1 `setup.sh` parameters

Due to the intention of overwriting some parameters in the `CONTROL` file with the passed command line arguments, the following parameters are already defined in Section 3: `ECUID`, `ECGID`, `GATEWAY`, `DESTINATION`, `CONTROLFILE`. All other parameters are listed below:

Name:	Type:	Value range:	Default:
`TARGET`	String	local, ecgate, cca, ccb	None
Description: Defines which flex_extract application mode will be used. Local = local mode; egate = Remote or Gateway mode, cca = Remote or Gateway mode. Whether the local mode is for public or member state users doesn't matter during the installation process.			
Name:	Type:	Value range:	Default:
`MAKEFILE`	String	any Makefile present, according to environment	Makefile.gfortran
Description: Name of the makefile in source/fortran directory to be used for compiling the Fortran program. "Makefile.gfortran" has the configuration for the ecgate environment. For local server versions there are templates "Makefile.local.gfortran" and "Makefile.local.ifort" which have to be adapted (pathes to the eccodes library) by the user.			
Name:	Type:	Value range:	Default:
`INSTALLDIR`	String	full path	`$HOME` on ECMWF server and current flex_extract root path on local server
Description: Root path where flex_extract should be installed. If it is not set it will be set to `$HOME` on ECMWF server and set to the current flex_extract root path on local servers.			
Name:	Type:	Value range:	Default:
`JOB_TEMPLAT`	String	job.template	job.template
Description: The rudimentary template file to create a batch job template for submission to ECMWF servers. Should not be changed since it is optimized for ECMWF server. (Remote and Gateway mode)			

2 `run.sh` parameters

Due to the intention of overwriting some parameters in the `CONTROL` file with the passed command line arguments, the following parameters are already defined in Section 3: `START_DATE`, `END_DATE`, `DATE_CHUNK`, `BASETIME`, `STEP`, `LEVELIST`, `DEBUG`, `OPER`, `REQUEST`, `RRINT`. All other parameters are listed below:

Name:	Type:	Value range:	Default:
JOB_CHUNK	Integer	depends on resolution	None

Description:
# of days to be retrieved within a single job. Can be selected to start the submit script once and let it automatically divide the time period in smaller job chunks. Might be very useful for example if one would like to retrieve one month with 0.1○ space resolution and 1h time resolution. Then only 1 day per job is possible.

Name:	Type:	Value range:	Default:
CONTROLFILE	String	any CONTROL file	CONTROL_ERA5

Description:
The file with all CONTROL parameters.

Name:	Type:	Value range:	Default:
AREA	Double [f/f/f/f]	any float within lat and lon boundaries	”“

Description:
Domain defined as north/west/south/east

Name:	Type:	Value range:	Default:
PUBLIC	Integer	[0,1]	0

Description:
Public mode - retrieves the public datasets. Important: This is necessary to select for each public user and member state user who wants to retrieve a public dataset.

Name:	Type:	Value range:	Default:
INPUTDIR	String	any path	None

Description:
Path to the temporary directory for the retrieval grib files and other processing files. The temporary directory will be created if it does not already exist.

Name:	Type:	Value range:	Default:
OUTPUTDIR	String	any path	None

Description:
Path to the final directory where the final FLEXPART ready input files are stored. The final output directory will be created if it does not already exist.

Name:	Type:	Value range:	Default:
PPID	Integer	-	None

Description:
This is the specify parent process id of a single flex_extract run to identify the files. It is the second number in the GRIB files. This is usually only necessary if the GRIB data were retrieved and a rerun of prepare_flexpart has to be done. Then PPID is used to select the files.

Name:	Type:	Value range:	Default:
JOB_TEMPLATE	String	job.temp	job.temp

Description:
The job template file which are adapted to be submitted to the batch system on ECMWF server.

Name:	Type:	Value range:	Default:
QUEUE	String	ecgate; cca; ccb	None

Description:
The ECMWF server name for submission of the job script to the batch system.

3 CONTROL file parameters

3.1 User section

Name:	Type:	Value range:	Default:
ECUID	String	-	None

Description:
-

Name:	Type:	Value range:	Default:
ECGID	String	-	None

Description:
-

Name:	Type:	Value range:	Default:
DESTINATION	String	<name>@ generic <scope>	None

Description:
The remote destination which is used to transfer files from ECMWF server to local gateway server. This has to be set up by users on the local gateway server. See installation instruction for more information of how to set it up.

Name:	Type:	Value range:	Default:
GATEWAY	String	IP address / name	None

Description:
The gateway server users is using.

3.2 Data section

Name:	Type:	Value range:	Default:
CLASS	String [xx]	EI, E4, EA, EP, OD	None

Description:
ECMWF data classification identifier for datasets. EI: ERA-Interim; E4: ERA-40; EA: ERA5; EP: CERA-20C; OD: operational;

Name:	Type:	Value range:	Default:
DATASET	String	CERA20C, INTERIM	None

Description:
This keyword has to be defined as soon as you want to retrieve public datasets. Public data are stored on a different server and are available for everyone after registration at ECMWF (see installation plan).

Name:	Type:	Value range:	Default:
STREAM	String [xxxx]	OPER, ENFO, ENDO	None

Description:
Identifies the forecasting system used to generate the data. Most times the operational data stream OPER is appropriate. Use ENFO for ensemble forecasts and ENDO for CERA data.

Name:	Type:	Value range:	Default:
EXPVER	Integer	-	1

Description:
Experiment number, necessary for R&D experiments, E-suites

Name:	Type:	Value range:	Default:
NUMBER	String [i/to/i]	depends on availability	"OFF"

Description:
In most cases this can be set to "OFF" but for access to individual ensemble members of ensemble forecasts it is necessary to select numbers. Note, however, that model level data are not stored in MARS for individual ensemble members except the control run. They exist only for a few days before they are discarded. For retrieving CERA data, a number has to be selected explicitly. Usually "0" is enough. For operational data just one number or a list of number can be selected. You can use the syntax "1/to/9" for example.

Name:	Type:	Value range:	Default:
FORMAT	String	[GRIB1, GRIB2]	GRIB1

Description:
Output format (either GRIB1 or GRIB2). Use GRIB2 only when using FLEXPART versions >9.2 or FLEXPART hast to be adapted for reading GRIB2. GRIB2 is only available for 3D model level fields, not for surface fields.

3.3 Data field section

Name:	Type:	Value range:	Default:
TYPE	list of String [xx xx ... xx]	[AN, FC, CV, CF, 4V, PF]	None

Description:
A list of field types for each retrieving hour per day. E.g. AN FC FC FC AN FC FC FC for a day with 3-hourly retrieval DTIME . At 0 and 12 UTC we retrieve analysis fields and at 3, 6, 9, 15, 18 and 21 UTC forecast fields. So far flex_extract is using ANalysis (AN), ForeCast (FC), Control Forecast (CF), Perturbed Forecast (PF), Calibration/Validation forecast (CV) and 4D Variational analysis (4V). Other types might be also possible but were not tested. The analysis fields are usually (depending on dataset) available at 00/06/12/18 UTC. For better temporal resolution, the time in-between the AN fields can be filled with forecasts (FC). Additionally, it is recommended to use analysis fields only at 00 and 12 UTC and fill the rest of the times with other field types, such as forecasts.

Name:	Type:	Value range:	Default:
TIME	list of String [xx xx ... xx]	[00 – 23]	None

Description:
The time of the corresponding field type (TYPE) in hours. It is important to set the correct forecast times. E.g. In ERA-Interim dataset, to get forecast fields between 1 and 11 hours, forecast times starting at 00 UTC are normally used. To get fields between 13 and 23 hours, forecast times starting at 12 UTC are used. Mostly there are two forecasts starting per day. Has to have the same amount of values as in TYPE! The start times of forecasts can vary from one dataset to another. It is also possible to have just one forecast per day, such as in the CERA dataset.

Name:	Type:	Value range:	Default:
STEP	list of String [xx xx ... xx]	[00 – max available STEP in dataset]	None

Description:
This is the forecast time step in hours for each corresponding field type (TYPE). Counting of the steps starts from the forecast times. E.g. In ERA-Interim, for forecasts at 3, 6, 9 UTC the STEPS 3,6 and 9 are used and the forecast TIME 00 UTC. Has to have the same amount of values as in TYPE and TIME! For analysis (AN) fields the STEP has to be 00 always!

Name:	Type:	Value range:	Default:
MAXSTEP	Integer	> 24	None

Description:
This parameter allows to retrieve data from forecasts longer than 24 hours. With MAXSTEP > 24 the forecasts from different days overlap, the naming scheme of the output files changes from <PREFIX>YYMMDD to <PREFIX>YYMMDD.HH.SSS where HH is the hour of the start of the forecast and SSS is the forecast step in hours. Optional parameter.

3.4 Time section

Name:	Type:	Value range:	Default:
START_DATE	String [YYYYMMDD]	depends on dataset	None

Description:
The first day of the retrieval period. If END_DATE is set, START_DATE must be greater than END_DATE.

Name:	Type:	Value range:	Default:
END_DATE	String [YYYYMMDD]	depends on dataset	None

Description:
The last day of the retrieval period. For a one day retrieval it has to be the same date as START_DATE. If not set, it is automatically equal to START_DATE. Do not have to be set. If set, it has to be greater or equal than START_DATE.

Name:	Type:	Value range:	Default:
DTIME	Integer	1,3,6	None

Description:
Time step of retrieved data. Detects TYPE, TIME, STEP, ACCTYPE, ACCTIME according to DTIME. Therefore CONTROL file can have more values than needed. Available resolution in time depends on availability in the dataset. Coarser resolution can always be selected.

Name:	Type:	Value range:	Default:
DATE_CHUNK	Integer	depends on resolution	3

Description:
Maximum number of days retrieved within one MARS request. This number is limited due to maximum allowed memory and time limit for one MARS request. Be careful in changing this number. It can be larger for reanalysis data but may be too large for very high resolution retrievals.

Name:	Type:	Value range:	Default:
BASETIME	Integer	[0;12]	None

Description:
This parameter is intended for half-day retrievals. Only half a day will be retrieved starting from BASETIME going back 12 hours. E.g. 20180510 with a BASETIME = 00 would lead to a data retrieval of 20180509 12h until 20180510 00h. Can be set to 00 or 12 only.

3.5 General section

Name:	Type:	Value range:	Default:
DEBUG	Integer	[0,1]	0

Description:
If set to "1" all temporary output files from the mars requests are kept and some extra information is written out to the log file. Usually all temporary files except the FLEXPART ready output files are deleted at the end of flex_extract .

Name:	Type:	Value range:	Default:
REQUEST	Integer	[0,1,2]	2

Description:
This parameter allows to write out the mars requests in a separate file. The requests are stored in the file "mars_requests.csv". Possible selections are: 1: normal data retrieval; 2: neglect data retrieval and just writes out the mars requests; 3: retrieve data and write out the MARS requests.

Name:	Type:	Value range:	Default:
PUBLIC	Integer	[0,1]	0

Description:
This specifies the selection of the kind of ECMWF Web Api access and therefore the kind of available datasets. Public datasets (1) and Member-state datasets (0). Selecting the public access method, the DATASET parameter has explicitly to be set to select the dataset. (CLASS is not enough.) ATTENTION: For public datasets, users have to accept the license of the dataset to be retrieved. See here for available datasets and their licenses: https://software.ecmwf.int/wiki/display/WEBAPI/Available+ECMWF+Public+Datasets

Name:	Type:	Value range:	Default:
OPER	Integer	[0,1]	0

Description:
Switch to prepare the operational job script. Start date, end date and BASETIME will be prepared with environment variables at ecgate. This is only necessary if extraction of half-day retrievals should be done automatically. Specific extra feature which is usually not used by normal flex_extract users.

Name:	Type:	Value range:	Default:
ECSTORAGE	Integer	[0,1]	0

Description:
Switch to store FLEXPART ready output files in the ECFS file system. Mind the data limit.

Name:	Type:	Value range:	Default:
ECTRANS	Integer	[0,1]	0

Description:
Switch to transfer FLEXPART ready input files to the gateway server. The gateway server has to be up and running. A destination has to be installed. See installation instructions for more information.

Name:	Type:	Value range:	Default:
PREFIX	String [xx]	anything	EN

Description:
Prefix of FLEXPART ready input files. The output files usually have the format <PREFIX>YYMMDDHH. Sometimes it also has some more information, e.g. to distinguish between numbers or in pure forecast mode.

Name:	Type:	Value range:	Default:
ECFSDIR	String	any path available	ectmp: /$USER/econdemand/

Description:
The destination directory on ECFS file system if the retrieved data should be stored on ECMWF servers. This is only used if parameter ECSTORAGE is set to "1".

Name:	Type:	Value range:	Default:
MAILOPS	list of String [m1, m2, . . .]	any number of mail addresses, seperated by comma	["$USER"]

Description:
Email list for operational log files on ECMWF servers. The email addresses should be seperated by a comma. For the ECMWF server it is enough to give $USER as input. On local system an actual email is preferred to operate correctly.

Name:	Type:	Value range:	Default:
MAILFAIL	list of String [m1, m2, . . .]	any number of mail addresses, seperated by comma	["$USER"]

Description:
Email list for operational log files on ECMWF servers. The email addresses should be seperated by a comma. For the ECMWF server it is enough to give $USER as input. On local system an actual email is prefered to operate correctly.

3.6 Flux data section

Name:	Type:	Value range:	Default:
ACCTYPE	String	[FC, CV, CF, 4V, PF] - just one of them	None

Description:
The type of field for accumulated data retrieval. The accumulated data fields are only available from forecast fields. Therefore it is seperated from the normal TYPE parameter to allow, for example, retrievment solely of hourly analysis fields in the Era5 dataset. For downward compatibility to older versions: if ACCTYPE is not specified, the default value is taken form the second position of the original TYPE parameter (TYPE[2]). NOTE: This is important at the moment the original TYPE parameter is changed from FC to AN for example.

Name:	Type:	Value range:	Default:
ACCTIME	String	E.g.: Operational, Era-Interim: 00/12 CERA: 18 ERA5 06/18	None

Description:
Start times of accumulated forecast fields (fluxes). The starting times of forecast fields varies between different datasets. For downward compatibility to older versions: if ACCTIME is not specified, the default value is 00/12 for ERA-Interim and operational data, 06/18 for ERA5 data and 18 for CERA-20C data.

Name:	Type:	Value range:	Default:
ACCMAXSTEP	Integer	12, 24 or larger	None

Description:
This parameter specifies the maximum step in hours for a specific accumulated forecast start time. For daily retrievals with one forecast time the step shouldn´t be greater than 24h. For two forecast times the ACCMAXSTEP should be 12. If the parameter MAXSTEP is specified to retrieve forecasts longer than 24 hours, this parameter must have the same value. For downward compatibility to older versions: if ACCMAXSTEP is not specified, the default value is set to 12 for ERA5, Era-Interim and operational data or 24 for CERA-20C data, according to one or two forecast times of the dataset.

Name:	Type:	Value range:	Default:
RRINT	Integer	[0,1]	None

Description:
Switch to select method of disaggregation for precipitation fields. Old method (0) with a simple linear disaggregation or new method (1) with 2 additional subsequent intervals per time interval. For more information see article (**?**)

3.7 Domain section

Name:	Type:	Value range:	Default:
UPPER	String	-90+GRID to 90	None

Description:
Latitude of upper right corner of grid area.

Name:	Type:	Value range:	Default:
LOWER	String	-90+GRID to 90	None

Description:
Latitude of lower left corner of grid area.

Name:	Type:	Value range:	Default:
LEFT	String	-180+GRID to 180	None

Description:
Longitude of lower left corner of grid area. For cyclic (global) grids, use e.g. LEFT = - 180 + GRID, RIGHT = 180. For noncyclic grids crossing the dateline (180W), RIGHT may be smaller than LEFT.

Name:	Type:	Value range:	Default:
RIGHT	String	-180+GRID to 180	None

Description:
Longitude of upper right corner of grid area. For cyclic (global) grids, use e.g. LEFT = - 180 + GRID, RIGHT = 180. For noncyclic grids crossing the dateline (180W), RIGHT may be smaller than LEFT.

Name:	Type:	Value range:	Default:
LEVEL	Integer	60, 91, 137 depends on dataset	None

Description:
Maximum number of vertical levels. ERA-Interim has 60 levels; ERA5 has 137 levels; CERA-20C has 91 levels; Operational data can have different number of model levels depending on the date. Check upfront in the MARS catalogue. If LEVELIST is set , this parameter is not needed.

Name:	Type:	Value range:	Default:
LEVELIST	String [start/to/end]	1/to/137	None

Description:
List of vertical levels for MARS request. It can be a subset of levels but it has to include the maximum level (end). If full list of levels is needed and parameter LEVEL is set, the LEVELIST parameter is not needed. “”end“” has to be the maximum number of possible levels and has to be the same as in LEVEL, if specified.

Name:	Type:	Value range:	Default:
GRID	String[i/i]	0.1○ - appropriate value e.g. 2○	None

Description:
Horizontal resolution of Latitude/Longitude grid. Best possible resolution varies for different datasets. E.g in operational data it´s 0.1○ whereas in Era-Interim it is 0.75○. It can be specified in tenth degrees (1○) or thousandth degrees (1000 for 1○).

Name:	Type:	Value range:	Default:
RESOL	String	depends on GRID	None

Description:
Horizontal resolution of spectral fields. Specifies the desired triangular truncation of retrieved data, before carrying out any other selected post-processing.

Name:	Type:	Value range:	Default:
SMOOTH	Integer	approriate number	0

Description:
Spectral truncation of ETADOT after calculation on Gaussian grid. For more information see: Sardeshmukh and Hoskins, (1984): Spatial Smoothing on the Sphere, Monthly Weather Review, Vol. 112, No.12, P. 2524-2529

3.8 Vertical wind section

Name:	Type:	Value range:	Default:
GAUSS	Integer	[0,1]	0

Description:
A switch to calculate ETADOT from Lat/Lon grid (0) or from Gaussian grid (1).

Name:	Type:	Value range:	Default:
ACCURACY	Integer	-	24

Description:
Specifies the number of bits per value to be used in the generated GRIB coded fields.

Name:	Type:	Value range:	Default:
OMEGA	Integer	[0,1]	0

Description:
Retrieve Omega from MARS and put it to file OMEGAyymmddhh. Only useful for debugging reasons.

Name:	Type:	Value range:	Default:
OMEGADIFF	Integer	[0,1]	0

Description:
Calculate Omega and Dps/Dt from continuity equation for diagnostic purposes and include it in file OMEGAyymmddhh. Only useful for debugging reasons.

Name:	Type:	Value range:	Default:
ETA	Integer	[0,1]	0

Description:
Switch to read ETADOT precalculated by ECMWF and multiply it with DPDETA to be compatible with ETADOT calculated from continuity equation. ETADOT calculation from continuity equation on either Gaussian or lat/lon grid is disabled unless ETADIFF is set to 1 as well. ETADOT is available in ERA5, CERA-20C and operational datasets. Precalculated ETADOT is operationally available from September 2008 onwards. However, it is not available in the ERA-40 and ERA-Interim dataset. If ETA is selected in the last two datasets, flex_extract fails.

Name:	Type:	Value range:	Default:
ETADIFF	Integer	[0,1]	0

Description:
Switch to calculate etadot and Dps/Dt from continuity equation for diagnostic purposes and include it in file `ETAyymmddhh`. Expensive option, only for debugging purposes.

Name:	Type:	Value range:	Default:
DPDETA	Integer	[0,1]	1

Description:
Switch to multiply etadot with dpdeta – this is the default. In some future version this may change.

Name:	Type:	Value range:	Default:
ETAPAR	Integer	77	77

Description:
GRIB parameter number for ETADOT/DPDETA.

3.9 Additional data section

Name:	Type:	Value range:	Default:
CWC	Integer	[0,1]	0

Description:
Switch to retrieve cloud water content (sum of cloud liquid water and cloud ice) (1) or not (0).

Name:	Type:	Value range:	Default:
ADDPAR	String [p1/p2/../pn]		None

Description:
Additional optional surface parameters (2D fields, non-accumulated) Mostly:
27/28/173/186/187/188/235/139/39. Parameters can be specified as the Integer IDs or with the short names.

Name:	Type:	Value range:	Default:
DOUBLEELDA	Integer	[0,1]	0

Description:
Switch to select the calculation of extra ensemble members for the ELDA stream. It doubles the amount of retrieved ensemble members. Each ensemble member is used to create a new synthesized ensemble member by subtracting 2 * (current time step value - last time step value) from the current time step value.