9 OBSERVATION PROCESSING

9.1 BASIC PRINCIPLES

The ECMWF Data Assimilation Observation Processing System prior to IFS Cy26r1 (April 2003) was roughly split in two parts.

- Non-IFS observation processing modules.
- IFS integrated observation processing module.

Originally, the main difference in function between these two parts was based on whether information about a field (e.g. first guess) was required or not. Thus, the observation processing functions for which field information was not required were dealt with by the non-IFS modules, whereas the IFS itself dealt with those observation processing functions for which field information was needed.

The non-IFS observation processing came in two main parts.

- Preparation and massaging of input BUFR data.
- Creation of two data structure; one acceptable by the IFS as input and the other one for archiving purposes (observation feedback).

The first part of the non-IFS observation processing, which is still intact, consists of a number of modules: `PRE1CRAD` (further split by instrument type), `PREOBS`, `PREOBS_WAVE`, `PREGEOS`, `PREREO3` and `PRESCAT`. Without going into too many details here, the main theme for all of them is to prepare input BUFR data in an appropriate form for further processing. This also involves performing preliminary data thinning. As such, this part is preserved even after the major change which occurred with IFS Cy26r1.

The second part of the non-IFS observation processing consisted of two modules: `OBSPROC` and `OBSORT`. The main task of `OBSPROC` was to prepare input BUFR data in a form to be used by the analysis, whereas `OBSORT` dealt with any issues related to parallel computing. In this context `OBSORT` was not doing anything on its own; it was normally called by `OBSPROC` to ensure efficient parallelisation. During an analysis cycle `OBSPROC` is executed twice: just before and just after the IFS. The task before the IFS, called `MAKECMA` or for short `MKCMA`, performed a number of observation processing functions.

- Read in and crack input BUFR data.
- Carry out preliminary data checks.
- Perform necessary variable changes.
- Assign observation errors.
- Create CMA data structure recognised by the IFS.
- Etc.

On the other hand the task of `OBSPROC` just after the IFS, called `FEEDBACK`, was to create BUFR feedback. This was done by appending the input BUFR data with analysis-related information (departures, flags, events, etc.).

IFS Cy26r1 saw a major revision in this area. Observation processing modules `OBSPROC` and `OBSORT`, as well as the CMA observation data structure, have been phased out. Hence, `MKCMA` and `FEEDBACK` tasks as we knew them were made obsolete. However, a new data structure, the ODB, as well as two new observation processing modules (`BUFRTOODB` and `ODBTOBUFR`) have been introduced. Most of the observation processing functions earlier performed by the `MKCMA` task within `OBSPROC` have now been included in the IFS. It is only purely BUFR related processing functions that have now been taken over by `BUFRTOODB` and `ODBTOBUFR`.

- BUFRTOODB, together with MERGEODB, runs just before the IFS and is called MAKEODB. Effectively what it does is to read input BUFR data and create initial ODB which is formally acceptable by the IFS.
- ODBTOBUFR together with MATCHUP runs just after the IFS and is called ODB2BUFR.

Both MAKEODB and ODB2BUFR have been developed and are handled by the Operations Department.

The OBSORT observation processing functions have now almost entirely been incorporated into the ODB software.

As mentioned earlier most of the observation processing functions of OBSPROC are now integrated in the IFS. These newly integrated IFS observation processing functions are now known as "MAKE CMA REPLACEMENT" or for short MKCMARPL.

Here we will mostly concentrate on the IFS integrated observation processing whereas the other parts of the ECMWF documentation will deal with the remaining aspects of observation processing.

9.2 MAIN MKCMARPL TASKS AND FUNCTIONS

9.2.1 Basic observation processing setup

In order to perform the observation processing functions, a number of basic observation processing setups are carried out at the very beginning of initialising the IFS. This is done by calling several routines in addition to all other routines needed to setup the IFS (see Figure 9.1).

- Program MASTER calls CNT0 which in turn calls SU0YOMA.
- SU0YOMA calls (among other routines) SUOAF from which SUCMOCTP, SUEVENTS, SUCODES, SUFLTXT and SUCMA are called. SUCMOCTP defines the ODB observation types and code types, and SUEVENTS, SUCODES and SUFLTXT define analysis events, various codes used and flags naming conventions.
- SUCMA calls SUCMAF which then calls several subroutines: SUCMAD1, SUCMAD2, SUCMAHFP, SUCMAHOP, SUCMBDFP and SUCMBDTP. These routines define the structure of ODB Data Descriptor Records (DDRs) as well as the ODB packing patterns (bit structure) employed for header and body respectively.

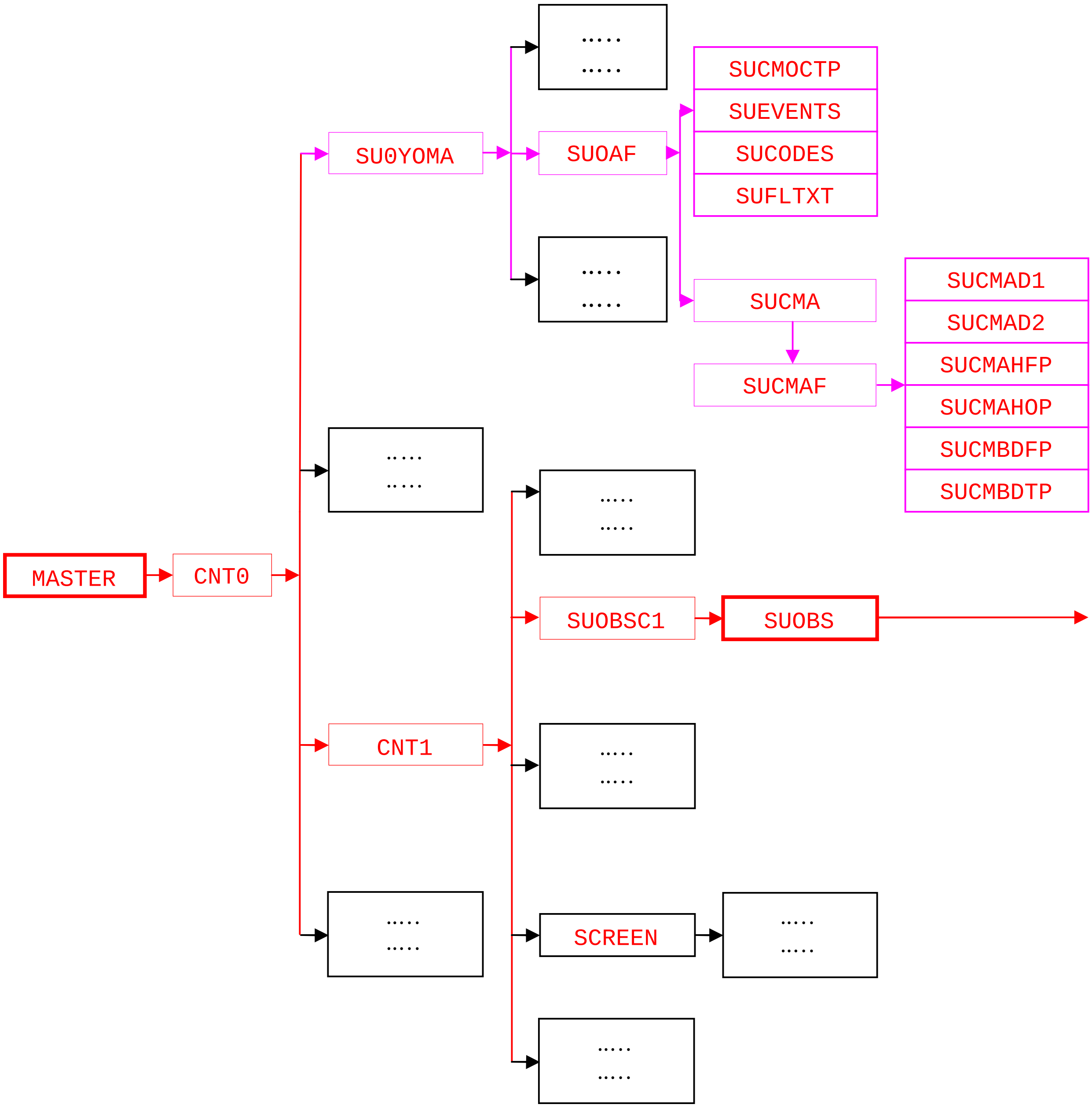

Figure 9.1 *Simplified IFS observation pre-processing flow diagram (*MASTER*). Colour coding scheme: (a) routines in red boxes perform observation pre-processing , (b) routines in pink boxes carry out observation pre-processing set up, and (c) routines in black boxes are not directly involved in observation pre-processing. Figure 9.2 continues the flow diagram from* SUOBS.

9.2.2 Invoking, initializing and controlling the MKCMARPL

The MKCMARPL run is initiated by the MKCMARPL subroutine (see Figure 9.2). This routine is only invoked in the SCREENING run of the IFS. It is called, together with some of its additional setup routines via subroutine SUOBS. The additional setup routines called at this level are: SUANCT, DEFRUN, SULIM, SULEVLAY, SUSATRET, SUVNMB, SUSCRE0, SUOBSORT, SETCOM, DEPERERR, SUERRORS, INIERSCA and INISSMIP.

- MKCMARPL is namelsit driven and in DEFRUN a logical variable LMKCMARPL is defined. By default LMKCMARPL=.T. but it can be overwritten via namelist NAMOBS. Furthermore, many other parameters and switches are defined in DEFRUN and some of them can also be overwritten via namelists.
- SUANCT and SULIM define some additional analysis constants and limits.
- SULEVLAY and SUSATRET define analysis related level/layer and satellite retrieval parameters, respectively.
- SUVNMB declares variable numbers.
- SUSCRE0, SUOBSORT and SETCOM define flag limits, identify ambiguous moving platforms, initialise observation sorting, and provide some general observation common variables.
- DEPERERR and SUERRORS deal with observation error statistics definitions. SUERRORS calls SUPERERR to define observation persistence errors and SUOBSERR to define prescribed observation errors.
- INIERSCA and INISSMIP deal with initialising SCATT and SSMI processing.

The next step is to find out if it is a SCREENING run and if so to check if it is a MKCMARPL run as well. In the case of a MKCMARPL run all aspects of the observation processing before the screening are dealt with by calling MKCMARPL (more about it in Section 9.2.3). After MKCMARPL has finished there are several ways to proceed. These depend on the status of LMKCMARPLO and LRPLSWAPOUT logical switches (NAMOBS namelist). If LRPLSWAPOUT=.T. the ODB is swapped out and if LMKCMARPLO=.T. the ODB is written out and the run terminated. Both of these options are not normally used and their use is for diagnostics/debugging purposes. Once the MKCMARPL work has been completed the remainder of SUOBS will execute as before. Thus, calls to WRITEOBA, WINDAUX, OBATABS, SUAMV, SURAD, SULIMB, SUOBAREA, MKGLOBSTAB and SUREO3 are issued.

In the context of operational running, the MKCMARPL related switches are set:

LMKCMARPL=.T.   LRPLSWAPOUT=.F.   LMKCAMRPLO=.F.

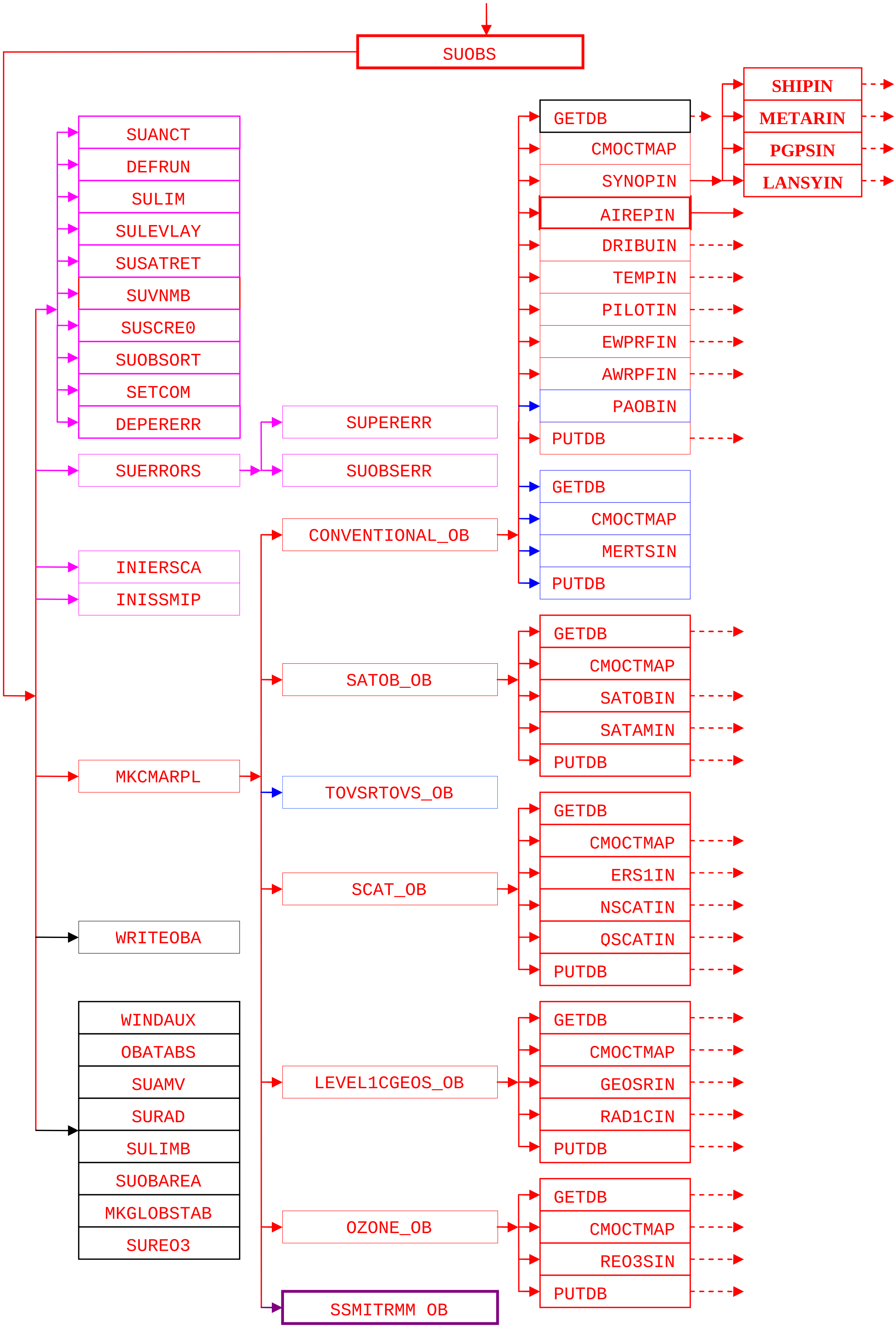
SUOBS
SUANCT
DEFRUN
SULIM
SULEVLAY
SUSATRET
SUVNMB
SUSCRE0
SUOBSORT
SETCOM
DEPERERR
SUERRORS
SUPERERR
SUOBSERR
INIERSCA
INISSMIP
MKCMARPL
WRITEOBA
WINDAUX
OBATABS
SUAMV
SURAD
SULIMB
SUOBAREA
MKGLOBSTAB
SUREO3
CONVENTIONAL_OB
SATOB_OB
TOVSRTOVS_OB
SCAT_OB
LEVEL1CGEOS_OB
OZONE_OB
SSMITRMM_OB
GETDB
CMOCTMAP
SYNOPIN
AIREPIN
DRIBUIN
TEMPIN
PILOTIN
EWPRFIN
AWRPFIN
PAOBIN
PUTDB
SHIPIN
METARIN
PGPSIN
LANSYIN
GETDB
CMOCTMAP
MERTSIN
PUTDB
GETDB
CMOCTMAP
SATOBIN
SATAMIN
PUTDB
GETDB
CMOCTMAP
ERS1IN
NSCATIN
QSCATIN
PUTDB
GETDB
CMOCTMAP
GEOSRIN
RAD1CIN
PUTDB
GETDB
CMOCTMAP
REO3SIN
PUTDB

Figure 9.2 (Continued from Figure 9.1) *Simplified IFS observation pre-processing flow diagram (*SUOBS*). Colour coding scheme: (a) routines in red boxes perfor observation pre-processing , (b) routines in pink boxes carry out observation pre-processing set up, (c) routines in black boxes are not directly involved in observation pre-processing, (d) routines in blue boxes are obsolete and (e) routines in plum boxes are awaiting revision. Figure 9.3 continues the flow diagram from* AIREPIN.

9.2.3 MKCMARPL

The main purpose of MKCMARPL is to control the IFS observation pre-processing. Observation pre-processing at this stage is done in groups of observations. At the moment there are seven groups: CONVENTIONAL, SATOB, TOVS/RTOVS, SCATT, LEVEL1C/GEOSS, OZONE and SSMI/TRMM observations. For each group a separate subroutine is called: CONVENTIONAL_OB, SATOB_OB, TOVSRTOVS_OB, SCAT_OB**,** LEVEL1CGOES_OB, OZONE_OB and SSMITRMM_OB. These routines are just cover or hat routines for the actual work to be carried out underneath. However, TOVSRTOVS_OB and SSMITRMM_OB are currently not called because TOVSRTOVS_OB is obsolete and SSMITRMM_OB is waiting for a major revision.

Each cover routine would call the ODB to get the observations it wants to process. This is done by calling the ODB GETDB subroutine. As the observations are brought, in one or more worker routines would be called to perform the observation processing functions. Once the worker routines have finished the control is handed back to the cover routine. The next step in the cover routine is to return observations back to the ODB database. This is done by calling the ODB PUTDB routine. In some of these cover routines several calls to GETDB/PUTDB might be issued. This is because there may be sufficient differences between similar data to justify a slightly different approach in their pre-processing. For example under the CONVENTIONAL_OB routine there are two calls to a GETDB and PUTDB pair. The first call deals with all conventional observations except SATEMs; the second call deals with the SATEMs. As indicated earlier, between each GETDB and PUTDB a number of observations type or code type designed worker routines are called.

- CONVENTIONAL_OB calls the following worker routines: SYNOPIN, AIREPIN, DRIBUIN, TEMPIN, PILOTIN, EWPRFIN, AWPRFIN, PAOBIN and MERTSIN. A worker routine name indicates which observations it is dealing with.
- SATOB_OB calls SATOBIN and SATAMIN.
- SCAT_OB calls ERS1IN, NSCATIN and QSCATIN.
- LEVEL1CGEOS_OB calls RAD1CIN and GOESRIN.
- OZONE_OB calls only REO3SIN.

9.2.4 Basic observation handling routines

The observation pre-processing worker routines referred to in Section 9.2.3, names of which always end with “IN”, are the basic observation handling routines. They all follow more or less the same logic. As an example consider AIREPIN which deals with AIREP observations (see Figure 9.3).

The first thing which is done is to define the instrument specification (OBINSTP) followed by preliminary quality control check both at the report level (PRLMCHK) as well as at the data level (GETSETE and AIREPBE).

- PRLMCHK calls REPSEL and TIMDIF to do report selection according to preset criteria and to find out time difference between analysis time and the actual observation time, respectively.
- GETSETE makes a local copy of a given observation variable and its related parameters from an ODB supplied array.
- After updating the local copy, AIREPBE is called to return the updated local copy back to the ODB supplied array.

The preliminary quality control at the report level consists of making sure that observation position, date and time are reasonable. Furthermore, as there is a possibility of excluding certain observations via the NAMOBS namelist, a check is made of whether the observation is actually wanted at this stage. Once the report level check is passed attention is turned to the data itself. Each datum is checked against predefined list of expected data. If not in the list, datum is rejected and a warning message issued. At this stage it is also ensured that missing indicators used are unique.

After the preliminary phase attention is turned to getting data in the right form and shape for further usage. Thus, in the case of an AIREP observation, this is done in sections of available variables: wind and temperature.

- **Wind**. There are four wind variables: wind direction (*DDD*), wind force (*FFF*), *u* and *v* components. For each of these variables the first thing which is done is to get a local copy of it together with its related parameters from an ODB supplied array (GETSETE). Once a variable is made available locally a check is made to ensure that the vertical coordinate is pressure; if instead of pressure a flight level is supplied it is converted into pressure by assuming a standard ICAO atmosphere (Z2PICAO). If the variable in question is either *u* or *v*, then *DDD* and *FFF* are converted into *u* and *v* wind components. Furthermore, for each of the four variables appropriate observation error statistics are assigned (ERRSTAT, FINOERR). Also, if any flags are set at this stage an appropriate word in the local copy is updated (PPVAFL). Finally, an updated local copy of an observed quantity and its related parameters are returned back into the ODB (AIREPBE).

- **Temperature**. In the case of temperature only one observed variable is dealt with. The pattern of making a local copy (GETSETE), ensuring that pressure is the vertical coordinate (Z2PICAO), assigning the observation error statistics (ERRSTAT), updating flags (PPVAFL) and returning an updated local copy back to the ODB (AIREPBE) is repeated.

As just mentioned ERRSTAT deals with assigning observation errors for a given observation variable. ERRSTAT first calls OBSPERR to assign observation persistence error; then it calls OBSERR which in turn calls FIXERR to assign prescribed observation error. It is worth mentioning that observation errors themselves are already predefined at an earlier stage (SUERRORS).

The pattern of activities outlined for AIREPIN is repeated more or less in the other worker routines. However, the SYNOPIN routine is first split further into SHIPIN, METATRIN, PGPSIN and LANSYIN. This is because SHIP, METAR, GPS and SYNOP LAND observations are sufficiently different to justify a separate worker routine. Furthermore, LANSYIN is somewhat more complicated than AIREPIN. One of the reasons for this is that we have to distinguish between low and high level stations.

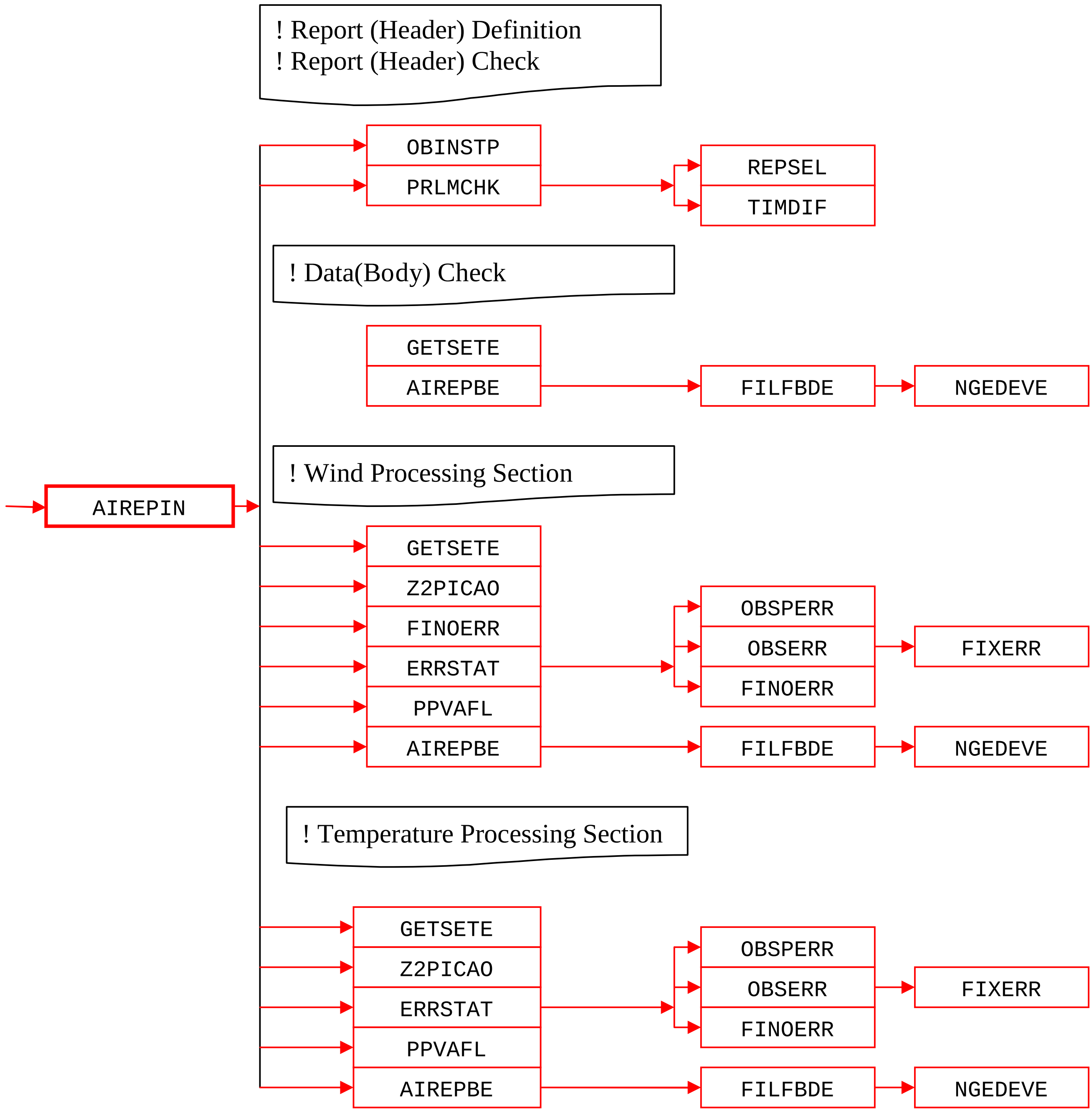

Figure 9.3 (Continued from Figure 9.2) *Simplified IFS observation pre-processing flow diagram (*AIREPIN*). Colour coding scheme: routines in red boxes perform observation pre-processing.*

9.3 OBSERVATION TYPES, SUBTYPES AND CODE TYPES

All observations, both in the BUFR and ODB contexts, are split into a number of observation types. The observation types are then further divided into observation code types (ODB) and observation subtypes (BUFR). Although BUFR observation types and subtypes are not directly used in the IFS they are defined here. BUFR observation types and subtypes are mapped into ODB observation types and code types before the IFS (i.e. the MERGEODB step).

9.3.1 BUFR observation types and subtypes

There are eight BUFR observation types. However, the number of subtypes differs between observation types; they are listed in Table 9.1.

Table 9.1 *BUFR observation types and subtypes*

Observation Type		Subtype	
Code	Name	Code	Name
0	Land Surface	1	Land SYNOP
		3	Automatic Land SYNOP
		9	Abbreviated Land SYNOP
		110	GPS
		140	METAR
1	Sea Surface	9	SHIP
		11	SHIP
		13	Automatic SHIP
		19	Reduced SHIP
		21	DRIBU
		22	BATHY
2	Upper Air Sounding	91	Land PILOT
		92	SHIP PILOT
		95	Wind Profiler (American)
		96	Wind Profiler (European/Japanese)
		101	Land TEMP
		102	SHIP TEMPS
		103	DROP TEMP
		104	ROCOB
		105	SHIP ROCOB
		106	Mobile TEMP
3	Satellite Sounding	0	High Resolution TOVS

		51	High Resolution TOVS
		53	RTOVS
		54	ATOVS
		55	ATOVS
		57	ATOVS
		61	Low Level Temperature SATEM
		62	High Level SATEM
		63	PWC SATEM
		65	Merged SATEM
		71	Low Level TOVS
		72	High Level TOVS
		73	PWC TOVS
		75	Merged TOVS
		129	TRMM
		130	TMI
		161	PAOB
		206	OZONE Retrieved Layers
4	AIREP	142	AIREP
		143	COLBA
		144	AMDAR
		145	ACARS
5	SATOB	82	Temperature and Wind
		83	Wind Only
		84	Temperature only
		85	Temperature only
		86	High Resolution VIS Wind
		87	AMV
		89	Geostationary Clear Sky Radiances (GRAD)
		189	Geostationary Clear Sky Radiances (GRAD)
12	ERS/SSMI	8	ERS 1
		122	ERS 1
		127	SSMI
		136	NSCAT
		137	
253	PAOB	161	PAOB

9.3.2 ODB observation and code types

There are ten ODB observation types and, as with BUFR, there are a different number of code types for each of them. It is a reasonable to question why the BUFR and ODB observation types and sub or code types are different. The answer is a historic one. The ODB observation types and code types have been used before BUFR came in to existence and as an international code it was difficult to impose our practice on the others. Also, there was not enough enthusiasm on our side to switch to the BUFR ones. The ODB observation types and code types are listed in Table 9.2.

Table 9.2 *ODB observation types and code types*

Observation Type		Code Type	
Code	Name	Code	Name
1	SYNOP	11	Land SYNOP
		14	Automatic Land SYNOP
		16	French RADOME
		21	SHIP
		22	Abbreviated SHIP
		23	SHRED
		24	Automatic SHIP
		140	METAR
		110	GPS
2	AIREP	41	CODAR
		141	AIREP
		142	Simulated AIREP
		144	AMDAR
		145	ACARS
		241	COLBA
3	SATOB	88	SATOB
		89	High Resolution VIS wind
		90	AMV
		188	SST
4	DRIBU	63	BATHY
		64	TESAC
		160	ERS as DRIBU
		165	DRIBU
5	TEMP	35	Land TEMP
		36	SHIP TEMP
		37	Mobile TEMP

		39	ROCOB
		40	SHIP ROCOB
		135	DROP TEMP
		137	Simulated TEMP
6	PILOT	32	Land PILOT
		33	SHIP PILOT
		34	American Wind Profiler
		131	Japanese Wind Profiler
		132	Mobile Wind Profiler
		134	European Wind Profiler
7	SATEM	86	GTS SATEM
		184	High Resolution Simulated TOVS
		185	High Resolution Simulated DWL SATEM
		186	High Resolution SATEM
		200	GTS BUFR SATEM 250km
		201	GTS BUFR Clear Radiances
		202	GTS BUFR Retrieved Profiles/Clear Radiances
		210	ATOVS/GRAD
		211	RTOVS
		212	TOVS
		215	SSMI
8	PAOB	164	PAOB
9	SCATTEROMETER	8	SCATTEROMETER
		122	SCATTEROMETER
		210	SCATTEROMETER
10	RAW RADIANCE		

9.3.3 Mapping between ODB and BUFR observation types, code types and subtypes

As indicated in Section 9.3.2 the coexistence of different codes used for BUFR and ODB observation types and the subtype and code type requires a mapping from one to another. This is given in Table 9.3.

Table 9.3 *Mapping between ODB and BUFR observation types, code types and subtypes*

ODB (Observation Type, Code type)		BUFR (Observation Type, Subtype)
ODB(1, 11)	↔	BUFR[(0,1);(0,9)]
ODB(1, 14)	↔	BUFR(0,3)

ODB(1, 21)	↔	BUFR(1,9)
ODB(1, 22)	↔	BUFR(?,?)
ODB(1, 23)	↔	BUFR(1,19)
ODB(1, 24)	↔	BUFR(1,13)
ODB(1,110)	↔	BUFR(0,110)
ODB(1,140)	↔	BUFR(0,140)
ODB(2,41)	↔	BUFR(?,?)
ODB(2,141)	↔	BUFR(4,142)
ODB(2,142)	↔	BUFR(?,?)
ODB(2,144)	↔	BUFR(4,144)
ODB(2,145)	↔	BUFR(4,145)
ODB(2,241)	↔	BUFR(4,143)
ODB(3,88)	↔	BUFR[(5,82);(5,83);(5,84);(5,85)]
ODB(3,89)	↔	BUFR(5,86)
ODB(3,90)	↔	BUFR(5,87)
ODB(3,188)	↔	BUFR(?,?)
ODB(4,63)	↔	BUFR(1,23)
ODB(4,64)	↔	BUFR(1,22)
ODB(4,160)	↔	BUFR(?,?)
ODB(4,165)	↔	BUFR(1,21)
ODB(5,35)	↔	BUFR(2,101)
ODB(5,36)	↔	BUFR(2,102)
ODB(5,37)	↔	BUFR(2,106)
ODB(5,39)	↔	BUFR(2,104)
ODB(5,40)	↔	BUFR(2,105)
ODB(5,135)	↔	BUFR(2,103)
ODB(5,137)	↔	BUFR(?,?)
ODB(6,32)	↔	BUFR(2,91)
ODB(6,33)	↔	BUFR(2,92)
ODB(6,34)	↔	BUFR(2,94)
ODB(6,131)	↔	BUFR(2,95)
ODB(6,134)	↔	BUFR(2,95)
ODB(7,86)	↔	BUFR[(3,61);(3,62);(3,63);(3,65)]
ODB(7,184)	↔	BUFR(?,?)
ODB(7,185)	↔	BUFR(?,?)
ODB(7,186)	↔	BUFR[(3,71);(3,72);(3,73);(3,75)]
ODB(7,200)	↔	BUFR(?,?)
ODB(7,201)	↔	BUFR(?,?)
ODB(7,202)	↔	BUFR(?,?)
ODB(7,206)	↔	BUFR(?,?)

ODB(7,210)	↔	BUFR[(3,54);(5,89)]
ODB(7,211)	↔	BUFR(3,53)
ODB(7,212)	↔	BUFR[(3,0);(3,51)]
ODB(7,215)	↔	BUFR(12,127)
ODB(8,180)	↔	BUFR(253,154)
ODB(9,8)	↔	BUFR(12,8)
ODB(9,122)	↔	BUFR(12,122)
ODB(9,300)	↔	BUFR(12,136)
ODB(9,301)	↔	BUFR(12,137)
ODB(9,511)	↔	BUFR(?,?)
ODB(10,1)	↔	BUFR(?,?)

9.4 VARIABLES

Different quantities are observed by different observing systems. It is only a subset of observed quantities that are used in the analysis and most of them are used in their original form. However, some of them are transformed into the ones actually used by the analysis. This transformation, or a change of variable, may also include retrieval from satellite data if they are independent from the background model fields. The original variables may be kept with the derived ones so that first guess departures can be assigned for both. Furthermore, if an observed variable is transformed then, if necessary, so is its observation error statistics. Also, in the case of an off-time SYNOP observation, the observed surface pressure may be adjusted.

9.4.1 Observed variables

The exact list of what is observed or present in the list of BUFR observation types and sub types (Table 9.3) is long. Therefore Table 9.4 just lists (per observation types) those variables which are of interest at present.

Table 9.4 *Observed variables*

Observation Type		Observed Variable
BUFR	**ODB**	
Land Surface	Land SYNOP	Surface Pressure (P_s)
		10 m Wind Direction/Force (*DDD*/*FFF*)
		2 m Temperature (T_{2m})
		2 m Dew Point (Td_{2m})
		Pressure Tendency (P_t)
		Cloud Information
		Precipitation Information
		Snow Depth (*Sd*)
		Etc.
Sea Surface	SHIP SYNOP, DRIBU	Surface Pressure (P_s)
		10 m Wind Direction/Force (*DDD*/*FFF*)
		2 m Temperature (T_{2m})
		2 m Dew Point (Td_{2m})

		Etc.
Upper Air Sounding	TEMP, PILOT	10m/Upper Air Wind Direction/Force (*DDD*/*FFF*)
		2 m/Upper Air Temperature (T_{2m}/T)
		2 m/Upper Air Dew Point (Td_{2m}/Td)
		Geopotential Height (*Z*)
		Etc.
Satellite Sounding	SATEM	Mean Layer Temperature
		Precipitable Water Content (*PWC*)
		Brightness Temperature (*Tb*)
AIREP	AIREP	Upper Air Wind Direction/Force (*DDD*/*FFF*)
		Temperature (*T*)
SATOB	SATOB	Upper Air Wind Direction/Force (*DDD*/*FFF*)
		Brightness Temperature (*Tb*)
ERS/SSMI	SCATTEROMETER	Backscatter (σ^0)
		Brightness Temperature (*Tb*)

9.4.2 Derived variables

Variables which are transformed for further use by the analysis are as follows.

- Wind direction (*DDD*) and force (*FFF*) are transformed into wind components (*u* and *v*) for SYNOP, AIREP, SATOB, DRIBU, TEMP and PILOT observations.
- Temperature (*T*) and dew point (*Td*) are transformed into relative humidity (*RH*) for SYNOP and TEMP observations, with a further transformation of the *RH* into specific humidity (*Q*) for TEMP observations.
- SCATTEROMETER backscatters (σ^0's) are transformed into a pair of ambiguous wind components (*u* and *v*); this actually involves a retrieval according to some model function describing the relationship between winds and σ^0's and requires a fair bit of computational work.
- Mean layer temperature is transformed into thickness (*DZ*) for SATEM and TOVS observations.

All these variable transformations, except for the σ^0's transformation, are more or less trivial ones. The wind components are worked out as:

$$u = -FFF \sin(DDD \frac{\pi}{180})$$

$$v = -FFF \cos(DDD \frac{\pi}{180})$$

The *RH* is derived by using the following relationship:

$$RH = \frac{F(Td)}{F(T)}$$

where function F of either *T* or *Td* is expressed as:

$$\mathbf{F}(T) = a\frac{\mathbf{R}_{dry}}{\mathbf{R}_{vap}}e^{b\frac{T-T_0}{T-c}}$$

where $T_0 = 273.16$ K, a = 611.21, b = 17.502, c = 32.19, $R_{dry} = 287.0597$ and $R_{vap} = 461.5250$ are constants, Specific humidity Q is worked out by using the following relationship:

$$Q = RH\frac{\mathbf{A}}{1 - RH\left(\frac{R_{vap}}{R_{dry}} - 1\right)A}$$

where function A is expressed as:

$$\mathbf{A} = \min\left[0.5, \frac{\mathbf{F}(T)}{P}\right]$$

where P is pressure. Q is assigned in the RH2Q subroutine.

Scatterometer wind retrieval is dealt with in Section 9.6.

9.4.3 Adjusted variables

The only observed quantity which is adjusted is the SYNOP's surface pressure (P_s). This is done by using pressure tendency (P_t) information, which in turn may be first adjusted. P_t is adjusted only in the case of SYNOP SHIP data for the ship movement.

The ship movement information is available from input data in terms of ship speed and direction, which are first converted into ship movement components U_s and V_s. The next step is to find pressure gradient ($\partial p/\partial x$ and $\partial p/\partial y$):

$$\frac{\partial p}{\partial x} = C(A_1 u - A_2 v)\frac{1}{2}$$

$$\frac{\partial p}{\partial y} = -C(A_1 u + A_2 v)$$

where u and v are observed wind components, and $A_1 = 0.94$ and $A_2 = 0.34$ are the sine and cosine of the angle between the actual and geostrophic winds. C is the Coriolis term multiplied by a drag coefficient (D):

$$C = 2\Omega D \sin\theta$$

where, θ is the latitude and $\Omega = 0.7292 \times 10^{-4} s^{-1}$ is the angular velocity of the earth and D is expressed as:

$$D = GZ$$

G = 1.25 is an assumed ratio between geostrophic and surface wind over sea and $Z = 0.11$ kgm^{-3} is an assumed air density. Now the adjusted pressure tendency (P_t^a) is found as:

$$P_t^a = P_t - \left(U_s\frac{\partial p}{\partial x} + V_s\frac{\partial p}{\partial y}\right)$$

Finally, the adjusted surface pressure (P_s^a) is found as:

$$P_s^a = P_s - P_t^a \Delta t$$

where, Δt is a time difference between analysis and observation time. Of course in the case of non-SHIP data $P_t^a \equiv P_t$. Subroutine PTENDCOR is used for this adjustment.

9.4.4 Variable's code

To provide easy recognition of 'observed' variables each of them is assigned a numerical code. These codes are then embedded in ODB reports. There are 76 codes used so far. These codes are defined in subroutine SUVNMB. For the sake of completeness these codes are listed in Table 9.5.

Table 9.5 *Numbering of variables in the ODB*

No.	Code	Name	Unit
1	3	Wind Component (u)	ms^{-1}
2	4	Wind Component (v)	ms^{-1}
3	1	Geopotential (Z)	m^2s^{-2}
4	57	Thickness (DZ)	m^2s^{-2}
5	29	Relative Humidity (RH)	numeric
6	9	Precipitable Water Content (PWC)	kgm^{-2}
7	58	2 m Relative Humidity (RH_{2m})	numeric
8	2	Temperature	K
9	59	Dew Point	K
10	39	2 m Temperature (T_{2m})	K
11	40	2 m Dew Point (Td_{2m})	K
12	11	Surface Temperature (Ts)	K
13	30	Pressure Tendency (P_t)	Pa/3h
14	60	Past Weather (W)	WMO Code 4561
15	61	Present Weather (WW)	WMO Code 4677
16	62	Visibility (V)	WMO Code 4300
17	63	Type of High Clouds (C_H)	WMO Code 0509
18	64	Type of Middle Clouds (C_M)	WMO Code 0515
19	65	Type of Low Clouds (C_L)	WMO Code 0513
20	66	Cloud Base Height (N_h)	m
21	67	Low Cloud Amount (N)	WMO Code 2700
22	68	Additional Cloud Group Height (h_sh_s)	m
23	69	Additional Cloud Group Type (C)	WMO Code 0500
24	70	Additional Cloud Group Amount (N_s)	WMO Code 2700
25	71	Snow Depth (Sd)	m
26	72	State of Ground (E)	WMO Code 0901
27	73	Ground Temperature (T_gT_g)	K
28	74	Special Phenomena (S_pS_p)	WMO Code 3778
29	75	Special Phenomena (s_ps_p)	WMO Code 3778
30	76	Ice Code Type (R_s)	WMO Code 3551
31	77	Ice Thickness (E_sE_s)	WMO Code 1751
32	78	Ice (I_s)	WMO Code 1751
33	79	Time Period of Rain Information (t_rt_r)	hour
34	80	6 Hour Rain Amount	kgm^{-2}
35	81	Maximum Temperature (JJ)	K
36	82	Ship Speed (V_s)	ms^{-1}
37	83	Ship Direction (D_s)	degree
38	84	Wave Height (H_wH_w)	m
39	85	Wave Period (P_wP_w)	s
40	86	Wave Direction (D_wD_w)	degree
41	87	General Cloud Group	WMO Code
42	88	Relative Humidity from Low Clouds	numeric
43	89	Relative Humidity from Middle Clouds	numeric
44	90	Relative Humidity from High Clouds	numeric
45	91	Total Amount of Clouds	WMO Code 20011
46	92	6 Hour Snowfall	m
47	110	Surface Pressure (P_s)	Pa
48	111	Wind Direction	degree
49	112	Wind Force	ms^{-1}
50	119	Brightness Temperature (Tb)	K
51	120	Raw Radiance	K
52	121	Cloud Amount from Satellite	%

53	122	Backscatter (σ^0)	dB
54	5	Wind Shear ($\partial u/\partial z$)	s^{-1}
55	6	Wind Shear ($\partial v / \partial z$)	s^{-1}
56	41	u_{10m}	ms^{-1}
57	42	v_{10m}	ms^{-1}
58	19	Layer Relative Humidity	numeric
59	200	Auxiliary Variable	numeric
60	123	Cloud Liquid Water (Q_l)	$kgkg^{-1}$
61	124	Ambiguous v	ms^{-1}
62	125	Ambiguous u	ms^{-1}
63	7	Specific Humidity (Q)	$kgkg^{-1}$
64	126	Ambiguous Wind Direction	degree
65	127	Ambiguous Wind Speed	ms^{-1}
66	8	Vertical Speed	ms^{-1}
67	56	Virtual Temperature (Tv)	K
68	206	Ozone	Dobson
69	156	Height	m
70	215	SSM/I Pseudo Variable	kgm^{-2}
71	160	Past Weather	numeric
72	130	Pressure Tendency Characteristics	numeric
73	12	Sea Water Temperature	K
74	192	Rader Reflectivity	Db
75	128	Atmospheric Path Delay in Satellite Signal	m
76	162	Radio Occultation Bending Angle	Rad

9.5 OBSERVATION ERROR STATISTICS

Three types of observation errors are dealt with at the observation pre-processing level.

- Persistence observation error.
- Prescribed observation error.
- Combination of the two above called the final observation error.

9.5.1 Persistence observation error

The persistence error is formulated in such a way to reflect its dependence on the following.

- Season.
- Actual geographical position of an observation.

Seasonal dependency is introduced by identifying three regimes.

- Winter hemisphere.
- Summer hemisphere.
- Tropics.

The positional dependency is then introduced to reflect the dependence on the precise latitude within these three regimes.

The persistence error calculation is split into two parts. In the first part the above dependencies are expressed in terms of factors a and b which are defined as:

$$a = \sin\left(2\pi \frac{d}{365.25} + \frac{\pi}{2}\right)$$

and

$$b = 1.5 + a\left\{0.5 \min\left[\max\left(\theta, 20\right)/20\right]\right\}$$

where d is a day of year and θ is latitude.

The persistence error for time difference between analysis and observation Δt is then expressed as a function of b with a further dependence on latitude and a maximum persistence error E_{maxpers} for 24 hour:

$$E_{pers} = \frac{Emaxpers}{6}\left[1 + 2\sin\left(|2\theta|\right) b\Delta t\right]$$

where Δt is expressed as a fraction of a day. The E_{maxpers} have the values shown in Table 9.6.

Table 9.6 *Observation persistence errors of maximum 24-hour wind (u,v), height (Z) and temperature (T)*

Variable (unit)	1000–700 hPa	699–250 hPa	249–0 hPa
u, v (ms^{-1})	6.4	12.7	19.1
Z (m)	48	60	72
T (K)	6	7	8

Subroutine SUPERERR is used to define all relevant points in order to carry out this calculation, and is called only once during the general system initialization. The calculation of the actual persistence error is dealt with by OBSPERR.

9.5.2 Prescribed observation errors

Prescribed observational errors have been derived by statistical evaluation of the performance of the observing systems, as components of the assimilation system, over a long period of operational use. The prescribed observational errors are given in the Tables 9.7, 9.8 and 9.9. Currently, observational errors are defined for each observation type that carries the following quantities.

- Wind components
- Height
- Temperature
- Humidity

As can be seen from the tables of prescribed observation errors, they are defined at standard pressure levels but the ones used are interpolated to the observed pressures. The interpolation is such that the observation error is kept constant below the lowest and above the highest levels, whereas in between it is interpolated linearly in $\ln p$. Several subroutines are used for working out the prescribed observation error: SUOBSERR, OBSERR, FIXERR, THIOERR and PWCOERR.

- SUOBSERR defines observation errors for standard pressure levels.
- OBSERR and FIXERR calculates the actual values.
- THIOERR and PWCOERR are two specialised subroutines to deal with thickness and PWC errors.

Relative humidity observation error RH_{err} is either prescribed or modelled. More will be said about the modelled RH_{err} in Section 9.5.3. RH_{err} is prescribed only for TEMP and SYNOP data. RH_{err} is preset to 0.17 for TEMP and 0.13 for SYNOP. However, if RH<0.2 it is increased to 0.23 and to 0.28 if T<233 K for both TEMP and SYNOP.

Table 9.7 *Prescribed RMS observation errors for the u and v wind components (ms^{-1})*

Observation Type	SYNOP	AIREP	AIREP	SATOB	DRIBU	TEMP	PILOT	SATEM	PAOB	SCATT	R RAD
Code Type	All	All but AIREP	AIREP	All	All	All	All	All	All	All	All
Levels (hPa)											
1000	3.00	2.46	2.86	2.00	1.80	1.80	1.80	n/a	n/a	2.00	n/a
850	3.00	2.51	2.91	2.00	1.80	1.80	1.80	n/a	n/a	2.00	n/a
700	3.00	2.56	2.96	2.00	1.80	1.90	1.90	n/a	n/a	2.00	n/a
500	3.40	2.71	3.11	3.50	1.80	2.10	2.10	n/a	n/a	2.00	n/a
400	3.60	2.81	3.21	4.30	1.80	2.50	2.50	n/a	n/a	2.00	n/a
300	3.80	2.86	3.26	5.00	1.80	2.60	2.60	n/a	n/a	2.00	n/a
250	3.20	2.91	3.31	5.00	1.80	2.50	2.50	n/a	n/a	2.00	n/a
200	3.20	2.96	3.36	5.00	1.80	2.50	2.50	n/a	n/a	2.00	n/a
150	2.40	2.91	3.31	5.00	1.80	2.40	2.40	n/a	n/a	2.00	n/a
100	2.20	2.76	3.16	5.00	1.80	2.20	2.20	n/a	n/a	2.00	n/a
70	2.00	2.66	3.06	5.00	1.80	2.10	2.10	n/a	n/a	2.00	n/a
50	2.00	2.66	3.06	5.00	1.80	2.00	2.00	n/a	n/a	2.00	n/a
30	2.00	2.86	3.26	5.00	1.80	2.10	2.10	n/a	n/a	2.00	n/a
20	2.50	3.06	3.46	5.00	1.80	2.30	2.30	n/a	n/a	2.00	n/a
10	3.00	3.36	3.76	5.70	1.80	3.00	3.00	n/a	n/a	2.00	n/a

Table 9.8 *Prescribed RMS height observation errors (m)*

Observation Type	Code Type	Levels (hPa)														
		1000	850	700	500	400	300	250	200	150	100	70	50	30	20	10
SYNOP	Manual Land	5.6	7.2	8.6	12.1	14.9	18.8	25.4	27.7	32.4	39.4	50.3	59.3	69.8	96.0	114.2
	Automatic Land	4.2	5.4	6.45	9.07	11.17	14.1	19.05	20.77	24.3	29.55	37.72	44.47	52.35	72.0	85.65
	Manual SHIP	7.1	7.1	7.1	7.1	7.1	7.1	7.1	7.1	7.1	7.1	7.1	7.1	7.1	7.1	7.1
	Automatic SHIP	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.2
AIREP	All	n/a	n/a.	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
SATOB	All	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
DRIBU	All	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97	4.97
TEMP	All	4.3	4.4	5.2	8.4	9.8	10.7	11.8	13.2	15.2	18.1	19.5	22.5	25.0	32.0	40.0
PILOT	All	4.3	4.4	5.2	8.4	9.8	10.7	11.8	13.2	15.2	18.1	19.5	22.5	25.0	32.0	40.0
SATEM	All	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
PAOB	All	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0

Observation Type	Code Type															
SCATT	All	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
R RAD	All	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Table 9.9 *Prescribed RMS temperature observations error (K)*

Observation Type	SYNOP		AIREP		SATOB	DRIBU	TEMP	PILOT
Code Type	Land	SHIP	All but AIREP	AIREP	All	All	All	All
Levels (hPa)								
1000	2.0	1.8	1.65	1.40	n/a	1.8	1.40	n/a
850	1.5	1.8	1.43	1.18	n/a	1.5	1.25	n/a
700	1.3	1.8	1.25	1.00	n/a	1.3	1.10	n/a
500	1.2	1.8	1.23	0.98	n/a	1.2	0.95	n/a
400	1.3	1.8	1.21	0.96	n/a	1.3	0.90	n/a
300	1.5	1.8	1.20	0.05	n/a	1.5	1.00	n/a
250	1.8	1.8	1.20	0.95	n/a	1.8	1.15	n/a
200	1.8	1.8	1.31	1.06	n/a	1.8	1.20	n/a
150	1.9	1.8	1.43	1.18	n/a	1.9	1.25	n/a
100	2.0	1.8	1.55	1.30	n/a	2.0	1.30	n/a
70	2.2	1.8	1.65	1.40	n/a	2.2	1.40	n/a
50	2.4	1.8	1.75	1.50	n/a	2.4	1.40	n/a
30	2.5	1.8	1.85	1.60	n/a	2.5	1.40	n/a
20	2.5	1.8	2.05	1.80	n/a	2.5	1.50	n/a
10	2.5	1.8	2.35	2.10	n/a	2.5	2.10	n/a

SATEM	All	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
PAOB	All	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
SCATT	All	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
R. RAD	All	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

9.5.3 Derived observation errors

Relative humidity observation error, RH_{err}, can also be expressed as function of temperature T:

$$RH_{\mathrm{err}} = \min[0.18, \min(0.06, -0.0015T + 0.54)]$$

This option is currently used for assigning RH_{err}.

Specific humidity observation error, Q_{err}, is a function of RH, RH_{err}, P, P_{err}, T and T_{err}, and formally can be expressed as:

$$Q_{\mathrm{err}} = Q_{\mathrm{err}}\,(RH, RH_{\mathrm{err}}, P, P_{\mathrm{err}}, T, T_{\mathrm{err}})$$

Or

$$Q_{err} = RH_{err}\mathbf{F}_1(RH,P,T) + \frac{RHP_{err}}{P}\mathbf{F}_2(RH,P,T) + RHT_{err}\mathbf{F}_3(RH,P,T)$$

where functions F_1, F_2 and F_3 are given as:

$$\mathbf{F}_1(RH,P,T) = \frac{A}{\left[1 - RH\left(\frac{R_{vap}}{R_{dry}} - 1\right)A\right]^2}$$

$$\mathbf{F_2}(RH,P,T) = -\frac{\left\{\left[1 - RH\left(\frac{R_{vap}}{R_{dry}} - 1\right)A\right] + \left(\frac{R_{vap}}{R_{dry}} - 1\right)A\right\}}{\left\{1 - \left[RH\left(\frac{R_{vap}}{R_{dry}} - 1\right)A\right]\right\}^2}$$

$$\mathbf{F}_3(RH,P,T) = \frac{Ab\left(T_0 - c\right)}{\left(T - c\right)^2}\left\{\left[1 - \left(\frac{R_{vap}}{R_{dry}} - 1\right)A\right]RHA\left(\frac{R_{vap}}{R_{dry}} - 1\right)\right\}$$

At present only the first term of the above expression for Q_{err} is taken into account (dependency on relative humidity). Subroutine RH2Q is used to evaluate Q_{err}.

Surface pressure observation error Ps_{err} is derived by multiplying the height observation error Z_{err} by a constant:

$$Ps_{err} = 1.225 Z_{err}$$

However, the Ps_{err} may be reduced if the pressure tendency correction is applied. For non-SHIP data the reduction factor is 4, whereas for SHIP data the reduction factor is either 2 or 4, depending on if the P_t is adjusted for SHIP movement or not.

The thickness observation error (DZ_{err}) is derived from Z_{err}.

9.5.4 Final (combined) observation error

In addition to the prescribed observation and persistence errors, the so called final observation error is assigned at this stage too. This is simply a combination of observation and persistence errors:

$$F_{OE} = \sqrt{O_E^2 + P_E^2}$$

W `here F_{OE}, O_E and P_E are final, prescribed and persistence observation errors, respectively. The subroutine used for this purpose is FINOERR.

9.6 SCATTEROMETER DATA

9.6.1 Overview

The processing for three different scatterometer instruments, each having its own geometrical and geophysical characteristic has been coded in IFS. These are scatterometer data from the AMI instrument on-board the ERS-1 and ERS-2 platforms, the NSCAT instrument on-board ADEOS-I, and the SeaWinds instrument on-board QuikSCAT.

As for a number of other observation types, scatterometer data need to be transformed into the variables used by the analysis within the IFS. This transformation converts the backscatter measurements acquired by the instrument (triplets for ERS and quadruplets for NSCAT and QuikSCAT) into the two ambiguous *u* and *v* wind components that will actually be assimilated into the IFS. The (empirical) relation between wind and backscatter is described by a geophysical model function (GMF).

At the implementation of IFS Cy28r1 (March 2004), data from ERS-2 (re-introduced after an interruption since January 2001, and using a new GMF), and from QuikSCAT (introduced in IFS Cy24r3) were actively assimilated.

9.6.2 ERS data

For ERS-1 and ERS-2 scatterometer data the inversion task is performed within the MKCMARPL run by the scatterometer data handling subroutine SCAT_OB, especially by its core subroutine ERS1IF (ERS-1 interface). Like SCAT_OB, ERS1IF deals only with one scatterometer report at a time.

The main purpose of ERS1IF is to retrieve the wind components by inverting the geophysical model function. Some quality control is also done in the process, based on the quality information provided with the raw data, as well as on the residual from the wind retrieval, reflecting the agreement between the measurements and their theoretical wind dependency. The procedure closely follows the PRESCAT wind retrieval and ambiguity removal scheme developed for the ERS-1 scatterometer (Stoffelen and Anderson, 1997), though the original geophysical model function CMOD4 has been replaced by CMOD5.

Backscatter data from the ERS-1 and ERS-2 platform are provided into wind-vector cells (WVC) with a spatial resolution of 25 km. Swath width is 500 km, defining 19 cells in across-track direction. In this direction, in IFS, ERS data is thinned by only selecting predefined cells (subroutine SCAQC). The reason for this are known across-track variations in data quality. Along-track, this is not an issue, and therefore, thinning is performed according to the standard procedure (THINN) as also used for other satellite data. Both SCAQC and THINN result in an average selection of cells determined by the parameter NTHINSCA; a default value of NTHINSCA = 4 results in a selection of every fourth cell giving a thinning resolution of 100 km.

Moreover, bias corrections are applied, both in terms of backscatter and wind speed, particularly to compensate for any change in the instrumental calibration and to ensure consistency between the retrieved and model winds.

9.6.3 QuikSCAT data

For data from the SeaWinds scatterometer on-board the QuikSCAT satellite, the task of wind inversion is performed at an earlier stage of the processing. The program performing this task, QSCAT25TO50KM, is a part of a scatterometer library called SCAT.

Backscatter data from the QuikSCAT platform is also provided at a resolution of 25 km. Across the 1,800 wide swath 76 cells are defined. Rather than data thinning, in SCAT, a 50 km product is created instead which contains information about the backscattering from the four underlying original sub-cells. For the 38 across-track cells defined in this way the outer 4 at either side of the swath are, due to their known reduced quality, not assimilated. The weight of the scatterometer cost function (defined in routine HJO) of each 50 km wind vector cell is reduced by a factor four, which effectively mimics the assimilation of a 100 km product.

9.6.4 NSCAT data

Due to its short lifetime of nine months, NSCAT data has never been part of the operational assimilation chain at ECMWF. Although, in IFS, NSCAT data can be handled, assimilation experiments are only possible after certain offline pre-processing of the data. The Research Department should be contacted for further information.

9.6.5 Wind retrieval

In general, the wind retrieval is performed by minimizing the following distance between observed backscatter values σ_o^0 and modelled backscatter values σ_{mi}^0:

$$\mathbf{D}(u) = \sum_i^n \frac{\left[\left(\sigma_{oi}^0\right)^p - \sigma_{mi}^0 (u)^p \right]^2}{k_p \left[\sum_j^n \sigma_{mj}^0 (u)^p \right]^2}$$

For ERS data, the sum is over triplets, while for QuikSCAT the sum may extend to 16 values (four 25 km sub-cells with each four observations). The quantity p is equal to unity for NSCAT and QuikSCAT. For ERS data, a value of $p = 0.625$ was introduced because it makes the underlying GMF more harmonic, which helps to avoid direction-trapping effects (Stoffelen and Anderson, 1997). The noise to signal ratio k_p provides an estimate for the relative accuracy of the observations.

The simulation of σ_m^0 is for ERS-2 data based on the CMOD5 model function (replacing CMOD4 used for earlier cycles; although for ERA-40 experiments CMOD4 is still selected). For NSCAT data the NSCAT-2 GMF has been utilized. For QuikSCAT data, the choice of GMF is handled by a logical switch LQTABLE in the SCAT library. By default LQTABLE=.T. and the QSCAT-1 model function is used, otherwise, modelled backscatter values are based on the NSCAT-2 GMF. The minimization is achieved using a tabular form of the GMF, giving the value of the backscatter coefficient for wind speeds, direction and incidence angles discretized with, for ERS data, steps of 0.5 ms^{-1}, 5^0 and 1^0, respectively. For NSCAT and QuikSCAT data the corresponding values are 0.2 ms^{-1}, 2.5° and 1°. For ERS the table is read in the initialisation subroutine of the scatterometer observation processing INIERSCA. For QuikSCAT, this takes place in the QSCAT25TO50KM program in the PRESCAT task.

Up to four minima are kept at first. The first wind-vector solution is then defined as the minimum with the smallest residual, and the second one as the secondary minimum, etc.

The retrieval process is the subject of a few dedicated subroutines. For ERS, routine S0TOWIND directly returns two ambiguous wind vectors associated with a given triplet. This subroutine itself uses two specialized procedures, SPEEEST and MINIMA, to get a first guess estimate for the wind speeds yielding the lowest residuals and to search for the relative minima in the table, respectively. For QuikSCAT, a similar routine, INVERT50, returns up to four ambiguous wind vectors. The two corresponding routines it calls for first guess estimation and minimization are WSFG and MINIMA. In addition, it calls MEDIAS, a routine that calculates, based on the used data, the centre-of-gravity of the 50 km vector cell, and FFT99, a routine that was introduced to suppress numerical noise.

The retrieved wind components obviously play a major role in the definition of the scatterometer ODB reports. In addition, a logical parameter indicating whether no solution has actually been found is also transmitted for the subsequent processing applied in the IFS.

9.6.6 Quality control

Before calling for the wind retrieval, a first quality control step consists of checking the BUFR supplied instrumental quality flags set by ESA for ERS or JPL for QuikSCAT data. The data are processed only if they are complete and free from transmission errors or land, sea and/or ice contaminations.

For ERS it is also checked whether no arcings are present for any of the three antennae. Also for ERS, for each antenna k_p must stay smaller than 10%, and the missing packet number must be less than 10 to ensure that enough individual backscatter measurements have been averaged for estimating the value.

For QuikSCAT, it is checked whether the data are likely to be contaminated by rain. Since February 2000, the BUFR product provides a rain flag. This flag, which was developed by NASA/JPL, is based on a multidimensional histogram (MUDH) incorporating various quantities that may be used for the detection of rain (Huddleston and Styles 2000). Examples of such parameters are *mp_rain_probability* (an empirically determined estimate for the probability of a columnar rain rate larger than 2 $m^2 hr^{-1}$; typically values larger than 0.1 indicate rain contamination) and *nof_rain_index* (a rescaled normalized objective function – values larger than 20 give a proxy for rain). Since at the time of implementation, the quality of the JPL rain flag had not been fully confirmed, an alternative flag was established in house. Based on a study in which QuikSCAT winds were compared to collocated ECMWF first guess winds, the following quality flag was introduced: (tested in the `REGROUP` subroutine):

$$L_{rain} = \{ nof\_rain\_index + 200 mp\_rain\_probability > 30 \}$$

Both *mp_rain_probability* and *nof_rain_index* are provided in the BUFR product (for details see Leidner *et al.* 2000). When one of these quantities is missing, the above mentioned condition for the remaining quantity is used.

In addition, for QuikSCAT, it is verified whether inverted winds are well-defined, i.e. whether minima D(u) are sufficiently sharp. In practise this is mainly an issue for cells in the central part of the swath. Data is rejected when the angle between the most likely solution and its most anti-parallel one is less than 135^o (routine `SCAQC`).

After wind inversion, a further check is then done on the backscatter residual associated to the rank-1 solution (or, more precisely, its square root called ‘distance to the cone’). This quantity, representing the misfit between the observed and modelled backscatter values contains both the effects of instrumental noise and of GMF errors. Locally, these errors can become large when the measurements are affected by geophysical parameters not taken into account by the GMF, such as sea-state or intense rainfall. For ERS the distance to the cone is normalized by its expected standard deviation, computed from the distance to the cone and an estimation of the geophysical noise as a function of wind speed and incidence angle. A triplet is considered rejected if the result exceeds a threshold of 3 (three-standard deviation test). For QuikSCAT data such a test is not performed.

Following these quality control checks, a flag is defined. This will be different from zero if any technical problem has been detected during the test of the ESA or JPL flag, or if either the distance to the cone has turned out to be too large (ERS) after wind retrieval or no solutions have even been found. This flag is used in the subsequent processing made in the screening, as described in the corresponding part of the IFS documentation.

In addition to a distance-to-cone test on single observations, a similar test is performed for averages for data within certain time slots. If these averages exceed certain values, all data within the considered time slot is suspected to be affected by an instrument anomaly, since geophysical fluctuations are expected to be averaged out when grouping together large numbers of data points. For ERS, node-wise averages are calculated for the default 4D-Var observation time slot (30 minutes since Cy24r3, 1 hour for older cycles) in the IFS routine `SCAQC`, and its rejection threshold (1.5 times average values) are defined in the IFS routine `SUFGLIM`. For QuikSCAT averages are considered over six-hourly data files and are calculated in the SCAT program `DCONE_OC`, using a threshold of 1.45 for any of WVC’s 4 to 35.

9.6.7 ERS bias corrections

For ERS, two separate bias corrections are prepared in `ERS1IF` to improve the agreement of the retrieved winds with ECMWF model winds. First, a σ^0 bias correction is performed before the wind retrieval, by subtracting constant bias estimates from the raw backscatter measurements as a function of their antenna and WVC numbers. These bias estimates, derived from a routine comparison between the σ^0 measured by the scatterometer and the σ^0 simulated by CMOD5 (or CMOD4 for ERA-40) from the first guess winds of the ECMWF model, are supposed to account both for the variations that may occur in the instrumental calibration in time and for the residual defaults affecting the fit of the geophysical model function in the backscatter space.

A wind-speed bias correction is then added following the wind retrieval, in the form of a table that is dependent on retrieved wind speed and the measurement across-cell number. Its purpose is to match the scatterometer and model wind speeds over the entire wind-speed range so as to avoid introducing any speed-up or slow-down tendency in the assimilation process. Like the σ^0 bias correction, this wind-speed dependent bias correction relies on a direct comparison between scatterometer and model data, in which the wind speeds retrieved with the σ^0 bias correction are fitted as a function of those deduced from the model first guess according to a Maximum Likelihood Estimation (MLE) procedure.

The σ^0 wind-speed bias corrections are defined by two dedicated files read in the initialization subroutine INIERSCA, and containing appropriate coefficients for ERS-1 and ERS-2. Files are both model-cycle and date dependent. However, at the implementation of IFS Cy28r1, the appropriate files had no effect for ERS-2 (i.e. containing only unity correction factors for σ^0, and zeros for wind bias corrections), since the CMOD5 geophysical model function had explicitly been tuned to result in, compared to ECMWF, unbiased winds.

It should be noted that the corrections made are not kept explicitly in the scatterometer ODB reports, where the main outputs are limited to the retrieved wind components as well as to the distances to the cone and the associated quality-control flags. Moreover, the original σ^0 measurements are also stored (and for ERS together with the ESA-retrieved wind speeds and directions), to allow subsequent data monitoring from the analysis feedback file.

9.6.8 QuikSCAT bias correction

For QuikSCAT data no bias corrections in σ^0 space is applied, though, wind-bias corrections are made. Such corrections are performed in three steps. First of all, wind speeds are reduced by 4%:

$$v' = 0.96\, v$$

Where V is wind speed obtained from the INVERT50 subroutine. It was observed that the residual bias between QuikSCAT winds and ECMWF first guess winds depends on the value of *mp_rain_probability* (see Section 9.6.7). The motivation is that, for higher amounts of precipitation, a larger part of the total backscatter is induced by rain, leaving a smaller part for the wind signal. The following correction is applied:

$$v'' = v' - 20 \langle mp\_rain\_probability \rangle$$

where, $\langle\ \rangle$ denotes the average value over the 25 km sub-cells that were taken into account in the inversion (i.e. not over rain-flagged sub-cells). The maximum allowed correction is 2.5 ms^{-1}, which is seldom reached. Finally, for strong winds, QuikSCAT winds were found to be quite higher than their ECMWF first guess counterparts. In order to accommodate this, for winds stronger than 19 ms^{-1} the following correction is applied:

$$v''' = v'' - 0.2(v'' - 19.0)$$

The wind-speed bias corrections are applied in the QSCAT25TO50KM program of the SCAT library.

9.7 GEOSTATIONARY CLEAR-SKY RADIANCES OR BRIGHTNESS TEMPERATURES

9.7.1 Data, data produces and reception at ECMWF

Radiances from geostationary imagers of the Meteosat and the GOES series are used at ECMWF in the form of clear-sky radiances and corresponding brightness temperatures (CSR or CSBT, below referred to as CSR for brevity). The CSRs are area averages of those image pixels of a segment that have been diagnosed as clear-sky. This data pre-processing, including the cloud-detection, is carried by the satellite data providers.

Meteosat data are processed by EUMETSAT (Darmstadt, Germany). CSRs are produced for the water vapour and the infrared channel from hourly images for averaging segments of 16×16 pixels (about 80×80 km^2 areas at sub-satellite point). Data are encoded as BUFR and delivered via the GTS.

Data from the GOES satellites are processed by CIMSS/NESDIS (Madison, USA). CSRs are derived for all channels (visible, water vapour and infrared) and also produced hourly. GOES segments are 11×17 pixels (about 45×45 km^2 areas at sub-satellite point). Data are also BUFR encoded, but currently received at ECMWF via internet/ftp.

The content of the CSRs from Meteosat and GOES comprise clear-sky radiances for the channels indicated above as well as additional information such as location, time, satellite zenith angle, and fraction of clear and cloudy sky in the averaging area. For a complete list, see the data descriptors of the BUFR format. There are differences between the data provided by Meteosat and GOES, and changes to data format and content have occurred during the period for which CSR data have been received and treated. A common BUFR format has been approved (descriptors 301023 for imagers with up to 12 channels, 301024 for imagers with up to 3 channels). It is used by EUMETSAT since 2 December 2002; CIMSS should have provided GOES CSR data in this format from early 2003. Then also the standard deviation of the pixels within the CSR mean is provided as quality indication for all satellites.

After reception, data are recoded at ECMWF into a common BUFR format for storage in MARS and insertion into assimilation (IFS). For GOES data, some simple checks on reasonable time and location specifications have been included at this stage in order to trap erroneous data. In case of occurrence of any incorrect values, the whole data set (corresponding to one image) is rejected.

9.7.2 Overview over METEOSAT and GEOS imager CSR in the ECMWF archives

Table 9.10 gives a short summary of the CSR data stored at ECMWF either in MARS or in ECFS, including the BUFR subtype of the data. For more information on the actual content of the data see BUFR templates, bearing in mind that not all data items which can be encoded according to the CSR BUFR template are actually always provided (i.e. missing values). Incoming data from Meteosat and GOES are currently recoded into one BUFR format being the interface to observation processing and assimilation in IFS. This BUFR was originally designed for the Meteosat CSR. For the GOES data, not all information from the original BUFR can be retained in this BUFR and a change may be therefore useful once the incoming GOES data are encoded in the agreed common BUFR format, using descriptor 301023.

Table 9.10 *ECMWF METEOSAT and GEOS CSR archives*

Satellite	**Time Period**	**Data Type**	**BUFR Subtype**	**Locatio n**
METEOSAT-5 METEOSAT-6 METEOSAT-7	15/05/1996 To ??/05/1997	Geostationary radiances, 32×32 pixel segments, 4 times daily	88	MARS
METEOSAT-5 METEOSAT-6 METEOSAT-7	02/05/1997 To 14/01/2002	Geostationary radiances, 32×32 pixel segments, Hourly	88	MARS
METOSAT-5 METEOSAT-6 METEOSAT-7	Since 25/01/1999	Geostationary clear-sky radiances, 16×16 pixel segments, hourly Including clear and cloudy sky fractions	89	MARS
METEOSAT-2 METEOSAT-3	Periods for ERA	Geostationary clear-sky radiances (as above)	89	ECFS(1)
GEOS-8 GEOS-10	Since 24/10/2001	Clear sky brightness temperatures, 11×17 pixel segments, hourly, Including clear and cloudy sky fractions	89 and original BUFR formats Several changes	ECFS(2)
GOES-8 GOES-10	Since 09/04/2002	Clear sky brightness temperatures, 11×17 pixel segments, hourly, Including clear and cloudy sky fractions	89 And original BUFR formats Several changes	MARS original data on ECFS(2)
ECFS(1): ec:/ERA/era40/obs/bufr/EUM_reproc/$yyyy/$mm/CSR${yyyymmddhh} ECFS(2): ec:/oparch/gicsbt/$yyyymm/$dd/gicsbt..				

9.7.3 Thinning and screening prior to insertion into the assimilation

In order to reduce the data load of the hourly CSR data, the data are screened in a separate task before insertion into assimilation (IFS). This is done by the program GEOS_PRESCREEN (SATRAD library). It decodes the BUFR and applies basic checks on latitude, longitude, time values, and on brightness temperatures being within a physical range. Also, data points are rejected where the value for the water vapour channel brightness temperature is missing. Based on specifications given through namelist input, a geographical thinning may (or may not) be applied for each individual satellite. If switched on, the thinning is performed separately for data falling into hourly timeslots. An overview of the number of remaining valid data points per hour and satellite is printed and the remaining data are encoded into BUFR using the same format as the input file.

9.8 DEFINITIONS

9.8.1 Observation characteristics: instrument specification, retrieval type

Tables 9.11 to 9.19 describe in details how the ODB's instrument specification word is structured. Tables provided are for different observation types.

Table 9.11 *SYNOP instrument specification*

Type	Bit Position	No. of Bits	Value – Description
Instrument Specification	0	10	32 – SYNOP Instrument Code Type
Not Defined	10–30	21	Reserved

Table 9.12 *AIREP instrument specification*

Type	Bit Position	No. of Bits	Value – Description
Instrument Specification	0	10	23 – AIREP Instrument Code Type
Flight Information	10	4	BUFR Code Table 8004 – Flight Phase
Not Defined	10–30	21	Reserved

Table 9.13 *SATOB instrument specification*

Type	Bit Position	No. of Bits	Value – Description
Instrument Specification	0	10	60 – GOES 62 – METEOSAT 63 – Indian SATOB 68 – Japan
I1 (Country Name)	10	4	0 – Europe 1 – Japan 2 – USA 3 – USSR 4 – India
I2I2 (Satellite Indicator Figure)	14	8	4 – METEOSAT 177 – Pretoria 0 – GEOS 3 – Japan 20 – India
Not Defined	22–30	8	Reserved

Table 9.14 *DRIBU instrument specification*

Type	Bit Position	No. of Bits	Value – Description

Instrument Specification	0	10	Not Defined
K1	10	4	Not Defined
K2	14	4	Not Defined
K3	18	4	Not Defined
Not Defined	22–30	8	Reserved

Table 9.15 *TEMP instrument specification*

Type	Bit Position	No. of Bits	Value – Description
Instrument Specification	0	10	Not Defined
Not Defined	10–30	21	Reserved

Table 9.16 *PILOT instrument specification*

Type	Bit Position	No. of Bits	Value – Description
Instrument Specification	0	10	Not Defined
A4	10	4	Not Defined
Not Defined	14–30	17	Reserved

Table 9.17 *SATEM instrument specification*

Type	Bit Position	No. of Bits	Value – Description
Instrument Specification	0	23	77 777 777B
I3	24	4	WMO Manual On Codes, vol II, section II-4-E-8
I4	28	4	Data processing technique. WMO Manual On Codes, vol II, section II-4-E-9
I2I2	32	7	Satellite name. WMO Manual on Codes, vol II, section II-4-E-7
I1	39	4	Country operating satellite. WMO code 1761
IS	43	7	Instrument specification code. Research Manual 5, Table 7.5
Not Defined	50	18	Reserved

Table 9.18 *TOVS instrument specification*

Type	Bit Position	No. of Bits	Value – Description
Instrument Specification	0	10	
A	10	2	0 – No HIRS/2 Data 1 – Clear Radiances are Derived from Clear Spots 2 – Clear Radiances are Derived from the N*

			method
B	12	2	0 – No HIRS/2 Data 1 – All HIRS/2 channels were used 2 – Tropospheric HIRS/2 channels were unusable due to clouds and only stratospheric channels were used
C	14	2	0 – Statistical retrieval method used 1 – Minimum information retrieval used 2 – Minimum information retrieval attempted but statistical retrieval used
V	16	3	0 – No retrieval 1 – HIRS+MSU 2 – HIRS
W	19	3	0 – No retrieval 1 – HIRS+MSU 2 – HIRS
X	22	3	0 – No retrieval 1 – HIRS(1, 2, 3, 8, 9, 16, 17)+MSU(4) 2 – HIRS(1, 2, 3, 8, 9, 16, 17) 3 – HIRS(1, 2, 3, 9, 17)+MSU(4) 4 – HIRS(1, 2, 3, 9, 17)
Y	25	3	0 – No retrieval 1 – HIRS+SSU+MSU(3, 4) 2 – HIRS+MSU(3, 4) 3 – SSU+MSU(3, 4)
Z	28	3	Not Defined
Not Defined	31	1	Reserved

Table 9.19 *SSMI instrument specification*

Type	Bit Position	No. of Bits	Value – Description
Instrument Specification	0	10	Not defined
Not Defined	10–31	22	Reserved

In Table 9.20 the ODB's header retrieval word codes are described.

Table 9.20 *Satellite retrieval codes*

Retrieval Codes	Description
1	Clear
2	Partly Clear
3	Cloudy

9.8.2 Vertical coordinate, pressure, satellite ID and level ID codes

In the ODB the vertical coordinate is expressed by various codes, and Table 9.21 describes those codes.

Table 9.21 *Vertical coordinate*

Vertical Coordinate Codes	Description
1	Pressure (Pa)
2	Height (GPM)
3	Satellite Channel (numeric)
4	Scatterometer Channel (numeric)

Also, the ODB pressure code word is expressed in terms of codes which are defined in Table 9.22.

Table 1 *Pressure codes*

Pressure Codes	Description
0	Sea Level
1	Station Level
2	850 hPa Geopotential
3	700 hPa Geopotential
4	500 hPa Geopotential
5	1000 GPM Pressure
6	2000 GPM Pressure
7	3000 GPM Pressure
8	4000 GPM Pressure
9	900 hPa Geopotential
10	1000 hPa Geopotential
11	500 hPa Geopotential
12	925 hPa Geopotential

Each satellite used in the assimilation has is identification attached to it. The satellite identification codes used are described in Table 9.23.

Table 9.23 *Satellite IDs*

Satellite ID Codes	Description
208/906	NOAA10 – TOVS
235	NOAA10 – SATEM
201/907	NOAA11 – TOVS
236	NOAA11 – SATEM
202/908	NOAA12 – TOVS
237	NOAA12 – SATEM
???/909	NOAA13 – TOVS
206/910	NOAA14 – TOVS
239	NOAA14 – SATEM

205/911	NOAA15
207	NOAA16
208	NOAA17
209	NOAA18
210	NOAA19
222	NOAA20
202/241	DMSP8
203/242	DMSP9
204/243	DMSP10
205/244	DMSP11
245	DMSP12
246	DMSP13
247	DMPS14
248	DMSP15
1022	DMSP16

Upper air observations (TEMP and PILOT) have the level at which the observation was taken defined in terms what it is and that information is stored in the ODB. Details are given in Table 9.24.

.

Table 9.24 *Level ID*

Bit Position	**No. of Bits**	**Value – Description**
0	1	1 – Max Wind Level
1	1	1 – Tropopause
2	1	1 – D Part
3	1	1 – C Part
4	1	1 – B Part
5	1	1 – A Part
6	1	1 – Surface Level
7	1	1 – Significant Wind Level
8	1	1 – Significant Temperature Level
9-31	24	Not Defined

9.8.3 ODB report status, events, flags, codes

The status of each ODB report is described in terms of being active, passive, rejected or blacklisted. The ODB report status word is packed with the 4 bits given in Table 9.25.

Table 9.25 *Report Status*

Bit Position	**No. of Bits**	**Value – Description**
0	1	1 – Report Active
1	1	1 – Passive Report
2	1	1 – Rejected Report

3	1	1 – Blacklisted Report

There is one, 31 bits packed, word for each ODB report to account for various blacklist events. Details are given in Table 9.26.

Table 9.26 *Blacklist Events*

Bit Position	No. of Bits	Value – Description
0	1	1 – Monthly Monitoring
1	1	1 – Constant Blacklisting
2	1	1 – Experimental Blacklisting
3	1	1 – Whitelisting
4	1	1 – Experimental Whitelisting
5	1	1 – Observation Type Blacklisting
6	1	1 – Station ID Blacklisted
7	1	1 – Code Type Blacklisted
8	1	1 – Instrument Type Blacklisted
9	1	1 – Date Blacklisted
10	1	1 – Time Blacklisted
11	1	1 – Latitude Blacklisted
12	1	1 – Longitude Blacklisted
13	1	1 – Station Altitude Blacklisted
14	1	1 – Blacklisted due to Land/Sea Mask
15	1	1 – Blacklisted due to Model Orography
16	1	1 – Blacklisted due to distance from reference point
17–30	14	Not Used

Each ODB report has two words to store report events. Each report event word uses 31 bits. These events are set during observation processing to describe in more details what happened with a report.

The first ODB report event word is described in Table 9.27.

Table 9.27 *Global report events*

Bit Position	No. of Bits	Description (Value)
0	1	1 – No Data in Report
1	1	1 – All Data Rejected
2	1	1 – Bad Reporting Practice
3	1	1 – Rejected due to RDB Flag
4	1	1 – Activated due to RDB Flag
5	1	1 – Activated by Whitelist
6	1	1 – Horizontal Position out of Range
7	1	1 – Vertical Position out of Range
8	1	1 – Time out of Range
9	1	1 – Redundant Report

10	1	1 – Over Land
11	1	1 – Over Sea
12	1	1 – Missing Station Altitude
13	1	1 – Model Surface too far from Station level
14	1	1 – Report Rejected via Namelist
15	1	1 – Failed Q/C
16–30	15	Not Used

The second ODB report event word holds an additional set of events which are now dependent on observation type. Details are given in Tables 9.28 to 9.37.

Table 9.28 *SYNOP report events*

Bit Position	No. of Bits	Value – Description
0–30	31	Not Used

Table 9.29 *AIREP report events*

Bit Position	No. of Bits	Value – Description
0–30	31	Not Used

Table 9.30 *SATOB report events*

Bit Position	No. of Bits	Value – Description
0–30	31	Not Used

Table 9.31 *DRIBU report events*

Bit Position	No. of Bits	Value – Description
0–30	31	Not Used

Table 9.32 *TEMP report events*

Bit Position	No. of Bits	Value – Description
0	1	1 – Old Style Z Bias Correction Applied
1	1	1 – New Style T Bias Correction Applied
2	1	1 – RH Bias Correction Applied
3–30	28	Not Used

Table 9.33 *PILOT report events*

Bit Position	No. of Bits	Value – Description
0	1	1 – American Wind Profiler
1	1	1 – European Wind Profiler
2–30	29	Not Used

Table 9.34 *SATEM report events*

Bit Position	No. of Bits	Value – Description
0	1	1 – Thinned Report
1–30	30	Not Used

Table 9.35 *PAOB report events*

Bit Position	No. of Bits	Value – Description
0–30	31	Not Used

Table 9.36 *SCAT report events*

Bit Position	No. of Bits	Value – Description
0	1	1 – Thinned Report
1	1	1 – Reported Wind Directions too Close
2	1	1 – Report not in QuikScat Sweet Spots
3	1	1 – Report Contaminated by Rain
4–30	29	Not Used

Table 9.37 *Raw radiance report events*

Bit Position	No. of Bits	Value – Description
0–30	31	Not Used

The ODB report RDB flag word is 30 bits packed which contains flags for five report parameters: latitude, longitude, date, time and altitude. Each parameter occupies 6 bits with further stratification which is identical for every parameter as indicated in Table 9.38.

Table 9.38 *RDB report (latitude, longitude, date, time and altitude) flags*

Parameter	No. of Bits	Bit Position			
			Bit Position	**No. of Bits**	**Value – Description**
Latitude Longitude Date Time Altitude	6 6 6 6 6	0+ 6+ 12+ 18+ 24+	0	1	0 – No Human Monitoring Substitution 1 – Human Monitoring Substitution
			+1	1	0 – No Q/C Substitution 1 – Q/C Substitution
			+2	1	0 – Override Flag not Set 1 – Override Flag Set
			+3	2	0 – Parameter Correct 1 – Parameter Probably Correct 2 – Parameter Probably Incorrect 3 – Parameter Incorrect
			+5	1	0 – Parameter Flag Set by Q/C or not Checked 1 – Parameter Flag Set by Human Monitoring

9.8.4 Datum status, events, RDB and analysis flags

The status of each datum, like report status, is described in terms of being: active, passive, rejected or blacklisted. Table 9.39 shows that the ODB datum status is a packed word with 4 bits used to describe its status.

Table 9.39 *Datum status*

Bit Position	No. of Bits	Value – Description
0	1	1 – Report Active
1	1	1 – Passive Report
2	1	1 – Rejected Report
3	1	1 – Blacklisted Report

There are two ODB words reserved for datum events. They both use 31 bits each to store relevant information. The first event word has the same structure for all observation types, whereas the second event word is observation type dependent. Tables 9.40 to 9.50 describe the event words structures.

Table 9.40 *Global datum events*

Bit Position	No. of Bits	Value – Description
0	1	1 – Missing Vertical Coordinate
1	1	1 – Missing Observed Value
2	1	1 – Missing Background (First Guess) Value
3	1	1 – Rejected due to RDB Flag
4	1	1 – Activated due to RDB Flag
5	1	1 – Activated by Whitelist
6	1	1 – Bad Reporting Practice
7	1	1 – Vertical Position out of range
8	1	1 – Reference Level Position out of Range
9	1	1 – Too Big First Guess Departure
10	1	1 – Too Big Departure in Assimilation
11	1	1 – Too Big Observation Error
12	1	1 – Redundant Datum
13	1	1 – Redundant Level
14	1	1 – Report Over Land
15	1	1 – Report Over Sea
16	1	1 – Not Analysis Variable
17	1	1 – Duplicate Datum/Level
18	1	1 – Too Many Surface Data
19	1	1 – Multi Level Check
20	1	1 – Level Selection
21	1	1 – Vertical Consistency Check
22	1	1 – Vertical Coordinate Changed from Z to P
23	1	1 – Datum Rejected via Namelist
24	1	1 – Combined Flagging
25	1	1 – Datum Rejected due to Rejected Report

26	1	1 – Variational QC Performed
27	1	1 – Observation Error Increased
28–30	3	Not Used

Table 9.41 *SYNOP datum events*

Bit Position	**No. of Bits**	**Value – Description**
0	1	1 – Bias Corrected Ps
1–30	30	Not Used

Table 9.42 *AIREP datum events*

Bit Position	**No. of Bits**	**Value – Description**
0–30	31	Not Used

Table 9.43 *SATOB datum events*

Bit Position	**No. of Bits**	**Value – Description**
0–30	31	Not Used

Table 9.44 *DRIBU datum events*

Bit Position	**No. of Bits**	**Value – Description**
0	1	1 – Bias Corrected Ps
1–30	30	Not Used

Table 9.45 *TEMP datum events*

Bit Position	**No. of Bits**	**Value – Description**
0	1	1 – Bias Corrected Value Used
1–30	30	Not Used

Table 9.46 *PILOT datum events*

Bit Position	**No. of Bits**	**Value – Description**
0–30	31	Not Used

Table 9.47 *SATEM datum events*

Bit Position	**No. of Bits**	**Value – Description**
0	1	1 – Not Predefined Layer
1	1	1 – Layer Formed by Thinning
2	1	1 – Layer Formed by Summing Up
3	1	1 – Channel Not Used in Analysis

4	1	1 – Overwritten by ADVAR
5–30	26	Not Used

Table 9.48 *PAOB datum events*

Bit Position	No. of Bits	Value – Description
0–30	31	Not Used

Table 9.49 *SCAT datum events*

Bit Position	No. of Bits	Value – Description
0–30	31	Not Used

Table 9.50 *Raw radiances datum events*

Bit Position	No. of Bits	Value – Description
0–30	31	Not Used

Furthermore, each datum in the ODB has a blacklist event word. This word uses 31 bits to describe various blacklist events as indicated in Table 9.51.

Table 9.51 *Datum blacklist events*

Bit Position	No. of Bits	Value – Description
0	1	1 – Pressure Blacklisted
1	1	1 – Variable Blacklisted
2	1	1 – Blacklisted due to Pressure Code
3	1	1 – Blacklisted due to Distance from Reference Point
4	1	1 – Blacklisted due to Type of Vertical Coordinate
5	1	1 – Blacklisted due to Observed Value
6	1	1 – Blacklisted due to First Guess departure
7–30	24	Not Used

For each datum in ODB there is an RDB flag word which holds flags for pressure (vertical coordinate) and the datum itself. This is packed word with 30 bits used – see Table 9.52. Pressure and datum RDB flags use 15 bits each. Thus pressure RDB flag starts at bit position 0, whereas the datum flag starts at bit position 15. Each 15 bits structure is further stratified in exactly the same way for both parameters:

Table 9.52 *RDB pressure (vertical coordinate) and datum flags*

Parameter	**No. of Bits**	**Bit Position**			
			Bit Position	**No. of Bits**	**Value – Description**
			0	1	0 – No Human Monitoring Substitution 1 – Human Monitoring Substitution
Pressure	15	0+	+1	1	0 – No Q/C Substitution 1 – Q/C Substitution

Datum	15	15+	+2	1	0 – Override Flag not Set 1 – Override Flag Set
			+3	2	0 – Correct 1 – Probably Correct 2 – Probably Incorrect 3 – Parameter Incorrect
			+5	1	0 – Flag Set by Q/C or not Checked 1 – Flag Set by Human Monitoring
			+6	2	0 – Previous Analysis judged it correct 1 – Previous Analysis judged it probably correct 2 – Previous Analysis judged it probably incorrect 3 – Previous Analysis judged it incorrect
			+8	1	0 – Not used by previous analysis 1 – Used by previous analysis
			+9	5	Not Used

In addition to RDB datum flags there is a word in ODB to store analysis flags. There are five types of analysis flags: final analysis, first guess, departure, variational q/c and blacklist flags. Each flag occupies 4 bits and the exact description is given in Table 9.53

Table 9.53 *Analysis flags*

Flag Type	Bit Position	No. of Bits	Value – Description
Final	0	4	0 – Correct 1 – Probably correct 2 – Probably incorrect 3 – Incorrect
First Guess	4	4	0 – Correct 1 – Probably correct 2 – Probably incorrect 3 – Incorrect
Departure	8	4	0 – Correct 1 – Probably correct 2 – Probably incorrect 3 – Incorrect
Variational Q/C	12	4	0 – Correct 1 – Probably correct 2 – Probably incorrect 3 – Incorrect
Blacklist	16	4	0 – Correct 1 – Probably correct 2 – Probably incorrect 3 – Incorrect
Not Defined	20	11	Reserved

REFERENCES

Gaffard, C. and Roquet, H., 1995: Impact of the ERS-1 scatterometer wind data on the ECMWF 3D-Var assimilation system. ECMWF Tech. Memo. 217.

Le Meur, D., Isaksen, L. and Stoffelen, A., 1997: Wind speed calibration of ERS scatterometer data for assimilation purposes at ECMWF. Proceedings of the ESA CEOS wind and wave validation workshop held at ESA/ESTEC, Noordwijk, the Netherlands, from June 3-5 1997.

Stoffelen, A. and Anderson, D., 1997: Ambiguity removal and assimilation of scatterometer data. Q. J. R. Meteorol. Soc., 123, 491-518.