# NWPSAF 1D-Var User Manual

## Software Version 1.2

### NWPSAF-MO-UD-032

| Revision History | | | |
|---|---|---|---|
| Document revision | Date | Author | Notes |
| Version 1.1 | 06/12/16 | F. Smith | Draft Version valid for NWPSAF 1D-Var v1.1 |
| Version 1.1.1 | 20/06/18 | S. Havemann | v1.1.1: new release after LWP bugfix |
| Version 1.2 | 24/02/20 | S. Havemann | v1.2: Version valid for NWPSAF 1D-Var v.1.2 |

*The documentation was developed within the context of the EUMETSAT Satellite Application Facility on Numerical Weather Prediction (NWP SAF), under the Cooperation Agreement dated 7 December 2016, between EUMETSAT and the Met Office, UK, by one or more partners within the NWP SAF. The partners in the NWP SAF are the Met Office, ECMWF, DWD and Meteo France.*



**Contents**



**1. Introduction**

This document describes a stand-alone 1DVar code for nadir-viewing passive sounding satellite instruments that has been developed at the Met Office as part of the NWP SAF. The code contains the merged capabilities of the obsolete Met Office, ECMWF and SSMIS 1D-

Var schemes. It was originally developed for the IASI instrument but it may be used with many different sounding instruments with minimal changes.

The philosophy behind the development of this code was to produce a flexible, stand-alone 1DVar retrieval system that may be used for a wide variety of satellite instruments. However, it is not exhaustive. Extra subroutines could be added by the user to meet their own requirements, such as code to add features such as cloud detection, or code may be modified to include extra retrieval variables.

The program functions first by reading in user supplied background and observation data (together with appropriate error covariance matrices). The routines that are provided to do this assume data is supplied in a predefined ASCII format, but these routines can be replaced with ones more specific to the format of data that is to be used.

Routines for channel selection and cloud detection are included but bias correction needs to be performed by the user, or a bias correction subroutine added at the appropriate placeholder in the code.

Retrievals are made through the optimal estimation methods of Rodgers (1976). Three minimisation routines are provided. The one that is to be used depends on the size and linearity of the problem.

A radiative transfer program is required that will supply both radiances and Jacobians and this is accessed through a general purpose interface routine. Currently RTTOV Versions 11.3 and 12 are implemented.

The code is written in European Standard Fortran 90 and advanced F90 features are used as much as possible. The code was originally designed to closely follow the structures in the Met Office's Observation Processing System (OPS).

[ Return to Top ]

## 2. Compiling and Running the Code.

Instructions for compiling and testing the code can be found in the `readme.txt` file in the top 1DVar directory. Further notes can be found in the same place in file `NWPSAF_1DVar_vn1.2_Release_Information.txt`

### 2.1 Compiling

The supplied makefile is set up to compile with RTTOV-12 by default. To compile with RTTOV-11, simply change the version number in the makefile.

Modify the makefile to match your compilation of RTTOV; it is recommended to compile RTTOV with the HDF libraries, though it is not a requirement. NetCDF 4.2+ is also supported for RTTOV11. If you wish to read in emissivities from an atlas then you will also need to make a few changes to the makefile. See the readme.txt file for more information.

Due to some preprocessing directives (e.g. #include) present in the code, a C preprocessor is run during the compilation.

One further thing to note (which has always been the case, it is not a change in behaviour for this realease), is that the makefile includes the bare minimum of dependencies to make the code compile. If you change the code yourself, in particular the interfaces between subroutines, you will often find you need to run 'make clean' before recompiling. Also, the interfaces are not auto-generated, so you will need to make sure you modify the interface files in the 'include' directory to match.

### 2.2. Running the Code

For this release, the code should be run from a subdirectory. The directory `WorkDir` is provided for your convenience with some test and example scripts. This directory can be copied and/or renamed. The `NWPSAF_1DVar` executable should be linked into this directory. Namelist and data files will also need to be linked in. The example scripts show how this should be done.

To run the code successfully, relevant radiative transfer coefficients files must also be located in the working directory (WorkDir or a copy thereof). The coefficients files must be compatible with the radiative transfer model chosen and the instrument whose data is being used. Coefficients files can be downloaded for free from the NWPSAF RTTOV website. Please read through the readme.txt file for instructions on ensuring your coefficient files are named correctly.

[ Return to Top ]

## 3. Input Files



There are two main control files, `ControlData.NL` and `Retrieval.NL`, and two main data files, `ObsFile.dat` and `Background.dat`.

Auxiliary data such as error covariance matrices and channel selection choices are stored for each data type in a directory called `xxx_COEFFS_DIR` where `xxx` refers to the data type being processed (e.g., IASI_COEFFS_DIR).

All input files are ASCII and all values are in free-format.

### 3a. Control Files: ControlData.NL

`ControlData.NL` is a Fortran90 namelist file containing the input parameters required for the running of the code. The table below lists the parameters that may be specified through this file and which ones would normally be required. As one can see most of these parameters have a default value. This file and the NWPSAF_Read_ControlData that reads it may be easily expanded in future if one wants to add further options. Note that there are comments included in this file as supplied, but some Fortran90 compilers may not allow this, in which case they would need to be removed.

**Namelist Control for IASI_1DVar**

| **Variable** | **Required?** | **Default Value** | **Type** | **Notes** |
|---|---|---|---|---|
| FirstOb, LastOb | No | 1, Number of obs. | INTEGER | |
| GeneralMode | No | 10='ProductionMode' | INTEGER | Allowed values are 0='Operational', 10='Production', 20='Diagnostic', 30='Debug' and 40='Verbose'. Note that Operational and Production modes are the same in practice. |
| DetectCloud | No | True | LOGICAL | Note this is a cloud detection test designed for use with IR sounders. |
| CostThresh_Land CostThresh_Sea | Yes | -9999. | REAL | If DetectCloud is TRUE |
| CostThresh_IRWindow_Land CostThresh_IRWindow_Sea | Yes | -9999. | REAL | If DetectCloud is TRUE |
| CloudAbsThresh_IRWindow | Yes | -9999. | REAL | If DetectCloud is TRUE |
| HighCloudAbsThresh_IRWindow | Yes | -9999. | REAL | If DetectCloud is TRUE |
| Perform1DVar | No | True | LOGICAL | Purpose of running with false might be to test the setup or IO before running with true |
| Minimisation_Method | No | 1='Newtonian' | INTEGER | 0=None (sets Perform1DVar to false), 1=Newtonian, 2=Marquardt-Levenberg |
| MaxIterations | No | 10 | INTEGER | |
| DoTExtrapolation | No | TRUE | LOGICAL | If using a background profile where the top level |

| | | | | |
|---|---|---|---|---|
| | | | | is lower than 0.01hPa then setting this option to true means the profile is extrapolated to 0.01hPa |
| DeltaJ | No | 0.01 | REAL | Maximum fractional change in cost function for convergence. |
| SmallJCost_Gradient | No | 1. | REAL | Maximum value of cost function gradient (in terms of total cost) for convergence. Marquardt-Levenberg minimisation only. |
| Additional_Cost_Function | No | 0=None | INTEGER | 0=None, 1=CloudBoundaries For retrieval of cloud |
| Max_ML_Iterations | No | 10 | INTEGER | Marquardt-Levenberg minimisation only |
| Gamma_Factor | No | 10 | REAL | Factor by which Gamma is changed between iterations. Marquardt-Levenberg minimisation only |
| GammaMax | No | $10^{25}$ | REAL | Marquardt-Levenberg minimisation only |
| Allow_Eqn_101 | No | .TRUE. | LOGICAL | Set to false if the Eqn 101 minimisation is never to be used. |
| Force_Eqn_101 | No | .FALSE. | LOGICAL | Set to true to use Eqn 101 in all situations (Marquardt-Levenberg and additional cost function terms |

| | | | | |
|---|---|---|---|---|
| | | | | cannot be used in this case). |
| MaxChanUsed | No | 10000 | INTEGER | |
| Cloud_Min_Pressure | No | 100hPa | REAL | Minimum allowed pressure for retrieved cloud top in hPa. Retrieval of cloud only. |
| Use_EmisAtlas | No | .FALSE. | LOGICAL | Switch to control whether an emissivity atlas is used or not. Note that you must also have compiled with the switch in the makefile EmissAtlas=1 to use this setting. |
| Atlas_Dir | No | 'EmisAtlas' | CHARACTER | Directory containing the emissivity atlas |
| Read_CLW_Background | No | FALSE | LOGICAL | False = No clw profile in the Background file and True=clw profile in the Background file |
| Retrieve_qtotal | No | .FALSE. = LWP Retrieval | LOGICAL | False=LWP retrieval, True=Qtotal retrieval |
| EnahncedDiagnostics | No | .FALSE. | LOGICAL | Produces extra output files (see Output Files) |
| Gas_Units | No | 2 | INTEGER | Units for humidity profile input to RTTOV. 0 = ppmv w.r.t. dry air (compatibility with previous NWPSAF-1DVar releases). 2 = ppmv w.r.t. moist air. |
| Legacy_Settings | No | .FALSE. | LOGICAL | Sets RTTOV options interp_mode=1 and use_q2m=.false. for |

| | | | | |
|---|---|---|---|---|
| | | | | testing using old datasets. |

Two directory locations, which in previous releases were set in the ControlData.NL file, are now set via environment variables. These are:

- `coeffsdir`: now environment variable `COEFFS_DIR`
- `outputdir`: now environment variable `OUTPUT_DIR`

These variables should be set and exported in the shell (or running script) before calling the 1D-Var. Please see the file `WorkDir/Run_1DVar_test.ksh` to see how this should be done.

Example ControlData.NL files can be found in the Sample_Namelists directory. When used, they should be copied, or linked to rename them exactly "ControlData.NL". An example ControlData.NL file is:

```
! AIRS Control file for IASI_1DVar

! ******** THESE COMMENTS MAY NEED TO BE REMOVED IF COMPILING ***********
! ******** WITH F90 (rather than F95) ***********************************

! This controls the verbosity of the output 0=Minimal, 40=verbose
&Control  GeneralMode=40
! For the minimisation option, 1 is Newtonian, 2 is Marquardt-Levenberg
          Minimisation_Method=2
          MaxIterations=10
! These are cloud cost thresholds:
          CostThresh_Land=20.
          CostThresh_Sea=20.
          CostThresh_IRWindow_Land=0.
          CostThresh_IRWindow_Sea=0.
          CloudAbsThresh_IRWindow=-10.
          HighCloudAbsThresh_IRWindow=-55. /
```

[ Return to Top of Section ]

### 3b. Control Files: Retrievals.NL

`Retrieval.NL` controls the variables that are to be retrieved and provides the mapping between the minimization vector, the RT model vector and the B-matrix (Figure 1 gives and example of how the three vectors are mapped for a typical IASI retrieval).

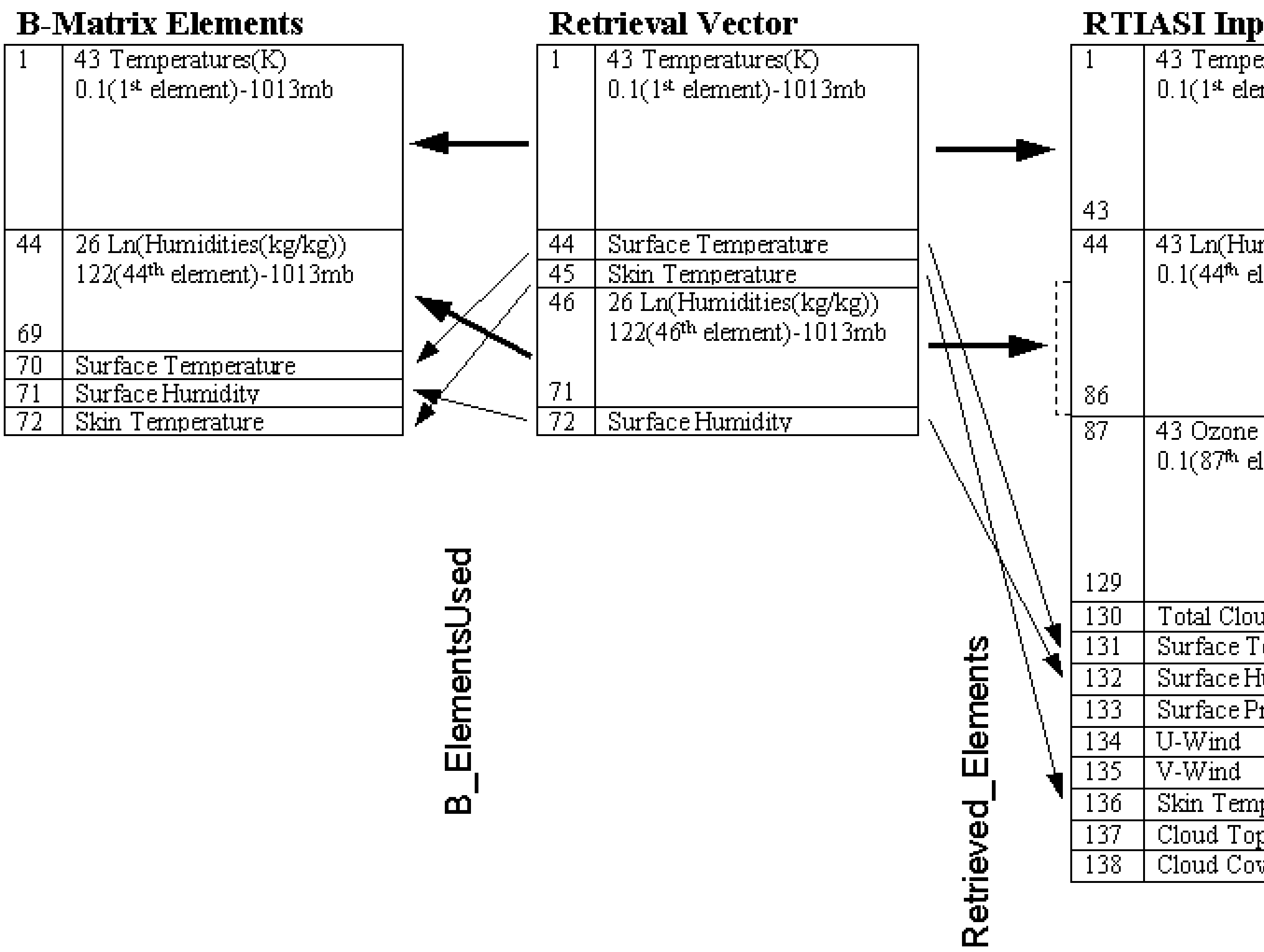


Fig. 1. Mapping of retrieval vector, RT vector and B-matrix

The `Retrieval.NL` file is a Fortran90 namelist file. The quantities to be retrieved are specified through the presence of two- or three-element matrices in this file. Absence of a quantity from this file means that this element is not to be retrieved.

Example Retrieval.NL files can be found in the Sample_Namelists directory and when copied or linked into the working directory should have any satellite instrument or number of levels removed from the file name and be called exactly "Retrieval.NL". A typical Retrieval.NL file for a 54 level retrieval is shown below. In addition there are sample Retrieval.NL files for 43, 54 and 70 level retrievals supplied with the package which are used for testing that your installation is working properly. All possible namelist entries are included here but, in practice, only those quantities that are actually being retrieved need to be included.

```
! Retrieval set up file for IASI
!
! ******** THESE COMMENTS MAY NEED TO BE REMOVED IF COMPILING ***********
! ******** WITH F90 (rather than F95) **********************************

 &Retrieval Temperature= 1,54,1
```

```
Humidity= 26,29,55
Surface_Temperature= 1,84
Surface_Humidity= 1,85
Surface_Pressure= 0,0
Skin_Temperature= 1,86
Cloud_Liquid_Water= 0,0
Cloud_Top_Pressure= 0,0
Cloud_Fraction= 0,0
Surface_UWind= 0,0
Surface_VWind= 0,0
Surface_Emissivity_PC= 1,0 /
```

The three-element lists are for profiles. The first two elements denote the top levels in the atmosphere and the number of levels that are to be retrieved respectively. The pressure profile of the atmosphere being defined in the `Background.dat` file. The final element is the position in the background error covariance matrix that corresponds to the top level to be retrieved. If any of these elements are zero, the profile is not retrieved.

In the example given, therefore, 54 levels of temperature are to be retrieved starting with the first (top) level. Only 26 humidity levels are to be retrieved starting at level 26 (and finishing with level 54). Ozone is not to be retrieved at all. In the background error covariance matrix, the first 54 elements are for temperature, while elements 55-84 correspond to the humidity on levels 26-54.

The two-element lists are for scalar (mostly surface) quantities. If the first element is non-zero the quantity is to be retrieved (provided the second entry is valid). The second entry points to the position in the background error covariance matrix for this quantity. For microwave cloud liquid water (or liquid water path (LWP)) retrievals, a covariance matrix not correlated to any other variables is assumed and this single value (0.2, inherited from the SSMIS 1D-Var code) is hardcoded into the routine NWPSAF_InitBmatrix.f90. Hence, if LWP is to be retrieved, while the first element has to be non-zero, the second element has to remain 0.

Cloud_Top_Pressure and Cloud_Fraction do not require a background error covariance entry as the default is to assume large background errors for these quantities if they are to be retrieved. The retrieval of these cloudy parameters uses some additional code to the normal retrieval. The methodology for this is documented in Appendix C.

In the example given, surface temperature, skin temperature and surface humidity are retrieved.

The last line triggers a retrieval of surface emissivity using principal components as retrieval vector state elements.

[ Return to Top of Section ]

### 3c. Data Files: Obsfile.dat

`Obsfile.dat` file contains the observation information. All data are specified in free format with mandatory colons delimiting the explanatory labels in the header.

The file starts with 10 lines that are reserved for users comments. Next comes header information which details the number of observations present in the file plus information on which channels have been used to make up a composite instrument (such as ATOVS).

There are are three optional header entries: "Units:", Composite Instruments:" and "Channels:".

"Units: " may be followed by one of "BT", "Radiance" or "PC Score". If no "Units: " line is found, the data will be assumed to be brightness temperature, so there should be no need to modify observation files used with previous versions of the 1D-Var software. Note that the "PC Score" option is only available for IASI or AIRS (only these are supported by PC-RTTOV at the present time) and can only be used in single instrument mode, not as part of a composite instrument.

"Composite Instruments:" is used where there is the possibility of more than one "composite instrument" (a composite instrument being a collection of one (e.g, IASI) or more (e.g., ATOVS) instruments that are treated as a single entity for retrieval purposes). The composite instrument entry will then specify the number of instruments followed on separate lines of text with the names of the composite instruments (these should be identical with the required entries in the RMatrix files). E.g. for ATOVS on NOAA satellites 15 and 16:

```
This is a test ATOVS observations file

This is based on a RTTOV simulated set of radiances from NOAA-15
which is computed from ADC's background profile surf emis=0
RWS 22nd March 2001
There are 10 header lines in total here!!!



So this is line 10.
Number of Observations in File:    10
No. of Chans per Observation:      40
Total Number of instruments making up observations : 6
*** In the following Series, Platform and Instrument are defined  ***
*** according to the relevant RT Model definitions (if required): ***
Units: BT
Composite Instruments: 2
NOAA-15 ATOVS
NOAA-16 ATOVS
Sat. Series   Platform   Instrument First_Channel   Last_Channel  Sat ID
    1           15          0            1                20       15
    1           15          3           21                35       15
    1           15          4           36                40       15
    1           16          0            1                20       16
    1           16          3           21                35       16
    1           16          4           36                40       16
Channels:
    1    2    3    4    5    6    7    8    9   10   11   12   13   14   15
16
   17   18   19   20   21   22   23   24   25   26   27   28   29   30   31
32
   33   34   35   36   37   38   39   40
----------------------------------------------------------------------
Obs ID:              1 Obs Type:          2 Satellite ID:    15
```

```
Latitude:       0.000 Longitude:     0.000 Elevation:        0.0
Surface Type:        1 Sat Zen Angle:   0.000 Solar Zen Ang:   38.000
Brightness Temperatures:
      226.717      218.949      219.645      224.452      236.374
247.804
      257.609      274.562      247.897      269.388      248.331
226.588
      261.211      252.591      241.597      236.591      266.911
271.983
      271.537     -9999.00      152.095      152.105      209.206
245.611
      244.334      233.650      226.083      220.954      216.761
215.937
      217.657      221.861      233.096      245.561      200.407
206.016
      235.261      242.110      252.165      261.941
Obs ID:              2 Obs Type:          2 Satellite ID:    16
Latitude:       0.000 Longitude:     0.000 Elevation:        0.0
Surface Type:        1 Sat Zen Angle:   0.000 Solar Zen Ang:   38.000
Brightness Temperatures:
      225.552      219.724      219.573      226.521      237.666
246.276
      255.913      267.355      250.477      259.886      251.149
229.323
      260.399      248.081      240.887      237.802      264.660
266.901
      266.492     -9999.00      152.369      156.488      217.802
243.712
      243.308      234.084      225.322      220.200      217.276
217.978
```

For the case where the "Composite Instruments" option isn't specified (e.g., for IASI) the section between "Number of Observations" and "-----------" above should look something like this (noting that the "Channels:" section is optional.):

```
Number of Observations in File :    980
No. of Chans per Observation:    8461
Number of instruments making up observations : 1
*** In the following Series, Platform and Instrument are defined  ***
*** according to the relevant RT Model definitions (if required): ***
Units: BT
Sat. Series    Platform    Instrument First_Channel   Last_Channel  Sat ID
    1             1           1            1              8461        300
----------------------------------------------------------------------
```

The "Sat ID" is used to keep track of the instruments where more than one is being used at a time (it was previously called the "WMO No." but this was misleading).

The "Channels" option is used for those cases where only a subset of available instrument channels are required. After this the instrument channels that are actually being used are listed. This is purely to allow the correct channels to be set up in the RT code and in all other places (except the R-matrix - see below) the *n* input channels are referred to as channels 1 ... *n*.

After the headers lines, each observation is listed, prefixed by a sub-header detailing auxiliary data for the observation in question such as observation number, latitude, longitude, satellite view angle, etc.

**Observation sub-headers in ObsFile.dat**

| **Entry** | **Notes** | **Type** | **Variable set in IASI_Read_Observations** |
|---|---|---|---|
| *First sub-header line:* | | | |
| Obs ID | Integer | Number identifying the current observation (usually 1-Number of Obs.) | `Observations % ID` |
| Obs Type | Integer | Obsolete but retained for backward compatibility | - |
| Satellite ID | Integer | Specifies the satellite to be used (ref. "Sat ID" in main header) | `Observations % SatID` |
| *Second sub-header line (Optional):* | | | |
| Year | Integer | Year of Observation (YYYY) | `Observations % Date(1)` |
| Month | Integer | Month of Observation (MM). | `Observations % Date(2)` |
| Day | Integer | Day of Observation (DD) | `Observations % Date(3)` |
| *Third sub-header line:* | | | |
| Latitude | Real | Latitude of Observations (°N) | `Observations % Latitude % Value` |
| Longitude | Real | Longitude of Observations (°E). | `Observations % Longitude % Value` |
| Elevation | Real | Height of surface (metres above sea level). (Not currently used) | `Observations % Elevation` |
| *Fourth sub-header line:* | | | |
| Surface Type | Integer | Type of surface. 1=Sea; 2=SeaIce; 3=Land; 4=Highland; 5=Mismatch. For types 3-5, land is assumed by the radiative transfer model. | `Observations % Surface` |
| Sat Zen Angle | Real | Satellite Zenith Angle at surface (°). (This was in previous versions erroneously labelled Sat View Angle). | `Observations % SatZenith` |
| Solar Zen. Ang. | Real | Solar Zenith Angle at surface (°) | `Observations % SolarZenith` |

Please note once more that the colons in the header and sub-header lines are used by the program as delimiters and must be included. Please also note that the second (optional) sub-header line is required when using an emissivity atlas. In all other situations it may be omitted. Following the sub-header lines, the observed brightness temperatures are listed (free-format).

This file and the subroutine that reads it may be replaced by one in a format defined by the user.

[ Return to Top of Section ]

### 3d. Data Files: Background.dat

The `Background.dat` file starts with a 10 header lines that are reserved for user's comments.

The next three lines are general information for the file:
The first of these lines is the number of profiles contained in the file. This may either be unity (in which case the same background is used for all observations) or be the same as the number of observations (i.e., one background per observation).
The second line is the number of levels for each profile. In the current implementation this should be set to 43.
The third and final line defines the choice for the units for water vapour abundance:

| **Value** | **Definition** |
|---|---|
| 1 | volume mixing ratio in ppmv (see more info below) |
| 2 | mass mixing ratio in kg/kg (see more info below) |
| 3 | relative humidity |

Where the humidity units = 1, in previous releases, volume mixing ratio and mass mixing ratio were specified relative to dry air. The conversion factor used between the two is `q_mixratio_to_ppmv = 1.60771704e+6`, which used to be specified in RTTOV but is no longer used there. In fact, it is as likely that users have been providing ppmv w.r.t. moist air, or specific humidity (kg/kg w.r.t. moist air as opposed to water vapour mass mixing ratio w.r.t. dry air) anyway. Indeed, the standard field available from the Met Office model is specific humidity, and the sample 1D-Var B-matrices calculated at the Met Office are also w.r.t. moist air. Furthermore, the NWPSAF Diverse Profile Set provides specific humidity. The default for RTTOV is now ppmv w.r.t moist air (gas_units=2), and so this option has been set by default for the 1D-Var code to match. This renders the unit conversion between kg/kg and ppmv incorrect, although it will probably make little difference to the retrieved profile. In a future release the conversion will be corrected. To switch to an input of gas_units=0 (ppmv w.r.t. dry air), as for the previous release, please set the namelist variable Gas_Units in the file Control.NL.

The minimisation itself is performed using a quantity specified as the natural logarithm of q in (kg/kg), Ln(q). In previous releases of the 1D-Var it was stated that the minimisation was in ln(mass mixing ratio), but as stated above the sample B-matrices are actually ln(specific humidity) so there has been an inconsistency that is not easy to resolve without changing the unit conversion, a task resolved for a future release of the 1D-Var.

Each profile is then listed with three comment lines preceeding. The profiles are presented in four columns in free-format (pressure(Pa or hPa ), temperature in Kelvin, Water Vapour in the predefined units as discussed above, and Ozone in ppmv with respect to dry or moist air, as above). Finally surface information is provided (each entry being preceeded by a label ending with a colon). The surface parameters provided, in order, are: Surface Temperature (K), Surface Humidity (in the pre-defined units), Skin Temperature(K), Surface Pressure (Pa or hPa to match profile), 10m U-Wind (m/s), 10m U-Wind (m/s).

If microwave cloud liquid water retrievals are to be performed (either LWP retrieval or qtotal retrieval), a fifth column containing cloud liquid water profiles (in kg/kg) has to be provided.

Note that the code can now read a profile file on levels expressed either in Pa or hPa, and they can run either top-of-atmosphere to surface, or surface to top-of-atmosphere.

[ Return to Top of Section ]

**Auxiliary Files:**

Auxiliary data found in the `xxx_COEFFS_DIR` and `Sample_Bmatrices` directories include the observation error covariance matrix (in `Rmatrix_orig`), four background error covariance matrices (in `Bmatrix_43L, Bmatrix_51L, Bmatrix_54L, & Bmatrix_70L`), and the channel selection file (`Channelchoice_orig.dat`). "`xxx`" refers to the observation type which is currently IASI, ATOVS, ATMS, CrIS, SSMIS or AIRS.

Additional directories are included for use with the 70L test profile to check set-up for the use of data in radiance and principal component score units.

**3e. Auxiliary Files: Rmatrix**

The `Rmatrix` file contains the observational error covariance matrix used by the 1DVar scheme. Currently one error covariance is used for all observations being processed (for a given instrument, of course) with the values being given in brightness temperature. There is no correction for scene temperature in the code, so the R-matrix values should be appropriate for the expected scene temperature of the channel in question. The forward model error is included in the observational error. This matrix must be positive definite or a fatal error will result. Sample R-matrices can be found in the xxx_COEFFS_DIR directories, but it is strongly recommended that you consider the contents of these files carefully for your application.

Three options are available for storing the observational errors. These are full matrix, band-diagonal (of which diagonal is a subset), and eigenvalue/eigenvector. In the case of a diagonal R-matrix, the values in the file represent the variances of the observation errors. The band-diagonal option has been included as one way of representing the errors of apodised interferometer observations, while the eigenvalue format is in anticipation of more sophisticated ways of representing the errors of instruments with many channels and has not been tested as yet.

The file is formatted as follows (all entries are in free-format):

The first line describes the instrument (or "Composite Instrument") for which the file is valid. If "Composite Instruments" were set up in the observation file, the names in the observation and R-matrix files should agree. If composite instruments were not set up, the first R-matrix is read in.

The second line is made up of four integers which are (in order):

- The "matrix type", i.e, the form in which the matrix is stored: 1=Full matrix, 2=band diagonal (of which diagonal is a subset) and 3=eigenvalue/eigenvector representation.
- The number of channels in the R-matrix
- The number of "elements": For the full matrix, this is the number of channels once more; for the band diagonal matrix this is the number of bands (1=diagonal, 2=tri-

diagonal, etc.); for the eigenvalue/eigenvector representation this is the number of eigenvectors used.

- "Inverse". If this fourth entry is one, the matrix is stored as an inverse. This would only be useful for a full matrix where the number of channels to be used (in retrievals and cloud detection) is fixed.

Next, before the errors themselves, the channels used are listed. This is the only place apart from the observation file where the absolute instrument channel numbers are used. This is done to ensure that the observation error covariances are indeed for the channels that are present in the observation file. See the example below for ATOVS where channels 1-20 for HIRS, 1-15 for AMSU-A, and 1-5 for AMSU-B are lsited.

Finally, the R-matrix itself is read in.

- In the full matrix case, the full matrix is simply placed in the file here.
- In the band diagonal case, the matrix is read in starting with the diagonal elements and then progressing through the bands further and further from the diaginal. Padding zeros are added to the ends of each off-diagonal band, so the entry for each band is as long as the number of channels. Of course, for the diagonal case only one band is required.
- In the eigenvalue/eigenvector case, the required number of eigenvectors are read from the file and then a vector of the associated eigenvalues are read in.

An example Rmatrix file is shown here:

```
NOAA-15 ATOVS
    2    40      1      0
       1        2        3        4        5        6        7        8        9
      10       11       12       13       14       15       16       17       18
      19       20        1        2        3        4        5        6        7
       8        9       10       11       12       13       14       15        1
       2        3        4        5
      1.22 0.5  0.5  0.5  0.5  0.5  1.22 6.32 1.4  6.32
      3.0  4.0  1.22 1.22 0.5  0.4  0.4  0.4  0.4  9.9
      2.0  2.0  2.0  1.26 0.25 0.25 0.25 0.25 0.4  0.4
      0.5  0.95 1.22 1.22 3.0  8.0  5.0  4.0  4.0  4.0
```

In this example the matrix type is band-diagonal with 40 channels and 1 band (so it is a diagonal matrix). The matrix is not already inverted. The R-matrix for HIRS 1-20, AMSU-A 1-15 and AMSU-B 1-5 is set up.
[ Return to Top of Section ]

### 3f. Auxiliary Files: Bmatrix

The `Bmatrix` file should contain the background error covariance matrices used by the 1DVar scheme. The full matrix is specified. A retrieval may be run on any number of levels with RTTOV making use of the vertical interpolation functionality within the RT model itself. This will happen automatically if you provide a profile on a number of levels different from the number of coefficient levels. The provided sample Bmatrix files in the directory `Sample_Bmatrices` (Bmatrix_43L, Bmatrix_51L, Bmatrix_54L, Bmatrix_70L) should be considered as example files. It is strongly recommended that you provide your own matrix to suit your own application. In particular, if your profile is not already provided on RTTOV levels, you will need to supply your own `Bmatrix` file on the required number of levels.

In this implementation, there are two B-Matrices specified in each file - one for land and one for sea - but there is no other variability in this matrix allowed for (e.g., no latitudinal variation).

The file itself simply consists of two comment lines followed by the dimension of the matrix (integer). The next lines are the B-Matrix in free format with one line per row/column (the matrix is symmetric).

For the second, land, matrix the same format is repeated (i.e., two comment lines, the number of elements and the matrix itself).

[ Return to Top of Section ]

### 3g. Auxiliary Files: ChannelChoice.dat

The `ChannelChoice.dat` file contains the choice of channels to be processed. The first line in the file contains *Num_Channels*, the number of channels for which selection codes (for cloud detection, retrieval, monitoring) are specified. The following *Num_Channels* lines in the file contain three columns of integers (plus a fourth, optional, column which can include characters). These columns are:

- channel number -- the index of the selected channel in the channel set specified in the observation file (if a reduced (pre-selected) set of *n* channels is present in the observation file, channel numbers specified in the `ChannelChoice_orig.dat` file correspond to indices (in the range 1 ... *n*) for the reduced channel set).
- a code for the situations in which this channel is used for retrievals
- a code indicating whether this channel is to be used for cloud detection and/or background monitoring.
- A fourth may be used to indicate the true channel numbers for those situations where the observation file contains a reduced set of channels. This is for informational purposes only and is not used by the program.

The codes employed in this file are as follows:

2nd column: **Retrievals Code**. The situations in which the channel is used are determined by the value of the integer in this column. The code is based on which bits are one in the binary representation of the integer. Bits for surface type and cloud situation are coded as follows:

```
Bit Number          Sounding Option
1                   Surface = Sea
2                   Surface = SeaIce
3                   Surface = Land
4                   Surface = Highland
5                   Surface = Surface Mismatch
6                   Cloud = Clear
7                   Cloud = IRCloudy
8                   Cloud = MWCloudy *
9                   Cloud = Rain *
10                  Cloud = High Cloud
```

So, for example, if the code is 1023 (=1111111111 in binary) the channel is used in all cases. If the code is 33 (=100001) then the channel is to be used only for clear skies over sea.

* The cloud codes corresponding to bits 8 & 9 are not used at present, but are reserved for microwave cloudy and rain situations.

3rd column: **Monitoring and Cloud Detection Code**. If the absolute integer value in this column is set to 1, the channel is to be monitored (i.e., the radiances for this channel are to be calculated for the background profile and compared to the observations. If the column value is 3, the channel is used for monitoring and cloud detection. If the value is 2, the channel is used in cloud detection but not monitoring. All channels that are to be used in retrievals are monitored by default.

If the value of the Monitoring and Cloud Detection Code is negative, the same codes as above apply to the absolute value plus the channel is the designated window channel (if more than one window channel is specified, the last one is used). The window channel is used in the window channel test where the field of view is designated as cloudy if the difference between the observed and background brightness temperatures for the window channel exceeds the value of `CostThresh_IRWindow_sea/land` which is set up through the ControlData.NL file.

In the following example file the selection codes for 11 channels are specified. HIRS 5-12 are used for retrievals in clear cases and AMSU-A 3-5 in all cases. HIRS-8 is also used as the window channel:

```
    11
           5     33     3   HIRS-5
           6     33     3   HIRS-6
           7     33     3   HIRS-7
           8     33    -3   HIRS-8
           9     33     3   HIRS-9
          10     33     3   HIRS-10
          11     33     3   HIRS-11
          12     33     3   HIRS-12
          23   1023     3   AMSU-3
          24   1023     3   AMSU-4
          25   1023     3   AMSU-5
```

[ Return to Top of Section ]

### 3h. Auxiliary File: EmisEigenVec

The `EmisEigenVec` file contains the surface emissivity eigenvectors required for the surface emissivity retrieval in principal component space. At the time of the release of version 1.2 such a file is only provided for IASI and uses the Collard (2007) selection of 314 channels. The file contains the leading 146 eigenvectors. The retrievals only use the first 12 eigenvectors. This achieves good efficiency without sacrifying accuracy.

[ Return to Top of Section ]

### 3i. Auxiliary File: EmisPCAtlas

The EmisPCAtlas file contains an atlas of principal component scores (i.e. weights) with is to be used together with the EmisEigenVec described in the previous subsection. The principal component scores in the atlas are stored on a latitude and longitude grid and there is one such set for each month. The specification of latitude, longitude and month by the user in the is used to select the appropriate principal component scores from the atlas.

[ Return to Top of Section ]

[ Return to Top]

**Output Files:**



There are two main output files, Retrieved_BTs.dat and Retrieved_Profiles.dat. These are ASCII files and their contents should be obvious from inspection. In addition, a file ProfileQC.dat is now provided with one line per input observation declaring whether the observation was processed or not and whether the 1D-Var converged.

Additional diagnostic files are produced if the program is run in DebugMode or higher (set GeneralMode to 30 or more in ControlData.NL). These files are Minimisation.log, and Minimisation_BT.log.

Finally, if the namelist option EnhancedDiagonstics=.true. at the higher verbosity levels (DebugMode or higher), then further output files are produced: A-Matrix.out, Am-Matrix.out, RetJacobian.out, BgJacobian.out, AveragingKernel.out

**4a. Output: Retrieved_Profiles.dat**

The Retrieved_Profiles.dat file contains the profiles retrieved from the 1DVar stage. If the 1DVar retrieval is not performed (e.g., the field of view is cloudy but only clear retrievals are allowed), the profiles are not written to this file; if the retrieval is performed but does not converge the final retrieved profile is written to this file, however (the number of iterations will be one higher than MaxIterations in this case).

The file is written in ASCII and should be self-explanatory. Both background and retrieved fields are supplied. The retrieved fields that are output are: Temperature Profile, Humidity Profile (in the same units as set up in the Background.dat file), Ozone profile, surface temperature, surface humidity (in same units as in the background file), skin temperature and surface pressure. If a Cloudy Retrieval is being done, cloud top pressure and cloud fraction are also output. If a microwave LWP retreival is being done, that will also be output. In addition the observation number, number of iterations, normalised cost function and normalised cost function gradient are reported.

An abridged example of the contents of this file is:

```
Observation =  1
                          Retrieval                                 Background
  Pressure (hPa) T (K)     q (ppmv)     Ozone         T (K)    q (ppmv)
Ozone
     1013.25   272.132  0.4296E+04  0.2456E-01   272.070  0.4280E+04
0.2456E-01
     1005.43   271.952  0.4227E+04  0.2507E-01   271.854  0.4220E+04
0.2507E-01
      985.88   271.509  0.4061E+04  0.2643E-01   271.311  0.4074E+04
0.2643E-01
          .         .         .          .          .          .          .
          .         .         .          .          .          .          .
          .         .         .          .          .          .          .
        0.69   265.252  0.5636E+01  0.6042E+01   265.741  0.5636E+01
0.6042E+01
        0.29   256.937  0.5410E+01  0.5971E+01   257.553  0.5410E+01
0.5971E+01
        0.10   241.249  0.4361E+01  0.5792E+01   241.635  0.4361E+01
0.5792E+01
Surface Temperature (K):          272.935   272.070
Surface Humidity (ppmv):         5202.160  4279.650
Skin Temperature (K):             272.935   272.070
Surface Pressure (Pa):            101325.  101325.
No. of Iterations:                2
Normalised Cost Function:       0.654 Normalised Gradient:       0.001
 ---------------------------------------------------------------
Observation =  2
                           Retrieval                                 Background
  Pressure (hPa) T (K)     q (ppmv)     Ozone         T (K)    q (ppmv)
Ozone
     1013.25   271.698  0.4169E+04  0.2456E-01   272.070  0.4280E+04
0.2456E-01
     1005.43   271.480  0.4110E+04  0.2507E-01   271.854  0.4220E+04
0.2507E-01
      985.88   271.008  0.3965E+04  0.2643E-01   271.311  0.4074E+04
0.2643E-01
          .         .         .          .          .          .          .
          .         .         .          .          .          .          .
          .         .         .          .          .          .          .
        2.61   242.475  0.4945E+01  0.6106E+01   242.500  0.4945E+01
0.6106E+01
        1.42   256.285  0.5429E+01  0.6080E+01   256.212  0.5429E+01
0.6080E+01
        0.69   265.855  0.5636E+01  0.6042E+01   265.741  0.5636E+01
0.6042E+01
        0.29   257.713  0.5410E+01  0.5971E+01   257.553  0.5410E+01
0.5971E+01
        0.10   241.273  0.4361E+01  0.5792E+01   241.635  0.4361E+01
0.5792E+01
Surface Temperature (K):          268.230   272.070
Surface Humidity (ppmv):         4559.714  4279.650
Skin Temperature (K):             268.230   272.070
Surface Pressure (Pa):            101325.  101325.
No. of Iterations:                2
Normalised Cost Function:       0.782 Normalised Gradient:       0.000
 ---------------------------------------------------------------
```

[ Return to Top of Section ]

### 4b. Output: Retrieved_BTs.dat

The `Retrieved_BTs.dat` file contains the brightness retrieved from the 1DVar stage together with the observed brightness temperatures and those calculated from the background profile. If the 1DVar retrieval is not performed (e.g., the field of view is cloudy but only clear retrievals are allowed), the brightness temperatures are not written to this file; if the retrieval is performed but does not converge the final retrieved brightness temperatures are written to this file, however.

Each observation has three headers: the observation number, the number of channels used in the retrieval and the column titles. Four columns are then reported: the channel number, the background brightness temperature, the observed brightness temperature and the retrieved brightness temperature. The channel numbers reported are those specified in the first column of the `ChannelChoice.dat` file. Only those channels used in the retrieval are reported.

The following is a typical `Retrieved_BTs.dat` produced from a retrieval using HIRS 1-19 and AMSU-13:

```
Observation =   1
 Number of Channels Used =   20
 Channel   Background  Observed    Retrieved
      1     226.579    226.717    226.167
      2     219.569    218.949    219.310
      3     219.321    219.645    219.153
      4     225.255    224.452    225.000
      5     237.027    236.374    236.813
      6     247.197    247.804    247.077
      7     256.419    257.609    256.574
      8     271.211    274.562    271.996
      9     250.658    247.897    251.042
     10     269.220    269.388    269.773
     11     251.843    248.331    250.209
     12     232.814    226.588    230.007
     13     261.277    261.211    261.721
     14     250.869    252.591    250.952
     15     241.685    241.597    241.568
     16     237.392    236.591    237.163
     17     266.360    266.911    266.937
     18     270.537    271.983    271.343
     19     271.088    271.537    271.877
     32     223.373    221.861    222.736
```

[ Return to Top of Section ]

### 4c. Additional Output: ProfileQC.dat, Minimisation.log and Minimisation_BT.log

These files are output when the program is run in DebugMode or VerboseMode.

`ProfileQC.dat` contains a single line per profile, consisting of the observation number and a code, which is as follows:

```
   0 = successful minimisation
   1 = 1D-Var did not reach convergence.
   2 = observation not processed (bad BT, failure in RTTOV etc)
```

The other files are used primarily to aid in debugging by tracking the iteration-to-iteration progress of the minimisation.

`Minimisation.log` outputs the elements of the retrieval vector (`RT_Params%RTGuess(:)` in `IASI_Minimize`) together with the Marquardt-Levenberg gamma factor (zero if not using Marquardt-Levenberg minimisation) and the (un-normalised) cost function for each iteration. The exact makeup and order of the retrieval vector will vary according to the variables that are to be retrieved, but Fig. 1 gives the makeup of this vector in the default set-up supplied.

`Minimisation_BT.log` outputs the observed-guess brightness temperature differences for the channels used in the retrieval together with the Marquardt-Levenberg gamma factor (zero if not using Marquardt-Levenberg minimisation) for each iteration.

[ Return to Top of Section ]

### 4d. Enhanced diagnostics: A-Matrix, Am-Matrix, Jacobians and Averaging Kernels

The analysis error covariance matrix and the propagated measurement noise matrix are output to `A-Matrix.out` and `Am-Matrix.out` respectively. These are the expected retrieval error and the expected contribution to the retrieval error from observation noise respectively as calculated through linear retrieval theory. They are output in the same order as the elements in the input B-matrix and are output for each iteration. The A-matrix is also calculated from the background profile for the first observation only. The matrices for the converged profiles are only output if EnhancedDiagnostics=.true., but this background A-matrix is output regardless of that setting.

For a full discussion of the derivation of the propagated measurement error matrix see Rodgers (1990) or Rodgers (2000) where it is referred to as the measurement error covariance, $\mathbf{S_M}$.

Two files containing Jacobians will also be produced: `BgJacobian.out` contains the Jacobian corresponding to the profile elements selected in the Bmatrix calculated from the background profile, and `RetJacobian.out` is output at the end of the minimisation. The Jacobians are output using the format statement '(10E12.4)' and the variable written out is sized (NumChans,NumElements), where NumChans refers to the number of channels used for that ob in the 1D-Var (which may be affected by cloud conditions) and NumElements is determined by the Bmatrix. NumChans is written out in `Retrieved_BTs.dat` as described above.

The final matrix written out is the Averaging Kernel matrix (in `AveragingKernel.out`) according to the formula presented in Chapter 2 of Rodgers (2000). The matrix is written out using format '(10E12.4)' and is of size (NumElements, NumElements).

[ Return to Top of Section ]

[ Return to Top ]

### 5. Setting Up a New Observation Type

To set up a new instrument requires the creation of new data and auxiliary files but should be relatively straightforward unless unsupported retrieval variables and/or radiative transfer models are required in which case code changes will be needed.

If a fastmodel other than RTTOV is required, the interface with the new fastmodel will need to be coded up and called from `NWPSAF_Fastmodel_Interface.f90`. Extra pre-processing codes will likely be required and the makefile changed accordingly.

Additional retrieval variables will require changes to the fastmodel interfaces, `NWPSAFMod_RTModel`, `NWPSAF_Read_Background`, `NWPSAF_SetUpBackground`, `NWPSAF_SetUpRetrievals`, `NWPSAF_TranslateDataOut` and probably `NWPSAF_CheckIteration`. Additional variables will of course need to be added to the Bmatrix.

For questions on changing the code, please submit a request to the NWPSAF helpdesk via the website.

[ Return to Top ]

### 6. Simulating observations

One important use of the NWPSAF-1DVAR code is to test proposed changes to assimilation or retrieval systems in a simple simulated environment. In order to do this, one may use model profiles as "truth", and simulate observations from these true profiles, adding noise to the simulated "true observations" to match the Rmatrix used in the 1D-Var. To facilitate this process, code has been provided in the directory src/sim_spec to simulate observations from RTTOV-11 or -12, to output observation files of the correct format for the NWPSAF-1DVar. The code can simulate brightness temperatures, radiances, and for IASI and AIRS, PC scores or reconstructed radiances. Routines are also provided to convert existing brightness temperature or radiance data into PC Scores and back into Reconstructed Spectra. Please see Appendix E for information on how to run this code.

[ Return to Top ]

### 7. Use of PC Scores

This release of the 1D-Var allows observations to be processed in the form of principal component scores. This option is currently only available for AIRS and IASI, as PC-RTTOV only supports these instruments. To pass in PC-Score observations, the observation file must contain a line

```
Units: PC Score
```

PC-RTTOV can only be used in single instrument mode, not as part of a composite instrument. To use PC-RTTOV, a separate pccoef file is required, and this must be compatible with the rtcoef file. See RTTOV documentation for more information.

The provision of PC Score retrievals is highly experimental. It is advised that the user makes themselves very confident with the PC-RTTOV options before attempting to run a 1D-Var. It may be necessary to modify the code when experimenting with PC-scores, with either real or simulated data. If you are using simulated data, be aware that the PC scores seem to be very sensitive to the settings used in RTTOV - you must make sure that your observations are simulated with the same settings as you subsequently use in the 1D-Var or the results will be very suspect if you even manage to get any passing QC! Note that there is a hard-coded value for QC-checking of the PC scores in NWPSAF_ProcessData.f90 - you will likely need to play with this number at least, and with the R-matrix, in order to get any sensible output.

Please do report successes and failures to the helpdesk, but be aware that it may be more difficult that usual to help you because few NWP-SAF staff have expertise in PC Score assimilation.

[ Return to Top ]

**8. Surface emissivity retrieval**

With version 1.2 the functionality to retrieve surface emissivity has been added. The implementation follows the methods set out in Pavelin and Candy (2014). In the retrieval state vector the surface emissivity is represented by principal components. This allows a compact representation of the surface emissivity and keeps the number of additional state vector elements low. The new capability is applied to IASI for up to 314 channels which have been selected as specified by Collard (2007).

For the surface emissivity retrieval, two additional auxiliary files are required. One of these files, EmisEigenVec contains the principal components and the other an atlas specifying the weights (or scores) of these EmisPCAtlas. The latter are a function of latitude, longitude and month, which are specified by the user in the ObsFile.dat.

All the user has to do to trigger a surface emissivity retrieval is to add a line in Retrievals.NL (more detail there).

[ Return to Top ]

**9. Notes, bugs and features**

Every attempt has been made to make this code as versatile and as bug-free as possible. Any problems should be reported by raising a ticket to the helpdesk.

There are many user-defined parameters in this code. Some of these parameters may not have been optimised for the users' requirements. This particularly applies to channel selection and and cloud detection channels and thresholds. The user is invited to critically review these.

[ Return to Top ]

**10. Getting Help**

Contact the NWPSAF helpdesk for all enquiries.

[ Return to Top ]

**11. References.**

Rodgers, C.D. (1990). Characterization and error analysis of profiles retrieved from remote sounding measurements. *J. Geophys. Res.*, **95**, 5587-5595.

Rodgers, C.D. (2000). Inverse Methods for Atmospheric Sounding: Theory and Practice *World Scientific, Singapore, New Jersey, London, Hong Kong*, ISBN-13: 978-981-02-2740-1

Pavelin E.G and Candy B. (2014). Assimilation of surface-sensitive infrared radiances over land: Estimation of land surface temperature and emissivity *Q.J.R.Meteorol.Soc. 140: 1198-1208, April 2014 B DOI:10.1002/qj.2218*

Collard A.D. Selection of IASI channels for use in numerical weather prediction *Q.J.R.Meteorol.Soc. 133: 1977-1991, 2007, DOI:10.1002/qj.178*

[ Return to Top ]

## 12. Acknowledgements

Many members of staff at the Met Office and beyond have worked on this code over the years. We are also very grateful to the work put in by our beta testers, and to everyone who reports bugs to the NWPSAF Helpdesk.

[ Return to Top ]

## Appendices

Appendix A. Top Level Design Note that this file is nwpsaf-mo-ds-026_top_level_design.pdf and supercedes the old html file.
Appendix B. Minimization
Appendix C. Cloudy Retrievals
Appendix D. Microwave cloud liquid water
Appendix E. Instructions for using the observation simulation code