MICROTOPS INVERSE

SOFTWARE PACKAGE FOR RETRIVING AEROSOL COLUMNAR SIZE DIUSTRIBUTIONS USING MICROTOPS II DATA

V.I. Tsanev [1,2], T.A. Mather [3]

[1] Department of Geography, University of Cambridge, Downing Place, Cambridge CB2 3EN, UK
[2] Now Department of Chemistry, University of Cambridge, Lensfiled Road, Cambridge CB2 1EW, UK
[3] Department of Earth Sciences, University of Oxford, Parks Road, Oxford OX1 3PR, UK.

0. Introduction

Sun photometers are specialized narrow field-of-view radiometers designed to measure solar irradiance (Shaw, 1893). They typically have between 6 and 10 well-defined spectral bands, each of the order of 10nm full width at half maximum (FWHM). Sun photometers were first developed during the early part of the 20th century taking advantage of the new electrical thermopile devices and developments in the glass industry, which led to cut-off filters. These instruments were primarily designed to measure the solar constant using the spectral extinction method developed by S.P. Langley. This 'long method' is based on measurements of the solar flux in narrow wavelength bands at varying solar zenith angles and yields (as a by-product) the atmospheric transmissivity. The Voltz hand-held photometer, which was originally developed in 1959, includes two narrow spectral bands specifically for measuring atmospheric turbidity and can be considered the precursor of modern sun-photometers.

Modern instruments vary little from early designs, but incorporate technological advances in optics and electronics and are generally more sensitive and much more stable. For example, the MICROTOPS II and CIMEL model 318 are electronically controlled, have on-board data storage capability and some of them incorporate an automated tracking system for accurate positioning and pointing the photometer (Holben *et al.*, 1998; Morys *et al.*, 2001).

Sunphotometer measurements can be used to recover atmospheric parameters including; spectral aerosol optical depth, precipitable water vapour, sky radiance distributions and ozone amount. Aerosol volume and size distribution are retrievable by inversion modelling from the spectral aerosol optical depth. The atmospheric data retrieved from sunphotometers is primarily used in meteorological and atmospheric applications, and this type of instrument has been used for many years by atmospheric scientists (King *et al.*, 1978; Schmid *et al.*, 1997, Holben *et al.*, 1998).

The basic sunphotometer design comprises a collimating tube defining a narrow angle (of the order of 1° to 3°), a series of interference filters and one (or more) solid-state detector (usually silicon photodiode) with an amplifier and voltmeter or analogue-to-digital converter. It is possible to have two filter-detector arrangements. In the filter-wheel arrangement, a series of filters is located on a rotating wheel and passed in turn in front of a single detector, resulting in sequential measurement of each band. Alternatively, in the multiple-detector arrangement each filter is fixed in front of a dedicated detector and all bands are measured simultaneously. Lenses may be present in the optical train, but this tends to be avoided because they are unnecessary and their transmission properties can change when exposed to UV radiation. Most modern instruments incorporate microprocessor control of the measurement sequence, using Sun-seeking and Sun-tracking devices, zenith and azimuth stepping motors for accurate pointing and positioning to within 0.1º. They also include on-board data storage and/or data transmission capabilities. For positioning purposes, the time must be recorded accurately and latitude and longitude must be input with known high precision.

The filter characteristics are critical since these must define a narrow band-pass and be well-blocked (i.e. not allow the transmission of light outside the wavelengths limits of the band). Filters must also be well sealed in their mounts to prevent their exposure to pollutants and any resulting deterioration. Modern instruments usually use thin film dielectric interference filters. Stability is fundamental to measurement accuracy and modern silicon photodiode detectors are well suited to the purpose.

At a wavelength λ, the total optical thickness recorded by a Sun photometer can be expressed as the sum

(1) $\tau(\lambda) = \tau_{\mathrm{A}}(\lambda) + \tau_{\mathrm{R}}(\lambda)$,

where $\tau_{\mathrm{A}}(\lambda)$ is the aerosol optical thickness (AOT) and $\tau_{\mathrm{R}}(\lambda)$ is the Rayleigh optical thickness (which depends on wavelength and local pressure at the measurement site). If we assume that aerosol particles can be modelled to a sufficient degree of accuracy by equivalent spheres with known and constant index of refraction over entire radii range then Mie theory leads to (Bohren and Huffman, 1983)

(2) $$\tau_{\mathrm{A}}(\lambda) = \int_0^\infty \int_0^\infty \pi r^2 \, Q_{\mathrm{ext}}(2\pi r/\lambda, \tilde{m}) \, n(r,z) \, dr \, dz \ .$$

Here $Q_{\mathrm{ext}}(2\pi r/\lambda, \tilde{m})$ is the Mie extinction efficiency factor, $\tilde{m} = n - i \times k$ is the known complex index of refraction and $n(r,z)$ is the altitude-dependent aerosol size distribution. The later equation is usually simplified by performing the altitude integration, which is equivalent to using the columnar size distribution $n_{\mathrm{c}}(r)$ instead of $n(r,z)$. The relation between both size distributions is

(3) $$n_{\mathrm{c}}(r) = \int_0^\infty n(r,z) \, dz$$

and the AOT expressed by $n_{\mathrm{c}}(r)$ is

(4) $$\tau_{\mathrm{A}}(\lambda) = \int_0^\infty \pi r^2 \, Q_{\mathrm{ext}}(2\pi r/\lambda, \tilde{m}) \, n_{\mathrm{c}}(r) \, dr \ .$$

The MICROTOPS II performs an automatic correction for Rayleigh scattering and outputs $\tau_{\mathrm{A}}(\lambda)$ at the five channels' peak wavelengths.

A comprehensive review of the methods used to retrieve aerosol size distributions based on measured AOT is given by Mather *et al.* (2004). One of the best algorithms is that proposed by King *et al.* (1978) and thoroughly investigated by King (1982) and Jorge *et al.* (1996).

This manual describes the theoretical background and usage of a software package for retrieval of columnar optical thickness of aerosols concentrated in plumes. To investigate the aerosol properties of the plume (originating from volcanoes, fires, industrial plants, etc.), it is necessary to make measurements observing the Sun through both the background atmosphere (clear from plume) and the plume. The optical thickness of the plume $\tau_{\mathrm{p}}(\lambda)$ can then be obtained by subtracting the optical thickness of the background atmosphere $\tau_{\mathrm{bg}}(\lambda)$ from the total through-plume optical thickness $\tau_{\mathrm{total}}(\lambda)$

(5) $\tau_{\mathrm{plume}}(\lambda) = \tau_{\mathrm{total}}(\lambda) - \tau_{\mathrm{bg}}(\lambda)$.

If it is assumed that the atmosphere remains homogeneous between the background and in-plume measurements (i.e. both measurements are made in quick succession and within close proximity), that the optical thickness of the 'clear' atmospheric layer filled by the plume is negligible compared to $\tau_{bg}(\lambda)$, and also that the Rayleigh optical thickness ($\tau_R(\lambda)$) is approximately the same for $\tau_{total}(\lambda)$ and $\tau_{bg}(\lambda)$, then $\tau_p(\lambda)$ is the plume aerosol optical thickness. Sunphotometers have previously been used to study aerosol emissions from the sustained degassing of Mount Etna (Watson and Oppenheimer, 2000, 2001) and Kilauea (Porter *et al.*, 2002) and Lascar and Villarrica volcanoes, Chile (Mather et al., 2004). More recently MICROTOPS II sunphotometers have been used to study the plume from the Buncefield oil depot fire (Mather et. al., 2007).

The package presented in this manual is based on the investigations reported by King *et al.* (1978) and King (1982). The first version of the computer programs was adapted from FORTTAN77 codes **CSIGT**, **PREPRCES** and **RADINV, originally located at** http://ftpwww.gsfc.nasa.gov/crg/software.html (no longer available). These three programs have been modified by adding calculations of the surface and volume distributions, including a set of effective radii (Seinfeld and Pandis, 1998) and improving the structure of input/output files etc.

Codes in the package are created using FORTRAN95 and Basic. The executable programs are DOS programs. The user should be familiar with formatted FORTRAN inputs, although the structure of all input files is explained in detail in this manual. The following is part of formatted FORTRAN input file :

```
_ _ _2
_ _1.250000_ _0.000000
```

Hereafter the symbol "_" stands for an empty position in the data stream. Empty positions have a crucial importance in formatted FORTRAN data files (inputs). The symbols used to explain FORTRAN formats are: **#** - digits before decimal point; **&** - digits after decimal points; **.** – decimal point; $$$$ - E±XX exponent presentation in scientific format; **±** - position of sign. Example: #### is an integer number with maximum four digits, the leading digits could be replaced by **"+"** or **"-"** sign.

The BASIC codes in the package (as used in Newtam.exe) can handle only short filenames constructed by letters and numbers but no longer than 8 characters. Therefore long path names should be avoided.

Table 1 contains a list of codes included in the software package **MICROTOPS INVERSE** accompanied with brief description of their purpose and associated test input and output files. Detailed description of all input/output files can be found in following paragraphs where codes are discussed. Most of the input and output files have fixed names that facilitates managing of code execution. Only the code **NEWTAM** uses input files with variable filenames. These names are printed in red in Table 1. Each executable code requires a corresponding configuration file (e.g. **newtam.cfg**, **eff_factors.cfg** etc.). In most of the cases the purpose of the configuration file is to point all other input and output files to a working folder selected by the user. The executable code and its configuration file can be placed in any folder.

A few of the references cited in this manual are available for downloading. These are printed in red.

The authors are very grateful to Jennifer Le-Blond ans Rachel Gaulton for doing an admirable job in reading and correction of the text. Their valuable comments greatly improved the manual.

Table 1

Code	Purpose	Input files	Output files
1	**2**	**3**	**4**
NEWTAM	Reads two* MICROTOPS II output files that have been preliminarily cleared from all non-necessary records. The first file contains groups of scans registered in background conditions. The second file contains groups of scans recorded beneath a plume or measured aerosol ensemble. Program **NEWTAM** performs averaging of AOT within scan groups and prepares input file for programmes **PRELIMINARY** and **INVERSION**.	**newtam.cfg** **mt_7346.cal** **testetna.csv**	**temp.dat** **inv_in.dat** (this file has to be renamed as **inversion_in.dat**)**
EFF_FACTORS	Computes Mie efficiency factors for a set of complex indices of refraction which are used by programmes **PRELIMINARY** and **INVERSION.**	**eff_factors.cfg** **eff_factors_in.dat**	**eff_factors_out.da** **invversion_in_ext.dat**
PRELIMINARY	Performs inversion for a set of radii intervals and for three values of Junge parameter (ν^*-0.5, ν^* and ν^*+0.5) and permits estimation of the most informative or optimum radii interval for the inversion.	**inversion.cfg** **inversion_in.dat** (this is renamed file **inv_in.dat**) **	**preliminary_out.dat** **preliminary_short.dat**

* These two files may coincide, i.e. the background and plume data could be taken from the same file; see section 4 where is explained the content of the file **newtam.cfg**.

** The renaming of the file is obligatory.

Table 1 (continuation)

1	2	3	4
INVERSE	Performs inversion of AOT measured my MICROTOPS II and obtains estimates of : (a) number, surface and volume columnar distributions for a predefined number of radii (radii intervals); (b) columnar burdens within predefined radii intervals; (c) set of effective radii for the investigated aerosol ensemble	**inversion.cfg** **inversion_in_ext.dat** **inversion_in.dat** (this is renamed file **inv_in.dat**) **	**inversion_out_king.dat** **inversion_out_short.dat** **inversion_gamma.dat** **inversion_eff.dat** **inversion_out.dat** **inversion_tau.dat** **inversion_distr.dat** **inversion_rad.dat** test_13nov1975.opj (applicable for ORIGIN users)
SORT_DISTR	Sorts the output file inversion_distr.dat and facilitates further use of the results.	**sort_distr.cfg** **inversion_distr.dat**	**protocol.dat** **number_disrtr.dat** **particles.dat** **surface_distr.dat** **volume_distr.dat**
SORT_TAU	Sorts the output file inversion_distr.dat and facilitates further use of the results	**sort_tau.cfg** **inversion_tau.dat**	**measured_tau.dat** **calculated_tau.dat**

** The renaming of the file is obligatory.

1. Theoretical background

The retrieval of the aerosol distributions is based on the solution of the Fredholm integral equation

(6) $$\tau_{\mathrm{M}}(\lambda) = \int_0^{\infty} \pi r^2 Q_{\mathrm{ext}}\left(\frac{2\pi r}{\lambda}, \tilde{m}\right) n_{\mathrm{c}}(r)\, dr \ ,$$

where

- $n_{\mathrm{c}}(r)$ is the unknown columnar aerosol size distribution $n_{\mathrm{c}}(r)$ (or columnar aerosol number density) in linear r-scale; $n_{\mathrm{c}}(r)$ is the number of particles per unit area per unit radius interval in vertical column through the atmosphere; units (particles $\times$ cm^{-2} $\times$ μm^{-1});
- $\tau_{\mathrm{M}}(\lambda)$ is the aerosol optical thickness (AOT), measured at some discrete set wavelengths $\lambda_1, \lambda_2, \ldots, \lambda_p$;
- $Q_{\mathrm{ext}}\left(\frac{2\pi r}{\lambda}, \tilde{m}\right)$ is the extinction efficiency factor calculated from Mie theory;
- $\tilde{m} = n - i \times k$ is the complex index of refraction (supposed to be known).

Utilization of Equation (6) implicitly assumes that the aerosol particles are spheres with constant chemical composition over all radii. Thus we automatically are excluding external aerosol mixtures from any considerations. An expression for $n_{\mathrm{c}}(\lambda)$ cannot be derived by analytical methods and is, therefore, necessary to apply an appropriate approximation. The standard approach is based on replacement of the integral in Equation (6) by a summation over a few intervals and use of finite limits instead of infinite, i.e.

(7) $$\tau_{\mathrm{M}}(\lambda) = \int_{r_a}^{r_b} \pi r^2 Q_{\mathrm{ext}}\left(\frac{2\pi r}{\lambda}, \tilde{m}\right) n_{\mathrm{c}}(r)\, dr = \sum_{j=1}^{q} \int_{r_j}^{r_{j+1}} \pi r^2 Q_{\mathrm{ext}}\left(\frac{2\pi r}{\lambda}, \tilde{m}\right) n_{\mathrm{c}}(r)\, dr \, .$$

Here $r_1 = r_a$ and $r_b = r_{q+1}$ and $r_1, r_2, \ldots, r_{q+1}$ are the $q+1$ boundaries of q coarse intervals of integration. These intervals have equal length $\Delta \log r = (\log(r_b) - \log(r_a))/q$ and boundaries $\log(r_k) = \log(r_{k-1}) + \Delta \log r$ in logarithmic scale. Each coarse interval is composed of several sub-intervals required for numerical calculation of terms $\int_{r_j}^{r_{j+1}} \pi r^2 Q_{\mathrm{ext}}\left(\frac{2\pi r}{\lambda}, \tilde{m}\right) n_{\mathrm{c}}(r)\, dr$. The solution of Equation (6) can be significantly simplified assuming $n_{\mathrm{c}}(\lambda) = h(r)\, f(r)$, where $h(r)$ is some rapidly varying function, whilst $f(r)$ is slowly varying and constant within each coarse interval.

Thus the Freehold equation is replaced by simultaneous equations

(8) $\vec{\mathbf{g}} = \hat{\mathbf{A}}\,\vec{\mathbf{f}} + \vec{\boldsymbol{\varepsilon}}$,

where

(8.a) $\mathbf{g}_i = \tau_{\mathrm{M}}(\lambda_i)$, $i = 1, 2, \ldots, p$,

(8.b) $\mathbf{A}_{ij} = \mathbf{A}_{ij}(\lambda_i) = \int\limits_{r_j}^{r_{j+1}} \pi r^2 Q_{\mathrm{ext}}\left(\dfrac{2\pi r}{\lambda_i}, \tilde{m}\right) h(r)\, dr$, $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, q$,

(8.c) $\mathbf{f}_j = \mathbf{f}(\bar{r}_j)$,

and

(8.d) $\bar{r}_j = \sqrt{r_j\, r_{j+1}}$

is the geometrical midpoint of j-th coarse interval $[r_j, r_{j+1}]$. Note, the solution of Equation (8) is obtained in logarithmic scale with respect to particle radius r.

The Junge size distribution

(9) $h(r) = r^{-(\nu^* + 1)}$

could be considered as the simplest approximation of the rapidly varying function $h(r)$ with $\nu^* = \alpha + 2$, where α is the exponent in Ångström's empirical formula $\tau_{\mathrm{M}}(\lambda) = \beta \lambda^{-\alpha}$. Some authors are refer to α as Ångström's turbidity coefficient.

The elements $\boldsymbol{\varepsilon}_i$ of the unknown vector $\vec{\boldsymbol{\varepsilon}}$ represent the deviations between measurement $(\mathbf{g}_i)$ and theoretical estimate $\tau_{\mathrm{C}}(\lambda_i) = \sum\limits_j \mathbf{A}_{ij}(\lambda_i)\mathbf{f}_j$. These deviations arise from measurement and quadrature (integration) errors and due to uncertainties of the kernel function $\pi r^2 Q_{\mathrm{ext}}\left(\dfrac{2\pi r}{\lambda}, \tilde{m}\right)$.

The solution vector $\vec{\mathbf{f}}$ can be obtained by minimizing a performance function

(10) $Q = Q_1 + \lambda Q_2$.

Here γ is some non-negative Lagrangian parameter,

(11) $Q_1 = \vec{\boldsymbol{\varepsilon}}^{\mathrm{T}}\, \hat{\mathbf{C}}^{-1}\, \vec{\boldsymbol{\varepsilon}}$,

and

(12) $Q_2 = \vec{\mathbf{f}}^{\mathrm{T}}\,\hat{\mathbf{H}}\,\vec{\mathbf{f}}$,

$\hat{\mathbf{H}}$ denotes Twomey's matrix (smoothing constraint on second derivatives of $\mathbf{f}_i$), $\hat{\mathbf{C}}$ is the measurement co-variance matrix and superscripts correspond to transposing and inversion operations. Now we can replace Equation (10) by

(13) $$Q = \sum_{i=1}^{p}\sum_{j=1}^{p} C_{ij}^{-1}\boldsymbol{\varepsilon}_i\,\boldsymbol{\varepsilon}_j + \lambda \sum_{j=2}^{q-1}(\mathbf{f}_{j-1} - 2\mathbf{f}_j + \mathbf{f}_{j+1}).$$

The minimum value of Q corresponds to the statistically optimum estimate of $\vec{\mathrm{f}}$. Differentiating Equation (13) with respect to unknowns $\mathbf{f}_i$ and using (8) one can obtain

(14) $\vec{\mathbf{f}} = (\hat{\mathbf{A}}^{\mathrm{T}}\,\hat{\mathbf{C}}^{-1}\,\hat{\mathbf{A}} + \gamma\,\hat{\mathbf{H}})^{-1}\,\hat{\mathbf{A}}^{\mathrm{T}}\,\hat{\mathbf{C}}^{-1}\,\vec{\mathbf{g}}$.

The co-variance matrix for the MICROTOPS data is

(15) $\mathbf{C}_{ij} = \delta_{ij}\,w_i$, $w_i = 1/(\delta\tau_{\mathrm{M}}(\lambda_i))^2$,

where $\delta\tau_{\mathrm{M}}(\lambda_i)$ are the standard deviation of $\tau_{\mathrm{M}}(\lambda_i)$ or measurement errors and δ_{ij} is Kronecker delta.

Two important advantages can result from the assumption $n_{\mathrm{c}}(r) = h(r)\,f(r)$. Firstly, when $h(r)$ represents size distribution $n_{\mathrm{c}}(\lambda)$ exactly, the solution vector $\vec{\mathbf{f}}$ will have all components equal to one. Secondly, the smoothing constraint guaranties minimum curvature of $\vec{\mathbf{f}}$ on a linear scale and it works better when $\mathbf{f}_j = \mathbf{f}(\bar{r}_j)$ are almost constants. The columnar size distribution $n_{\mathrm{c}}(r)$ varies over some orders of magnitude and has an explicitly large curvature, thus it is difficult to constraint.

2. Solution algorithm

We have measured (using the Sun photometer) the columnar AOT $\tau_{\mathrm{M}}(\lambda_i)$, $i=1,2,\ldots,p$, and the corresponding standard deviations $\delta\tau_{\mathrm{M}}(\lambda_i)$. We suppose that aerosol particles have complex index of refraction $\tilde{m}=n-i\times k$, i.e., we know extinction efficiency factor $Q_{\mathrm{ext}}\left(\frac{2\pi r}{\lambda},\tilde{m}\right)$. We want to retrieve columnar aerosol size distribution $n_{\mathrm{c}}(r)$ over the radii interval $[r_a, r_b]$. This is a typical inverse problem and we can solve it using an iteration process. This means we have to start with some zero-approximation $h^{(0)}(r)$ of the rapidly varying multiplier in $n_{\mathrm{c}}(r)=h(r)\,f(r)$ and use it to evaluate a first-order approximation of solution vector $\vec{\mathbf{f}}^{(1)}$. Then we have to utilize $\vec{\mathbf{f}}^{(1)}$ and obtain some reasonable first order approximation $h^{(1)}(r)$ of rapidly varying multiplier and repeat these steps k-times until $\vec{\mathbf{f}}^{(k)}$ converges to unit vector $\vec{\mathbf{1}}$, i.e. until $h^{(k)}(r)$ approaches $n_{\mathrm{c}}(r)$.

The Junge size distribution given by Equation (4) is a good guess for $h^{(0)}(r)$. Aimed to obtain the first approximation $\vec{\mathbf{f}}^{(1)}$ of the solution vector we have to evaluate components of $\hat{\mathbf{A}}$ matrix using the Equation (8.b). The integral over coarse interval $[r_j, r_{j+1}]$ can be replaced by sum of integrals over sub-intervals covering it, i.e.,

$$(16)\qquad \mathbf{A}_{ij}=\int_{r_j}^{r_{j+1}}\pi r^2 Q_{\mathrm{ext}}\left(\frac{2\pi r}{\lambda_i},\tilde{m}\right)h(r)\,dr=\sum_{k=N(j)}^{N(j+1)-1}\int_{R_j}^{R_{j+1}}\pi r^2 Q_{\mathrm{ext}}\left(\frac{2\pi r}{\lambda_i},\tilde{m}\right)h(r)\,dr\cong\sum_{k=N(j)}^{N(j+1)-1}\pi\overline{R}_k^2 Q_{\mathrm{ext}}\left(\frac{2\pi\overline{R}_k}{\lambda_i},\tilde{m}\right)\int_{R_k}^{R_{k+1}}h(r)\,dr\,.$$

Here R_k are the boundaries of the sub-intervals that satisfy: $R_1=r_1=r_a$, $R_K=r_q=r_b$, $k=1,2,\ldots,K$, $R_{N(j)}=r_j$, $R_{N(j+1)}=r_{j+1}$, $\overline{R}_k=\sqrt{R_k\,R_{k+1}}$ (geometrical midpoint of k-th sub-interval $[R_k, R_{k+1}]$). As far as $n_{\mathrm{c}}(r)$ has a very sharp shape and the extinction cross-section $\pi r^2 Q_{\mathrm{ext}}\left(\frac{2\pi r}{\lambda},\tilde{m}\right)$ is relatively smooth compared with $n_{\mathrm{c}}(r)$, we can select a large enough value of K so that each coarse interval contains at least ten sub-intervals within which $\pi r^2 Q_{\mathrm{ext}}\left(\frac{2\pi r}{\lambda},\tilde{m}\right)$ is a constant. This means that the approximation in the last term of Equation (16) holds true and we can consider $W^{(0)}(\overline{R}_k)=\int_{R_k}^{R_{k+1}}h(r)\,dr$ as weighting function in the numerical integration

$$(17)\qquad \int_{r_j}^{r_{j+1}} \pi r^2 Q_{\mathrm{ext}}\left(\frac{2\pi r}{\lambda_i}, \tilde{m}\right) h^{(0)}(r)\, dr \cong \sum_{k=N(j)}^{N(j+1)-1} \pi \bar{R}_k^2 Q_{\mathrm{ext}}\left(\frac{2\pi \bar{R}_k}{\lambda_i}, \tilde{m}\right) W^{(0)}(\bar{R}_k).$$

Having calculated components $\mathbf{A}_{ij}$ we can obtain the first approximation of the solution vector $\vec{\mathbf{f}}^{(1)}$ using Equation (14). Actually we have estimates $\mathbf{f}^{(1)}(\bar{r}_j)$ and thus we can calculate only $W^{(1)}(\bar{r}_j) = \mathbf{f}_j^{(1)}(\bar{r}_j)\, W^{(0)}(\bar{r}_j)$ and then use linear approximation to evaluate $W^{(1)}(\bar{R}_k)$ for all geometrical midpoints $\bar{R}_k \neq \bar{r}_j$. Now we are ready to perform second iteration, i.e. to estimate $\mathbf{f}^{(2)}(\bar{r}_j)$ and so on. In principle this process as mentioned above has to continue until $\vec{\mathbf{f}}^{(k)}$ converges to unit vector $\vec{\mathbf{1}}$, but M. King (1982) proved that usually no more that eight iterations are needed.

The selection of the value of Lagrange multiplier is a very specific issue in the considered constrained linear inversion of AOT $\tau_{\mathrm{M}}(\lambda_i)$. King (1982) proved that is favourable to consider relative value $\gamma_{\mathrm{rel}} = \gamma \mathbf{H}_{11} / (\hat{\mathbf{A}}^T \hat{\mathbf{C}}^{-1} \hat{\mathbf{A}})_{11}$ of Lagrange parameter γ and that γ_{rel} varies in the range 10^{-3} to 5 until range of values are reached for which all element of solution vector $\vec{\mathbf{f}}$ are positive (negative values constitute an unphysical solution). More over in the case of covariance matrix given by Equation (15) it can be proved that

$$(18)\qquad Q_1 \leq E\{Q_1\} = p\ ,$$

where $E\{\ \}$ denotes expectation operator. Thus the algorithm of γ-value selection during each iteration is as follows:

- ♣ start with $\gamma_{\mathrm{rel},0} = 0$ and evaluate $\mathbf{f}^{(k)}(r_j \,|\, \gamma_{\mathrm{rel},0})$, $Q_1(\gamma_{\mathrm{rel},0})$, $Q_2(\gamma_{\mathrm{rel},0})$, respective $Q(\gamma_{\mathrm{rel},0})$, and count the number $\mathrm{M}(\gamma_{\mathrm{rel},0})$ of negative components $\mathbf{f}^{(k)}(r_j \,|\, \gamma_{\mathrm{rel},0})$.
- ♣ repeat this for 14 values of γ_{rel} starting with $\gamma_{\mathrm{rel},1} = 0.001$ and then doubling each subsequent value, i.e. $\gamma_{\mathrm{rel},2} = 0.002$, $\gamma_{\mathrm{rel},3} = 0.004$, …, $\gamma_{\mathrm{rel},14} = 4.096$.
- ♣ find the smallest γ_{rel}-value $\gamma_{\mathrm{rel},m}$ $(m = 0,1,2,\ldots,14)$ for which $\mathrm{M}(\gamma_{\mathrm{rel},m}) = 0$ and $Q_1(\gamma_{\mathrm{rel},m}) \leq p$ and accept $\mathbf{f}^{(k)}(r_j \,|\, \gamma_{\mathrm{rel},m})$ as solution during k -th iteration.
- ♣ if some of solution components $\mathbf{f}^{(k)}(r_j \,|\, \gamma_{\mathrm{rel},14})$ are still negative at $\gamma_{\mathrm{rel},14} = 4.096$ then other non-negative components $\mathbf{f}^{(k)}(r_j \,|\, \gamma_{\mathrm{rel},14})$ are used to obtain some non-negative extrapolation on a $\log \mathbf{f}(\bar{r}_j)$ *vs.* j scale.

Appendix 2 contains a few examples explaining the above-described algorithms in more details.

Now we can summarize King's algorithm for inversion of AOT data and retrieving aerosol columnar size distributions. The inversion starts with an initial approximation accepted to be the Junge distribution with parameter ν^* estimated using Angström exponent α. The initial approximation is later iteratively modified until the final solution satisfies both the coincidence of measured and calculated optical thickness within experimental noise level and the positivity constraint. The goodness of inversion is characterized by five interrelated quantities:

♦ Covariance matrix of solution or corresponding mean relative error E_{rel} of the solution vector $\vec{\mathbf{f}}$ components

The covariance matrix of solutions is calculated as

(19) $$\hat{\mathbf{S}} = (\hat{\mathbf{A}}^{\mathrm{T}}\hat{\mathbf{C}}^{-1}\hat{\mathbf{A}} + \lambda\hat{\mathbf{H}})^{-1}$$

if the measurement covariance matrix $\hat{\mathbf{C}}$ is known or otherwise as

(20) $$\hat{\mathbf{S}} = s^2(\hat{\mathbf{A}}^{\mathrm{T}}\hat{\mathbf{A}} + \lambda\hat{\mathbf{H}})^{-1},$$

where

(21) $$s^2 = \frac{1}{p-q}\sum_{i=1}^{p}\varepsilon_i^2$$

is the sample variance of the regression fit. As it is not convenient to print and, even less, to compare matrices we can use the mean relative error E_{rel} of the solution vector $\vec{\mathbf{f}}$ components instead. This quantity is defined as

(22) $$\mathrm{E}_{rel} = 100\frac{1}{p}\sum_{i=1}^{p}\sqrt{\mathbf{S}_{ii}}\,/\mathbf{f}_i \ \text{(percent)}.$$

♦ The sum

(23) $$\sum_{i=1}^{p}\boldsymbol{\varepsilon}_i^2 \ ,$$

where $\boldsymbol{\varepsilon}_i = \tau_{\mathrm{M}}(\lambda_i) - \tau_{\mathrm{C}}(\lambda_i)$. This quantity is a measure of sample variance s^2 of the regression fit.

♦ The first term in performance function, i.e.

(24) $$Q_1 = \vec{\boldsymbol{\varepsilon}}^{\mathrm{T}}\hat{\mathbf{C}}^{-1}\vec{\boldsymbol{\varepsilon}} = \sum_{i=1}^{p} w_i\,\boldsymbol{\varepsilon}_i^2 \ .$$

This quantity is a useful measure of goodness of inversion due to the relation $Q_1 \le E\{Q_1\} = p$, which gives a natural upper limit of its value.

♦ The number of coincidences M_c of retrieved aerosol optical depth with measured accounting for experimental error. This is the number of times when the inequality $\left|\tau_C(\lambda_i) - \tau_M(\lambda_i)\right| \le \delta\tau_M(\lambda_i)$ holds true or the number of retrieved measured AOTs or the number of coincidences between measured and calculated AOTs.

♦ The coincidence of solution obtained with ν^*-0.5, ν^* and ν^*+0.5. This requirements means that solution has to be relative non-sensitive to the shape of initial approximation $h^{(0)}(r)$ of $h(r)$.

In other words, King's algorithm performs constrained linear inversion of the measured spectral AOTs to obtain columnar aerosol size distribution and other related distributions and quantities (effective ensemble radii). King's algorithm relies on two constraints – a positive solution vector that constitutes a physical size distribution and satisfaction of the original integral equation to within the noise of the measurements. A more sophisticated and powerful algorithm is based on measurement of direct and scattered solar radiation in almucantar plane (Dubovik and King, 2000) but is not applicable to our case due to strong heterogeneity of plumes

A few practical difficulties appear when applying King's method:

(a) *A priori* known Mie extinction efficiency factor $Q(2\pi r/\lambda, \tilde{m})$ within considered radii interval $[r_a, r_b]$ is required, i.e. aerosol is assumed as a polydispersion of spherical particles with a known single index of refraction and thus external mixtures are excluded from the consideration (cf. Lesins *et al.*, 2002).

(b) (b) It is necessary to evaluate the radii limits of maximum sensitivity and guess the particulate index of refraction. These problems and their solutions will be discussed in more detail in the next paragraphs.

It is crucial to comment on the impact of the experimental errors $\delta\tau_M(\lambda_i)$ on the solution. If the uncertainties of the measurements are unusually large, then it is relatively easy to obtain an inversion solution which satisfies both criteria (positivity and coincidence of measured and calculated AOTs) but the uncertainties in the solution will also be large. On the other hand, measurement errors which are estimated as unrealistically small may preclude one from being able to obtain a final solution which satisfies both constraints. It is therefore important that realistic uncertainties be assigned to the data before performing the inversion

3. Preliminary handling of MICROTOPS II data

The MICROTOPS II is a five-channel, handheld Sun photometer that can be configured to measure total ozone, total water vapour, or aerosol optical thickness at various wavelengths. The instrument measures $10 \times 20 \times 4.3$ cm and weighs 600 g. The principal design goal of the instrument has been the measurement of total atmospheric ozone content with accuracy about 1%. The intension was to provide an alternative to the much larger, heavier, and more expensive Dobson and Brewer spectrophotometers. The photometer utilizes highly stable ultraviolet filters manufactured with an ion deposition process and full width at half maximum (FWHM) 10 nm. The optical collimators (2.5° full field of view), low noise electronics and a high-performance 20-bit analogue-to-digital converter (ADC) of the instrument are carefully designed to optimize pointing accuracy, stray light rejection, thermal and long-term stability, signal-to-noise ratio, high linearity, resolution and dynamic range, and on board precise data analysis. The overall the dynamic range achieved in the instrument is about 300000, which corresponds to ≈8.3 μV resolution at 2.5 V full-scale of ADC. An internal microcomputer automatically calculates the total ozone column, aerosol optical thickness and precipitable water vapour. These calculations are based on measurements at the optical channels, the site's geographic coordinates, altitude, atmospheric pressure, universal time and date. The coordinates can be entered manually or by a Global Positioning System (GPS) receiver. A built-in pressure transducer automatically measures pressure. The MICROTOPS II saves up to 800 scans as raw and calculated data in nonvolatile memory. Measurements can be read from a liquid crystal display or transferred to an external computer via a RS interface.

The design, calibration and performance of operation of MICROTOPS II hand-held Sun photometers is described in detail in the instrument's manual and by Morys *et al.* (2001).

Experimental data are grouped in scans. Each scan consists of up to 64 samples from each of the five channels. The samples are taken in a rapid sequence at a rate of over 3 samples/second (one sample contains readings from all five channels). Consequently, the maximum time for a single scan is about 20 seconds. During this time interval the user has to firmly hold the Sun photometer, pointing it towards the Sun by keeping the image of the Sun is centered in the bull's-eye of the sun target. The number of samples in a scan (scan length) can be set by the user from 1 to 64. The default value is 32 and it is suitable for virtually all conditions. Only a selected number of samples, defined by the user (note, the default value 4 is a quite small for reliable statistical estimation) and having higher-ranking signal strengths, are used to calculate the signal's mean value and its standard deviation. These both estimates are passed for further processing. It is desirable to increase the number of the averaged scans but this imposes difficulties when pointing the MICROTOPS II towards the Sun.

The content of a MICROTOPS II data file downloaded to the computer may vary significantly depending on the software version used. A few examples are given in files: **example1_etna_22julyMTOPS_MT7346.txt** ,
example2_15_12-2005_Madhavi.xls ,
example3_villarrica sunset 1102.xls ,
example4_Langley2.xls.

The text files are comma-delimited ASCII files. In each case, each line contains the data related with just one scan. Excel, Origin, Kaleidagraph and many other software packages can access the output MICROTOPS II files. The columns in these contain different quantities

and are labeled correspondingly. **We only use the quantities listed in Table 2. All other columns must be deleted (cleared) and the cleared file has to be saved in Excel comma-delimited format (csv-file) or as comma-delimited ASCII file with unspecified (arbitrary) extension.** See for example the file **testetna.csv** obtained from **example1_etna_22julyMTOPS_MT7346.txt** by deleting/clearing all un-necessary quantities. The first two lines (line 1 - REC#0070, line 2 - FIELDS:) and the last line (END.) have been deleted as well. Further, columns ID, R440_675, R675_870, R870_936, R936_1020 are also deleted. They contain: User specified identification code (ID) and ratios (R) of signals $S(\lambda)$ in pairs of channels (S(440 nm)/S(675 nm), S(675 nm)/S(870 nm), S(870 nm)/S(936 nm) and S(936 nm)/S(1020nm)). If we have to tidy up a file like: **example2_15_12-2005_Madhavi.xls**, many more columns must be removed. The cleared file is the subject of subsequent analysis and data handling. It must only contain the columns listed in Table 2.

The next task is to check the cleared file for unsatisfactory scan records. These are uncompleted scans (usually the corresponding signal cells in the worksheets are filled up with "–999.00" or "###") or scans where some of measured quantities can be considered as outliers compared with neighbours. These scan records must be excluded from further consideration and the corresponding lines deleted. In this way we end up with a set of scans containing reliable data. Now we must split scan data into groups recorded in different conditions, i.e., presenting different aerosol ensembles. The latter can be done using the notes made during the measurements or by analyzing dependencies "signals *vs.* consecutive number", "signals *vs.* time", "AOT *vs.* time" etc. For example, the file **testetna.xls** (this is transformed to Excel format file **example1_etna_22julyMTOPS_MT7346.txt**) contains data recorded near an active degassing vent on Mt. Etna. Scan number 68 has longitude equal to 14.1° that differs significantly from 14.992° of all neighbouring scans and it is better that it be removed. Scan number 65 is characterized by very low values of the AOT in all five channels whilst its PWV value is not an obvious outlier. Probably this scan was performed when the plume was absent from the line of sight between the Sun photometer and the Sun. It is desirable also to exclude it. The analysis of plots in **testetna.xls** reveals that we can distinguish five groups of scans:

- ♦ scans 26-50 correspond to background aerosols – the plume was absent from the line of sight;
- ♦ scans 1-10 and 11-25 correspond to thin plume characterized by relatively weak fluctuations of AOT but with reasonably strong absorption at 440 nm and almost the same absorption in the other four channels; scans 1-10 and 11-25 are forming two groups because there exists a time gap between them;
- ♦ scans 51-60 and 61-70 are characterized by strong but different fluctuations of the AOT in all five channels.

Table 2

Column label	Content of the column or measured quantity
SN	4- or 5-digit serial number of the photometer
No	Consecutive number – this column has to be added by the user after the "clearing" the unsatisfactory scan records
DATE	The universal date of the measurement in format mm/dd/yyyy
TIME	The universal time of the measurement in format hh:mm:ss
LATITUDE	The latitude of the measurement site in degrees; latitude is positive North from equator
LONGITUDE	The longitude of the measurement site in degrees; longitude is positive East from Greenwich
ALTITUDE	The altitude of the measurement site in meters above see level
PRESSURE	Atmospheric pressure in millibars; either the pressure at a meteorological station situated nearby (manually entered) or the pressure measured by the internal sensor.
SZA	Solar zenith angle in degrees calculated by the internal microcomputer on the basis of TIME, LATITUDE, LONGITUDE and ALTITUDE. The SZA is 0° for overhead Sun and 90° for Sun on the horizon.
AM	Optical airmass factor, which is dimensionless quantity and is calculated by the internal microcomputer on the basis of SZA.
SDCORR	Earth-Sun distance correction, which is dimensionless quantity and is calculated by the internal microcomputer on the basis of DATE.
TEMP	Temperature of the optical block inside the photometer in degree Celsius.
SIG1 - **SIG440** SIG2 - **SIG675** SIG3 - **SIG870** SIG4 – **SIG936** SIG5 – **SIG1020**	Signals in millivolts from all five channels. The numbers 1 to 5 denote channel's wavelengths, which depend on the particular photometer. The blue labels correspond to wavelengths of MICROTOPS II S/N 7346.
STD1 - **STD440** STD2 - **STD 675** STD3 - **STD870** STD4 - **STD936** STD5 - **STD1020**	Standard deviations from all 5 optical channels, evaluated using all sampled samples within the current scan.
AOT1 - **AOT440** AOT2 - **AOT675** AOT3 - **AOT870** AOT4 - **AOT936** AOT5 - **AOT1020**	Aerosol optical thickness from all 5 optical channels. NOTE: These are aerosol optical thickness for the vertical atmospheric content, whilst the signals are measured along the inclined line of sight defined by the SZA.
WATER	Precipitable vertical columnar water (PWV) content in centimeters

4. Usage of program newtam – preparation of input data for inversion programs

As explained sections 1 and 2, the retrieval of columnar size distributions $n_c(r)$ requires:

(1) measured columnar AOT and corresponding standard deviations, i.e. $\tau_M(\lambda_i)$ and $\delta\tau_M(\lambda_i)$ for a set of wavelengths $i = 1, 2, \ldots, p$;

(2) the value of the exponent α in Ångström's empirical formula $\tau_M(\lambda) = \beta\lambda^{-\alpha}$.

The registration process of MICROTOPS II signals (each scan consists of enough large number of samples but only a few that have higher-ranking signal strengths are averaged in order to evaluate mean values and standard deviation during the particular scan) aims to obtain precise estimates of signals but with very low standard deviations (std).

The standard deviations of the signals originate from the natural randomness of the measured aerosol ensemble and from the errors of pointing to the Sun. We can distinguish two limiting cases of strongly variable plume and relatively constant plume during the time interval when samples are collected for calculation the mean and std values associated with a particular scan. Usually the errors of pointing are significantly reduced (almost negligible) by the measuring algorithm except in the case of very dense and fluctuating plume when sometimes is even difficult to observe the Sun. In this case signal standard deviations could increase. The typical level of the signal-to-noise values in all channels about 10^{+6} in the firs case. But beneath strong plume sometimes values of standard deviations could be comparable end even prevail mean values evaluated during intervals of 10-20 seconds.

The MICROTOPS II calculates the AOT value at each wavelength λ_i based on the channel's signal $V(\lambda_i)$ (millivolts), its extraterrestrial calibration constant $\ln(V_o(\lambda_i))$, atmospheric pressure p (millibarr), airmass factor (AMF) and Earth-Sun distance correction (SDCORR). The last two quantities are calculated using the geographical coordinates, altitude, time and date, which are provided by the GPS or entered by the user through instrument keyboard. The optical thickness calculations are based on the Bouguer-Lambert-Beer law (for more details see MICROTOPS II User Guide and Morys *et al.*, 2001)

$$(25) \qquad \mathrm{AOT}(\lambda_i) = \tau_M(\lambda_i) = \frac{\ln(V_o(\lambda_i)) - \ln(\mathrm{SDCORR}\ln(\mathrm{V}(\lambda_i)))}{\mathrm{AMF}} - k_R(\lambda_i)\frac{p}{p_0} \ .$$

Note that $\mathrm{AOT}(\lambda_i)$ characterizes the total vertical atmospheric column. The second term in above formula accounts for the contribution of Rayleigh scattering. The Rayleigh coefficient is evaluated according to (private communication by E.M. Rollin)

$$(26) \qquad k_R(\lambda_i) = \left[117.2594\,\lambda_i^4 - 1.3215\,\lambda_i^2 + 0.00032073 - 0.000076842\ \Big/ \lambda_i^2\right]^{-1},$$

where λ_i is in micrometer and the ratio p/p_0 accounts for the altitude of the measurement site ($p_0 = 1013.25\,\mathrm{mB}$ is the standard atmospheric pressure at see level).

Finally, AMF is the airmass factor is calculated according to

(27) $\mathrm{AMF} = \sec(Z) - 0.0018167\,(\sec(Z)-1)-1) - 0.002875\,(\sec(Z)-1)^2 - 0.0008083\,(\sec(Z)-1)^3$,

where Z is the Solar zenith angle. Morris *et al.* (2001) reported that in MICROTOPS II the maximum error ΔZ of Z -calculation is ±0.03° and it arises from simplified algorithms and partially from the use of single-precision arithmetic by the onboard MICROTOPS II microcomputer.

Using the above set of formulae it is easy to calculate $\mathrm{AOT}(\lambda_i)$ -values and compare them with the $\mathrm{AOT}(\lambda_i)$ -values, calculated by MICROTOPS II microcomputer and available through MICROTOPS II organizer (software).

The error (standard deviations) of MICROTOPS II $\mathrm{AOT}(\lambda_i)$ -values, which are not calculated by the MICROTOPS II, can be evaluated using the error propagation formulae

(28) $\delta(\mathrm{AOT}(\lambda_i)) = \sqrt{\sigma^2(\mathrm{AOT}(\lambda_i))}$,

(29) $$\sigma^2(AOT(\lambda_i)) = \left(\frac{\partial AOT(\lambda_i)}{\partial V(\lambda_i)}\right)^2 \times \sigma^2(V(\lambda_i)) + \left(\frac{\partial AOT(\lambda_i)}{\partial \mathrm{AMF}}\right)^2 \times \sigma^2(\mathrm{AMF}) + \left(\frac{\partial AOT(\lambda_i)}{\partial \mathrm{p}}\right)^2 \times \sigma^2(\mathrm{p})\ .$$

Here variances $\sigma^2(V(\lambda_i))$ are squared standard deviations of signals $V(\lambda_i)$ in i -th channel, variance $\sigma^2(\mathrm{p})$ is unknown and thus is accepted to be a constant ($\sigma^2(\mathrm{p}) = (\Delta p)^2$) independent on pressure value and airmass factor variance $\sigma^2(\mathrm{AMF})$ can be evaluated as

(30) $$\sqrt{\sigma^2(\mathrm{AMF})} = \Delta Z \times \frac{d\sec(\mathrm{Z})}{d\,Z} \times [1 - 0.0018167 - 0.002875\,(\sec(\mathrm{Z}) - 1) - 0.0008083\,(\sec(\mathrm{Z}) - 1)^2]\ .$$

Further, the formulae for calculation of squared partial derivatives are

(31) $(\partial AOT(\lambda_i)/\partial V(\lambda_i))^2 = (\mathrm{SDCORR}/(\mathrm{AMF}\,V(\lambda_i)))^2$,

(32) $(\partial AOT(\lambda_i)/\partial \mathrm{AMF})^2 = (\tau_{\mathrm{R}}(\lambda_i)/p_0)^2$,

(33) $$\left(\frac{\partial AOT(\lambda_i)}{\partial \mathrm{p}}\right)^2 = \left(\frac{\ln(V_o(\lambda_i)) - \ln(\mathrm{SDCORR}\ln(\mathrm{V}(\lambda_i)))}{\mathrm{AMF}^2}\right)^2\ .$$

In most of the cases for all channels standard deviations are about 10^{+6} times smaller than signals (i.e. signal-to-noise ratios are of the order of

10^{+6}) and thus they do not contribute to the evaluated values of $\delta(\mathrm{AOT}(\lambda_i))$. The impact of airmass factor error on $\delta(\mathrm{AOT}(\lambda_i))$ becomes comparable with the impact of pressure error mainly in two cases:(a) when signal-to-noise is low, i.e. when aerosol is dense and $\mathrm{AOT}(\lambda_i)$-values are high; (b) at solar zenith angles about 80° or more.

Now let us consider again Equation (25) re-writing it in the form

$$\mathrm{AOT}(\lambda_i) = \tau_{\mathrm{M}}(\lambda_i) = \frac{\ln\left(\dfrac{V(\lambda_i)}{V_o(\lambda_i)/\mathrm{SDCORR}}\right)}{\mathrm{AMF}} - k_{\mathrm{R}}(\lambda_i)\frac{p}{p_0} \quad . \tag{34}$$

The values of airmass factor AMF and the total atmosphere Rayleigh optical thickness $k_{\mathrm{R}}(\lambda_i)\, p/p_0$ can be considered constants during any time interval shorter that 20 seconds (this is the longest time interval for performing samples required to calculate scans average and std). The denominator $\tau(\lambda_i) = \ln\left(\dfrac{\mathrm{V}(\lambda_i)}{V_o(\lambda_i)/\mathrm{SDCORR}}\right)$ is the total atmospheric optical thickens along the line of the observation. It has the form $\tau = \ln(V_o/V)$, where wavelength dependency and SDCORR are ignored for simplicity. Further, $T = V/V_o$ is the transmittance of the whole atmosphere along the inclined path (line of view). It is very well known from quantitative spectroscopy that the non-linear transformation (logarithm of the ratio $T = V/V_o$) leads to the following expression for the relative error of calculated optical depth

$$\frac{\Delta\tau_{10}}{\tau_{10}} = \frac{\sqrt{10^{2\tau_{10}}+1}}{\tau_{10}}\frac{\Delta V_o}{V_o} = \phi(\tau_{10})\frac{\Delta V_o}{V_o}. \tag{35}$$

The relation between $\tau = \ln(V_o/V) = \tau_e$ and the optical thickness τ_{10} used in spectroscopy is $\tau_e = \ln(10)\tau_{10}$. In Equation (35) is assumed that the signals V_o and V, i.e. the signal entering and exiting the absorber (the whole inclined atmospheric path) are measured with the equal errors ΔV_o and ΔV. In this case the Equation (35) holds true independently on the wavelength. The function $\phi(\tau_{10})$ is known as the coefficient determining amplification of the measurement error after performing the non-linear logarithmic transformation. The behaviour of this function manifested in Table 3 and Figure 1. The "level 1" corresponds to ≈33% transmittance and it is optimal for performing measurements in quantitative spectroscopy. Usually it is accepted that the interval of transmittances providing double or triple excess of the minimum levels is the interval where is possible to perform reliable measurements. At lower $T = V/V_o$ values the signal-to-noise ration is quite small and error unacceptable. At higher $T = V/V_o$ values the amplitude resolution becomes crucial. In the case of MICROTOPS II the utilized low noise electronics and 20-bit analogue-to-digital converter provide 300000 overall dynamic range of the signal and thus permit to perform reliable measurements of $T = V/V_o$ up to ≈93% (level 7). This means that we cam measure the AOTs in all MICROTOPS II cannels with acceptable

accuracy if the denominator in Equations (13) and (22), i.e. $\tau(\lambda_i) = \ln\left(\frac{\mathrm{V}(\lambda_i)}{V_o(\lambda_i)/\mathrm{SDCORR}}\right)$ satisfies the inequality

(36) $0.07 \le \tau \le 3.77$.

Of coarse, the later conclusion should be considered with some precautions, as we do not know the error of the calibration constants $\ln(V_o(\lambda_i))$.

Thus

♠ if the plume is stable we will work with groups of scans, splitting our scans into relatively homogenous groups, average AOT within each of selected group and thus to obtain estimates of $\tau_{\mathrm{M}}(\lambda_i)$ and $\delta\tau_{\mathrm{M}}(\lambda_i)$

or

♠ if the plume is strongly fluctuating we will work with individual scans applying Equations (28-33) aimed to calculate a theoretical estimate of $\delta(\mathrm{AOT}(\lambda_i))$ and use it instead $\delta\tau_{\mathrm{M}}(\lambda_i)$.

Table 3

$T = e^{-\tau_e} = 10^{-\tau_{10}}$	τ_{10}	τ_e	$\phi(\tau_{10})$	Level
0.011	1.959	4.510	46.389	7× minimum
0.013	1.886	1.886	39.762	6× minimum
0.017	1.770	4.075	33.331	5× minimum
0.023	1.638	3.772	26.508	4× minimum
0.037	1.432	3.297	18.889	3× minimum
0.063	1.201	2.765	13.246	2 × minimum
0.330	0.481	1.109	6.627	1 – minimum
0.749	0.126	0.289	13.290	2× minimum
0.834	0.079	0.182	19.805	3× minimum
0.877	0.057	0.131	26.508	4× minimum
0.902	0.045	0.103	33.331	5× minimum
0.918	0.037	0.086	39.762	6× minimum
0.930	0.032	0.073	46.389	7× minimum

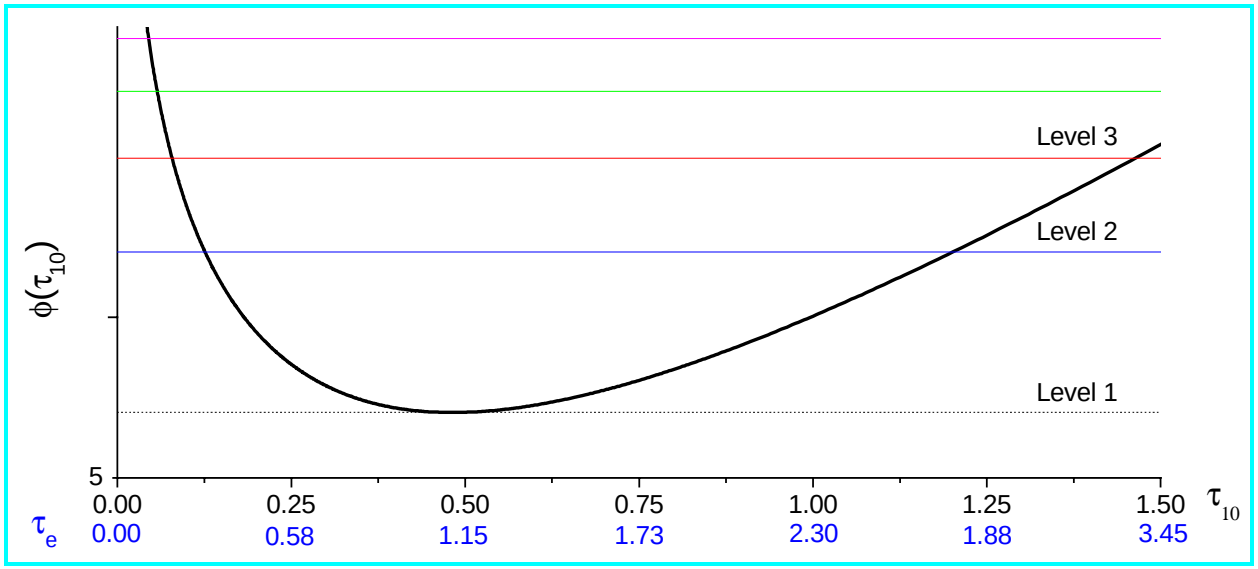

Fig. 1

Table 4

Group scans and description	Group's AOT mean (std)					Plume's AOT mean (std)					Angstrom parameters	
	440 nm	675 nm	870 nm	936 nm	1020 nm	440 nm	675 nm	870 nm	936 nm	1020 nm	α	β
1-10 thin plume #1	0.3968 (0.1485)	0.1290 (0.0378)	0.0869 (0.0212)	0.1002 (0.0172)	0.1139 (0.0138)	0.2818 (0.1526)	0.0910 (0.0396)	0.0521 (0.0224)	0.0576 (0.0193)	0.0633 (0.0174)	1.92	0.051
11-25 thin plume #2	0.3397 (0.1336)	0.1127 (0.0357)	0.0801 (0.0178)	0.0967 (0.0152)	0.1132 (0.0131)	0.2247 (0.1377)	0.0746 (0.0376)	0.0453 (0.0190)	0.0541 (0.0172)	0.0626 (0.0167)	1.57	0.049
26-50 background	0.1150 (0.0040)	0.0380 (0.0018)	0.0348 (0.0012)	0.0426 (0.0020)	0.0506 (0.0036)	-	-	-	-	-	1.09	0.037
51-60 thick plume #2	4.9219 (1.8896)	3.9200 (1.7497)	3.2365 (1.5496)	3.0221 (1.4602)	2.8077 (1.3731)	4.9179 (1.8936)	3.9182 (1.7515)	3.2353 (1.5508)	3.0201 (1.4622)	2.8041 (1.3766)	0.62	2.90
61-70 thick plume #2	1.3249 (0.6186)	0.8609 (0.4216)	0.6262 (0.3082)	0.5852 (0.2738)	0.5439 (0.2395)	1.2099 (0.6226)	0.8229 (0.4234)	0.5914 (0.3094)	0.5426 (0.2758)	0.4933 (0.2431)	0.03	0.69

If we apply the first approach to the scans recorded in file **example1_etna_22julyMTOPS_MT7346.txt** we can obtain five sets of input

data for the inversion. The averaging can be performed easily by means of Excel or Origin. An example of this averaging procedure is given in file **testetna.xls**. The Ångström's coefficients are then obtained by constructing a linear regression of $\log(\tau_{\mathrm{M}}(\lambda_i))$ on $\log(\lambda_i)$. The results are summarised in Table 4.

The process described above for obtaining input data for $n_{\mathrm{c}}(r)$-retrieval is quite time consuming and it is very easy to make mistakes during the data manipulation. An additional complication is the need to arrange the obtained values as input files for the programs performing calculation of $n_{\mathrm{c}}(r)$. To avoid these technical difficulties an assistant program has been created (**newtam**). Its purpose is to handle two files with data – first one (background file) containing scans (background scans) recorded in background conditions, whilst the second one (data file) – scans (data scans) recorded beneath the plume. The program uses a configuration file (**newtam.cfg**) to set the task to be performed. Below one configuration file is listed and its content explained.

"Path to all used files : " , "D:\J2\NT_007\"	**First group of lines printed on the screen**	Path to all input and output files, NOTE: Paths and file names must be no longer than 8 characters.
"MCROTOPS II calibration file : " , "MT_7346.cal"		The name of the files with calibration constants of the utilized photometer
"Output file : " , "inv_in.dat"		File name of the main output file, later one this file serves as input file for programs performing retrieval of $n_{\mathrm{c}}(r)$
"Date transformation key : " , 1		Key constant used to specify the format of date, recorded by MICROTOPS II – see below
"File with background data : " , "testetna.csv"		Name of background file, containing background scans
26 50		Numbers of first and last background scans (lines in the file) within the background file
"Background Etna July 22, 2006 11:11:09-11:16:34"		Some text (up to 80 characters including spaces), describing the background

"File with data : " , "testetna.csv"	**Second group of lines printed on the screen**	Name of data file, containing data scans
"Number of groups to handle : " , 4		Number of groups of scans, corresponding to some specific aerosols
1 10	**Third group of lines printed on the screen**	Numbers of first and last data scans (lines in the file) forming the first group of scans within the data file
"Set 1 Etna July 22, 2006"		Some text (up to 80 characters including spaces), describing the first group of scans
11 25		The same information but about second group of scans, etc.
"Set 2 Etna July 22, 2006"		
51 60		
"Set 3 Etna July 22, 2006"		
61 70		
"Set 4 Etna July 22, 2006"		
57 57		These six lines will indicate handling individually scans numbers 57, 63 and 68. Note that for this purpose the line/scan number has to be entered twice.
"Scan 57 Etna July 22, 2006"		
63 63		
"Scan 63 Etna July 22, 2006"		
68 68		
"Scan 68 Etna July 22, 2006"		
	These last 13 lines are not printed on the screen; do not modify them	**Two empty lines**
*** Date transformation key has to be set as follows :		Values of key constant used depending on the format of date, recorded by MICROTOPS II – see below
key.date=0 when date format is dd/mm/yyyy		
key.date=1 when date format is mm/dd/yyyy		

	One empty line
*** Background subtraction key has to be set as follows :	Values of key constant used to determine either the studied object is a plume aerosol (than we have to subtract background AOTs) or not
key.background=0 when background AOT are not subtracted	
key.background=1 when background AOT are subtracted	
	One empty line
*** Intermadiate print key has to be set as follows :	Values of key constant used permit or not printing on the screen error of AOTs estimated using error propagation formula
key.print=0 when intermediate print is forbidden	
key.print=1 when intermediate print is permitted	

Please, note that an input parameter could be some text or number. The input parameters are arranged in lines. Each line consists of text surrounded by quotation marks, and/or numbers and commas, used as delimiter (if a line contains two input parameters). All text inputs have to be surrounded by quotation marks. **The user should only modify the text or number after the comma.** The first text in lines (if any exists) is an explanation. Any number of spaces could be typed between the texts, numbers and commas.

Program **newtam** uses calibration constants that are specific for every photometer. These constants are written in a calibration file named as **MT_XXXX.cal**, where **XXXX** stands for the 4-digit serial number of the photometer (some newly produced instruments have 5-digit serial numbers). The structure of the lines within the calibration file is either text (only the title line) or number, i.e. the value of a parameter, comma (delimiter) and text (explanation), surrounded by quotation marks. The values of all calibration constants are available using the MICROTOPS II organizer (see the manuals provided by Solar Light Company, Inc. or NERC FSF). Below is an example calibration file (**MT_7346.cal**).

```
"Current calibration constants for MICROTOPS"
"7346"            , "S/N :"
0.4400            , "WVL1   - wavelength channel #1"
0.6750            , "WVL2   - wavelength channel #2"
0.8700            , "WVL3   - wavelength channel #3"
0.9360            , "WVL4   - wavelength channel #4"
1.0200            , "WVL5   - wavelength channel #5"
6.283E+00         , "LNV01  - solar constant channel #1"
7.221E+00         , "LNV02  - solar constant channel #2"
6.566E+00         , "LNV03  - solar constant channel #3"
7.259E+00         , "LNV04  - solar constant channel #4"
7.024E+00         , "LNV05  - solar constant channel #5"
3.450E-02         , "C1     - irradiation constant channel #1"
1.107E-02         , "C2     - irradiation constant channel #1"
1.332E-02         , "C3     - irradiation constant channel #1"
5.841E-03         , "C4     - irradiation constant channel #1"
```

```
6.404E-03         , "C5     - irradiation constant channel #1"
7.847E-01         , "K      - PWV constant K"
5.945E-01         , "B      - PVW constant B"
0.00              , "C      - PCW constant C"
-1.279E+01        , "POFFS  - pressure offset coefficient"
1.643E+01         , "PSCALE - pressure scale coefficient"
```

The user must put the calibration, background and data files in a selected folder (e.g. D:\J2\NT_007\). Both the executable (newtam.exe) and the calibration (newtam.cfg) files can be in any other folder. After running the code **newtam**, the user will see on the screen printed one after the other the four groups of inputs, printed in the configuration file and described above. Each time he/she must check for errors and agree to continue the code execution by pressing Y or y keys followed by Return. The first of these four screens is shown below in Figure 2. One of the next screens (Figure 3) lists all photometer calibration constants

The handling of the background and data groups of scans starts afterwards. Depending on the value of the key constant **key.print** programs prints or not the calculated estimates for each scan:

- wavelengths (WVL), mean values (SIG) and std (STD_SIG) of signals in the five MACROTOPS II channels
- AOTs (AOT_MTPS) in the five channels evaluated by MACROTOPS II microcomputer;
- AOTs (CALC_AOT) in the five MACROTOPS II channels evaluated using Equations (25-27) ;
- standard deviations (ATD_AOT) in the five MACROTOPS II channels evaluated using Equations (28-33) ;
- airmass factors (AMF), Rayleigh optical thickness $k_R(\lambda_i)\, p/p_0$ (RAY_OT)), estimation of minimum value of $\tau(\lambda_i) = \ln(V(\lambda_i)/[V_o(\lambda_i)/SDCORR])$ (TAU_MIN), the value of $\tau(\lambda_i) = \ln(V(\lambda_i)/[V_o(\lambda_i)/SDCORR])$ (TAU) and the estimation of maximum value of $\tau(\lambda_i) = \ln(V(\lambda_i)/[V_o(\lambda_i)/SDCORR])$ (TAU_MAX) in the five MACROTOPS II channels.

Two examples of these intermediate screens are given below in Figures 4 and 5. In the first case (Figure 4) the signal-to-noise value is high in all channels about 10^{+6} and the calculated AOT standard deviations (STD_AOT) are negligible small and the main impact comes from pressure error. The last three columns (SIGNAL, RAYLEIGH, AIRMASS) present calculated AOT standard deviations if only one contributing factor is accounted for, i.e. there are printed quantities $\sqrt{(\partial AOT(\lambda_i)/\partial V(\lambda_i))^2\, \sigma^2(V(\lambda_i))}$ or $\sqrt{(\partial AOT(\lambda_i)/\partial \mathrm{AMF})^2\, \sigma^2(\mathrm{AMF})}$ or $\sqrt{(\partial AOT(\lambda_i)/\partial \mathrm{p})^2\, \sigma^2(\mathrm{p})}$. Note that their sum is not equal to $\sqrt{\sigma^2(AOT(\lambda_i))}$, printed in columns STD_AOT. In the second case the signal-to-noise value is much lower and as result the impact of signal and solar zenith angle errors prevail. Nevertheless the calculated $\sqrt{\sigma^2(AOT(\lambda_i))}$ are small compared with corresponding AOTs. The coincidence between AOTs calculated by MICROTOPS II micro computer and those by program **newtam** is always satisfactory accounting for the single precision arithmetic in MICROTOPS II and the accuracy of calibration constants. In the second case (Figure 5) the values signal-to-noise ratios are quite low and the main contribution to evaluated AOT's std

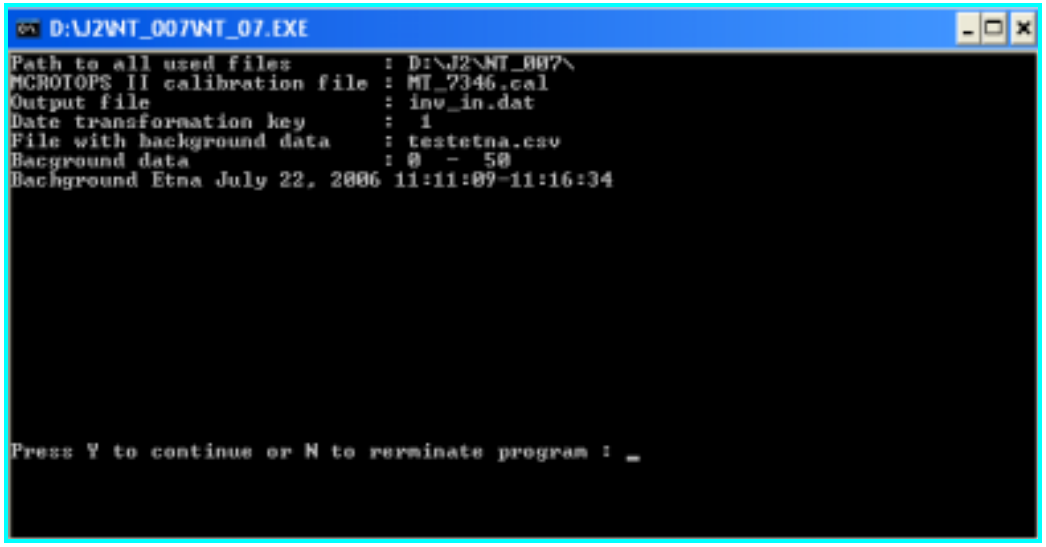

Figure 2.

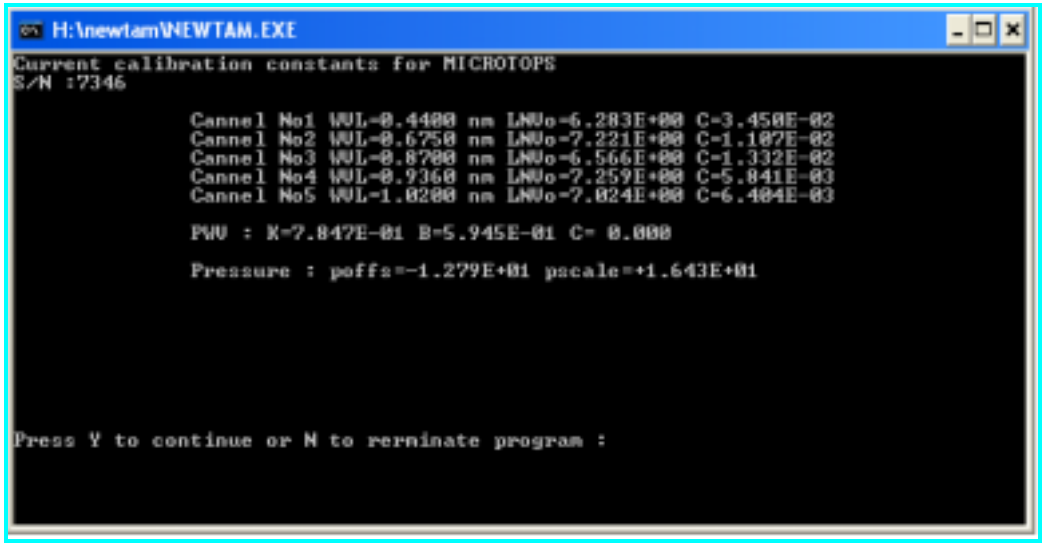

Figure 3.

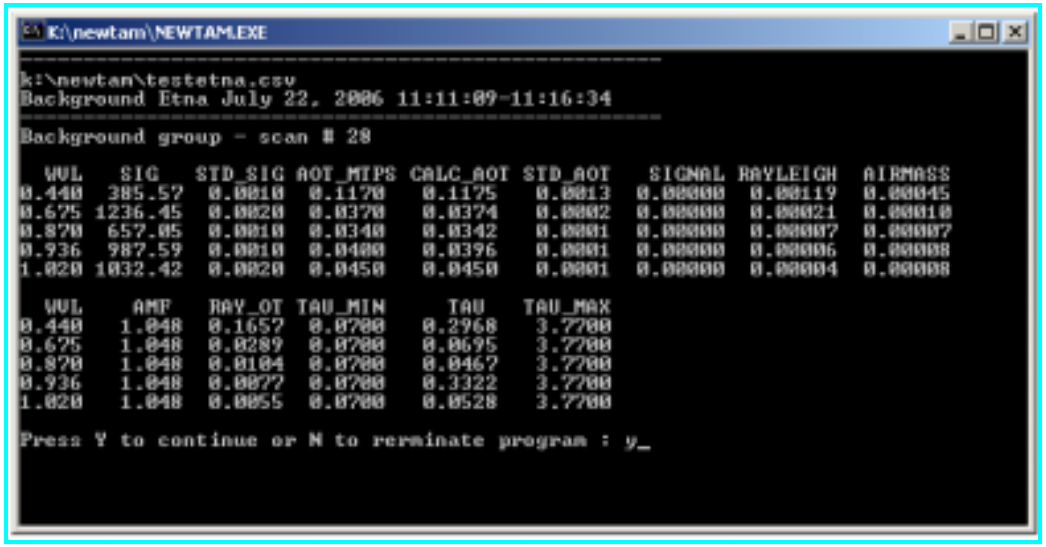

Figure 4.

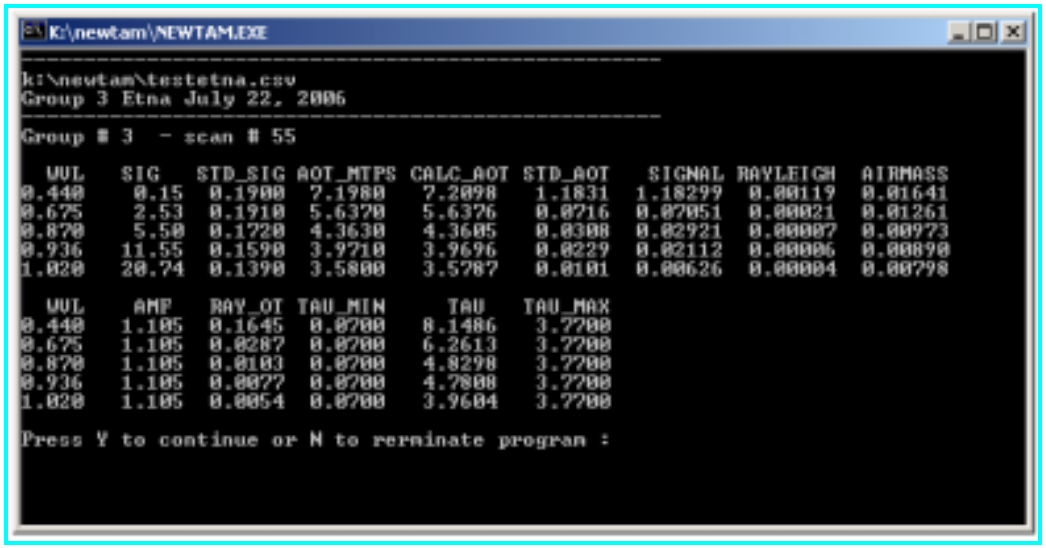

Figure 5.

(STD_AOT) originates from $\sqrt{(\partial AOT(\lambda_i)/\partial V(\lambda_i))^2\ \sigma^2(V(\lambda_i))}$ whilst the contribution of $\sqrt{(\partial AOT(\lambda_i)/\partial \mathrm{p})^2\ \sigma^2(\mathrm{p})}$ is negligible. More over, the signals at 440 nm and 675 nm law compared with corresponding standard deviations and as result the corresponding values of TAU are so high that it is better to exclude scan #55 from group #3.

Afterwards program **newtam** evaluates and remembers the mean values and the standard deviations of AOT for all background scans. These values are used later on to calculate plume mean and standard deviations by subtracting the background quantities from total, i.e. by using Equation (5)

$$\tau_{\mathrm{p}}lume(\lambda) = \tau_{\mathrm{total}}(\lambda) - \tau_{\mathrm{bg}}(\lambda)\ .$$

The program also evaluates the mean universal time, latitude, longitude, altitude, SZA, airmass factor, pressure, temperature and PWV for each group of scans. Having calculated a group's AOT mean values (see Table 2) the program has to evaluate parameters α and β in Ångström's empirical formula $\tau_{\mathrm{M}}(\lambda) = \beta\lambda^{-\alpha}$. The standard approach is to perform a logarithmic transformation $\ln(\lambda_i) = \ln(x_i)$ and $\ln\tau_{\mathrm{M}}(\lambda_i) = \ln(y_i)$ to build the linear regression of $\ln(y_i)$ on $\ln(x_i)$. This task is equivalent to minimization of sum of squares

$$(37) \qquad S_0 = \sum_i \left(\ln(y_i) - b - a\ln(x_i)\right)^2$$

with respect to unknowns a and b with accounting for $\alpha = -a$ and $\beta = \exp(b)$. Unfortunately, values obtained by this method are not precise due to the nonlinear transformation performed. Consider an example with data

i	x_i	$\ln(x_i)$	y_i	$\ln(y_i)$
1	0.4400	-0.8210	0.1150	-2.1628
2	0.6750	-0.3930	0.0650	-2.7334
3	0.8700	-0.1393	0.0500	-2.9957
4	0.9360	-0.0661	0.0426	-3.1559
5	1.0200	0.0198	0.0420	-3.1701

The minimization of sum of squares (13) leads to simultaneous equations

$$(38) \qquad \begin{vmatrix} a\sum_i \left(\ln(x_i)\right)^2 + b\sum_i \ln(x_i) = \sum_i \left(\ln(x_i)\right)\left(\ln(y_i)\right) \\ a\sum_i \left(\ln(x_i)\right) + bq = \sum_i \left(\ln(y_i)\right) \end{vmatrix}$$

with solution $\alpha = 1.2299$ and $\beta = 0.0413$. The regression $\ln(y_i)$ on $\ln(x_i)$ is presented in Figure 6. If we use the program Table Curve, which is able to perform non-linear fits then the estimates of the Ångström's parameters are $\alpha_{TC} = 1.2547$ and $\beta_{TC} = 0.04087$. To overcome this discrepancy and obtain more precise estimates of Ångström's parameters we assume that estimates obtained using regression of $\ln(y_i)$ on $\ln(x_i)$ and the first approximations $\alpha_0 = 1.2299$ and $\beta_0 = 0.0413$ of the true values $\alpha = \alpha_0 + \Delta\alpha$ and $\beta = \beta_0 + \Delta\beta$. The unknown non-linear corrections $\Delta\alpha$ and $\Delta\beta$ can be obtained using a Taylor's expansion of $y = (\beta_0 + \Delta\beta) \times x^{-(\alpha_0 + \Delta\alpha)}$. The required corrections have to minimize the sum of squares

(39) $$S = \sum_i \{ y_i - \varphi(x_i)[1 + \Delta\beta/\beta_0 - \Delta\alpha \ln(x_i)] \}^2 ,$$

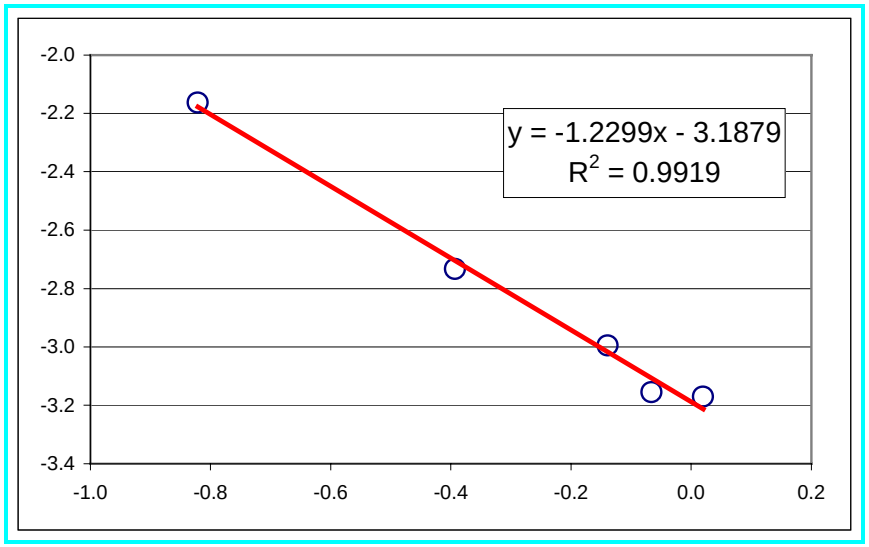

Figure 6

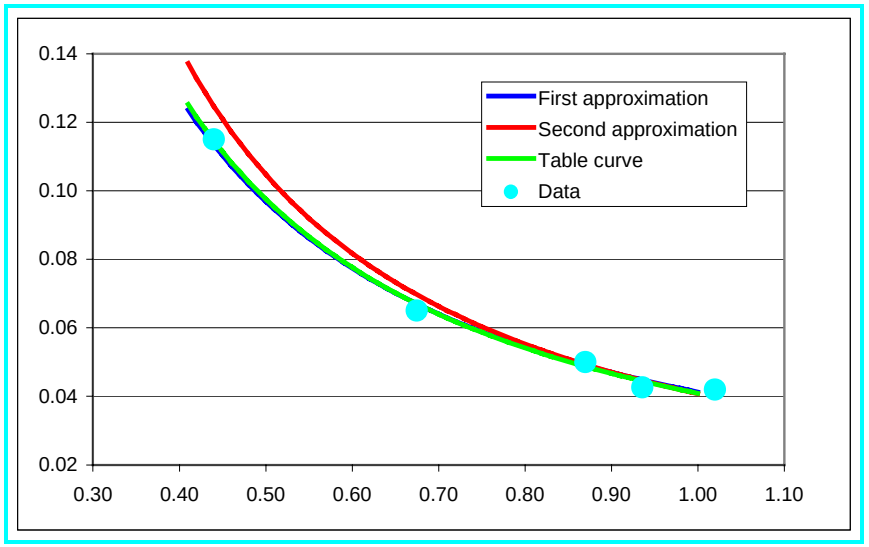

Figure 7

where $\varphi(x) = \beta_0 \times x^{-\alpha_0}$, i.e., they are solutions of simultaneous equations

(40) $$\begin{vmatrix} \Delta\alpha\, a_{11} + \Delta\beta\, a_{12} = c_1 \\ \Delta\alpha\, a_{12} + \Delta\beta\, a_{22} = c_2 \end{vmatrix}$$

with coefficients defined by

$$(41)\quad \begin{cases} a_{11} = \sum_i [\ln(x_i)\varphi(x_i)]^2 & a_{12} = -\frac{1}{\beta_0}\sum_i \ln(x_i)\varphi^2(x_i) \\ a_{21} = \sum_i \ln(x_i)\varphi^2(x_i) & a_{22} = -\frac{1}{\beta_0}\sum_i \varphi^2(x_i) \\ c_1 = -\sum_i \ln(x_i)\varphi(x_i)[y_i - \varphi(x_i)] & c_2 = -\sum_i \varphi(x_i)[y_i - \varphi(x_i)] \end{cases}.$$

Thus obtained results are $\Delta\alpha = 0.129617$ and $\Delta\beta = -0.000438$ or $\alpha = 1.3596$ and $\beta = 0.04082$. A comparison between first and second approximation and the results obtained by Table Curve is presented in Figure 7.

After evaluation of the corrected Ångström coefficients for the current group of scans program **newtam** prints the following screen (Figure 8).

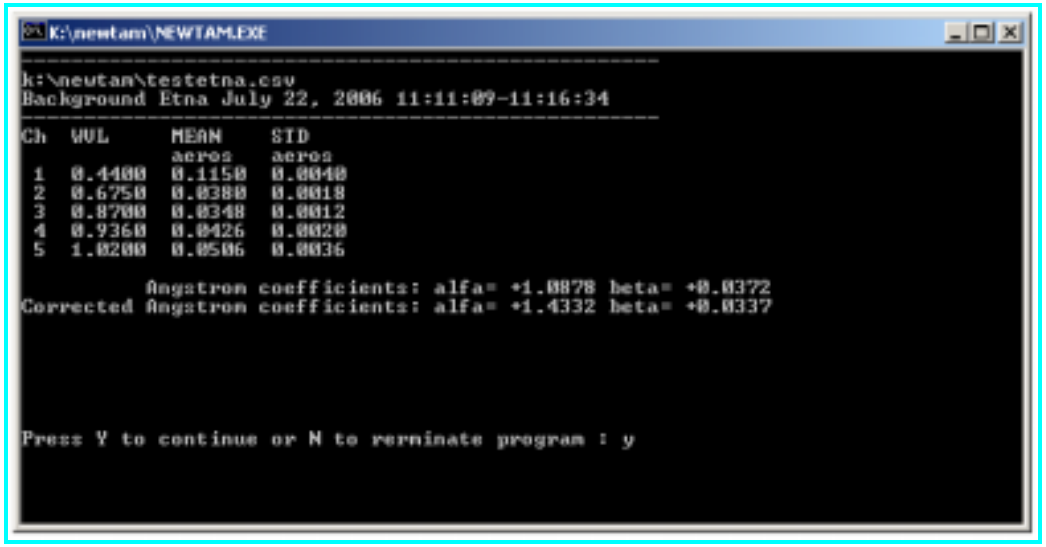

Figure 8

The program outputs two files with permanent names:

♠ **inv_in.dat** - ASCII file that **later must be renamed** and will be used as the input file by programs performing inversion. The following is the content of the **inv_in.dat** for the example **testetna.csv** discussed above. Note that the background and the through-plume data

may be contained in the same file but it must still be listed twice in the configuration file **newtam.cfg**. The content of the file **inv_in.dat** will be explained in sections 6 and 7. Below is listed the content of the output file **inv_in.dat**.

```
5 7 1.45 0.00 0.08 4.00
 0.4400 0.6750 0.8700 0.9360 1.0200

 3.43 0 0
  0, 22/07/2006 11:13:53, 11:13:53, +37.747000,  +15.000000, 3230
Background Etna July 22, 2006 11:11:09-11:16:34
 0.1150 0.0380 0.0348 0.0426 0.0506
 0.0040 0.0018 0.0012 0.0020 0.0036

 4.24 0 0
  1, 22/07/2006 09:11:50, 09:11:50, +37.742000,  +15.000000, 3041
Set 1 Etna July 22, 2006
 0.2818 0.0910 0.0521 0.0576 0.0633
 0.1485 0.0378 0.0212 0.0172 0.0138

 4.04 0 0
  2, 22/07/2006 09:17:08, 09:17:08, +37.693067,  +15.000000, 3041
Set 2 Etna July 22, 2006
 0.2247 0.0746 0.0453 0.0541 0.0626
 0.1336 0.0357 0.0178 0.0152 0.0131

 2.63 0 0
  3, 22/07/2006 12:30:27, 12:30:27, +37.747000,  +14.996000, 3268
Set 3 Etna July 22, 2006
 4.8069 3.8820 3.2017 2.9795 2.7571
 1.8896 1.7497 1.5496 1.4602 1.3731

 3.25 0 0
  4, 22/07/2006 12:59:50, 12:59:50, +37.737000,  +14.902800, 2921
Set 4 Etna July 22, 2006
 1.2099 0.8229 0.5914 0.5426 0.4933
 0.6186 0.4216 0.3082 0.2738 0.2395
```

♠ **temp.dat** - ASCII file that could be ignored if the user does not want to check the averaging and fitting processes and does not want to use mean values of temperature, pressure, SZA, airmass factor and PWV. The following is a print of part of **temp.dat** for the considered example. The full file is named **temp_testetna.dat**.

```
-----------------------------------------------------------------------

D:\J2\NT_007\testetna.csv
Background Etna July 22, 2006 11:11:09-11:16:34
-----------------------------------------------------------------------
 1  0.1200   0.0420   0.0350   0.0410   0.0480
 2  0.1200   0.0370   0.0340   0.0390   0.0450
 3  0.1170   0.0370   0.0340   0.0400   0.0450
 4  0.1180   0.0360   0.0340   0.0400   0.0460
 5  0.1160   0.0380   0.0340   0.0410   0.0470
 6  0.1180   0.0360   0.0330   0.0400   0.0470
 7  0.1180   0.0380   0.0330   0.0400   0.0470
 8  0.1200   0.0360   0.0350   0.0410   0.0480
 9  0.1180   0.0370   0.0340   0.0410   0.0470
10  0.1190   0.0370   0.0340   0.0420   0.0490
11  0.1170   0.0400   0.0350   0.0420   0.0500
12  0.1180   0.0380   0.0350   0.0430   0.0510
13  0.1160   0.0420   0.0360   0.0430   0.0490
14  0.1170   0.0380   0.0350   0.0440   0.0530
15  0.1140   0.0410   0.0370   0.0440   0.0520
16  0.1140   0.0380   0.0370   0.0450   0.0530
17  0.1170   0.0370   0.0370   0.0460   0.0550
18  0.1090   0.0400   0.0330   0.0430   0.0530
19  0.1100   0.0400   0.0350   0.0440   0.0530
20  0.1100   0.0370   0.0340   0.0440   0.0530
21  0.1100   0.0380   0.0350   0.0440   0.0540
22  0.1090   0.0380   0.0360   0.0440   0.0530
23  0.1090   0.0380   0.0340   0.0440   0.0540
24  0.1130   0.0350   0.0340   0.0460   0.0580
25  0.1080   0.0370   0.0360   0.0450   0.0550
-----------------------------------------------------------------------
Ch  WVL     MEAN    STD
            bgnd    bgnd
 1  0.4400  0.1150  0.0040
 2  0.6750  0.0380  0.0018
 3  0.8700  0.0348  0.0012
 4  0.9360  0.0426  0.0020
 5  1.0200  0.0506  0.0036
-----------------------------------------------------------------------
Simultaneous equations :
 +0.8527   -1.3996   +3.6779
```

```
 -1.3996   +5.0000  -14.9300
     DET:         +2.3043E+00     -2.5067E+00     -7.5824E+00
SOLUTION:   alfa= -1.0878   beta= -3.2905
-----------------------------------------------------------------
Simultaneous equations :
+6.1231E-03  +2.2596E-01  +1.3130E-03
-8.4137E-03  -4.3880E-01  -1.3489E-03
     DET:         -7.8560E-04     -2.7134E-04     +2.7878E-06
SOLUTION: d.alfa= +1.0878 d.beta= +0.0372
-----------------------------------------------------------------
          Angstrom coefficients: alfa= +1.0878 beta= +0.0372
Corrected Angstrom coefficients: alfa= +1.4332 beta= +0.0337
-----------------------------------------------------------------
Table curve results:
 Parm      Value        Std Error    t-value      95% Confidence Limits
 beta=a    0.033000901  0.008720316  3.784370071  0.005454837  0.060546965
 alfa=b    1.459903873  0.395906391  3.687497617  0.209299736  2.710508009
-----------------------------------------------------------------
  T mean = 34.720
  P mean =  694.8
SZA mean = 17.574
AMF mean =  1.049
PWV mean =  0.182
-----------------------------------------------------------------
 5 7 1.4500 0.0000 0.0800 4.0000
 0.4400 0.6750 0.8700 0.9360 1.0200

 3.43 0 0
  0, 07/22/2006 11:13:53, 11:13:53, +37.747000,  +15.000000, 3230
Background Etna July 22, 2006 11:11:09-11:16:34
 0.1150 0.0380 0.0348 0.0426 0.0506
 0.0040 0.0018 0.0012 0.0020 0.0036
-----------------------------------------------------------------
```

The next two listings illustrate the action of the program depending on the value of key constant key.background. When key.background=1 the we have to evaluate plume AOT according to Equation (5). Otherwise it is assumed that al background AOTs have zero values.

```
-----------------------------------------------------------------
```

```
h:\J2\newtam\testetna.csv
Set 1 Etna July 22, 2006
-----------------------------------------------------------------------
 1  0.5280   0.1420   0.0970   0.1090   0.1220
 2  0.4350   0.1490   0.1020   0.1140   0.1270
 3  0.2090   0.1120   0.0700   0.0820   0.0950
 4  0.5370   0.1780   0.1150   0.1200   0.1250
 5  0.5160   0.1460   0.0920   0.1100   0.1270
 6  0.4930   0.1450   0.0910   0.1030   0.1150
 7  0.1630   0.0590   0.0490   0.0700   0.0920
 8  0.3630   0.1120   0.0770   0.0930   0.1100
 9  0.2150   0.0800   0.0650   0.0830   0.1000
10  0.5090   0.1670   0.1110   0.1180   0.1260
-----------------------------------------------------------------------
Ch  WVL     MEAN    STD     MEAN    STD     MEAN    STD
            bgnd    bgnd    meas    meas    aeros   aeors
 1  0.4400  0.0000  0.0000  0.3968  0.1485  0.3968  0.1485
 2  0.6750  0.0000  0.0000  0.1290  0.0378  0.1290  0.0378
 3  0.8700  0.0000  0.0000  0.0869  0.0212  0.0869  0.0212
 4  0.9360  0.0000  0.0000  0.1002  0.0172  0.1002  0.0172
 5  1.0200  0.0000  0.0000  0.1139  0.0138  0.1139  0.0138
-----------------------------------------------------------------------
```

```
-----------------------------------------------------------------------
h:\J2\newtam\testetna.csv
Set 1 Etna July 22, 2006
-----------------------------------------------------------------------
 1  0.5280   0.1420   0.0970   0.1090   0.1220
 2  0.4350   0.1490   0.1020   0.1140   0.1270
 3  0.2090   0.1120   0.0700   0.0820   0.0950
 4  0.5370   0.1780   0.1150   0.1200   0.1250
 5  0.5160   0.1460   0.0920   0.1100   0.1270
 6  0.4930   0.1450   0.0910   0.1030   0.1150
 7  0.1630   0.0590   0.0490   0.0700   0.0920
 8  0.3630   0.1120   0.0770   0.0930   0.1100
 9  0.2150   0.0800   0.0650   0.0830   0.1000
10  0.5090   0.1670   0.1110   0.1180   0.1260
-----------------------------------------------------------------------
Ch  WVL     MEAN    STD     MEAN    STD     MEAN    STD
            bgnd    bgnd    meas    meas    aeros   aeors
```

```
 1  0.4400  0.1150  0.0040  0.3968  0.1485  0.2818  0.1526
 2  0.6750  0.0380  0.0018  0.1290  0.0378  0.0910  0.0396
 3  0.8700  0.0348  0.0012  0.0869  0.0212  0.0521  0.0224
 4  0.9360  0.0426  0.0020  0.1002  0.0172  0.0576  0.0193
 5  1.0200  0.0506  0.0036  0.1139  0.0138  0.0633  0.0174
-----------------------------------------------------------------------
```

5. The importance of using the correct index of refraction and calculation of extinction efficiency factors – program eff_factors

The calculation of efficiency factors $Q_{\text{ext}}(2\pi r/\lambda,\tilde{m})$, $Q_{\text{sc}}(2\pi r/\lambda,\tilde{m})$ and $Q_{\text{abst}}(2\pi r/\lambda,\tilde{m})$ for extinction, scattering and absorption is discussed in many text books (e.g., Bohren and Huffman, 1983).

The program **eff_factors.exe** computes the extinction, scattering, and absorption for 2400 values of the size parameter $\rho = (2\pi r)/\lambda$ $(0 \le \rho \le 150)$ and saves the extinction efficiency factors for inversion purposes. Program **eff_factors.exe** needs to be rerun whenever a new refractive index is needed for inversion purposes.

The program **eff_factors** requires two input files:

♠ **eff_factors.cfg** - this a configuration file; it has to be placed together with the executable **eff_factors.exe** in any folder selected by the user. This file contains just one line – text (typed starting at the leftmost position of the line, **not longer then 80 characters** and not enclosed in quotation marks) indicating the folder where has to be placed second input file (**eff_factors_in.dat**) and where code will write both output files (**eff_factors_out.dat** and **inversion_in_ext.dat**). If the content of **eff_factors.cfg** is
c:\temp\
this means that **c:\temp** is the working folder for the code **eff_factors**, i.e. the code will look there for the input file **eff_factors_in.dat** and during the execution will write both output files **eff_factors_out.dat** and **inversion_in_ext.dat** there.

♠ **eff_factors_in.dat** - this the input file. The following is example of its content:

```
_ _ _2
_ _1.250000_ _0.000000
_ _1.450000_ _0.000000
_ _1.550000_ _0.001000
```

Line number	Format (Expalnation)	Variable	Content
1	I4 (####)	NSETS	number of complex indices of refraction in the input file
2	2F10.6 (###.&&&&&&###.&&&&&&)	RFR, RFI	real and imaginary parts array of the of complex indices of refraction
3	2F10.6 (###.&&&&&&###.&&&&&&)	RFR, RFI	real and imaginary parts array of the of complex indices of refraction
…	…	…	…
NSET+1	2F10.6 (###.&&&&&&###.&&&&&&)	RFR, RFI	real and imaginary parts array of the of complex indices of refraction

The program **eff_factors** prints into two output files:

♠ **eff_factors_out.dat** - this a ASCII text file with space delimited numbers. It contains NSET tables $\{\rho_i, Q_{\text{ext}}(\rho_i), Q_{\text{scat}}(\rho_i), Q_{\text{bs}}(\rho_i), i = 1, 2, \ldots, 2400\}$; besides each table starts with two header lines; first is the consecutive number of the table and the second depicts index of refraction.

♠ **inversion_in_ext.dat** - this files is used by inversion codes.

For example, the execution of **eff_factors** above mentioned file **eff_factors_in.dat** will produce two tables. The first one corresponds to $\tilde{m} = 1.25 - i \times 0.00$

```
   1
VALUES OF MIE FACTORS FOR INDEX OF REFRACTION = 1.25 -0.0000I
   0.010          0.00000          0.00000          0.00000
   0.020          0.00000          0.00000          0.00000

.........................................................................
 149.800          2.12469          2.12469          0.00000
 149.900          2.11277          2.11277          0.00000
 150.000          2.09641          2.09641          0.00000
```

whilst the second one corresponds to $\tilde{m} = 1.55 - i \times 0.00$

```
   2
VALUES OF MIE FACTORS FOR INDEX OF REFRACTION = 1.45 -0.0000I
   0.010          0.00000          0.00000          0.00000
   0.020          0.00000          0.00000          0.00000

.........................................................................
 149.900          2.06908          1.13403          0.93505
 150.000          2.06905          1.13402          0.93503
```

The Figure 9 is made using a file **eff_factors_out.dat** and plots of $Q_{\text{ext}}(\rho, \tilde{m})$ for $0 \le \rho \le 12$ and four real values of $\tilde{m}$ (1.26, 1.45, 1.55 and 2.00). These values of $\tilde{m}$ correspond to quite a wide variation the chemical composition of the aerosol particles: 1.33 (pure water), 1.43 (silicate particles), 1.40-1.45 (H_2SO_4+H_20). The common tendency that the increase of real part of $\tilde{m}$ causes smoothening of the oscillating curve obtained at $\tilde{m} = 1.23 - i\,0.00$. All curves have strongest first maximum and decaying consecutive maxima with asymptote equal to 2. The

knowledge of $Q_{\text{ext}}(\rho,\tilde{m})$ is necessary if we want to estimate the radii interval contributing to measured AOT at some discrete set of wavelengths. This could be easily done by simple numerical calculations using the following Equations

(42) $$\tau_{\text{A}}(\lambda) = \int_0^{\infty} \pi r^2 \, Q_{\text{ext}}(2\pi r/\lambda,\tilde{m}) n_{\text{c}}(r)\, dr = \int_0^{\infty} \Gamma(\lambda, r)\, dr \ ,$$

(43) $$\Gamma(\lambda, r) = \pi r^2 \, Q_{\text{ext}}(2\pi r/\lambda,\tilde{m}) n_{\text{c}}(r) \ ,$$

(44) $$n_{\text{c}}(r) = \sum_{k=1}^{3} \frac{N_k}{\sqrt{2\pi}\, r \ln\sigma_k} \exp\left(-\frac{(\ln(r) - \ln(r_m))^2}{2\ln^2(\sigma_k)} \right) .$$

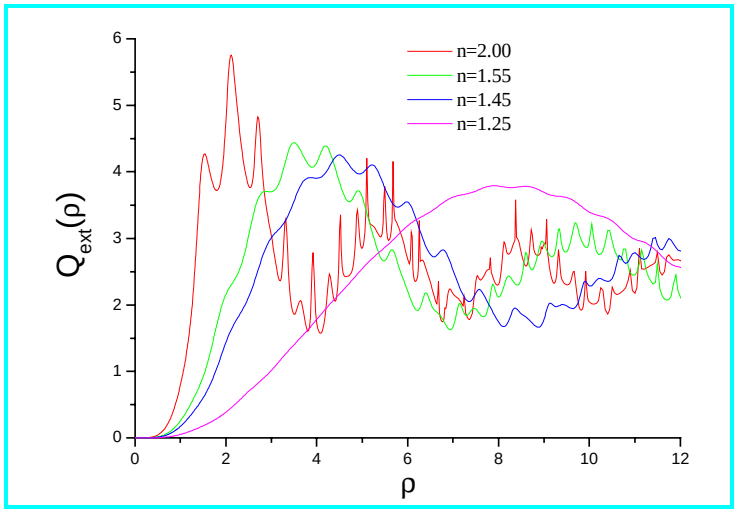

Figure 9.

According King (1982) and Jorge *et al.* (1996) the increase of the real part of $\tilde{m}$ (if $\text{Im}(\tilde{m}) \le (0.01 - 0.05)$) is tending to save the shape of retrieved columnar aerosol size distribution $n_{\text{c}}(r)$ but shifts the overall curve towards smaller radii.

Completely different is the situation if we consider aerosols containing black carbon (BC), which is an extremely strong absorber. Carbonaceous smokes are produced by a variety of combustion sources such as chimney stack furnaces, industrial flames, biomass burning, aircrafts and rockets, all motor vehicles and especially diesel engines (e.g. Seinfeld and Pandis, 1998). In general, near the source soot consists of

small spherical primary particles (also known as spherules or monomers) combined into branched aggregates. The monomers are composed of amorphous BC mixed with some amount of organic carbon (OC) and other elements. The number concentration varies widely from a few particles/cm^3 to perhaps 10^6 particles/cm^3 whilst sizes are predominantly within the nucleation mode (0.005-0.15 μm). Soot particles can have markedly different shapes depending on the way and the source of their formation. Pure BC particles are mainly hydrophobic but coating and mixing of BC particles with some organic compounds may change their hygroscopic properties. The aging processes lead to forming of external and different types of internal (homogeneous or layered particles) mixtures, growth in size and transformation to accumulation mode. Another type of internal mixture commonly found in smoke aerosols are long chain aggregates of BC particles formed at high temperatures close to fire. These chain aggregates can also be coated with nonabsorbing materials to form an internally mixed heterogeneous structure. After aging and interaction with water vapour and clouds, these opened clusters usually collapse to form closely packed spherical-like structures in accumulation mode.

The explosions and subsequent fire at the Buncefield oil depot in December 2005 afforded a rare opportunity to study the atmospheric consequences of a major oil fire at close range, using MICROTOPS II Sun-photometer (Mather *et al.*, 2007). Near-source measurements suggest that plume particles were ~50% black carbon (BC) with refractive index 1.73- i 0.42, effective radius (R_{eff}) 0.45-0.85μm and mass loading ~2000μg×.m^{-3}. About 50km downwind, particles were ~60-75% BC with refractive index between 1.80-0.52i and 1.89-0.69i, R_{eff} ~1.0μm and mass loadings 320-780μg×m^{-3}. Number distributions were almost all monomodal with peak at r<0.1μm.

At the time of the explosion, local temperatures were around freezing, wind-speeds were low and anti-cyclonic conditions prevailed. There was a strong inversion layer in the atmosphere, which trapped the lofted plume and its products at a moderate altitude. The first day following the explosion was marked by the formation of a dark plume which spread slowly across a wide area of southern England, as viewed from the ground and weather satellites. It appears that the high plume buoyancy and favorable meteorological conditions meant that the plume was trapped aloft, with minimal mixing to the ground.

In the absence of data concerning the exact composition of the plume's particles some testing had to be undertaken in order to establish the appropriate index of refraction to use. As it resulted from fuel combustion, the major constituents of the Buncefield plume were assumed to be black carbon and water. Thus we assumed that Buncefield smoke could be modelled as water droplets with embedded within black carbon particles (inclusions). The Maxwell-Garnett effective dielectric constant ε_m (complex number) is given by (Bohren and Huffman, 1983)

$$\varepsilon_m = \varepsilon_W \frac{\varepsilon_{BC} + 2\,\varepsilon_W + 2\,f_{BC}\,(\varepsilon_{BC} - \varepsilon_W)}{\varepsilon_{BC} + 2\,\varepsilon_W - f_{BC}\,(\varepsilon_{BC} - \varepsilon_M)}\,, \tag{45}$$

where $\varepsilon_W = 1.33 - i\,0.0$ and $\varepsilon_{BC} = 2.0 - i\,1.0$ are complex dielectric constants of water and BC inclusions and f_{BC} volume fraction of inclusions. Note that the refractive index is the square root of the dielectric constant. Thus the values of $Q(r, \lambda, m)$ can be evaluated using King's software. The radii retrieval interval of maximum sensitivity spanned from 0.08 to 4.0 μm. In any case we obtained coincidence (within range of retrieval error) of aerosol columnar distributions $n_c(r)$ evaluated with three values (ν^*-0.5, ν^* and ν^*+0.5) of initial Junge distributions. Thus the

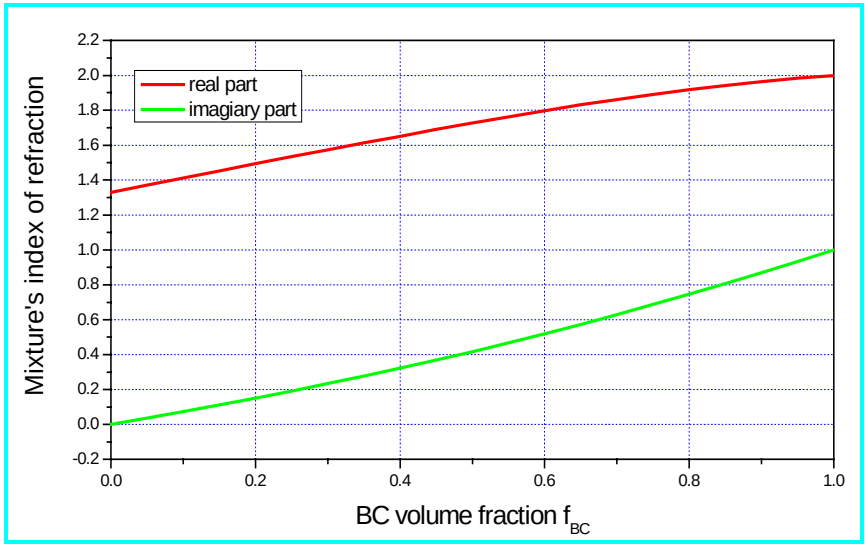

Figure 10.

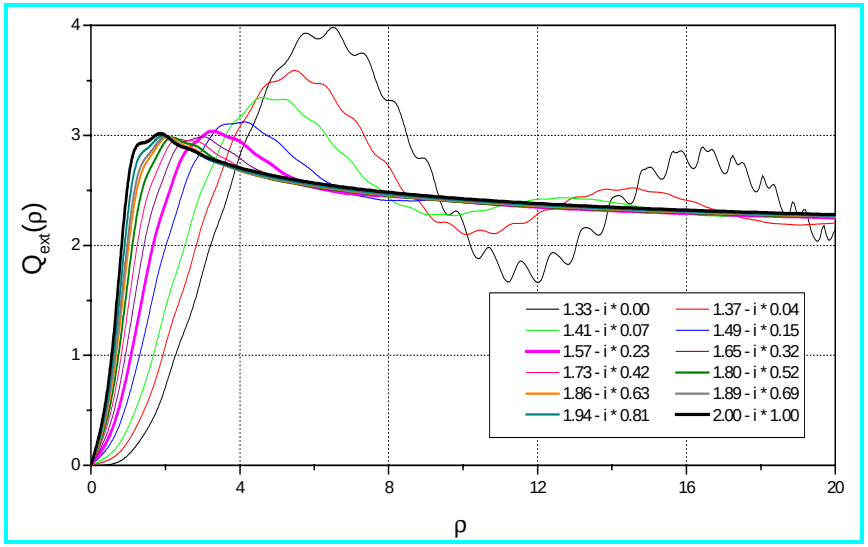

Figure 11.

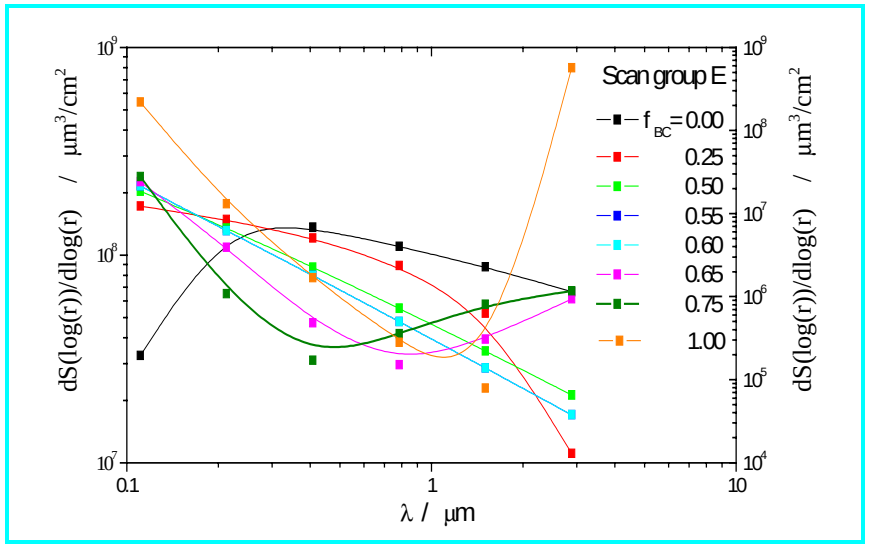

Figure 12.

only one significant uncertainty remained the proper selection of f_{BC}-value. We accepted that it has to minimise simultaneously (or provide reasonable values) for all three merit parameters already discussed in section 2. The inspection (see below) of $\mathrm{E}_{rel}(f_{BC})$, $\varepsilon^2(f_{BC})$ and $M_c(f_{BC})$ revealed that both of them have minima at $f_{BC} \approx 0.5$ whilst $M_c(0.5) \approx 4$. Lesins *et al.* (2002) proved by numerical calculations that for $f_{BC} \geq 0.5$ all internal mixture models are equivalent, i.e. this is an indirect proof that accepted by us model of Buncefield smoke is reasonable.

The next three figures in this paragraph manifest the significance of the proper choice of complex index of refraction. Figure 10shows how mixture index of refraction changes with increase of BC volume fraction: at $f_{BC} = 0.0$ we have pure water and $\tilde{m} = 1.33 - i\,0.00$; at $f_{BC} = 0.0$ - $\tilde{m} = 1.53 - i\,0.19$; at $f_{BC} = 0.5$ - $\tilde{m} = 1.73 - i\,0.42$; at $f_{BC} = 0.75$ - $\tilde{m} = 1.80 - i\,0.69$; at $f_{BC} = 1.0$ (pure BC) - $\tilde{m} = 2.00 - i1.00$. Figure 11 presents the variation of Mie extinction efficiency factors $Q_{\mathrm{ext}}(\rho, \tilde{m})$ depending on the complex index of refraction. Calculations have been performed by code **eff_factors.exe**. The efficiency curves for $f_{BC} = 0.5$ and $f_{BC} = 1.0$ are plotted with bold lines. They significantly differ form $Q_{\mathrm{ext}}(\rho, \tilde{m})$ oscillating curve for $\tilde{m} = 1.33 - i\,0.00$. The bold curves are extremely smooth, they have the shape similar to that of the step function and tend to shift to smaller ρ values with f_{BC} increase. This corresponds to the fact that BC is a gray (uniform) absorber at al visible wavelengths. This means that with increase of f_{BC} particles with different radii will be responsible for the absorption at peak wavelengths of MOCROTOPS II channels. This is manifested in Figure 12 where is outlined the modification of the columnar surface distribution when increasing f_{BC} value for one of scan groups registered beneath the plume. The right-hand ordinate corresponds to distribution evaluated for f_{BC} =1

6. Selection of retrieval radius interval – program preliminary

Optical thickness measurements are performed by MICROTOPS II at five different wavelengths ranging between 0.38 and 1.02 μm and we want to use them for the size distribution determinations using algorithm already described in sections 1 and 3. This spectral region of measurements limits the range of maximum sensitivity to the accumulation and coarse aerosol modes only ($0.1 \le r \le 4.0\,\mu m$) due to both the extinction cross sections (which increase significantly with radius) and the number densities of natural aerosol particles (which normally decrease with radius). Unfortunately it is not possible to propose an absolute rule, which determines the radii limits having the most significant contribution to the MICROTOPS II measurements. This radius range is dependent on both the form of the size distribution function and the values of the Mie extinction cross section over the radius range. Since the size distribution function is not known in advance, it is apparent that occasional trial and error is required in order to determine the radius range over which the inversion can be performed.

The program **preliminary** performs a constrained linear inversion with second derivative smoothing constraint on spectral optical depth data in order to obtain a columnar aerosol size distribution for a range of upper and lower particle radii limits and for three values of Junge parameter (ν^*-0.5, ν^* and ν^*+0.5). This is achieved by inverting one or more sets of observations for which the spectral variation of Mie (aerosol) optical depth is a smooth, nearly linear, function with modest amount of curvature when plotted on a $\log(\tau_M(\lambda))$ vs. $\log(\lambda)$ scale. The main purpose of the program is to estimate the most informative or optimum radii interval for the inversion.

One should run program **inversion** following program **preliminary** for a detailed listing of the inversion results for the desired data set and the selected radius range.

The executable code **preliminary.exe** and the configuration file **inversion.cfg** could be placed together in any folder. The file **inversion.cfg** contains only one line – text (typed starting at the leftmost position of the line, not longer then 80 characters and not enclosed in quotation marks) indicating the folder where both input files are placed (**inversion_in.dat** and **inversion_in_ext.dat**) and where the code will write both output files (**preliminary_out.dat** and **preliminary_short.dat**). If the content of **inversion.cfg** is

c:\temp\

this means that **c:\temp** is the working folder for the code **preliminary**, i.e., the code will look there for the input files **inversion_in.dat** and **inversion_in_ext.dat**. Remember that the file **inversion_in.dat** is the renamed file **inv_in.dat**, that was created by code **newtam**. Both output files **preliminary_out.dat** and **preliminary_short.dat** will be written in the same folder.

Here is an example of the input data stream written in **inversion_in.dat** :

```
_5_7_1.45_0.00_0.08_4.00
_0.4400_0.6750_0.8700_0.9360_1.0200
_
3.43_0_0
```

```
_ _0,_22/07/2006_11:13:53,_11:13:53,_+37.747000,_ _+15.000000,_3230
Background_Etna_July 22,_2006_11:11:09-11:16:34
_0.1150_0.0380_0.0348_0.0426_0.0506
_0.0040_0.0018_0.0012_0.0020_0.0036
_
_4.24_0_0
_ 1,_22/07/2006_09:11:50,_09:11:50,_+37.742000,_ _+15.000000,_3041
Set_1_Etna_July_22,_2006
_0.2818_0.0910_0.0521_0.0576_0.0633
_0.1485_0.0378_0.0212_0.0172_0.0138
_
_ _4.04_0_0
_ _2,_22/07/2006_09:17:08,_09:17:08,_+37.693067,_ _+15.000000,_3041
Set_2_Etna_July_22,_2006
_0.2247_0.0746_0.0453_0.0541_0.0626
_0.1336_0.0357_0.0178_0.0152_0.0131
```

The first two lines constitute the retrieval task defining the number of wavelengths p, the number of q radii where inversion will be performed, real and imaginary parts of the complex index of refraction, smallest r_a and largest r_b radii used for inversion and the list $\lambda_1, \lambda_2, \ldots, \lambda_p$ of wavelengths where AOTs have been measured. These coincide with peak wavelengths of MICROTOPS II channels. Dimension statements in the code are valid for NWVL and NRAD up to 16 but corresponding values are subject to input.

The next group of six lines (note, first one is obligatory empty) contains the input data obtained by the handling of one group of scans by code **newtam**. This group of lines could be repeated as often as desired in order to inverse a few scan groups.

The format of the first line is **2I2,4F5.2** which means two integer numbers ## written with maximum two digits and five floating point numbers **##.&&** with no more than two digits before decimal points and two digits in the fraction part. Remember that one position of digits ## could be used to type negative sign if necessary.

The content of the first line is as follows: NWVL (p), NRAD (q), RFR (n), RFI (k), RI (r_a), RF (r_b) and in the example listed above we have as follows: $p = 5$, $q = 7$, $\tilde{m} = 1.45 - i \times 0.00$, $r_a = 0.08\,\mu m$ and $r_b = 4.0\,\mu m$.

The format of the second line is **8F7.4** which means eight floating point numbers **##.&&&&** with no more two digits before decimal points and four digits in the fraction part.

The content of the first line is $\lambda_1, \lambda_2, \ldots, \lambda_p$. In the cases $p \geq 9$, wavelengths will be typed in two lines, i.e. we may have three or more lines before the empty line. In the example listed above we have as follows: $\lambda_1 = 0.4400\ \mu m$, $\lambda_2 = 0.6750\ \mu m$, $\lambda_3 = 0.8700\ \mu m$,

$\lambda_4 = 0.9360\ \mu m$ and $\lambda_5 = 1.0200\ \mu m$.

The first group of six lines is:

```
_ _ _
_ 3.43_0_0
_ _0,_22/07/2006_11:13:53,_11:13:53,_+37.747000,_ _+15.000000,_3230
Background_Etna_July 22,_2006_11:11:09-11:16:34
_0.1150_0.0380_0.0348_0.0426_0.0506
_0.0040_0.0018_0.0012_0.0020_0.0036
```

The first line in the group is always empty.

The second line in the group contains the estimate of Junge parameter ν^*, and two keys (KEYWNU and KEYIT), determining the outputs of the program **inversion**. The action of these two parameters is explained in the next section and in Appendix 1. The format of the second line is **F5.2,2I2** (**##.&& ## ##**).

The third line is a text, characterizing background or data group of scans handled by code **newtam**. Its structure is :

NNN,_DD/MM/YYYY_hh:mm:ss,_hh:mm:ss,_±XX.XXXXXX,_±YYY.YYYYYY,AAAA

where

NNN	number in the set; zero stands for background ;	
DD/MM/YYYY_hh:mm:ss	date and time for EXCEL p[lots ;	
hh:mm:ss	time ;	
±XX.XXXXXX	latitude in degree	positive when North ;
±YYY.YYYYYY	longitude in degree	positive when East ;
AAAA	altitude in meters .	

The content of this line is used later by programs **sort_tau** and **sort_distr** to sort the outputs of program **inversion** .

The fourth line contains some text information (up to 80 characters) describing the group of scans. This text is taken from file **newtam.cfg** and is not handled in anyway.

The fifth line contains the measured AOTs by wavelength, i.e. $\tau_M(\lambda_1), \tau_M(\lambda_2), \ldots, \tau_M(\lambda_p)$, typed in format **10F7.4**.

The sixth line contains standard deviations of measured AOTs by wavelength, i.e. $\delta\tau_M(\lambda_1), \delta\tau_M(\lambda_2), \ldots, \delta\tau_M(\lambda_p)$, typed in format **10F7.4**.

Note that if $p \geq 9$ both input sequences (arrays) will be typed in two lines.

The values of parameters r_a, r_b and ν^* and keys KEYWNU and KEYIT are default. They have been entered into **inversion_in.dat** by program **newtam** and do not influence the performance of the program **preliminary** but they are of crucial importance for the performance of the main program **inversion** in the package.

The program **preliminary** outputs two files – **preliminary_out.dat** and **preliminary_short.dat**. An extensive explanation of notations in the output files from **preliminary** and **inversion** is given in Appendix 1

Below are given two examples of file **preliminary_out.dat**, obtained using two different input data sets. These files first repeat the inversion task, i.e. the content of **inversion_in.dat**, and then they summarize in a table the retrievals for 28 radius ranges $[\, r_a \,, r_b \,]$ (four values for $\mathrm{R}_{\min} = r_a$ and seven values of $\mathrm{R}_{\max} = r_b$) and three values of the Junge parameter (ν^*-0.5, ν^*, ν^*+0.5).

$r_{\min}$ µm		$r_{\max}$ µm						
		1.0	1.5	2.0	2.5	3.0	3.5	4.0
0.08	ν^*-0.5							
	ν^*							
	ν^*+0.5							
0.10	ν^*-0.5							
	ν^*							
	ν^*+0.5							
0.15	ν^*-0.5							
	ν^*							
	ν^*+0.5							
0.20	ν^*-0.5							
	ν^*							
	ν^*+0.5							

In each cell of the table are presented three quantities (numbers). The first number is the value in the table is the first term in performance function , i.e. Q_1.

The smaller this value, the better the retrieval. Note that for a successful retrieval it is required that $Q_1 \leq E\{Q\} = p$. The second and third quantities are the maximum iteration number achieved by the inversion (maximum of 8, note – this number is given in parentheses) and the number of observations fit by the inversion within the error bars. Note that the last number is related with the last successful iteration and it is possible that some of the previous iterations are "better".

This table can be used to eliminate radii ranges that produce completely unacceptable inversions (e.g., few or no iterations for some ν^*'s, no or poor fit to the data, or Q_1's that are much larger than the number of wavelengths). For example, the in the first of the two examples given we have successful retrieval only for a few combinations $[\, r_a \,,\, r_b \,]$ and we may conclude bat the best retrieval is for $[\, 0.2\,\mu\text{m} \,,\, 2.5\,\mu\text{m} \,]$ and $\nu^* = 2.23$ when $Q_1 = 0.122.$ and $\text{IN} = 7$, whilst in the second example the best interval is also $[\, 0.2\,\mu\text{m} \,,\, 2.5\,\mu\text{m} \,]$ but the Q_1-value is relatively independent on Junge parameter ν^*.

```
------------------------------------------------------------------------
 DATE:  DEC  8, 1775
 111, 08/12/1775 09:10:22, 09:10:22, +22.113344, -123.001122, 1234
 TEST 1 === ARIZONA DREAMS === Fig.3 Paper #1
------------------------------------------------------------------------

 WAVELENGTH(MICRONS)          TAU(AEROSOLS)
 -------------------        ------------------
       0.4400               0.0180  +/- 0.0009
       0.5217               0.0208  +/- 0.0013
       0.6120               0.0211  +/- 0.0027
       0.6893               0.0216  +/- 0.0019
       0.7120               0.0218  +/- 0.0015
       0.7797               0.0213  +/- 0.0010
       0.8717               0.0218  +/- 0.0010

                 Sum of squared measurement errors  = 1.765E-05
In the inversion which follows: Index of refraction =  1.45 - 0.00 I
                                          NWVL  =  7
                                          NRAD  =  7
                                          SNU   = 1.73
```

Rmin	NU star	Rmax 1.0	1.5	2.0	2.5	3.0	3.5	4.0
	1.23	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)
0.08	1.73	0.000E+00 (0)	5.569E-01 (5)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)
	2.23	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)
	1.23	0.000E+00 (0)	6.306E-01 (8)	0.000E+00 (0)	1.006E+00 (5)	1.557E+00 (1)	0.000E+00 (0)	0.000E+00 (0)
0.10	1.73	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)
	2.23	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)
	1.23	0.000E+00 (0)	3.648E-01 (3)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)
0.15	1.73	1.195E+00 (1)	5.457E-01 (8) 7	1.483E-01 (8)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)
	2.23	1.374E+00 (1)	5.106E-01 (7)	0.000E+00 (0)	7.959E-01 (8)	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)
	1.23	1.905E-01 (8) 7	1.338E-01 (8) 7	5.076E-01 (7)	1.504E-01 (8) 7	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)
0.20	1.73	8.273E-01 (3)	4.025E-01 (8) 7	4.866E-01 (8) 7	1.544E-01 (8) 7	5.710E-01 (8)	0.000E+00 (0)	0.000E+00 (0)
	2.23	1.918E-01 (8) 7	1.919E-01 (8) 7	4.136E-01 (8) 7	1.224E-01 (8) 7	0.000E+00 (0)	0.000E+00 (0)	0.000E+00 (0)

```
------------------------------------------------------------------------------
 DATE:  DEC  8, 1775
 222, 08/12/1775 09:10:22, 09:10:22, +22.113344, -123.001122, 1234
 TEST 2 ===  ARIZONA DREAMS === Fig.X Paper #1
-----------------------------------------------------------------------

 WAVELENGTH(MICRONS)          TAU(AEROSOLS)
 -------------------       ------------------
       0.4400              0.0453  +/- 0.0010
       0.5217              0.0388  +/- 0.0012
       0.6120              0.0382  +/- 0.0020
       0.6893              0.0371  +/- 0.0013
       0.7120              0.0372  +/- 0.0013
       0.7797              0.0382  +/- 0.0020
       0.8717              0.0396  +/- 0.0010
       1.0303              0.0428  +/- 0.0011

                 Sum of squared measurement errors  = 1.603E-05
In the inversion which follows: Index of refraction =  1.45 - 0.00 I
                                              NWVL  =  8
                                              NRAD  =  7
                                              SNU   = 2.07

                                                                      Rmax
        NU
 Rmin   star       1.0               1.5               2.0               2.5               3.0               3.5               4.0
 ----   ----  ---------------  ---------------  ---------------  ---------------  ---------------  ---------------  ---------------
        1.57  3.948E+00 (8) 7  5.263E-01 (8) 8  1.371E+00 (8) 8  2.419E+01 (8)    3.257E+01 (8)    5.564E+00 (8) 6  4.569E+00 (8) 7
 0.08   2.07  4.084E+00 (8) 7  5.748E-01 (8) 8  6.087E-01 (8) 8  8.939E-01 (8)    2.053E+00 (8)    3.480E+00 (8) 7  2.877E+00 (8) 7
        2.57  4.081E+00 (8) 6  5.760E-01 (8) 8  8.545E-01 (8) 8  6.683E-01 (8)    1.832E+00 (8)    1.992E+00 (8) 7  2.777E+00 (8) 7
        1.57  4.026E+00 (8) 6  5.841E-01 (8) 8  8.493E-01 (8) 8  1.146E+00 (8) 8  1.573E+00 (8) 7  3.794E+00 (8) 7  5.683E+00 (8) 6
 0.10   2.07  4.037E+00 (8) 6  5.974E-01 (8) 8  7.548E-01 (8) 8  9.363E-01 (8) 8  1.411E+00 (8) 7  3.195E+00 (8) 7  5.458E+00 (8) 6
        2.57  4.051E+00 (8) 6  6.474E-01 (8) 8  8.383E-01 (8) 8  8.342E-01 (8) 8  1.495E+00 (8) 7  2.689E+00 (8) 7  4.742E+00 (8) 6
        1.57  3.938E+00 (8) 6  8.353E-01 (8) 8  2.485E+00 (8) 7  2.179E+00 (8) 7  4.446E+00 (8) 6  1.868E+00 (8) 8  4.439E+00 (8) 6
 0.15   2.07  3.964E+00 (8) 6  9.217E-01 (8) 8  2.206E+00 (8) 7  2.254E+00 (8) 7  3.773E+00 (8) 6  3.049E+00 (8) 7  2.627E+00 (8) 7
        2.57  3.973E+00 (8) 6  9.583E-01 (8) 8  2.267E+00 (8) 7  2.484E+00 (8) 7  2.957E+00 (8) 7  2.184E+00 (8) 8  2.089E+00 (8) 7
        1.57  3.760E+00 (8) 6  6.242E-01 (8) 8  2.774E+00 (8) 7  1.537E+00 (8) 8  5.436E+00 (8) 6  5.417E+00 (8) 7  4.615E+00 (8) 6
 0.20   2.07  3.690E+00 (8) 6  6.399E-01 (8) 8  2.559E+00 (8) 7  1.546E+00 (8) 8  4.283E+00 (8) 6  5.098E+00 (8) 6  5.074E+00 (8) 6
        2.57  3.634E+00 (8) 6  7.177E-01 (8) 8  2.509E+00 (8) 7  1.649E+00 (8) 8  2.824E+00 (8) 7  4.530E+00 (8) 6  3.764E+00 (8) 7

 ----   ----  ---------------  ---------------  ---------------  ---------------  ---------------  ---------------  ---------------
```

There is an important question that cannot be answered using the tables described above – is there a chance that better solutions can be obtained by performing less iterations than eight? The output file **preliminary_short.dat** was specifically designed to answer this. The file consists of a set of tables – one table correspond to each cell in the table printed to the output file **preliminary_out.dat**. The lines now describe the results of iterations. Below, an example is given – only two of the tables from file **preliminary_short.dat** obtained when inverting experimental data from TEST 2 (for more details see Appendix 2). Given fixed/selected values of ν^* and $[\, r_a , r_b \,]$ the best solutions are

characterized by the smallest Q_1-value, smallest γ_{rel}-values (the smaller γ_{rel} the smoother the solution, because weaker smoothing has been applied) and the bigger the number of retrieved AOTs. Note that 8-ht iteration is not always the best choice.

The content of the lines is as follows (see the first group of lines in the table below):

- ♣ ν^*-value (1.57) – first number in the raw;
- ♣ smallest and largest radii used for retrieval (- 0.008 1.00 -);
- ♣ iteration number (IT=1);
- ♣ number of retrieved AOTs (IN=6);
- ♣ number of calls of adjust subroutine after fall of convergence process over entire range of γ_{rel}-values (IC=0);
- ♣ number of steps trough γ-values, when all components of the solution vector $\vec{\mathbf{f}}$ are non-negative (Itest=-10 - here "-10" indicates that the iteration was completed successfully; see later in section 7 how in this counter varies during the search of the optimal γ_{rel}-value);
- ♣ selected optimal γ_{rel}-values in the current iteration (G(Rel)=0.128, see section 2 where selection algorithm is explained);
- ♣ obtained values of first term Q_1 in performance function (e.g. Q1=4.237E+00).

```
-------------------------------------------------------------------
 DATE:  DEC  8, 1775
 222, 08/12/1775 09:10:22, 09:10:22, +22.113344, -123.001122, 1234
 TEST 2
-------------------------------------------------------------------

-------------------------------------------------------------------------------------
 1.57 - 0.08  1.00 - IT=1 - IN= 6 - IC=0 - Itest=-10 - G(Rel)= 0.128 - Q1= 4.237E+00
 1.57 - 0.08  1.00 - IT=2 - IN= 7 - IC=0 - Itest=-10 - G(Rel)= 0.002 - Q1= 3.963E+00
 1.57 - 0.08  1.00 - IT=3 - IN= 6 - IC=0 - Itest=-10 - G(Rel)= 0.002 - Q1= 4.073E+00
 1.57 - 0.08  1.00 - IT=4 - IN= 7 - IC=0 - Itest=-10 - G(Rel)= 0.002 - Q1= 4.050E+00
 1.57 - 0.08  1.00 - IT=5 - IN= 7 - IC=0 - Itest=-10 - G(Rel)= 0.002 - Q1= 4.041E+00
 1.57 - 0.08  1.00 - IT=6 - IN= 7 - IC=0 - Itest=-10 - G(Rel)= 0.002 - Q1= 4.028E+00
 1.57 - 0.08  1.00 - IT=7 - IN= 7 - IC=0 - Itest=-10 - G(Rel)= 0.002 - Q1= 3.998E+00
 1.57 - 0.08  1.00 - IT=8 - IN= 7 - IC=0 - Itest=-10 - G(Rel)= 0.004 - Q1= 3.948E+00
-------------------------------------------------------------------------------------
. . . . . .
-------------------------------------------------------------------------------------
```

```
 2.07 - 0.08  3.50 - IT=1 - IN= 3 - IC=0 - Itest=-10 - G(Rel)= 2.048 - Q1= 2.087E+01
 2.07 - 0.08  3.50 - IT=2 - IN= 4 - IC=0 - Itest=-10 - G(Rel)= 0.512 - Q1= 1.285E+01
 2.07 - 0.08  3.50 - IT=3 - IN= 6 - IC=0 - Itest=-10 - G(Rel)= 0.128 - Q1= 6.503E+00
 2.07 - 0.08  3.50 - IT=4 - IN= 6 - IC=0 - Itest=-10 - G(Rel)= 0.064 - Q1= 4.688E+00
 2.07 - 0.08  3.50 - IT=5 - IN= 7 - IC=0 - Itest=-10 - G(Rel)= 0.032 - Q1= 4.009E+00
 2.07 - 0.08  3.50 - IT=6 - IN= 7 - IC=0 - Itest=-10 - G(Rel)= 0.016 - Q1= 3.776E+00
 2.07 - 0.08  3.50 - IT=7 - IN= 7 - IC=0 - Itest=-10 - G(Rel)= 0.004 - Q1= 3.436E+00
 2.07 - 0.08  3.50 - IT=8 - IN= 7 - IC=0 - Itest=-10 - G(Rel)= 0.004 - Q1= 3.480E+00
-------------------------------------------------------------------------------
```

7. Retrieval of aerosol distribution functions – program inversion

The program **inversion** is the main program in the package "MICROTOPS INVERSE". It performs a constrained linear inversion with second derivative smoothing constraint on spectral optical depth data in order to obtain a columnar aerosol size distribution for a pre-selected upper and lower particle radii limits and a values of Junge parameter ν^*. The pre-selection is result of the exploring results of the program **preliminary**.

The executable code **inversion.exe** and the configuration file **inversion.cfg** could be placed together in any folder. The file **inversion.cfg** contains only one line – text (typed starting at the leftmost position of the line, not longer then 80 characters and not enclosed in quotation marks) indicating the folder where both input files are placed (**inversion_in.dat** and **inversion_in_ext.dat**) and where code will write all output files, which are discussed later on. If the content of **inversion.cfg** is

c:\temp\

this means that **c:\temp** is the working folder for the code **inversion**, i.e., the code will look there for the input files **inversion_in.dat** and **inversion_in_ext.dat**. All output files will be written in the same folder. Remember that the file **inversion_in.dat** is renamed file **inv_in.dat**, when created by code **newtam**.

The program **inversion** uses the same input file **inversion_in.dat** as the program **preliminary**. An example of the input data stream written in **inversion_in.dat** was given in the previous section.

Note that the content of **inversion_in.dat** now has to be modified depending on the decision made exploring the results of program **preliminary**, i.e. the default values of parameters r_a, r_b and ν^* entered into **inversion_in.dat** by program **newtam** and not used by program **preliminary** have to be replaced by the selected values.

Dimension statements in the code **inversion** are valid for NWVL and NRAD up to 16 but corresponding values are subject to input.

The values of both keys KEYWNU and KEYIT are used to determine outputs from the program **inversion** according the conventions given in Table 5.

The values "1", "2" or "3" of KEYWNU determine which of the Junge parameters values ν^*-0.5, ν^*, ν^*+0.5 to use for inversion. For example, if KEYWNU=1 the inversion will be performed with Junge parameter ν^*-0.5. The positive values of KEYIT determine the number of the last iteration performed during the inversion, as in some cases it is better to perform less than eight iterations.

The output file **inversion_out_king.dat** is the most detailed of the output files. It follows the structure of the original King's code **RADINV**. The content of this file is as follows (please, check Appendix A for the notations used):

[1] Boundaries of sub-intervals used for inversion ;
[2] Boundaries of the coarse intervals in logarithmic scale ;
[3] Number of wavelengths, where aerosol optical depth is measured, number of radii, used during inversion (this is also the number of coarse

intervals) and common length, of all coarse intervals in logarithmic scale ;

[4] Date when AOTs have been measured ;

[5] A line with parameters describing the measurement conditions; this record has been created by program **newtam** and is used later on for sorting the results

[6] A text entered in the input file of program **newtam** and describing the group of scans or the individual scan that are handled

[7] A table with wavelengths, measured AOTs and corresponding standard deviations ;

[8] Sum of squared measurement errors ;

[9] A set parameters characterizing currently performed inversion: the complex index of refraction, expectation of the first term in performance function, number of wavelengths, where aerosol optical depth is measured, smallest and largest radii, used for inversion, number of radii, used during inversion, and parameter in Junge distribution function, used for inversion ;

Table 5

KPRINT	KEYWNU	KEYIT	Output files
1	-1	-1	inversion_out_king.dat
2	0	Any value	inversion_out_short.dat inversion _gamma.dat inversion _eff.dat
3	1,2,3	1,2,3,4,5,6,7,8	inversion _out.dat inversion _distr.dat inversion _tau.dat inversion_rad.dat

[10] Twenty four triad of titles, tables and parameters, manifesting the minimization process and result of eight performed inversion for three values of Junge parameter (ν^*-0.5, ν^*, ν^*+0.5) :

[10.4] The first table outlines the selection of the optimum Lagrange parameter ;

[10.4] The second table presents the results of the current iteration: columnar number of particles in coarse intervals, columnar aerosol size distribution and the corresponding standard deviation in logarithmic and linear r -scale; relative error or variation coefficient or ratio of the standard deviation end expectation for a particular component of the solution vector;

[10.4] The third table presents estimated AOTs on the basis of the retrieval, measured AOTs and corresponding measurement and presence of coincidences between measured and calculated AOTs ;

[10.4] The list of the parameters include: parameter in Junge distribution function, used for inversion, number of the current iteration, value of a matrix element used for normalization of Lagrange parameter, value of the selected normalized optimal Lagrange parameter, number of calls of adjust subroutine after fall of convergence process over entire range of normalized Lagrange parameter, total columnar number of particles in the vertical atmospheric column, number of coincidences between measured and

calculated AOTs and mean relative error of the solution vector components.

The output file **inversion_out_king.dat** is a perfect tool for exploring the features of King's algorithm and selecting the best retrieval conditions (value of ν^* and the number of the iteration producing the most convincing results) but at the same time it is difficult to use it for presenting results in the case when we have to consider results of sets of experimental data.

The output files **inversion_out_short.dat** and **inversion_gamma.dat** was created as an alternative of **inversion_out_king.dat**. These files are much shorter and after a few exercises the users will find that they are an effective tool for selecting the best retrieval.

The output file **inversion_eff.dat** contains the array of data $\pi \times \overline{R}_j^2 \times Q_{ext}(2\pi \overline{R}_j / \lambda_i)$, $i = 1,2,\ldots, p$, $j = 1,2,\ldots,136$, which can be used to study the most informative radii for retrieval.

The output file **inversion_out.dat** is the main output file. It contains all numerical information that can be obtained from the inversion – three types of distributions, column amounts, effective radii and mean relative error of the solution vector.

The output file **inversion_distrt.dat** and **inversion_tau.dat** represent only parts of the main output file **inversion_out.dat** – the tables with column amounts and distributions in the first file and the tables with AOTs in the second one.

Remember that input file **inversion_in.dat** may contain input data (groups of six lines, see previous section six) for as many as desired by the user experimental data sets, calculated by programs **newtam** for groups of scans and/or individual scans. As a result the output files will become a little bit complicated because they contain the corresponding amount of output data sets.

The content of the last output file **inversion_distrt.dat** requires more detailed explanation. In this file a line holds the data obtained from handling one experimental data set. The content lines are as follows:

(1) consecutive number in the input data sets ;
(2) date and time, when experiment has been performed – may be used as abscissa in Excel graphs ;
(3) time, when experiment has been performed ;
(4) latitude of experimental site ;
(5) longitude of experimental site ;
(6) altitude of experimental site ;
(7) ν^*-value used for inversion ;
(8) smallest and largest radii used for inversion ;
(9) ensemble mean radius ;
(10) ensemble geometrical mean radius ;
(11) ensemble radius of average surface ;
(12) ensemble radius of average volume ;
(13) ensemble surface weighted mean radius ;

(14) ensemble volume weighted mean radius ;
(15) total columnar number of particles.

8. Reordering retrieval results for further analysis – programs sort_tau and sort_distr

The further analysis of results obtained by program **inversion** requires intercomparison of distributions, column amounts and effective radii obtained at different locations, time instances and so on. All these actions imply capacious and time consuming reordering of **inversion**'s output files. Two assistant program (**sort_distr** and **sort_tau**) have been created, aiming to overcome these technical difficulties.

The purpose of program **sort_distr** is to read one file of the type **inversion_distr.dat** and to reorder the content into the following files:

(a) **particles.dat** - this files contains a table with **N+2** columns, where **N** is the number of the group input data in the file **inversion_distr.dat**. The first column is the consecutive number of the coarse radii interval. The second column contains corresponding geometrical midpoints. The remaining **N** columns are filled up with the columnar number of particles in the current coarse interval – one column for each group of data. These columns are titled by the consecutive number in the data set (in the file **inversion_distr.dat**), date and time, time, latitude, longitude or altitude of the measurement site. These titles can be used later on for labeling plots built by graphing software, e.g. Origin or Excel.

(b) **number_distr.dat** - the content of this file is similar of the content of the file **particles.dat**, but instead **N+2** columns it consists of **2×N+2** columns – two columns for each group input data containing columnar aerosol number distribution in logarithmic and linear r -scale. These columns are titled in the same manner.

(c) **surface_distr.dat** - the content of this file is similar of the content of the file **number_distr.datb**, but here subject of re-ordering are the surface distributions in logarithmic and linear r -scale.

(d) **volume_distr.dat** - the content of this file is similar of the content of the file **number_distr.datb**, but here subject of re-ordering are the volume distributions in logarithmic and linear r -scale.

(e) **protocol.dat** - the content of this file is self explainable after exploring it.

Program **sort_distr** uses a configuration file (**sort_distr.cfg**) to set the task to be performed. The executable code **sort_distr.exe** and the configuration file **sort_distr.cfg** have to be placed together in any folder selected by the user. The input file **inversion_distr.dat** has to be positioned in the working folder, where all the output files discussed above will also be created. The input file **inversion_distr.dat** could be constructed from any number of joined files of the same type **inversion_distr.dat**, but in order to be comparable the inversions have to be performed over the same radii range and number of radii. Below one configuration file is listed and its content explained.

```
c:\temp\
inversion_distr.dat
particles.dat
numner_distr.dat
surfase_distr.dat
volune_distr.dat
```

```
 5  4 1
--------------------------------------------------
Do not modify lines below!

First  line  - path to input/output data lines
               max 80 characters
Second line  - input filename with extension but w/o path
               this is the radinv_distr.dat file
               max 80 characters
Third  line  - first output filename with extension but w/o path
               max 80 characters
               this file will contain sorted retrieved
               aerosol number distributions
Forth  line  - second output filename with extension but w/o path
               max 80 characters
               this file will contain sorted retrieved
               aerosol surface distributions
Fifth  line  - third output filename with extension but w/o path
               max 80 characters
               this file will contain sorted retrieved
               aerosol volume distributions
Sixth  line  - forth output filename with extension but w/o path
               max 80 characters
               this file will contain sorted retrieved
               aerosol ??????
Seventh line - NWVL, NMEAS, NTITLE
               format(I2,I3,I2)
               NRAD   - number of radii (max 16)
               NMEAS  - number of retrieval results collected in the input
                        file (max 500)
               NTITLE - number of sequential record in formatted string used
                        for labeling the sorted outputs
                        1 - number in the set
                        2 - date and time
                        3 - time
                        4 - latitude
                        5 - longitude
                        6 - altitude
                        7 - date
```

During the execution of the code, the content of the input file **inversion_distr.dat is printed on the screen**. The sprint is scrollable and the user has to terminate execution by pressing the RETURN key.

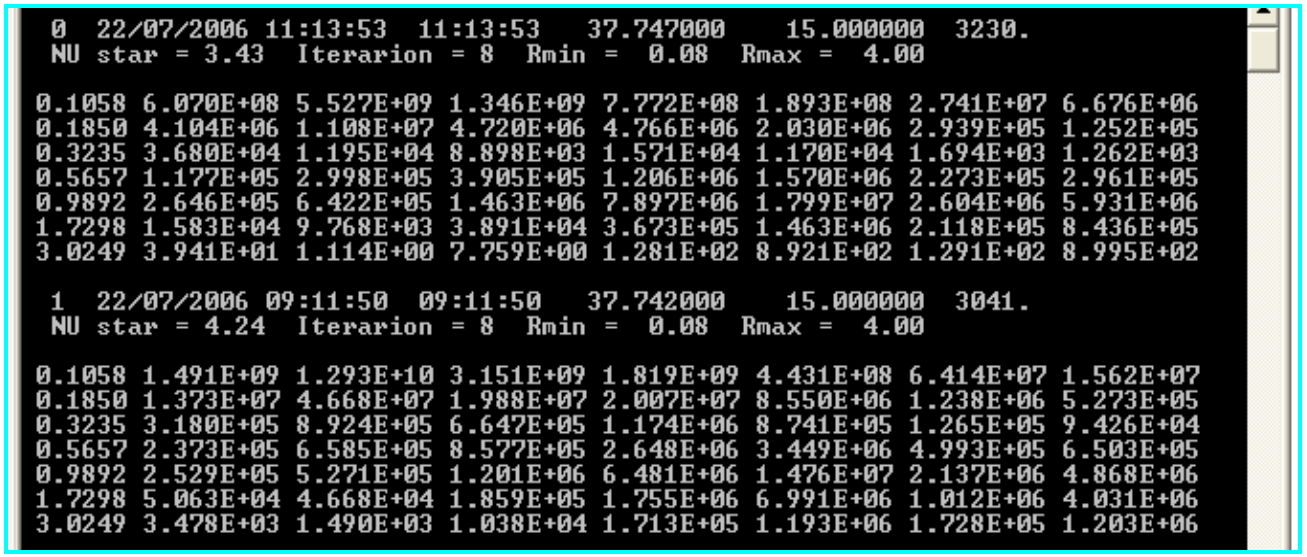

Figure 13.

The purpose of program **sort_tau** is to read one file of the type **inversion_tau.dat** and to reorder its content into the following files:

(a) **tau_measured.dat** - this files contains a table with 2×**N+2** columns, where **N** is the number of the group input data in the file **inversion_distr.dat**. The first column is the consecutive number of photometer channel, i.e. wavelength. The second column contains peak channels' wavelengths. The rest 2×**N** columns are filled up with the measured AOTs and their standard deviations. These columns are titled in the same manner as corresponding columns in the file **inversion_distr.dat**.

(b) **tau_calculated.dat** - the content of this file is similar to that of **tau_measured.dat** but here, the subjects of reordering are AOTs and the indicator of coincidences between measured and calculated AOTs ("1" for coincidence and "0" otherwise).

(c) **protocol.dat** - the content of this file is self explanatory after exploring it.

Program **sort_tau** uses a configuration file (**sort_tau.cfg**) to set the task to be performed. The executable code **sort_tau.exe** and the

configuration file **sort_tau.cfg** have to be placed together in any folder selected by the user. The input file **inversion_tau.dat** has to be positioned in the working folder, where all output files discussed above will also be created. The input file **inversion_tau.dat** could be constructed from any number of joined files of the same type **inversion_tau.dat**, but in order to be comparable the inversions have to be performed on data collected by the same Sun photometer, i.e. wavelengths have to coincide in all data sets. Below one configuration file is listed and its content explained.

```
c:\temp\
inversion_tau.dat
tau_measured.dat
tau_retrieved.dat
 5  8 3
-----------------------------------------------------
Do not modify lines below!

First  line - path to input/output data lines
              max 80 characters
Second line - input filename with extension but w/o path
              this is the radinv_tau.dat file
              max 80 characters
Third  line - first output filename with extension but w/o path
              max 80 characters
              this file will contain sorted measured aerosol optical
              depths and corresponding errors
Forth  line - second output filename with extension but w/o path
              max 80 characters
              this file will contain sorted retrieved aerosol optical
              depths
Fifth  line - NWVL, NMEAS, NTITLE
              format(I2,I3,I2)
              NWVL   - number of wavelengths (max 16)
              NMEAS  - number of retrieval results
                       collected in the input file (max 500)
              NTITLE - number of sequential record in formatted
                       string used for labeling the sorted outputs
                       1 - number in the set
                       2 - date and time
                       3 - time
                       4 - latitude
                       5 - longitude
                       6 - altitude
```

```
7 - date
```

During the execution of the code, the content of the input file **inversion_distr.dat is printed on the screen**. The sprint is scrollable and the user has to terminate execution by pressing the RETURN key.

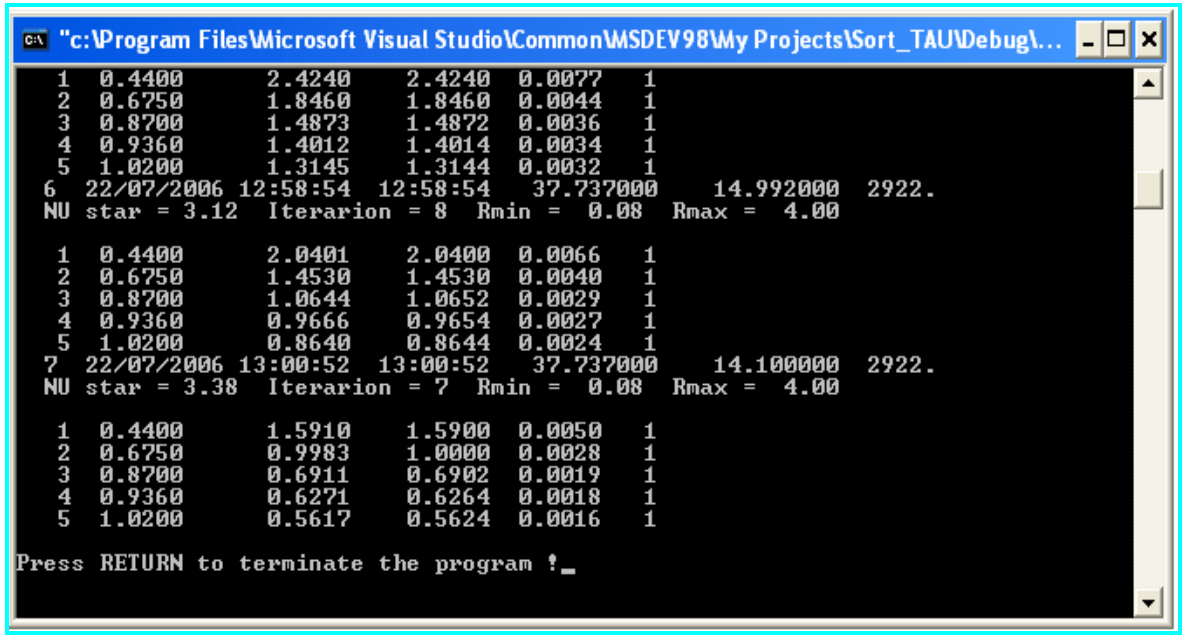

Figure 14.

Appendix 1 Nomenclature of notations used in output files

Table A1.1

Output variable or array	Used in output file	FORTRAN notation	Explanation *
1	**2**	**3**	**4**
NWVL	preliminary_out.dat inversion_out_king.dat inversion_out.dat	NWVL	Number of wavelengths, p, where aerosol optical depth is measured
WAVELENGTH(MICRONS) or WVL	preliminary_out.dat inversion_out_king.dat inversion_out.dat	WVL(I), I=1,NWVL	Array of wavelengths (microns), λ_i, $i = 1,2,\ldots, p$, where aerosol optical depths are measured; these wavelengths coincide with peak wavelengths of the five MICROTOPS II channels
NRAD	preliminary_out.dat inversion_out_king.dat inversion_out.dat	NRAD	Number of radii, q, used during inversion; number of coarse intervals
Rmin , Rmax	inversion_out_king.dat inversion_out_short.dat inversion_out.dat	RI , RF	Smallest radius (microns), r_a (RI), used for inversion Largest radius (microns), r_b (RF), used for inversion
RE(I):	inversion_out_king.dat	RE(I), I=1,137	Boundaries R_k of sub-intervals used for inversion
DX	inversion_out_king.dat	DX	Common length, $\Delta \log r = (\log(r_b) - \log(r_a))/q$, of all coarse intervals in logarithmic scale
NK(I)	inversion_out_king.dat	NK(I), I=1,NRAD	Indices $N(k)$ of sub-intervals (RE(I)), corresponding to the coarse interval boundaries; $R_{N(k)} = r_k$
X(I)	inversion_out_king.dat	X(I), I=1, NRAD	Boundaries of the coarse intervals $[r_j, r_{j+1}]$ in logarithmic scale; all these intervals have equal length $\Delta \log r = (\log(r_b) - \log(r_a))/q$, the values of the matrix $\mathbf{A}_{ij}$ are calculated by integration over these intervals

1	2	3	4
BOUNDARIES [MICRONS] **or** Rleft , Rright	inversion_out_king.dat inversion_out.dat	RE(NK(I)), RE(NK(I+1)), I=1,NREAD+1	Both boundaries (left and right) of each coarse intervals are printed under this caption
R [MICRON] **or** Rmean	inversion_out_king.dat inversion_out.dat	RR(J), J=1, NRAD	Geometrical mean $\bar{r}_j = \sqrt{r_j\, r_{j+1}}$ of boundaries of j-th coarse interval $[r_j, r_{j+1}]$
Index of refraction	preliminary_out.dat inversion_out_king.dat inversion_out.dat	RFR, RFI	Real and imaginary parts of complex index of refraction - $\tilde{m} = n - i \times k$
SNU **or** Junge parameter NU star **or** NU star	preliminary_out.dat inversion_out_king.dat inversion_out_short.dat inversion_out.dat inversion_tau.dat	SNU	Parameter ν^* in Junge distribution function $h(r) = r^{-(\nu^*+1)}$; $\nu^* = \alpha + 2$, where α is the exponent in Ångström's empirical formula $\tau_{\mathrm{M}}(\lambda) = \beta \lambda^{-\alpha}$
TAU(AEROSOLS) **or** MEAS TAU	preliminary_out.dat inversion_out_king.dat inversion_out.dat	TAU(I), I=1,NWVL ERR(I), I=1,NWVL	Under this caption are printed the values of the measured AOTs and corresponding measurement errors by wavelength, i.e. $\tau_{\mathrm{M}}(\lambda_i), i = 1,2,\ldots, p$ and $\delta\tau(\lambda_i), i = 1,2,\ldots, p$
Sum of squared measurement errors	preliminary_out.dat inversion_out_king.dat inversion_out	SUMSQ	Sum of squared measurement errors - $\sum_{i=1}^{p} (\delta\tau_{\mathrm{M}}(\lambda_i))^2$
COMP TAU	inversion_out_king.dat inversion_out.dat	TC(I), I=1,NWVL	Arrays of theoretical AOTs, $\tau_{\mathrm{C}}(\lambda_i), i = 1,2,\ldots, p$, or estimated AOTs on the basis of the retrieval
IN **or** Number of coincidences between measured and calculated AOTs is	preliminary_short.dat inversion_out_king.dat inversion_out.dat	IWVLTP NME	Number of retrieved measured AOTs or number of coincidences between measured and calculated AOTs , $\mathrm{M_c}$, i.e. the number of times when the inequality $\lvert \tau_{\mathrm{C}}(\lambda_i) - \tau_{\mathrm{M}}(\lambda_i) \rvert \le \delta\tau_{\mathrm{M}}(\lambda_i)$ holds true

1	2	3	4
F(1) … F(NRAD)	inversion_out_king.dat	F(1) … F(NRAD)	First, second, … and q-th components ($\mathbf{f}_1$, …, $\mathbf{f}_q$) of the solution vector $\vec{\mathbf{f}}$
Q	inversion_out_king.dat	Q	Performance function defined as $Q = Q_1 + \gamma\, Q_2$, where γ is the non-negative Lagrange multiplier
Q1	preliminary_short.dat inversion_out_king.dat	Q1	First term in performance function, i.e. $Q_1 = \vec{\boldsymbol{\varepsilon}}^{\mathrm{T}} \hat{\mathbf{C}}^{-1} \vec{\boldsymbol{\varepsilon}} = \sum_{i=1}^{p} w_i\, \boldsymbol{\varepsilon}_i^2$, where $\hat{\mathbf{C}}$ is the measurement covariance matrix $\mathbf{C}_{ij} = \delta_{ij}\, w_i$ (δ_{ij} is Kronecker delta), $w_i = 1/(\delta\tau(\lambda_i))^2$, and $\vec{\boldsymbol{\varepsilon}}$ is an unknown vector whose components $\boldsymbol{\varepsilon}_i$ represent the difference between measured $\tau_{\mathrm{M}}(\lambda_i)$ and theoretical $\tau_{\mathrm{C}}(\lambda_i) = \sum_j \mathbf{A}_{ij}(\lambda_i)\, f_j$ AOTs
Q1est	inversion_out_king.dat inversion_out.dat	Q1EST	Expectation $E\{Q_1\}$ of the first term in performance function, in our case $E\{Q\} = p$
Q2	inversion_out_king.dat	Q2	Second term in performance function, i.e. $Q_2 = \vec{\mathbf{f}}^{T}\, \hat{\mathbf{H}}\, \vec{\mathbf{f}}$, where $\hat{\mathbf{H}}$ is Twomey's smoothing matrix
Sigma epsilon squared **or** Eps Sq	inversion_out_king.dat inversion_out_short.dat inversion_out.dat inversion_gamma.dat	EPSQ	The sum $\sum_{i=1}^{p} \varepsilon_i^2$, where $\varepsilon_i = \tau_{\mathrm{M}}(\lambda_i) - \tau_{\mathrm{C}}(\lambda_i)$, this quantity is a measure of sample variance $s^2 = \frac{1}{p-q} \sum_{i=1}^{p} \varepsilon_i^2$ of the regression fit

1	2	3	4
IT **or** Iteration number	preliminary_short.dat inversion_out_king.dat inversion_out.dat inversion_gama.dat inversion_short.dat inversion_eff.dat inversion_tau.dat inversion_distr.dat	NODE	Number of the current iteration in **preliminary_inv** and **inversion**.
Gamma	Inversion_out_king.dat	GAM	Lagrange multiplier γ
G(Rel) **or** Gamma/ATA(1,1)	preliminary_short.dat inversion_out_king.dat inversion_out.dat	GGAM	Relative value $\gamma_{\text{rel}} = \gamma \mathbf{H}_{11} / (\hat{\mathbf{A}}^{\mathrm{T}} \hat{\mathbf{C}}^{-1} \hat{\mathbf{A}})_{11}$ of Lagrange multiplier γ; note that γ_{rel} varies in the range 10^{-3} to 5 until range of values are reached for which all element of solution vector $\vec{\mathbf{f}}$ are positive (negative values constitute an unphysical solution); if $\hat{\mathbf{H}}$ is Twomey's matrix, then $\mathbf{H}_{11} = 1$
(ATA)11	preliminary_short.dat inversion_out_king.dat inversion_out.dat	ATA(1,1)	Value of the matrix element $(\hat{\mathbf{A}}^{\mathrm{T}} \hat{\mathbf{C}}^{-1} \hat{\mathbf{A}})_{11}$
Itest	preliminary_short.dat inversion_gamma.dat		The number of steps trough γ-values, when all components of the solution vector $\vec{\mathbf{f}}$ are non-negative
Icnt **or** IC **or** Adjust routine used # times	preliminary_short.dat inversion_out_king.dat inversion_out.dat	ICOUNT	The number of calls of adjust subroutine after fall of convergence process over entire range of γ-values , i.e. when some of $\vec{\mathbf{f}}$-components are negative; I do not recommend to trust inversions where this happened more then one-two times or during the last three iterations (# 6, 7 or 8).

1	2	3	4
Rgeom_[μm]	inversion_eff.dat	R(I), I=1, 137	Array $\overline{R}_k = \sqrt{R_k\, R_{k+1}}$, $k = 1, 2, \ldots, 136$, of the geometrical means (geometrical midpoints) of the boundaries of 136-sub-intervals used for inversion
Q(0.4400nm) Q(0.6750 nm) Q(0.8700nm) Q(0.9360nm) Q(1.0200nm)	inversion_eff.dat	SIGEXT(I,J), I=1, NWVL J=1, 137	Array of data $\pi \times \overline{R}_j^2 \times Q_{ext}(2\pi\, \overline{R}_j / \lambda_i)$, $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, 136$, where λ_i are the wavelengths (microns), where AOTs aremeasured, and $\overline{R}_j$ are geometrical midpoints of 136-sub-intervals used for inversion
INT	inversion_out_king.dat inversion_out.dat	J	Number of the current coarse interval $[r_j, r_{j+1}]$
CD(R) **or** N partial	nversion_out_king.dat inversion_out.dat	CNP(J), J=1, NRAD	Columnar number of particles in the current coarse interval $[r_j, r_{j+1}]$, i.e. $\Delta n_c(\overline{r}_j) = \int_{r_j}^{r_{j+1}} n_c(r)\, dr$; units (particles $\times$ cm^{-2})
Total columnar number of particles	inversion_out_king.dat inversion_out.dat	CN	Total columnar number of particles $\mathrm{CN} = n_c = \int_{r_{\min}}^{r_{\max}} n_c(r)\, dr = \sum_{i=1}^{q} \Delta n_c(\overline{r}_i)$; units (particles $\times$ cm^{-2})
dN(r)	inversion_out.dat	DNR(I)	Columnar aerosol size distribution $n_c(r)$ (or columnar aerosol number density) in in linear r -scale. The $n_c(r)$ is the number of particles per unit area per unit radius interval in vertical column through the atmosphere; units (particles $\times$ cm^{-2} $\times$ μm^{-1}).
DN(R)/D(R) [PART/(CM**2-MICRON)]	inversion_out.dat	DNR(I), ERDNR=DNR(I) * SIGMAF(I) / F(I)	Both $n_c(r)$ and the corresponding standard deviation are printed under this caption; both quantities are in linear r -scale; units (particles $\times$ cm^{-2} $\times$ μm^{-1}).

1	2	3	4
dN(lgr)	inversion_out.dat	DNLGR(I)	Columnar aerosol size distribution $n_c(\log r)$ (or columnar aerosol number density) in logarithmic r-scale The $n_c(\log r)$ is the number of particles per unit area per unit log-radius interval in vertical column through the atmosphere; units (particles $\times$ cm^{-2}) **NOTE**: $n_c(\log r) = \ln 10 \times r \times n_c(r)$
DN(R)/D(LOGR) [PART/CM**2]	inversion_out.dat	DNLGR(I), ERDNLG=DNLGR(I) * SIGMAF(I) / F(I)	Both quantities columnar aerosol size distribution $n_c(\log r)$ and the corresponding standard deviation are printed under this caption; both quantities are in logarithmic r-scale; units (particles $\times$ cm^{-2} $\times$ μm^{-1}).
dS(r)	inversion_out.dat	DSR(I)	Columnar aerosol surface distribution $s_c(r)$ (or columnar aerosol surface density) in linear r-scale. The $s_c(r)$ is the surface of particles per unit area per unit radius interval in vertical column through the atmosphere; units (particles $\times$ cm^{-2} $\times$ μm^2). **NOTE**: $s_c(r) = \pi \times r^2 \times n_c(r)$
dS(lgr)	inversion_out.dat	DSLGR(I)	Columnar aerosol surface distribution $s_c(\log r)$ (or columnar aerosol surface density) in logarithmic r-scale. The $s_c(\log r)$ is the surface of particles per unit area per unit log radius interval in vertical column through the atmosphere; units (particles $\times$ cm^{-2} $\times$ μm). **NOTE**: $s_c(\log r) = \ln 10 \times r \times s_c(r)$

1	2	3	4
dV(r)	inversion_out.dat	DVR(I)	Columnar aerosol volume distribution $v_c(r)$ (or columnar aerosol volume density) in linear r-scale. The $v_c(r)$ is the volume of particles per unit area per unit radius interval in vertical column through the atmosphere; units (particles $\times$ cm^{-2} $\times$ μm^2). **NOTE**: $v_c(r) = \frac{4}{3}\pi \times r^3 \times n_c(r)$
dV(lgr)	inversion_out.dat	DVLGR(I)	Columnar aerosol volume distribution $v_c(\log r)$ (or columnar aerosol surface density) in logarithmic r-scale. The $v_c(\log r)$ is the volume of particles per unit area per unit log radius interval in vertical column through the atmosphere; units (particles $\times$ cm^{-2} $\times$ μm^3). **NOTE**: $v_c(\log r) = \ln 10 \times r \times v_c(r)$
PERCENT ERROR	inversion_out.dat	PERERR=100* SIGMAF(I) / F(I)	Relative error, variation coefficient or ratio of the standard deviation end expectation for a particular component of the solution vector $\vec{\mathbf{f}}$; units (percent)
Mean Relative Error **or** M Rel Err	inversion_out_king.dat inversion_out.dat	SUMERR	Mean relative error E_{rel} of the solution vector $\vec{\mathbf{f}}$ components
Mean radius	inversion_out.dat	RMEAN	Ensemble mean radius (microns) $r_{\text{mean}} = \sum_{j=1}^{q} \Delta n_c(\bar{r}_j)\bar{r}_j \Big/ \sum_{j=1}^{q} \Delta n_c(\bar{r}_j) = \sum_{j=1}^{q} \Delta n_c(\bar{r}_j)\bar{r}_j \Big/ \text{NC}$
Geometrical mean radius	inversion_out.dat	RGEOM	Ensemble geometrical mean radius (microns) $r_{\text{mean}} = \exp\left(\sum_{j=1}^{q} \Delta n_c(\bar{r}_j)\bar{r}_j \Big/ \text{NC} \right) = \left(\prod_{j=1}^{q} \bar{r}_j^{\,\Delta n_c(\bar{r}_j)} \right)^{1/\text{NC}}$

1	2	3	4
Radius of average surface	inversion_out.dat	RSURF	Ensemble radius of average surface (microns) $r_{\mathrm{s}} = \left(\sum_{j=1}^{q} \Delta n_{\mathrm{c}}(\bar{r}_j) \bar{r}_j^2 \Big/ \mathrm{NC} \right)^{1/2}$
Radius of average volume	inversion_out.dat	RVOL	Ensemble radius of average volume (microns) $r_{\mathrm{v}} = \left(\sum_{j=1}^{q} \Delta n_{\mathrm{c}}(\bar{r}_j) \bar{r}_j^3 \Big/ \mathrm{NC} \right)^{1/3}$
Surface weighted mean radius	inversion_out.dat	RSM	Ensemble surface weighted mean radius $r_{\mathrm{sm}} = \sum_{j=1}^{q} \Delta n_{\mathrm{c}}(\bar{r}_j) \bar{r}_j^3 \Big/ \sum_{j=1}^{q} \Delta n_{\mathrm{c}}(\bar{r}_j) \bar{r}_j^2$
Volume weighted mean radius	inversion_out.dat	RVM	Ensemble volume weighted mean radius $r_{\mathrm{vm}} = \sum_{j=1}^{q} \Delta n_{\mathrm{c}}(\bar{r}_j) \bar{r}_j^4 \Big/ \sum_{j=1}^{q} \Delta n_{\mathrm{c}}(\bar{r}_j) \bar{r}_j^3$

Comments: * Here notations used coincide with that in King's papers

Appendix 2 Test data and results

1. Description of provided test tasks

Each program included in software package **MICROTOPS INVERSE** is accompanied by at least one test task. The tasks descriptions and files included are listed in Table A2.1

Table A2.1

Program and task ID	Input files (0)	Output files (0)	Description
1	2	3	4
newtam Test Etna	newtam_testetna.cfg testetna.csv	temp_testetna.dat (1) inv_in_testetna.dat (2)	Set of 70 scans recorded near the active vent on Mount Etna vulcano in July 2006. This set can be easily separated into five groups of scans. One of them is accepted as caused by background aerosols (`Background – records 26-50`). Four other groups (`records 1-10, 11-25, 51-60, 61-70`) are selected by inspectrion of the time variations of registered AOTs. Three individual scan, members of fourth group characterized with very strong fluctuations, are also listed and hadled as an additional example. The "manual" exploration, which has to precede execution of code **newtam** is performed in the file testetna.xls. The worksheet **groups_and_fits** in the same file manifests difficulties in the manual evaluation of Ångström parameters and calculation of group mean and standard deviations.
eff_factors Test	eff_factors_test.cfg eff_factors_in_test.cfg	eff_factors_out_test.cfg inversion_in_ext_test.dat	This test task is set to calculate Mie efficiency factors for nine different values of the complex index of refraction.

1	2	3	4
preliminary Test 1	inversion.cfg inversion_in_ext.dat (3) inversion_in_test1.dat	preliminary_out_test1.dat preliminary_short_test1.dat	This test illustrates retrieval of monodisperse distribution.
preliminary Test 2	inversion.cfg inversion_in_ext.dat (3) inversion_in_test2.dat	preliminary_out_test2.dat preliminary_short_test2.dat	This test illustrates retrieval of bi-modal distribution.
Preliminary Test Etna	inversion.cfg inversion_in_ext.dat (3) inversion_in_testetna.dat	preliminary_out_testetna.dat preliminary_short_short.dat	This is a continuation of test Etna. The final conclusion is that all four groups and the three individual scans can be retrieved within radii range from 0.08 to 4.00 μm. For simplicity is accepted that the complex index of refraction is $1.45 - i \times 0.00$
Inversion Test 1 *	inversion.cfg inversion_in_ext.dat (3) inversion_in_test1.dat (4) inversion_in_test1_final.dat	inversion_out_king_test1.dat inversion_out_short_test1.dat inversion_gamma_test1.dat inversion_eff_test1.dat inversion_out_test1.dat inversion_distr_test1.dat inversion_tau_test1.dat inversion_rad_test1.dat	This test illustrates retrieval of monodisperse distribution.
Inversion Test 2 *	inversion.cfg inversion_in_ext.dat (3) inversion_in_test2.dat (4) inversion_in_test2_final.dat	inversion_out_king_test2.dat inversion_out_short_test2.dat inversion_gamma_test2.dat inversion_eff_test2.dat inversion_out_test2.dat inversion_distr_test2.dat inversion_tau_test2.dat inversion_rad_test2.dat	This test illustrates retrieval of bi-modal distribution.

1	2	3	4
Inversion Test Etna *	inversion.cfg inversion_in_ext.dat (3) inversion_in_testetna.dat (4) inversion_in_testetna_final.dat	inversion_out_king_testetna.dat inversion_out_short_ testetna.dat inversion_gamma_ testetna.dat inversion_eff_ testetna.dat inversion_out_ testetna.dat inversion_distr_ testetna.dat inversion_tau_ testetna.dat inversion_rad_ testetna.dat	This is a continuation of test Etna. The final conclusion is that all four groups and the three individual scans can be retrieved within radii range from 0.08 to 4.00 μm. For simplicity is accepted that the complex index of refraction is $1.45 - i \times 0.00$
sort_distr Test Etna	sort_DISTR_testetna.cfg inversion_distr_testetna.dat (5)	prtotocol_1.dat (6) particles_1.dat (6) numner_distr_1.dat (6) surface_distr_1.dat (6) volume_distr_1.dat (6) prtotocol_3.dat (7) particles_3.dat (7) numner_distr_3.dat (7) surface_distr_3.dat (7) volume_distr_3.dat (7)	This is a continuation of test Etna.
sort_tau Test Etna	sort_TAU_testetna.cfg inversion_tau_testetna.dat (5)	tau_measured_1.dat (6) tau_calculated_1.dat (6) tau_measured_3.dat (7) tau_calculated_3.dat (7)	This is a continuation of test Etna.

Comments: (*) Users of Origin will find three bonus files **test1.opj**, **test2.opj** and **testetna.opj** where they could explore these tasks in more details.

(0) File names listed in both columns have to be changed before running tests in a way, which comply with file name convention. For example **newtam_testetna.cfg** has to be renamed **to newtam.cfg**.

(1) The file temp_testetna_TC.dat contains and the Ångström parameters estimated using the non-linear fit performed by program Table Curve

(2) Later this file is used as an input file for codes **preliminary** and **inversion** and for this purpose it has been renamed to **inversion_in_testetna.dat.**

(3) This is the renamed output file **inversion_in_ext_test.dat** calculated by code **eff_factors**.

(4) Remember that content of the file **inversion_in.dat** varies depending on values of keys KEYWNU and KEYIT. Final versions of these files have suffix "**final**" and there values of KEYWNU and KEYIT are selected to provide optimal retrieval.
(5) These two files are obtained as results of test Etna using code **inverse**.
(6) In these files data columns are labelled by the sequential number in the task solved.
(7) In these files data columns are labelled by time of the measurement.

2. Classification of experimental data and expected solution type

In most cases the columnar size distributions can be classified in terms of three different types of distributions, although gradations between two different types are occasionally observed making this classification somewhat arbitrary (King, 1982).

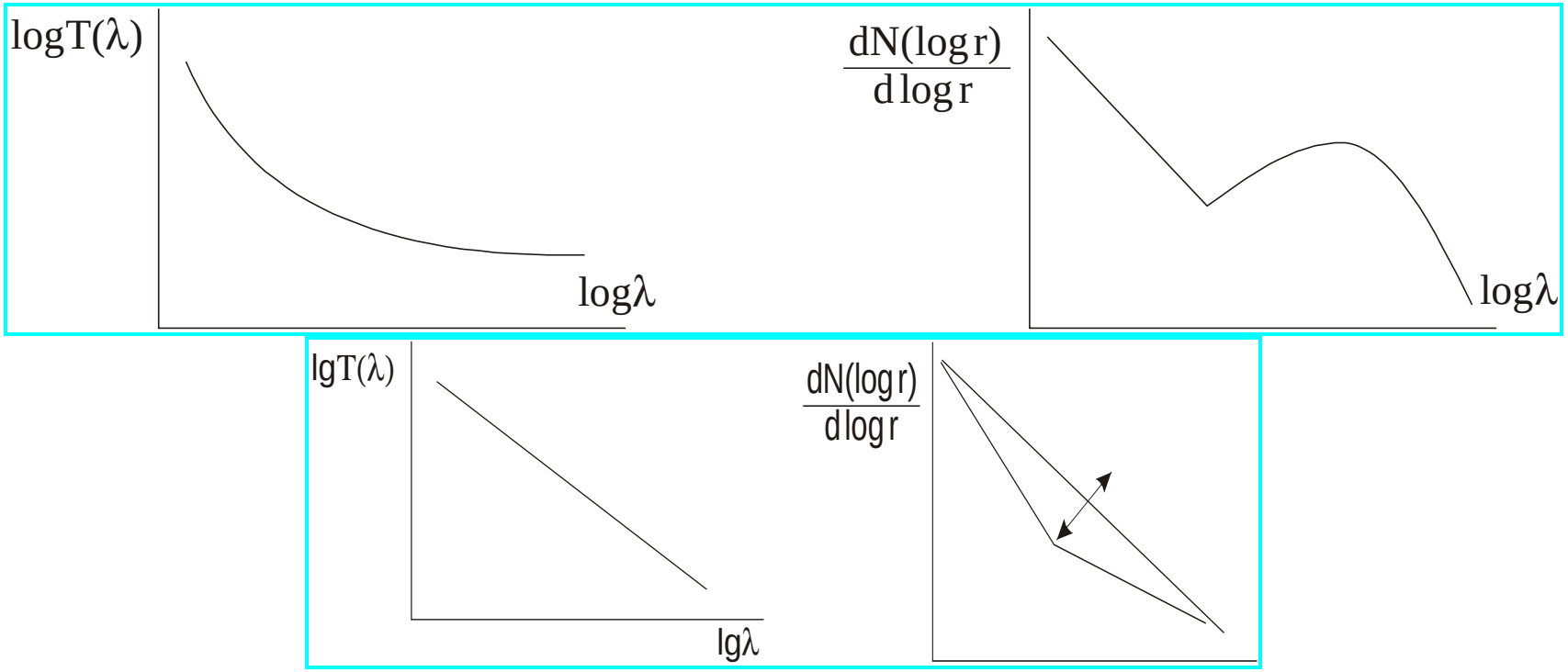

Figure A2.1. Type I measured AOTs and corresponding columnar size distributions

An example of the first type (**Type I**) is illustrated in Figure A2.1. The measured AOTs nearly follow Ångström's formula given by equation $\tau_{\mathrm{M}}(\lambda) = \beta \lambda^{-\alpha}$. The aerosol size distributions illustrated in the same figure can be best described as constructing Junge or two slopes type of distributions.

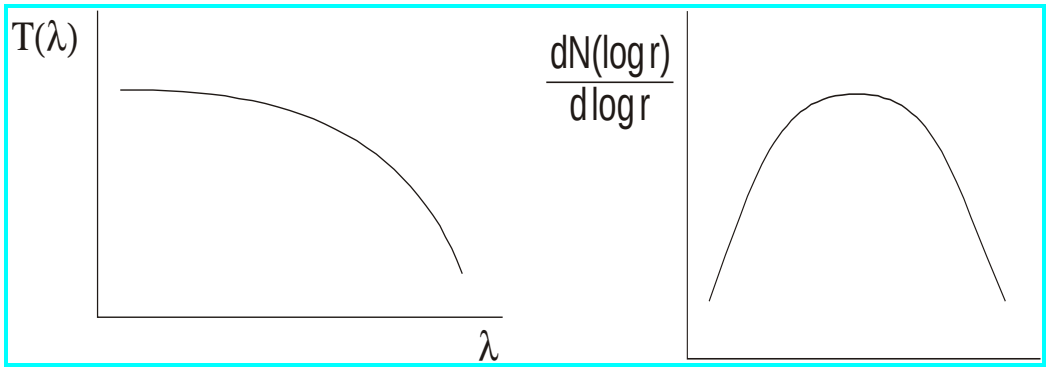

Figure A2.2. Type II measured AOTs and corresponding columnar size distributions

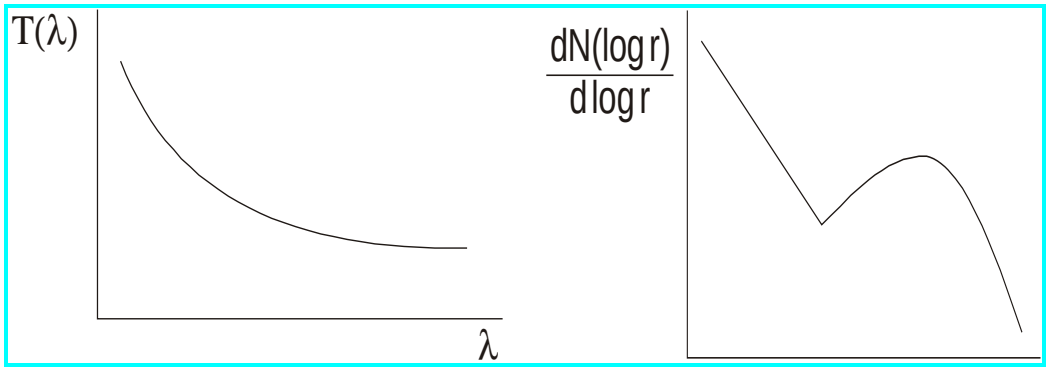

Figure A2.3. Type III measured AOTs and corresponding columnar size distributions

When measured AOTs exhibit small negative curvature (see Figure A2.2) the solutions tend to be monodisperse. This is not unexpected because the tendency for negative curvature suggests an absence of both small and large particles. Experimental data of this type (**Type II**) are

often difficult to invert due to problems associated with determining the radius range having the major contribution to the measurements. Sometimes it is desirable to invert such a data set several times with slight alterations in radius range. An important modification is the case when experimental data constitute two overlapping monodisperse distributions.

The most interesting distribution type is one for which the AOTs intermediate between these of type I and type II. In this case $\tau_M(\lambda)$ tends generally to have positive curvature. An example of this type (**Type III**) is illustrated in Figure A2.3. The solution is usually a bi-modal distribution.

Test 1 and Test 2 are examples of Type II and Type III distributions. Origin users may explore projects test1.opj and test2.opj for more details.

3. Example 1 – convergence of solution

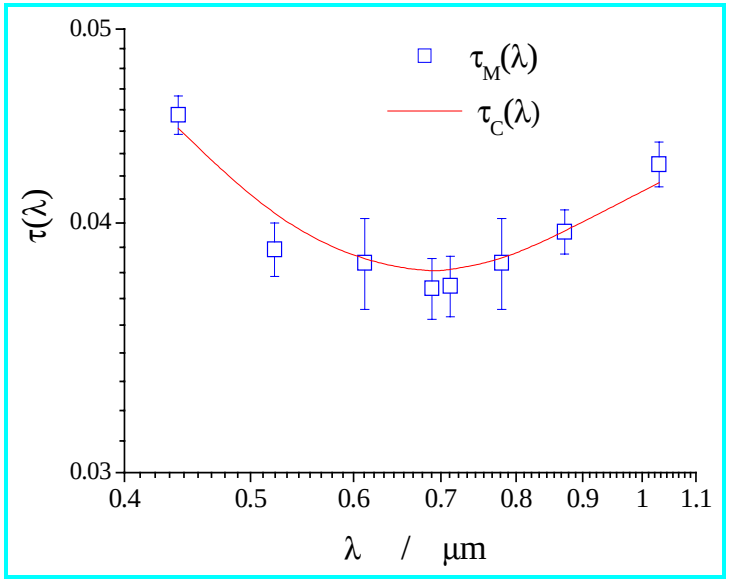

Figure A2.4. Measured and retrieved AOTs in Test 2

Test 2 is a typical example of measured AOTs with positive curvature of the plotted $\log(\tau_M(\lambda))$ *vs.* $\log(\lambda)$. The measured AOTs and their experimental errors together with retrieved AOTs are presented in Figure A2.4. Solution is obviously very good and only one outlier is

(AOT at 0.5217μm) is not reproduced by the obtained solution.

Figures A2.5- A2.8 manifest the selection of optimal value of γ_{rel} during four iterations. It will be useful to explore the content of the file **inversion_out_king_test2.dat** together with these four plots. During first iteration the biggest possible value of γ_{rel} is selected because only it guaranties positiveness of all components of the solution vector. The most problematic components during this iteration are $\mathbf{f}_2$ and $\mathbf{f}_3$ as they tend to be negative longer that the rest of the components. During the second iteration the situation is similar but now $\mathbf{f}_4$ is approaching slowly the asymptotic unit value. This iteration ends up with very small positive value (0.001) of $\mathbf{f}_3$. During next six iterations all solution components (except $\mathbf{f}_6$ and $\mathbf{f}_7$) are stable and almost equal to unit. The mentioned two components are responsible for the gradually decrease of the number of particles with largest radii and pumping particles into the maximum at 1 μm. The behavior of $\mathbf{f}_2$ and $\mathbf{f}_3$ is due to the forming of a minimum at 0.3734 μm during first iterations. After iteration number 2 this minimum remains relatively stable. The same tendencies are observable in Figure A2.9 where the dependency of performance function terms on value of γ_{rel} during iteration numbers 2 and 6 is plotted. Dotted vertical horizontal value denotes the limit value $Q_1 = 8$ of the first term. Finally Figure A2.10 shows the modification of the columnar size distribution in all five performed iterations in Test 2. The red bold line is the initial guess for $h^{(0)}(r)$, i.e. the Junge distribution.

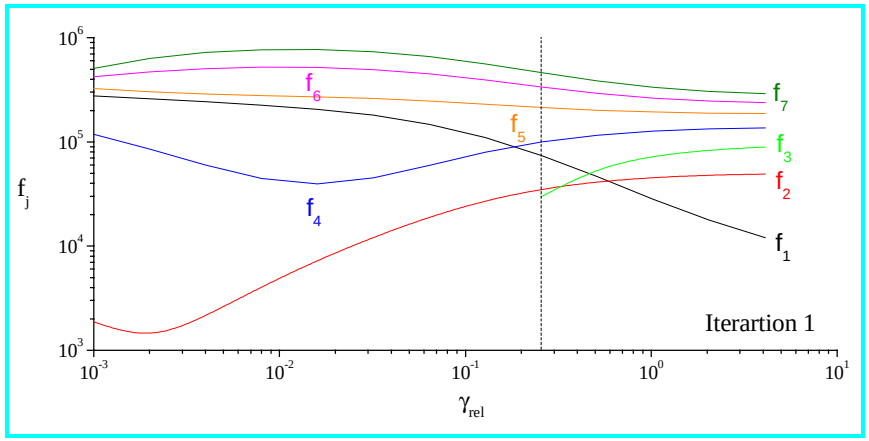

Figure A2.5. Dependency of components solution vectors on value of γ_{rel} during iteration number 1 in Test 2. Dotted vertical line denotes selected γ_{rel}-value

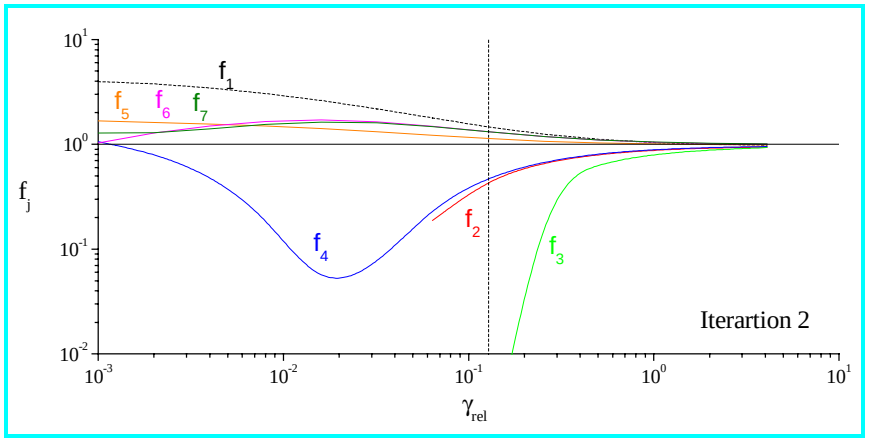

Figure A2.6. Dependency of components solution vectors on value of γ_{rel} during iteration number 2 in Test 2. Dotted vertical line denotes selected γ_{rel}-value. The horizontal black line outlines asymptotic unit value of all solution components.

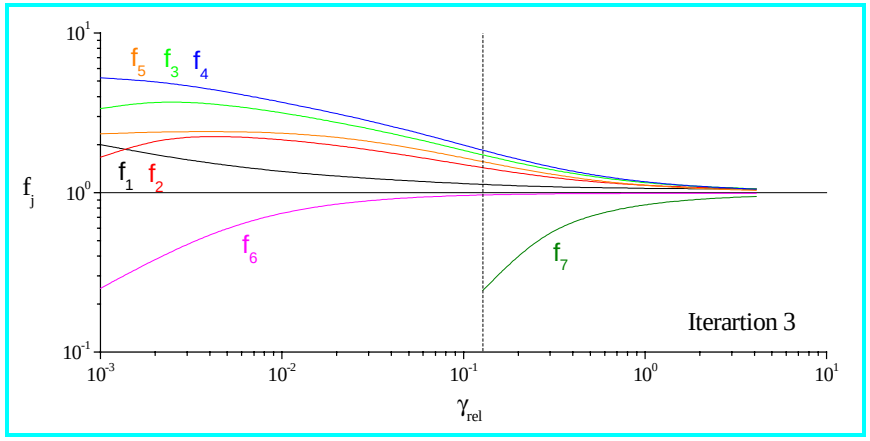

Figure A2.7. Dependency of components solution vectors on value of γ_{rel} during iteration number 3 in Test 2. Dotted vertical line denotes selected γ_{rel}-value. The horizontal black line outlines asymptotic unit value of all solution components.

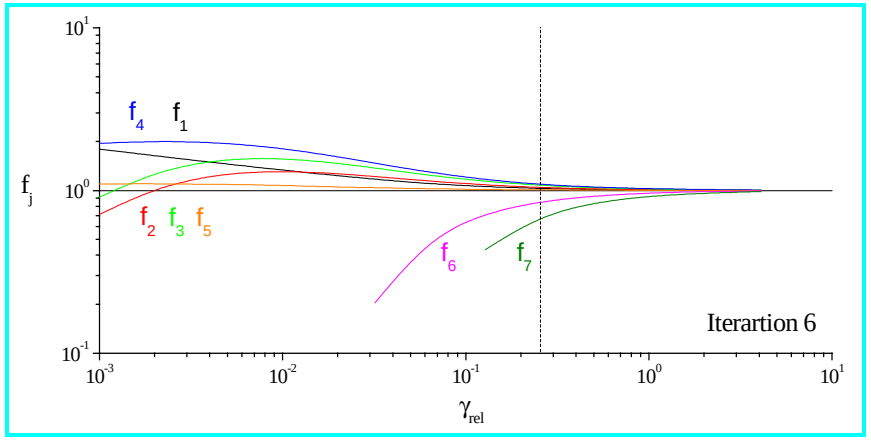

Figure A2.8. Dependency of components solution vectors on value of γ_{rel} during iteration number 6 in Test 2. Dotted vertical line denotes selected γ_{rel}-value. The horizontal black line outlines asymptotic unit value of all solution components.

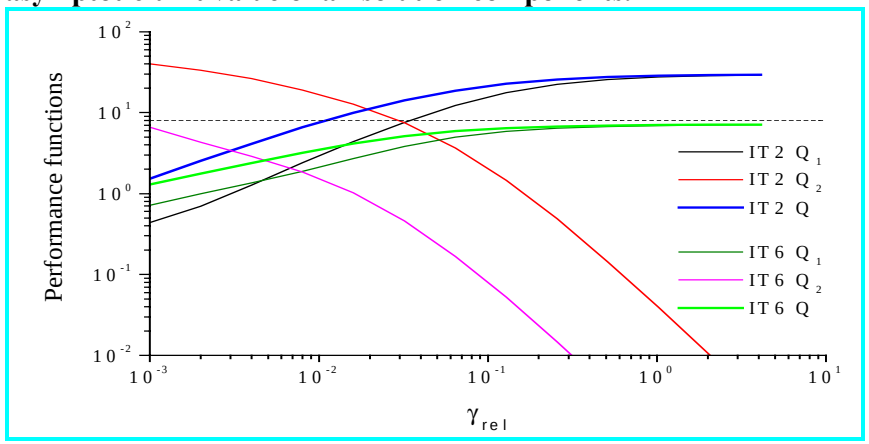

Figure A2.9. Dependency of performance functions terms on value of γ_{rel} during iterations number 2 and 6 in Test 2. Dotted vertical horizontal value denotes the limit value $Q_1 = 8$ of the first term.

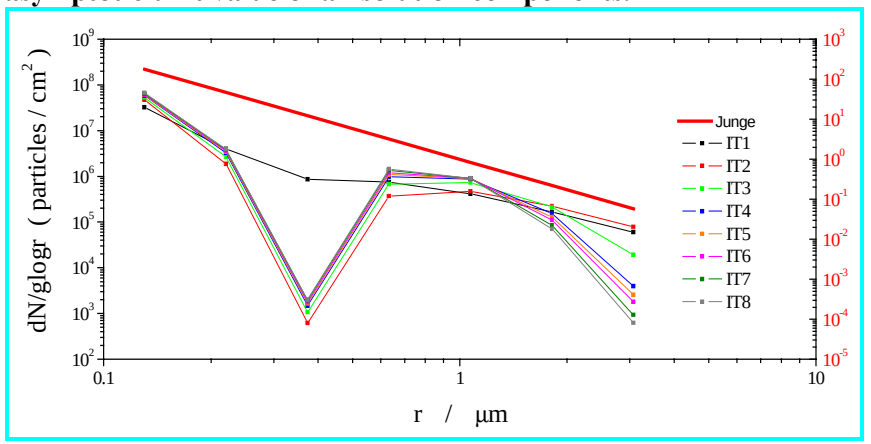

Figure A2.10. Modification of the columnar size distribution in all five performed iterations in Test 2. The red bold line is the initial guess, i.e. the Junge distribution. Right hand Y-axis relates to this quantity.

3. Example 2 – retrieval aerosol properties of Etna's plume data

The inspection of plots of $\log(\tau_M(\lambda))$ *vs.* $\log(\lambda)$ (see Figure A2.11) reveals the presence of at least two kinds of aerosols with possible bi-modal and monodisperse distributions because the background and first two groups have positive curvature and all other curves have small negative curvature. The retrieved columnar number distributions (see Figure A2.11) are quite interesting. The background aerosol ensemble has bi-modal distribution with a large amount of small particles, deep minimum at 0.3235 μm and maximum at 0.9892 μm. This distribution is practically the lowest at all radii. Distributions of groups 1 and 2 have similar but smoother shape – minimum at 0.3235 μm is much more shallow and about two orders more particles are present at largest radii. As a result the background aerosol and the aerosol causing AOTs in groups 1 and 2 are characterized with relatively high burdens (total columnar number of particles), caused mainly by smaller particles, but with small effective radius $r_{\text{eff}} = r_{\text{sm}}$. The ensemble surface weighted mean radius or the effective radius is a measure of the particles to cause extinction (see for example Watson and Oppenheimer (2000 and 2001)). Aerosol ensemble related with group 3 has a wide sprawling mode with maximum between 0.1850 and 0.3235 μm, shallow minimum at 0.9892 μm and probably a strong mode with radii more then 1 μm. This is the only one distribution, which reveals increase at radii above 1 μm. This ensemble is characterized with high burden and largest effective radius, that is responsible for the strongest extinction. Group 4 and scans 57, 63 and 68 are constituted by aerosols with quite similar properties. Distributions are composed by two monodisperse which could be fine ash and water droplets.

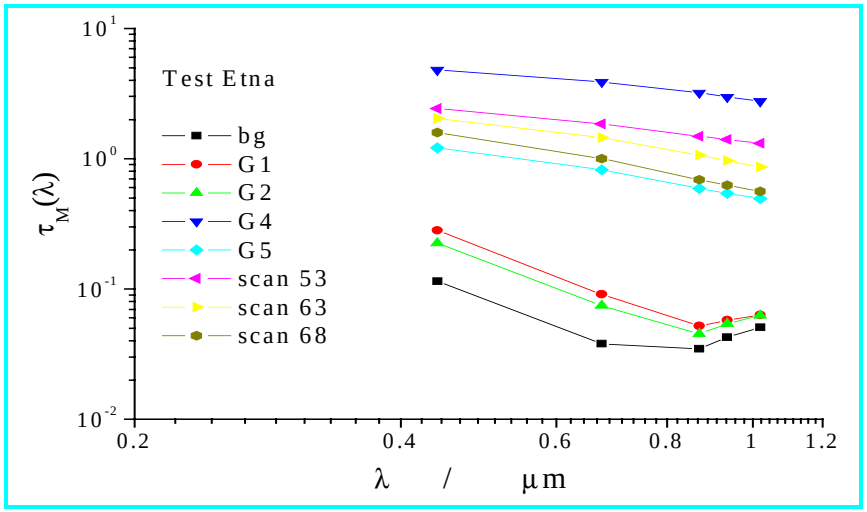

Figure A2.11. Measured AOTs in Test 2

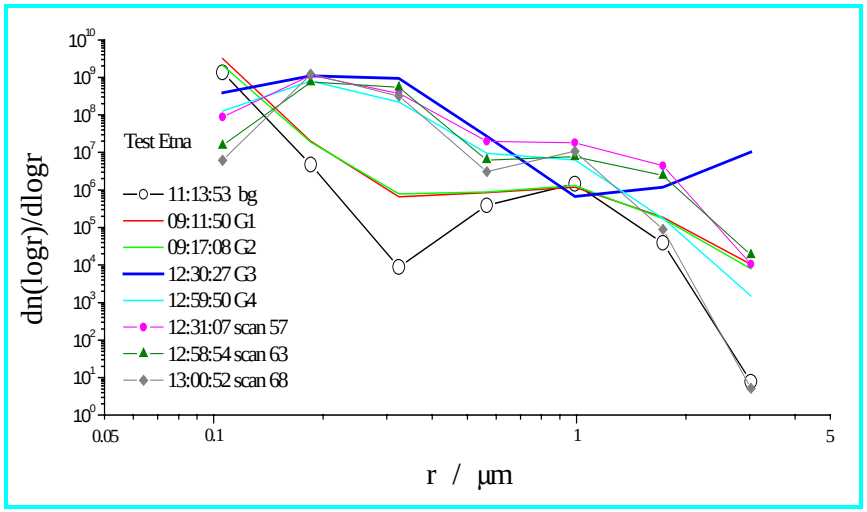

Figure A2.12. Retrieved columnar number distributions in Test 2

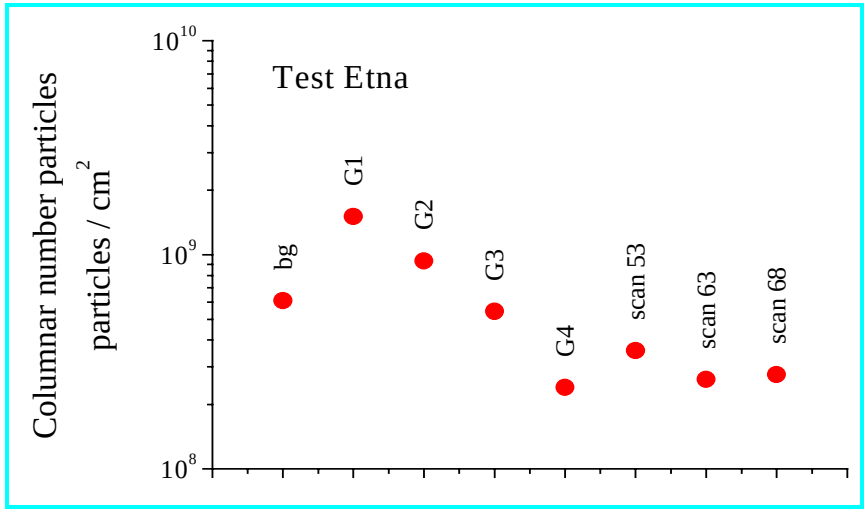

Figure A2.13. Retrieved columnar number particles in Test 2

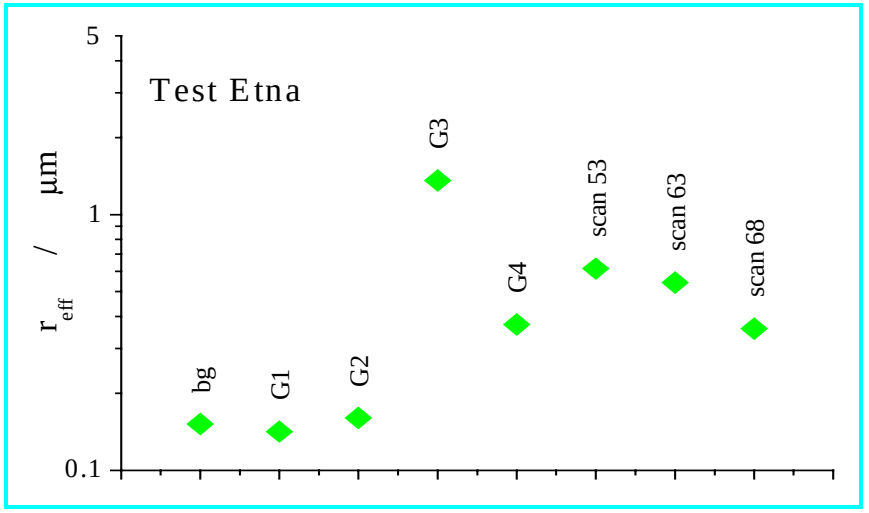

Figure A2.14. Retrieved effective radii in Test 2

References:

Bohren, C.F., and D.R. Huffman, 1983. Absorption and Scattering of Light by Small Particles, John Wiley, Hoboken, N. J.

Dubovik, O. and M.D. King, 2000. A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements, J. Geophys. Res., 105, 20673–20696.

Hallet, J., J.G. Hudson, and C.F. Rogers, 1989. Characterization of combustion aerosols for haze and cloud formation, Aerosol Sci. Technol., 10, 70-83

Holben, B.N., T.F. Eck, I. Slutsker, D. Tanre, J.P. Buis, A. Setzer, E. Vermote, J.A. Reagan, Y.J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov, 1998. AERONET - A Federated Instrument Network and Data Archive for Aerosol Characterization, Rem. Sens. Environ., Vol 66, 1–16.

Jorge, H.G. and J.A. Ogren, 1996. Sensitivity of Retrieved Aerosol Properties to Assumptions in the Inversion of Spectral Optical Depths., J. Atm. Sci., Vol. 53, No. 24, 3669–3683.

King, M.D., D.M. Byrne, B.M. Herman and J.A. Reagan, 1978. Aerosol Size Distribution Obtained by Inversion of Spectral Optical Depth Measurements. J. Atm. Science, Vol. 35, No. 11, 2153-2167.

King, M.D., 1982. Sensityvity of Constrained Linear Inversion to Selection of the Lagrange Multiplier. J. Atm. Science, 1982, Vol. 39, 1356-1369.

Lesins, G, P. Chylek, and U. Lohmann,2002. A study of internal and external mixing scenarios and its effect on aerosol optical properties and direct radiative forcing. J. Ggeophys. Res, 107, doi:10.1029/2001JD000973.

Mather, T.A., V.I. Tsanev, D.M. Pyle, A.J.S. McGonigle, C. Oppenheimer, and A.G. Allen, 2004. Characterization and evolution of tropospheric plumes from Lascar and Villarrica volcanoes, Chile. J. Geophys. Res., VOL. 109, D21303, doi:10.1029/2004JD004934.

Mather, T.A., R.G. Harrison, V.I. Tsanev, D.M. Pyle, M.L. Karumudi, A.J. Bennett, G.M. Sawyer and E.J. Highwood, 2007. Observations of the plume generated by the December 2005 oil depot explosions and prolonged fire at Buncefield (Hertfordshire, UK) and associated atmospheric changes. Proc. Roy. Soc. A, in press, doi: 10.1098/rspa.2006.1810.

Morys, M., F.M. Mims III, and S.E. Anderson, 2001. Design, calibration and performance of MICROTOPS II hand-held ozonometer. J. Geophys. Res., Vol. 106, No. D13, 14573-14582.

Porter, J.N., K.A. Horton, P.J. Mouginis-Mark, B. Lienert, S.K. Sharma, E. Lau, A.J. Sutton, T. Elias, and C. Oppenheimer, 2002. Sun Photometer and Lidar Measurements of the Plume from the Hawaii Kilauea Volcano Pu'u 'O'o Vent: Aerosol Flux and SO_2 Lifetime. Geophys. Res. Letters, Vol. 29, No. 16, 10.1029/2002GL014744.

Shaw, G.E, 1893, Sun Photometry, Bull. Amer. Meteorol. Soc., 1983, Vol. 64, No.1, 4-10.

Schmid, B., C. Maetzer, A.Meimo, and N. Kaempfer, 1997. Retrieval of Optical Dept and Particle Size Distribution of Tropospheric and Stratospheric Aerosols by Means of Sun Photometry. IEEE Trans. Geosci. and Rem. Sensing, Vol. 35, No. 1, 172-183

Seinfeld, J.H., and S.N. Pandis, 1998. Atmospheric Chemistry and Physics, John Wiley, Hoboken, N. J.

Watson, I.M. and C. Oppenheimer, 2000. Particle Size Distribution of Etna's Aerosol Plume Constarined by Sun Photometry. J. Geophys. Res., Vol. 105, No. D8, 9823-9829

Watson, I.M. and C. Oppenheimer, 2001. Photometric Observations of Mt. Etna's Different Aerosol Plumes. Atmospheric Environment, Vol. 35, No. 21, 3561-3572.