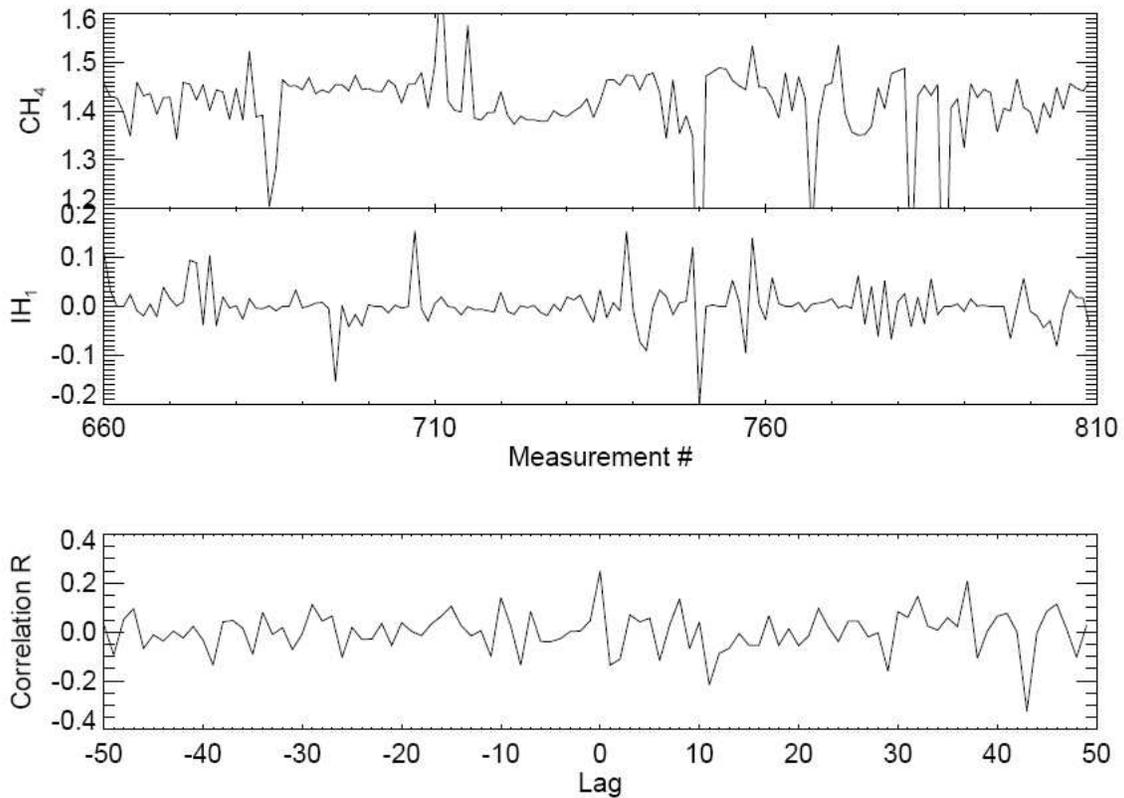


Figure 11a: Cross correlation between CH₄ and inhomogeneity coefficient IH₁, with lag from -50 to 50, performed for the flight from 26.04.2008 over Zarnekow wetlands at specific measurements (from number 660 to 810). The correlation coefficient is at R = 0.25 with IH₁.

Cross correlation between CO₂ ratio and inhomogeneity at smaller measurement range

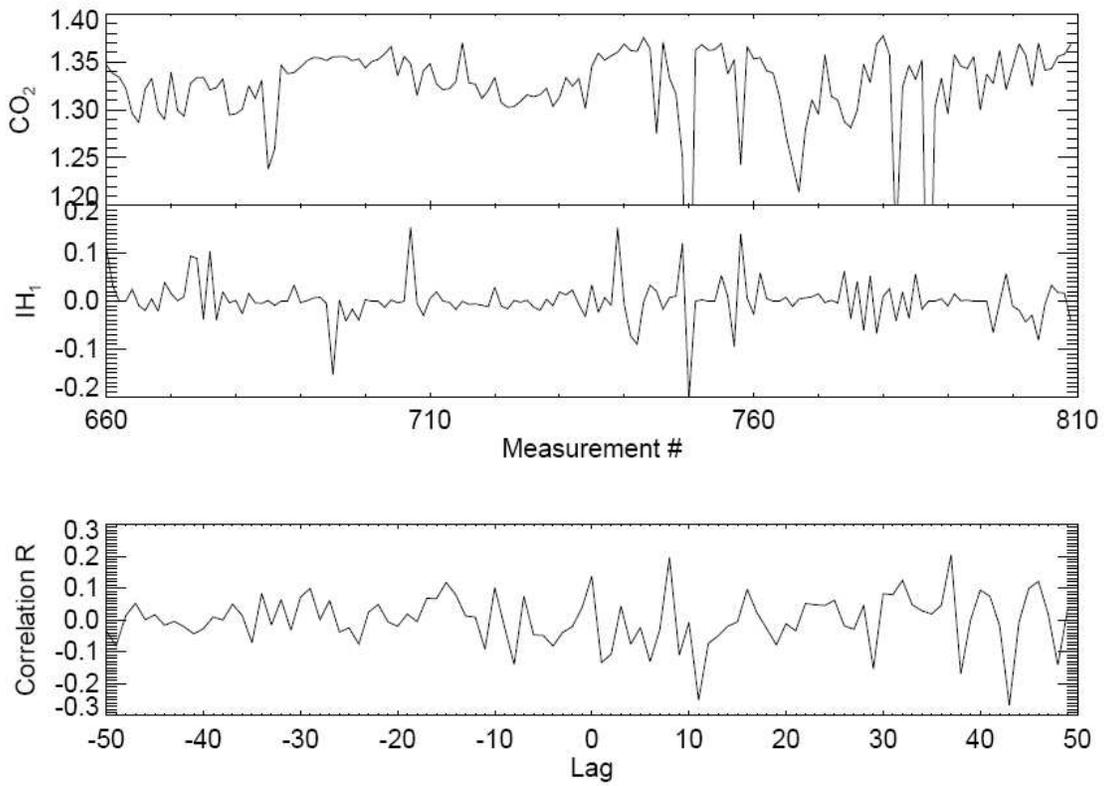


Figure 11b: As 11a but for CO₂ and IH₁. Correlation is at R = 0.14 with IH₁.

*Cross correlation between CH₄ / CO₂ ratio and inhomogeneity
at smaller measurement range*

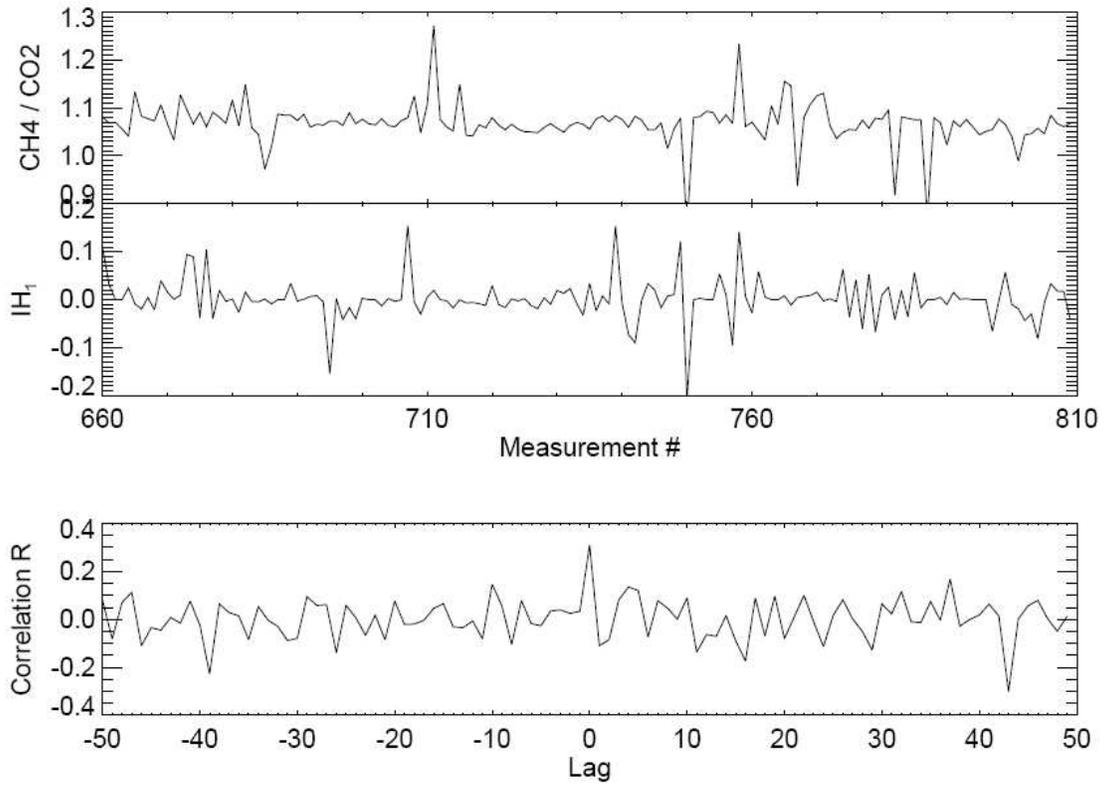


Figure 11c: As 11b but for the CH₄ / CO₂ ratio. Correlation is at R = 0.31 with IH₁.

5.5 Discussion of the Results

The results of the correlation analyses are presented in Table 3. As it can be seen, all correlation coefficients are quite small ($|R| < 0.32$ in all cases).

Cases 1-3 are valid for the entire flight performed by MAMap on 24.06.2008 over the Zarnekow wetlands. Shown are the results for the retrieved CO_2 columns and their correlation with the two Inhomogeneity Indices IH_1 and IH_2 (Case 1). The correlation analysis results are illustrated in Figure 8a. Case 2 shows the corresponding results for the retrieved CH_4 columns, and Case 3 for the CO_2 to CH_4 column ratio.

As it is possible that retrieval artifacts from, e.g., too low albedo, adverse affect the analysis, filtering criteria have been applied to ensure that only those data are used where the signal is above a given threshold (measured signal $SIG > 2000$) and a good spectral fit has been obtained (RMS of the fit residuum better than 3.3% for CO_2 and 3.54% for CH_4). The corresponding results are shown as Cases 4-6. The filtering result in slight change of correlation coefficients, but there still is nearly no significant correlation between the Inhomogeneity Indices and the retrieved GHG.

Further investigations were performed to see if specific randomly selected areas have a higher correlation. Cases 7-9 are analyzes of the measurements from number 460 to number 610. For this measurement area cross correlation was performed, for CH_4 , CO_2 , and the CH_4/CO_2 ratio. Case 7 shows an extension of the correlation for that area to $R=0.28$, which is shown in Figure 10a. In Case 8 the cross correlation was also performed for CO_2 with a correlation coefficient result of $R=0.27$ with IH_1 . Case 9 shows the correlation between the CH_4/CO_2 ratio and inhomogeneity coefficient IH_1 with a correlation coefficient of $R=0.25$.

In Case 10-12 the same proceeding was applied as in Case 7-9 with nearly the same result (see Table 3).

The specific areas (like in Cases 7-9 and 10-12) show that there are areas where a raise of correlation is observable (highest correlation with $R=0.31$) but the resulting correlation coefficients are all much smaller ($|R| < 0.32$) than the expected maximum correlation of $R=0.94$, which would be needed to explain the factor of 3 higher variability of the retrieved GHG columns and column ratios.

MAMap correlation analysis results						
			Correlation coefficient R			
<i>Case</i>	<i>Target</i>	<i>Filter</i>	<i>IH₁</i>	<i>IH₂</i>	<i>Specific Range</i>	<i>Figure</i>
1	CO ₂	-	0.0213	0.0318	-	8a
2	CH ₄	-	0.0120	0.0300	-	8a
3	CO ₂ /CH ₄	-	-0.0071	-0.0410	-	8a
4	CO ₂	<i>SIG > 2000, RMS < 3.30</i>	0.0250	0.0163	-	8b
5	CH ₄	<i>SIG > 2000, RMS < 3.54</i>	0.0192	0.0074	-	8b
6	CH ₄ /CO ₂	<i>SIG > 2000, RMS CO₂ < 3.3, RMS CH₄ < 3.45</i>	-0.0131	-0.0351	-	8b
7	CH ₄	-	0.2769	-0.3094	460-610	10a
8	CO ₂	-	0.2654	-0.3056	460-610	10b
9	CH ₄ /CO ₂	-	0.2512	-0.2674	460-610	10c
10	CH ₄	-	0.2499	-0.1366	660-810	11a
11	CO ₂	-	0.1374	-0.1934	660-810	11b
12	CH ₄ /CO ₂	-	0.3103	-0.0639	660-810	11c

Table 3: Summary of the MAMAP correlation analysis between the Inhomogeneity Indices and the retrieved GHG columns. Several correlation analyzes have been performed (Case 1-12) each with different setups. Cases 1-6 are correlation analyzes for entire flight from 24.06.2008 over Zarnekow wetlands with and without signal filters (SIG=Signal, RMS=Root Mean Square). Cases 7-12 are results for sub-sets of the same flight.

In summary, it was investigated in this thesis if scene inhomogeneity can explain a large fraction of the observed GHG variability, but the expected high correlation between the inhomogeneity indices and the GHG retrieval has not been found. In the following potential reasons for this finding are discussed.

Possible reasons why no significant correlation has been found:

Synchronization errors: A reason for the low correlation coefficient could be a wrong synchronization of the compared data. To exclude this, cross correlation was used to make possible better configurations visible. As Figures 10a-10c show, the best correlation is reached at a lag of 0. Figures 11a (CH_4) and 11b (CO_2) have better configurations at lag 43 but the ratio CH_4 / CO_2 (Figure 11c) has its best correlation again at lag 0. Therefore it can be assumed that the synchronization is correct.

Analysis of all data and filtered data: Both filtered and not filtered data were used in order to make sure that, for example, scenes with very low reflectivity and therefore high noise, do not spoil the analysis. It has however been found that the filtered and the not filtered data give similar low correlation coefficients.

Analysis of the entire flight and parts of it: The degree of scene inhomogeneity depends on the scene with varies along the flight track. Therefore, the entire flight as well as specific parts of it had been analyzed. The result is a significant raise of the correlation in some areas (see Figure 10 and 11). Nevertheless, all correlation coefficients found are much lower than expected.

Other contributions to the observed GHG variability: There might be additional “retrieval noise” due to unconsidered variability of many parameters which are not considered by the current MAMap retrieval algorithm, such as surface reflectivity variability, variability of temperature and pressure profiles, variability of aerosols and clouds, etc. These errors have not yet been quantified in detail but may also contribute to some extent to the observed retrieval noise. In this thesis it is assumed that the additional retrieval noise is essentially zero (see equation (1)). If the additional retrieval noise is significant, equation (1) is not valid and needs to be modified by adding an additional variance variable

$$\sigma_{Obs}^2 = \sigma_{Noise}^2 + \sigma_{IH}^2 + \sigma_{Other}^2 .$$

Including this additional term would lower the influence of σ_{IH} and may explain a significant fraction of the observed variability. This is clearly an aspect, which needs to be investigated in detail in the future.

Definition of Inhomogeneity Indices: It cannot be excluded that the definition of the inhomogeneity indices is not appropriate. The areas, that were chosen to define the indices, represent the area of a scene, which is mapped on the CH_4/CO_2 detector and accord to the simulation of the slit function, shown in Figure 5c. Therefore this is a strong argument to define the indices as they are. Nevertheless, it may be interesting to continue this study using differently defined inhomogeneity indices.

Different inhomogeneities at different wavelengths: The different spectral ranges (1590 – 1690nm for CH_4 / CO_2 and 754 – 774nm for O_2) could be the reason for the low correlation coefficients, which have been found. Figure 12 shows pictures taken at different wavelengths. As can be seen, the intensity and the homogeneity of a scene depend on wavelength. For this thesis it was not assumed that variation of the average albedos of the scenes in the two spectral regions are correlated. It was however assumed that the different spectral ranges, used in both spectrometer systems, do not affect the inhomogeneity of a scene. It was assumed that if a scene is inhomogeneous at one wavelength that this scene is also inhomogeneous at the other wavelength. This is however not true under all circumstances. Different terrain types have different albedos and therefore different reflectivities. This could affect the Inhomogeneity Indices. It might be possible to investigate this using the photos MAMAP takes during the flight of each scene. This requires characterizing in detail the composition of the surface of each scene and its albedo. This is also a task, which is beyond the scope of this thesis.



Figure 12: A photo of a grass area next to a street, taken in different wavelengths. Upper picture: Black/White at 850 nm. Lower picture: ~380-710 nm (with permission from Roth (2008)).

6. Conclusion and outlook

The Methane Airborne Mapper (MAMap) instrument, which has been built at IUP, University of Bremen, in cooperation with the GFZ Potsdam, has been specified to retrieve CH_4 and CO_2 at a horizontal resolution of ~100 m. First test flights have been performed and the in-flight data have been analyzed using an initial retrieval algorithm, which permits the retrieval of total columns of CH_4 and CO_2 with sufficient accuracy to assess the instrument performance.

It has been found that the instrument is very stable and that the noise of the spectral measurements is as specified. It has however also been found that the standard deviation of the retrieved CH_4 and CO_2 columns and column ratios (without the normal co-adding of 10 single readouts) are often around 3-4% for the single, not co-added, detector readouts (exposure time 58 ms), which is about a factor of 3 larger than expected. A possible explanation for this could be retrieval errors caused by, for example, unaccounted variations of the instrument line-shape (ILS) function due to scene inhomogeneity caused by optical effects such as the “Smile Effect”, which has been discussed in chapter 4.2 and 4.3.

As the SWIR spectrometer (~1600 nm) used for the GHG retrievals is only equipped with a linear detector array, it is not possible to obtain across flight track scene information. MAMap is however also equipped with a similarly designed second spectrometer covering the 760 nm spectral region (O_2 A-band) with a 2D CCD array which provides this information. Therefore, an analysis of the O_2 A-band spectra was used to assess to what extent the GHG retrievals are affected by scene inhomogeneity.

This has been investigated in this thesis. “Inhomogeneity Indices” have been defined and derived from the O_2 -A-band spectra. The indices have been correlated with the independent GHG retrievals. It was expected that because of the 3 times larger variability up to about 89% of the retrieved GHG variability could be explained by scene inhomogeneity. This would imply a correlation coefficient of $R = 0.94$ (as $R^2 = 0.89$ can be interpreted as “explained variance”).

The analysis has shown that the correlation is too small to explain the 3-4% precision of the measurement results, which were found by Bertagnolio (see Figure 4c). Also areas where a slight raise of correlation could be observed cannot explain the 2-3% higher variability. Possible reasons for non-correlation were excluded and discussed (see previous chapter).

It could be shown that the inhomogeneity of a target scene and therefore the Smile-Effect contributes to the standard deviation of 3-4%, but can only possibly explain a small fraction of the observed variability. The possible reason for this is that it was assumed that any additional “retrieval noise” σ_{Other} is zero, because these additional errors have not yet been quantified in detail yet. This has to be investigated in detail in the near future so that it can be quantified, for example, how fluctuations in the terrain reflectivity affect the retrieval.

This work cannot exclude that the observed deviations are significantly caused by scene inhomogeneity. However this work also cannot prove this. Another possible reason for the finding is that the inhomogeneity of a scene at 1600nm is not well enough correlated with the inhomogeneity of the same scene at 760nm. Possibly this can be investigated by analyzing the photos of each scene MAMap provides. This requires that it is possible to classify each terrain with high resolution and the reflection characteristics of the terrain types are known with sufficient accuracy. This, however, is out of the scope of this thesis.

A. References

- Bertagnolio, P. P., Mapping of tropospheric greenhouse gases using airborne near-infrared/shortwave-infrared spectroscopy, Master Thesis University of Bremen (supervisor Dr. M. Buchwitz), Institute of Environmental Physics, August 19, 2008.
- Buchwitz, M. (2000). Strahlungstransport- und Inversions-Algorithmen zur Ableitung atmosphärischer Spurengasinformationen aus Erdfernerkundungsmessungen in Nadirgeometrie im ultravioletten bis nahinfraroten Spektralbereich am Beispiel SCIAMACHY. PhD thesis, IUP/IFE Universität Bremen
- Bronstein, I.N., Semendjajew, K.A., Taschenbuch der Mathematik, 7. Auflage, Harri Deutsch Verlag, 2008
- Gerilowski, K. Passive airborne remote sensing of atmospheric methane CH_4 and carbon dioxide CO_2 , IUP intern talk, March 2009
- Fischer, J., Baumback, M., Bowles, J., Grossmann, J., and Antoniadis, J. 1998. Comparison of low-cost hyperspectral sensors, Proc SPIE, Vol. 3438, pp. 23-30.
- Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D., Haywood, J., Lean, J., Lowe, D., Myhre, G., Nganga, J., Prinn, R., Raga, G., Schulz, J. G., and Dorland, R. V. (2007). Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, chapter 2007: Changes in Atmospheric Constituents and in Radiative Forcing. Cambridge University Press
- Petit, J. R., Jouzel, J., Raynaud, D., Barkov, N. I., Barnola, J.-M., Basile, I., Bender, M., Chappellaz, J., Davis, M., Delaygue, G., Delmotte, M., Kotlyakov, I. V., Legrand, M., Lipenkov, V. Y., Lorius, C., Pépin, L., Ritz, C., Saltzman, E., and Stievenard, M. (1999). Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. Nature, 399:429–436.

Rodgers, C. D. (2000). Inverse methods for atmospheric sounding: theory and practice, volume 2 of Series on atmospheric oceanic and planetary physics. World Scientific, XVI edition.

Roth, W.D., Kniffliger Filter Zauber, Article at <http://www.fokkussiert.com/> from May 20th 2008.

Schneising, O., Analysis and interpretation of satellite measurements in the near-infrared spectral region: Atmospheric carbon dioxide and methane, PhD thesis, Germany, University of Bremen FB1, Institute of Environmental Physics (IUP), 24.Nov.2008 (day of PhD defense), pp. 200, 2009

Wallace, J. M. and Hobbs, P. V. (2006). Atmospheric science: an introductory survey. Number 92 in International Geophysics Series. Elsevier, 2nd edition.