Chapter 2: Description of the Reanalysis Systems

Chapter lead authors

Jonathon S. Wright	Department of Earth System Science, Tsinghua University	China
Masatomo Fujiwara	Faculty of Environmental Earth Science, Hokkaido University	Japan
Craig Long	Climate Prediction Center, National Oceanic and Atmospheric Administration *(retired)*	USA

Co-authors

James Anstey	Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada	Canada
Simon Chabrillat	Royal Belgian Institute for Space Aeronomy	Belgium
Gilbert P. Compo	(1) Cooperative Institute for Research in Environmental Sciences, University of Colorado (2) Physical Sciences Laboratory, National Oceanic and Atmospheric Administration	USA
Rossana Dragani	European Centre for Medium-Range Weather Forecasts	UK
Wesley Ebisuzaki	National Oceanic and Atmospheric Administration	USA
Yayoi Harada	Japan Meteorological Agency	Japan
Chiaki Kobayashi	Japan Meteorological Agency	Japan
Will McCarty	National Aeronautics and Space Administration	USA
Andrea Molod	National Aeronautics and Space Administration	USA
Kazutoshi Onogi	Research Center for Advanced Science and Technology, the University of Tokyo *previously at Japan Meteorological Agency*	Japan
Steven Pawson	National Aeronautics and Space Administration	USA
Adrian Simmons	European Centre for Medium-Range Weather Forecasts	UK
David G.H. Tan	European Centre for Medium-Range Weather Forecasts *(retired)*	UK
Susann Tegtmeier	University of Saskatchewan	Canada
Krzysztof Wargan	(1) National Aeronautics and Space Administration (2) Science Systems and Applications, Inc.	USA
Jeffrey S. Whitaker	National Oceanic and Atmospheric Administration	USA
Cheng-Zhi Zou	National Oceanic and Atmospheric Administration	USA

Abstract. Information on key components of twelve global atmospheric reanalysis systems with output data available in 2018 is summarized, including brief descriptions of the forecast models, assimilation schemes, and observational data used in these systems. Details of the execution streams and archived data products are also provided. Tables are used extensively to facilitate comparison of different reanalysis systems, and are arranged so that readers interested in one or more systems can easily find and compare relevant information. The information in this chapter will be referred to in the interpretation of results presented in the other chapters of this S-RIP report. This chapter is not intended to provide a complete description of the reanalysis systems; readers requiring further details are encouraged to refer to the cited literature and the online documentation provided for each system. A condensed version of the material in this chapter has been provided by Fujiwara *et al.* (2017). A longer and more detailed version (denoted Chapter 2E) is provided as an electronic file on the S-RIP website at **https://s-rip.ees.hokudai.ac.jp** (being migrated to **https://s-rip.github.io**).

Fujiwara *et al.* (2017) have published a shortened version of this chapter.

Content

2.1 Introduction

An atmospheric reanalysis system consists of a global forecast model, input observations, and an assimilation scheme that blends input observations with short-range forecasts. These systems produce global atmospheric data that represents best estimates (analyses) of past atmospheric states. The information collected in these analyses is then propagated forward in time and space by subsequent forecasts. In this chapter, we provide summary descriptions of the key components of the twelve global atmospheric reanalysis systems listed in **Table 2.1**. Our descriptions of these systems are by necessity incomplete. Further details may be found in the cited literature, particularly the publications listed in **Table 2.1**, or in the technical documentation compiled and provided by the reanalysis centres. A list of the acronyms used in this chapter is provided in the *Appendix* at the end of this chapter.

We classify reanalysis systems according to their observational inputs and temporal coverage. The three classes of reanalysis systems include "full input" systems (which assimilate surface and upper-air conventional and satellite data), "conventional input" systems (which assimilate surface and upper air conventional data but do not assimilate satellite data), and "surface input" systems (which assimilate surface data only). Some reanalysis centres also provide companion "AMIP-type" simulations, which do not assimilate any observational data and are constrained by applying observed sea surface temperatures, sea ice, and other boundary or forcing conditions on the atmospheric forecast model. We also broadly distinguish reanalyses of the "satellite era" (1979 - present) and reanalyses that provide data for dates before January 1979, with the latter referred to as "extended" reanalyses. All reanalyses are affected by changes in assimilated observations, as discussed below, but such temporal inconsistencies are especially important to keep in mind for extended reanalyses that assimilate satellite data during the later part of the record.

Four reanalyses produced by ECMWF are considered: ERA-40, ERA-Interim, ERA-20C, and ERA5. ERA-40 (*Uppala et al.*, 2005) is an extended full input reanalysis covering 45 years from September 1957 through August 2002. No satellite data were assimilated for dates prior to January 1973; ERA-40 is therefore a conventional input reanalysis from September 1957 through December 1972. ERA-40 represented an important improvement relative to the first generation of modern reanalysis systems and continues to be used in many studies that require long-term atmospheric data. ERA-Interim (*Dee et al.*, 2011) is a full input reanalysis of the satellite era (1979 -present) that applies several corrections and modifications to the system used for ERA-40. Major focus areas during the production of ERA-Interim included improving the representations of the hydrologic cycle and the stratospheric circulation relative to ERA-40, as well as improving the consistency of the reanalysis products in time. ERA5 (*Hersbach et al.*, 2020) is intended as the full input replacement for ERA-Interim, with finer resolution in time and space (see also *Section 2.2* and *Appendix A*) and the ability to assimilate several new types of observational data (see also *Section 2.4*). ERA5 is an extended reanalysis covering 1950 to present, and the first full input reanalysis to be conducted together with an ensemble of data assimilations, which allows for a more robust characterization of uncertainty in the analysis state. Some ERA5 data have been available since July 2018, ERA5 data from 1979 onward have been available since January 2019, and a preliminary version of ERA5 1950 - 1978 data have been available since November 2020. Products from ERA5 are evaluated in some chapters of this report. While ERA5 could not be included in the interim version of this chapter (*Fujiwara et al.*, 2017), we document its structure here in tandem with the other reanalysis systems considered by S-RIP. ERA-20C (*Poli et al.*, 2016) is a surface input reanalysis of the twentieth century (1900 - 2010). ERA-20C directly assimilates only surface pressure and surface wind observations, and can therefore generate reanalyses of the atmospheric state that extend further backward in time. Data from ERA-20C extend up to 0.01 hPa, but the lack of upper-air observational constraints means that these data should be used with caution in the upper troposphere and above. We omit the earlier ECMWF reanalysis products FGGE (*Bengtsson et al.*, 1982) and ERA-15 (*Gibson et al.*, 1997), as well as recent coupled atmosphere–ocean reanalysis efforts at ECMWF using the CERA data assimilation system (*Laloyaux et al.*, 2016).

Two reanalyses produced by JMA and cooperating institutions are considered: JRA-25/JCDAS and JRA-55. JRA-25 (*Onogi et al.,* 2007), a joint effort by JMA and CRIEPI, was the first reanalysis produced using the JMA forecast model and data assimilation system. This reanalysis originally covered 25 years from 1979 through 2004, and was extended an additional 10 years (through the end of January 2014) as JCDAS using an identical system. JRA-55 (*Kobayashi et al.*, 2015) is an extended full input reanalysis with coverage from 1958 through the present. JRA-55 is the first reanalysis system to apply a 4D-Var data assimilation scheme (see *Section 2.3*) to upper-air data during the pre-satellite era (note however that ERA-20C has also used 4D-Var to assimilate surface observations during the pre-satellite era, while extension of ERA5 backward in time to 1950 has recently been completed). Along with the JRA-55 reanalysis, JMA has provided two companion products: JRA-55C (*Kobayashi et al.*, 2014), a conventional input reanalysis that excludes satellite observations from the assimilation, and JRA-55AMIP, an ensemble of AMIP-type forecast model simulations without data assimilation.

Two full input reanalyses produced by NASA GMAO are considered: MERRA and MERRA-2. MERRA (*Rienecker et al.*, 2011) was conceived by NASA GMAO as a reanalysis of the satellite era (starting in January 1979), with particular focus on leveraging the large amounts of data produced by NASA's Earth Observing System (EOS) satellite constellation and improving the representations of the water and energy cycles relative to earlier reanalyses. MERRA production was discontinued after February 2016. Motivated by the inability of the MERRA system to ingest some recent data types, GMAO has developed the follow-on reanalysis MERRA-2 (*Gelaro et al.*, 2017). MERRA-2, which covers 1980-present, includes substantial upgrades to the model (*Molod et al.*, 2015) and changes to the data assimilation system and input data (*McCarty et al.*, 2016). Several new data sources are used that were not assimilated by MERRA, including hyperspectral radiances from IASI and CrIS, microwave radiances from ATMS, MLS temperature and ozone profiles, and GNSS-RO bending angles. One significant and unique feature of MERRA-2 is the assimilation of aerosol optical depth observations (*Randles et al.*, 2017; *Buchard et al.*, 2017), with analysed aerosols fed back to the forecast model radiation scheme. An earlier NASA reanalysis (*Schubert et al.*, 1993; *Schubert et al.*, 1995) covering 1980–1995 was produced by NASA's DAO (now GMAO) using the GEOS-1 data assimilation system; this reanalysis is no longer publicly available and is not included in the S-RIP intercomparison.

Four reanalyses produced by NOAA and cooperating organizations are considered: NCEP–NCAR R1, NCEP–DOE R2, CFSR/CFSv2, and NOAA–CIRES 20CR. NCEP–NCAR R1 (*Kalnay et al.*, 1996; *Kistler et al.*, 2001) was the first modern reanalysis system with extended temporal coverage (1948-present). This system, which uses a modified 1995 version of the NCEP forecast model, remains in widespread use. NCEP–DOE R2 covers the satellite era (1979-present) using essentially the same model, but corrects some important errors and limitations (*Kanamitsu et al.*, 2002). More recently, NCEP has produced CFSR using a 2007 version of the NCEP forecast model (*Saha et al.*, 2010). CFSR contains a number of improvements relative to *R1* and *R2* in both the forecast model and data assimilation system, including higher horizontal and vertical resolutions, more sophisticated model physics, and the ability to assimilate satellite radiances directly (rather than temperature retrievals). CFSR was also the first coupled global reanalysis of the atmosphere–ocean–sea ice system. Production of CFSR was transitioned to a newer version of the NCEP data assimilation system (CFSv2; *Saha et al.*, 2014) on 1 January 2011. This transition from CFSR to CFSv2 should not be confused with the transfer of CFSv2 production from NCEP EMC to NCEP operations, which occurred at the start of April 2011. The model used for CFSv2 has a different horizontal resolution and includes minor changes to physical parameterizations. Because CFSv2 has been touted as a continuation of CFSR, we treat CFSR and CFSv2 as a paired system in this chapter, including brief descriptions of differences between the original and updated systems where relevant. However, we note that subsequent chapters of this report document many significant differences between CFSR and CFSv2, and suggest that users of these products should be cautious when conducting studies that span the 1 January 2011 transition date (see also *Section 2.5*). NOAA–CIRES 20CR (*Compo et al.*, 2011) is the first reanalysis to span more than 100 years. Like ERA-20C, 20CR is a surface input reanalysis. Unlike ERA-20C, which uses a 4D-Var approach to assimilate both surface pressure and surface winds, 20CR uses an EnKF approach (see *Section 2.3*) and assimilates only surface pressure data. The forecast model used in 20CR is similar in many ways to that used in CFSR, but with much coarser vertical and horizontal grids. Because of its relatively coarse vertical resolution (see *Appendix A*) and the lack of direct observational inputs in the upper atmosphere, output from 20CR should be used with care, particularly in the upper troposphere and above. Although two updated versions of 20CR (20CRv2c and 20CRv3; see *Slivinski et al.*, 2019) have been released since the beginning of the S-RIP activity, this report focuses on the earlier 20CRv2 (*Compo et al.*, 2011) unless otherwise indicated.

The influence of observational data on reanalysis products differs not only by the type of reanalysis (*e.g.*, "full input" versus "surface input"), but also by variable (see, *e.g.*, the variable classification proposed by *Kistler et al.*, 2001). Atmospheric temperatures, horizontal winds, and geopotential heights are strongly influenced by the assimilation of observational data even in earlier reanalysis systems, although these variables may be determined mainly by the forecast model in regions or periods where observations are sparse or uncertain. Observational constraints on tropospheric water vapour are weaker but still influential, and some recent reanalysis systems assimilate data that establish constraints on ozone, total water, precipitation, and/or aerosol optical depth. Variables that are largely determined by the forecast model or surface boundary conditions (such as surface fluxes and tendency terms for heat, moisture, and momentum) are considered less reliable and should be used with caution and/or validated against independent estimates.

The SPARC community has particular interest in upper tropospheric and stratospheric ozone and water vapour. This chapter touches briefly on the treatment of these variables, with detailed intercomparisons deferred to *Chapter 4*. Many reanalysis systems simulate ozone using photochemistry schemes of varying complexity and assimilate satellite ozone retrievals during the period after 1979. Some reanalysis systems provide an ozone analysis but use a climatological ozone distribution for radiation calculations in the forecast model. Additional details regarding the treatment of ozone are provided in **Table 2.11**. Reanalysis estimates of stratospheric water vapour are rudimentary and often unreliable. Adjustments due to data assimilation are typically suppressed above a specified upper boundary that varies by reanalysis system, and are in several cases replaced by relaxation to a constant value or zonal mean climatology. Stratospheric air is dehydrated mainly at the tropical tropopause and transported and diffused from there, with only a few systems attempting to represent the source of water vapour due to methane oxidation (see **Table 2.24** for further details).

Table 2.1: *List of global atmospheric reanalysis systems considered in this report.*

Reanalysis system	Reference	Description
ERA-40	*Uppala et al.* (2005)	Class: full input; extended Centre: ECMWF Coverage: September 1957 to August 2002
ERA-Interim	*Dee et al.* (2011)	Class: full input; satellite era Centre: ECMWF Coverage: January 1979 to August 2019.
ERA-20C	*Poli et al.* (2016)	Class: surface input; extended Centre: ECMWF Coverage: January 1900 to December 2010 *Note: A companion ensemble of AMIP-style simulations (ERA-20CM; Hersbach et al., 2015) is also available.*
ERA5	*Hersbach et al.* (2020)	Class: full input; extended Centre: ECMWF Coverage: currently January 1979 to present; a preliminary version of extension backward in time to January 1950 has also been released. *Note: ERA5.1, a rerun covering 2000–2006, has been conducted to address a cold bias in the lower stratosphere during this period.*
JRA-25 / JCDAS	*Onogi et al.* (2007)	Class: full input; satellite era Centre: JMA and CRIEPI Coverage: January 1979 to January 2014 *Note: January 2005 through January 2014 are from JCDAS, a real-time extension of JRA-25.*
JRA-55	*Kobayashi et al.* (2015); *Harada et al.* (2016)	Class: full input; extended Centre: JMA Coverage: January 1958 to present *Note: Two ancillary products are also available: JRA-55C (a conventional input reanalysis covering November 1972 to December 2012; see Kobayashi et al., 2014) and JRA-55AMIP (which assimilates no observational data but uses the same boundary conditions as JRA-55).*
MERRA	*Rienecker et al.* (2011)	Class: full input; satellite era Centre: NASA GMAO Coverage: January 1979 to February 2016
MERRA-2	*Gelaro et al.* (2017)	Class: full input; satellite era Centre: NASA GMAO Coverage: January 1980 to present
NCEP-NCAR R1	*Kalnay et al.* (1996); *Kistler et al.* (2001)	Class: full input; extended Centre: NOAA/NCEP and NCAR Coverage: January 1948 to present
NCEP-DOE R2	*Kanamitsu et al.* (2002)	Class: full input; satellite era Centre: NOAA/NCEP and the DOE AMIP-II project Coverage: January 1979 to present
CFSR / CFSv2	*Saha et al.* (2010); *Saha et al.* (2014)	Class: full input; satellite era Centre: NOAA/NCEP Coverage: January 1979 to present *Note: Official data coverage by CFSR (CDAS-T382) extends through December 2010; production was migrated to the CFSv2 (CDAS-T574) analysis system starting from 1 January 2011. Although it has a different horizontal resolution (Table 2.2) and includes minor changes to physical parameterizations, CFSv2 can be considered as a continuation of CFSR for most purposes.*
NOAA-CIRES 20CR v2	*Compo et al.* (2011)	Class: surface input; extended Centre: NOAA and the University of Colorado CIRES Coverage: November 1869 to December 2012 *Note: Updated versions of 20CR covering 1851–2011 (20CR version 2c, released in 2015) and 1836–2015 (20CR version 3, released in 2019) have been completed and made available, but are not documented in this chapter. See Slivinski et al. (2019) for details.*

2.2 Forecast models

2.2.1 Summary of basic information

Table 2.2 provides a summary of key information regarding the forecast models used in each reanalysis, including the analysis system, the horizontal grid, and the number of levels in the vertical coordinate. The forecast models and data assimilation systems used in reanalyses are typically frozen versions of operational systems for numerical weather prediction. The atmospheric model used in a reanalysis thus often has much in common with the model used for operational numerical weather forecasting at the same forecasting centre around the time that reanalysis was started. Model names and generations are listed in the second column of **Table 2.2**.

The information on horizontal grids provides a rough idea of the finest horizontal scales represented by the models. We describe the horizontal grid structures of models that use spectral dynamical cores (*e.g.*, *Machenhauer*, 1979) using two separate notations. All of the models considered here use spectral dynamical cores except for MERRA and MERRA-2. Regular Gaussian grids are denoted by Fn and Tk. Fn refers to a regular Gaussian grid with 2n latitude bands and (in most cases) 4n longitude bands, while Tk indicates horizontal truncation at wave number k in the spectral dynamical core. The longitude grid spacing in a standard Fn regular Gaussian grid is 90°/n, so that the geographical distance between neighbouring grid cells in the east–west direction shrinks toward the poles. *R1*, *R2*, and 20CR use modified regular Gaussian grids with 4(n+1) longitude bands and longitude spacings of 90°/(n+1). Linear reduced Gaussian grids (*Hortal and Simmons*, 1991; *Courtier and Naughton*, 1994) are denoted by Nn and TLk, where the latter again indicates truncation at horizontal wave number k. The number of latitude bands in the Nn reduced Gaussian grid is also 2n, but the number of longitudes per latitude circle decreases from the equator (where it is 4n) toward the poles. Longitude grid spacing in reduced Gaussian grids is therefore quasi-regular in distance rather than degrees (**Table 2.2**). More details on Gaussian grids are available at **https://confluence.ecmwf.int/display/FCST/Gaussian+grids** (accessed 5 June 2020). Unlike the other reanalysis systems discussed in this chapter, the MERRA and MERRA-2 atmospheric models use finite volume dynamical cores. MERRA applied this dynamical core on a regular latitude–longitude grid (*Lin*, 2004), while MERRA-2 uses a cubed-sphere grid (*Putman and Lin*, 2007). The latter type of grid is denoted by Cn, following a similar convention as Fn and Nn *(i.e.*, approximately 4n longitude bands along the equator).

Table 2.3 lists the vertical locations of the model tops and describes special treatments applied in the uppermost layers of each model. Common special treatments include the use of a diffusive 'sponge layer' near the model top. Sponge layers mitigate the effects of the finite 'lid height' that must be assumed in numerical models of the atmosphere. The application of enhanced diffusion in a sponge layer damps upward propagating waves as they near the model top, thereby preventing unphysical reflection of wave energy at the model top that would in turn introduce unrealistic resonance in the model atmosphere (*Lindzen et al.*, 1968). It is worth noting, however, that diabatic heating and momentum transfer associated with the absorption of wave energy by sponge layers and other simplified representations of momentum damping (such as Rayleigh friction; see, *e.g.*, *Holton and Wehrbein*, 1980) may still introduce spurious behaviour in model representations of middle atmospheric dynamics (*Shepherd and Shaw*, 2004; *Shepherd et al.*, 1996). Most of the forecast models used by reanalysis systems include a sponge layer, but the formulation of this layer varies. The models that do not, such as that used to produce NCEP-NCAR R1, are known to include spurious wave reflection from the model top that affects their performance in the upper atmosphere.

Table 2.2: *Basic details of the forecast models used in the reanalyses. Horizontal grid spacing is expressed in degrees for regular grids and in kilometres for reduced grids.*

Reanalysis system	Model	Horizontal grid	Vertical grid
ERA-40	IFS Cycle 23r4 (2001)	N80: ~125 km (TL159)	60 (hybrid σ–p)
ERA-Interim	IFS Cycle 31r2 (2007)	N128: ~79 km (TL255)	60 (hybrid σ–p)
ERA-20C	IFS Cycle 38r1 (2012)	N80: ~125 km (TL159)	91 (hybrid σ–p)
ERA5	IFS Cycle 41r2 (2016)	N320: ~31 km (TL639)	137 (hybrid σ–p)
JRA-25 / JCDAS	JMA GSM (2004)	F80: 1.125°(T106)	40 (hybrid σ–p)
JRA-55	JMA GSM (2009)	N160: ~55 km (TL319)	60 (hybrid σ–p)
MERRA	GEOS 5.0.2 (2008)	1/2° latitude, 2/3° longitude	72 (hybrid σ–p)
MERRA-2	GEOS 5.12.4 (2015)	C180: ~50 km (cubed sphere)	72 (hybrid σ–p)
NCEP-NCAR R1	NCEP MRF (1995)	F47: 1.875° (T62)	28 (σ)
NCEP-DOE R2	Modified MRF (1998)	F47: 1.875° (T62)	28 (σ)
CFSR **CFSv2**	NCEP CFS (2007) NCEP CFS (2011)	F288: 0.3125° (T382) F440: 0.2045° (T574)	64 (hybrid σ–p) 64 (hybrid σ–p)
NOAA-CIRES 20CR v2	NCEP GFS (2008)	F47: 1.875° (T62)	28 (hybrid σ–p)

All of the reanalysis systems discussed in this chapter use hybrid σ–p vertical coordinates (*Simmons and Burridge*, 1981), with the exception of NCEP-NCAR R1 and NCEP-DOE R2, which use σ vertical coordinates. The number of vertical levels ranges from 28 (*R1*, *R2*, and 20CR) to 137 (ERA5), and top levels range from 3 hPa (*R1* and *R2*) to 0.01 hPa (MERRA, MERRA-2, ERA5, and ERA-20C).

Table 2.3: *Model top levels and special dynamical treatments applied in the uppermost model levels.*

Reanalysis system	Top level	Special treatment of uppermost levels
ERA-40	0.1 hPa	A sponge layer is applied at pressures less than 10 hPa by adding an additional function to the horizontal diffusion terms. This function, which varies with wavenumber and model level, acts as an effective absorber of vertically-propagating gravity waves. Rayleigh friction is also implemented at pressures less than 10 hPa.
ERA-Interim	0.1 hPa	Same as ERA-40.
ERA-20C	0.01 hPa	Similar to ERA-Interim, but an additional first order 'mesospheric' sponge layer is implemented at pressures less than 1 hPa. As in ERA-40 and ERA-Interim, Rayleigh friction is still applied at pressures less than 10 hPa, but the coefficient is reduced to account for the inclusion of parameterized non-orographic gravity wave drag (**Table 2.6**).
ERA5	0.01 hPa	Similar to ERA-20C, but Rayleigh friction is no longer applied.
JRA-25 / JCDAS	0.4 hPa	A sponge layer is applied by gradually enhancing horizontal diffusion coefficients with increasing height at pressures less than 100 hPa. Rayleigh damping is applied to temperature deviations from the global average on each of the uppermost three levels.
JRA-55	0.1 hPa	Sponge layer treatment is similar to JRA-25, but with Rayleigh friction implemented at pressures less than 50 hPa.
MERRA	0.01 hPa	A sponge layer consisting of the nine uppermost model levels (pressures less than ~0.24 hPa) is implemented by increasing the horizontal divergence damping coefficient (see also **Table 2.7**). Advection at the top model level is reduced to first order.
MERRA-2	0.01 hPa	Same as MERRA.
NCEP-NCAR R1	3 hPa	No sponge layer or other special treatment.
NCEP-DOE R2	3 hPa	No sponge layer or other special treatment.
CFSR / CFSv2	~0.266 hPa	Linear Rayleigh damping with a time scale of 5 days is applied at pressures less than ~2 hPa. The horizontal diffusion coefficient also increases with scale height throughout the atmosphere.
NOAA-CIRES 20CR v2	~2.511 hPa	No sponge layer or other special treatment.

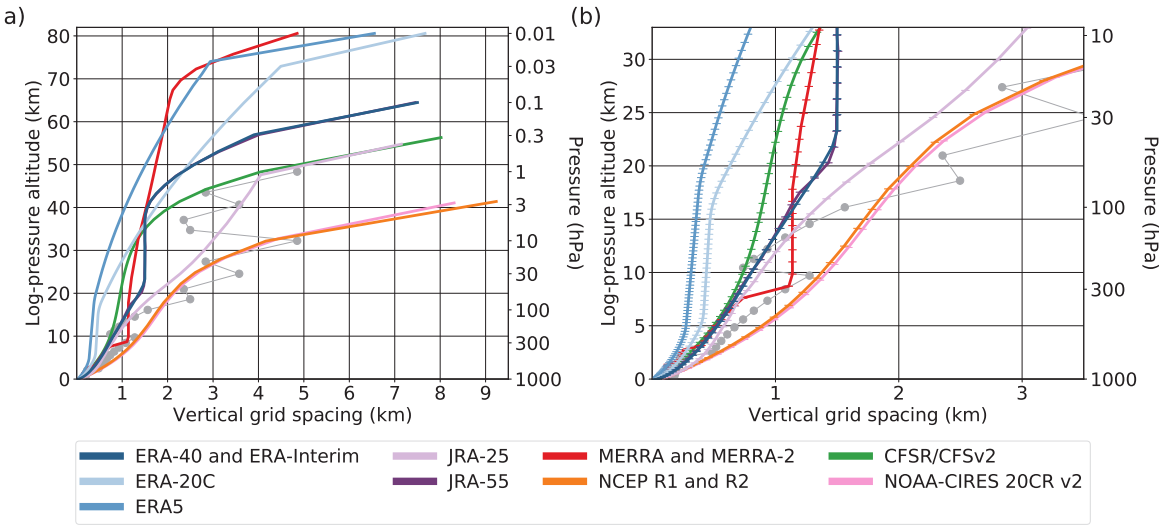


Figure 2.1: *Approximate vertical resolutions of the reanalysis forecast models for (a) the full vertical range of the reanalyses and (b) the surface to 33 km (~10 hPa). Altitude and vertical grid spacing are estimated using log-pressure altitudes ($z^* = H \ln[p0/p]$), where the surface pressure p0 is set to 1000 hPa and the scale height H is set to 7 km. The grid spacing indicating the separation of two levels is plotted at the altitude of the upper of the two levels, so that the highest altitude shown in (a) indicates the height of the top level. Some reanalyses use identical vertical resolutions; these systems are listed together in the legend. Other reanalyses have very similar vertical resolutions when compared with other systems, including JRA-55 (similar but not identical to ERA-40 and ERA-Interim) and 20CR (similar but not identical to R1 and R2). Approximate vertical spacing associated with the isobaric levels on which ERA-40 and ERA-Interim reanalysis products are provided (grey discs) is shown in both panels for context. Reproduced from Fujiwara et al. (2017).*

Figure 2.1 shows approximate vertical resolutions for the reanalysis systems in log-pressure altitude, assuming a scale height of 7 km and a surface pressure of 1000 hPa. A number of key differences are evident, including large discrepancies in the height of the top level (**Figure 2.1a**) and variations in vertical resolution through the upper troposphere and lower stratosphere (**Figure 2.1b**). These model grids differ from the isobaric levels on which many reanalysis products are provided. Vertical spacing associated with an example set of these isobaric levels (corresponding to ERA-40 and ERA-Interim) is included in **Figure 2.1** for context. See *Appendix A* for lists of model levels and further details of the vertical grid.

2.2.2 Major physical parameterizations

In this section we briefly describe some influential physical parameterizations used in the reanalysis forecast models, including those for longwave and shortwave radiation (**Table 2.4**), stratiform clouds (**Table 2.5**), moist convection (**Table 2.6**), gravity wave drag (**Table 2.7**), and horizontal and vertical diffusion (**Table 2.8**). Further details and additional references for many of these parameterizations are provided in the extended digital version of this chapter (*Chapter 2E*).

Table 2.4: *Radiative transfer schemes used in the forecast models of the reanalysis systems. A more complete discussion is provided in Chapter 2E.*

Reanalysis system	Radiative transfer scheme
ERA-40	Shortwave: *Fouquart and Bonnel* (1980) with four spectral intervals. Longwave: RRTM (*Mlawer et al.*, 1997). Radiation calculations are performed every 3 hours on a T63 horizontal grid.
ERA-Interim	Shortwave: Updated version of *Fouquart and Bonnel* (1980). Longwave: RRTM (*Mlawer et al.*, 1997). The scheme is a revised version of that used in ERA-40 with hourly radiation calculations on a T95 horizontal grid (*Dee et al.*, 2011).
ERA-20C	Shortwave: RRTM-G (*Mlawer et al.*, 1997; *Iacono et al.*, 2008). Longwave: RRTM-G (*Mlawer et al.*, 1997; *Iacono et al.*, 2008). Radiation calculations are performed every 3 hours on a T63 horizontal grid. A McICA approach with generalized overlap is used to represent the radiative effects of clouds (*Morcrette et al.*, 2008).
ERA5	Similar to ERA-20C, but with radiation calculations performed hourly on a T319 horizontal grid.
JRA-25 / JCDAS	Shortwave: *Briegleb* (1992) Longwave: line absorption based on the random band model of *Goody* (1952). Radiation calculations are performed on the full model grid, with calculations every hour for shortwave radiation and every three hours for longwave radiation.
JRA-55	Shortwave: *Briegleb* (1992), updated to use the formulation of *Freidenreich and Ramaswamy* (1999) for shortwave absorption by O_2, O_3, and CO_2. Longwave: *Murai et al.* (2005). Radiation calculations are performed on the full model grid, with calculations every hour for shortwave radiation and every three hours for longwave radiation.
MERRA	Shortwave: *Chou and Suarez* (1999). Longwave: *Chou et al.* (2001). Radiation calculations are performed hourly on the full model grid.
MERRA-2	Same as MERRA.
NCEP-NCAR R1	Shortwave: GFDL (*Lacis and Hansen*, 1974). Longwave: GFDL (*Schwarzkopf and Fels*, 1991; *Fels and Schwarzkopf*, 1975). Radiation calculations are performed every 3 hours on a 128×64 linear grid.
NCEP-DOE R2	Shortwave: *Chou and Lee* (1996). Longwave: GFDL (*Schwarzkopf and Fels*, 1991; *Fels and Schwarzkopf*, 1975; same as *R1*). Radiation calculations are performed hourly on the full model grid.
CFSR / CFSv2	Shortwave: Modified RRTM-G (*Clough et al.*, 2005). Longwave: Modified RRTM-G (*Clough et al.*, 2005). Radiation calculations are performed hourly on the full model grid. A McICA approach with maximum–random overlap is used for representing the radiative effects of clouds in CFSv2, but not in CFSR.
NOAA-CIRES 20CR v2	Shortwave: Modified RRTM-G (*Clough et al.*, 2005). Longwave: Modified RRTM-G (*Clough et al.*, 2005). Radiation calculations are performed hourly on the full model grid.

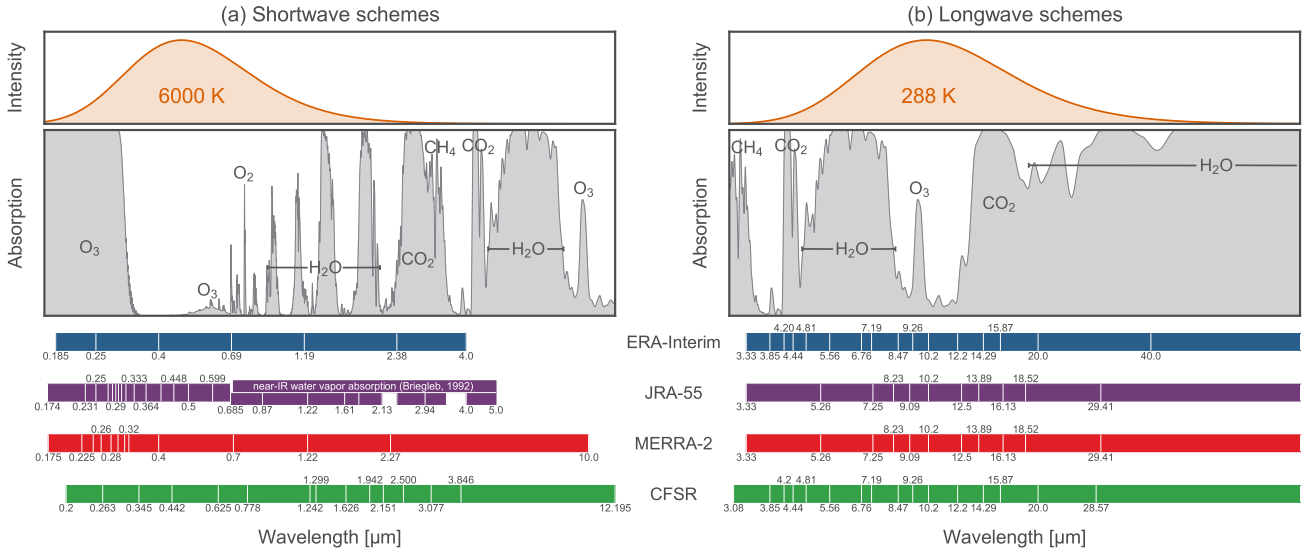


Figure 2.2: *Spectral bands in the radiation schemes used in four recent reanalyses.*

Other pertinent items include the treatment of incoming solar radiation, surface boundary conditions, and radiatively active gases and aerosols, which are summarized in *Section 2.2.3* (see also references in **Table 2.1**), as well as representations of land surface properties, which are described very briefly in **Section 2.2.4**.

The radiative parameterisations used in the forecast model components of reanalysis systems are broadband schemes, in which the radiative spectrum is discretized into a small set of spectral intervals or bands. The form of this discretization is dictated primarily by the presence of radiatively active constituents in the atmosphere and the wavelengths at which these constituents are active. Radiative fluxes and heating rates are computed by integrating across all spectral bands. Note that the radiative transfer schemes used in the atmospheric forecast models (**Table 2.4**) differ from the radiative transfer schemes used to process satellite radiances for data assimilation (**Table 2.19**).

Assumptions on cloud overlapping during radiation calculations are described in *Chapter 2E*.

Parameterizations of stratiform or "large-scale" clouds

Table 2.5: *Non-convective (stratiform) cloud parameterizations used in the forecast models of the reanalysis systems. A more complete discussion is provided in Chapter 2E.*

Reanalysis System	Cloud Parameterization
ERA-40	A prognostic cloud scheme (*Tiedtke*, 1993), in which cloud fraction and cloud water content both evolve according to physical sources and sinks.
ERA-Interim	Similar to ERA-40, but updated to include a treatment for ice supersaturation at temperatures less than 250 K (*Tompkins et al.*, 2007).
ERA-20C	Similar to ERA-Interim, but updated to permit separate estimates of liquid and ice water in non-convective clouds.
ERA5	Same as ERA-20C.
JRA-25 / JCDAS	A modified version of the parameterization proposed by *Smith* (1990), but with stratocumulus cloud fractions following *Kawai and Inoue* (2006).
JRA-55	Same as JRA-25.
MERRA	A prognostic scheme developed by *Bacmeister et al.*, (2006). Convectively-detrained "anvil" condensate is tracked separately from condensate formed in situ, with the former converted to the latter over a specified e-folding timescale.
MERRA-2	As in MERRA, but with new constraints on distributions of total water following *Molod* (2012) and a modified function governing the partitioning of cloud water into liquid and ice during cloud formation.
NCEP-NCAR R1	Diagnosed as a function of grid-scale relative humidity; known to produce discontinuities around 0°E and 180°E longitude (*Kanamitsu et al.*, 2002).
NCEP-DOE R2	Diagnosed as a function of grid-scale RH; modified from that used by R1 to eliminate the discontinuities around 0°E and 180°E.
CFSR / CFSv2	A simple cloud physics parameterization with prognostic cloud condensate (*Zhao and Carr*, 1997). Cloud fraction is diagnosed as a function of cloud water content and relative humidity (*Xu and Randall*, 1996).
NOAA-CIRES 20CR v2	Same as CFSR.

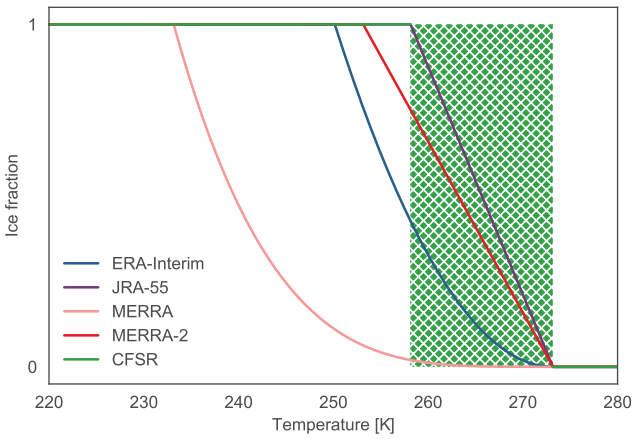


Figure 2.3: *Partitioning of prognostic cloud condensate between the ice and liquid phases as a function of temperature in five recent reanalysis systems. See Chapter 2E for details.*

in the reanalysis systems (listed in **Table 2.5**) influence surface fluxes and the atmospheric state via couplings with radiative transfer, precipitation, and convection. The simplest parameterisations diagnose stratiform cloud cover at each time step as a function of the difference between the grid-scale relative humidity and a critical relative humidity. The existence of clouds in the model atmosphere thus depends on the relative humidity exceeding this critical threshold. NCEP-NCAR R1 and NCEP-DOE R2 use this type of "diagnostic" parameterization. Although computationally inexpensive, diagnostic cloud parameterizations have a number of intrinsic flaws (see, *e.g.*, *Xu and Krueger,* 1991), and have been replaced in more recent reanalyses by variations on the "prognostic" approach pioneered by *Sundqvist* (1978). Prognostic parameterizations simulate the evolution of key cloud variables, such as cloud fraction, cloud water content, and precipitation, and allow for the persistence and advection of convectively-detrained anvil clouds across multiple time steps, as well as the inclusion of more sophisticated approaches to simulating the autoconversion of cloud condensate to rain and snow. The prognostic cloud parameterizations used in reanalyses consider two primary sources of stratiform clouds. The first of these, detrainment of cloud condensate from moist convection, depends on the formulation of the convection schemes documented in **Table 2.6**. The second source, in situ condensation resulting from large-scale cooling, may be represented either via empirically-based PDFs (*e.g.*, *Molod*, 2012; *Smith*, 1990) or by prognostic equations that track the physical sources and sinks of stratiform cloud (*e.g.*, *Tiedtke*, 1993).

Another potentially influential difference among the prognostic cloud schemes used in reanalysis systems is the approach to partitioning cloud condensate into ice and liquid phases (**Figure 2.3**), which affects both the optical properties (and hence radiative transfer) and microphysical properties (and hence autoconversion and precipitation) of the

Table 2.6: *Convective parameterizations used in the forecast models of the reanalysis systems. A more complete discussion is provided in Chapter 2E.*

Reanalysis System	Convective Parameterization
ERA-40	Deep, shallow, and mid-level cumulus convection are parameterized using a bulk mass flux scheme based on that proposed by *Tiedtke* (1989). Each simulated convective cloud consists of a single pair of entraining/detraining plumes that represent updraught and downdraught processes.
ERA-Interim	Similar to ERA-40, but modified in several respects to improve the diurnal cycle of convection, increase convective precipitation efficiency, and make more explicit distinctions among shallow, mid-level and deep convective clouds (*Dee et al.*, 2011).
ERA-20C	Similar to ERA-Interim but with modified representations of entrainment and detrainment rates and a revised convective adjustment time scale.
ERA5	Similar to ERA-20C but with a new closure that better accounts for coupling between the boundary layer and free troposphere, improving the diurnal cycle of convection (*Bechtold et al.*, 2014).
JRA-25 / JCDAS	An 'economical prognostic' mass-flux type Arakawa–Schubert cumulus scheme (*JMA*, 2007; *Arakawa and Schubert*, 1974).
JRA-55	Similar to JRA-25 but with a new triggering mechanism (*Xie and Zhang*, 2000).
MERRA	A version of the relaxed Arakawa–Schubert cumulus scheme (*Moorthi and Suarez*, 1992).
MERRA-2	Same as MERRA, but with a new stochastic Tokioka-type entrainment condition that limits the occurrence of plumes with very small entrainment rates (*Molod et al.*, 2015).
NCEP-NCAR R1	Deep convective clouds are simulated using a simplified Arakawa–Schubert convection scheme (*Pan and Wu*, 1995; *Arakawa and Schubert*, 1974); shallow convective clouds are simulated using a Tiedtke-type scheme (*Tiedtke*, 1989).
NCEP-DOE R2	Similar to NCEP-NCAR R1, but with minor tuning applied.
CFSR / CFSv2	Same underlying schemes as *R1* and *R2*, but with substantial updates as described by *Moorthi et al.* (2001, 2010) and *Saha et al.* (2010).
NOAA-CIRES 20CR v2	Same as CFSR.

simulated clouds. As with the cloud schemes themselves, this partitioning may be either diagnostic or prognostic. See *Chapter 2E* for further details.

Moist convection is another critical subgrid-scale process that must be parameterized in atmospheric models (*Arakawa*, 2004). All of the reanalyses described in this chapter represent moist convection using versions of bulk mass-flux parameterizations (*Tiedtke*, 1989; *Arakawa and Schubert*, 1974), which have as their conceptual basis the "hot tower" hypothesis of *Riehl and Malkus* (1958). These parameterizations represent the statistical effects of convection in a given grid cell via one or more updraft and downdraft plumes, which are in turn coupled to the background environment via entrainment and detrainment, diabatic heating, and the vertical transport of tracers and momentum. Key differences in the convective parameterizations used by the reanalysis systems include the trigger function, the principal closure, whether and to what extent momentum and tracer transport are included, restrictions on the properties of the individual plumes (*e.g.*, entrainment, detrainment, cloud base, and cloud top), and assumptions governing the production and partitioning of rainfall and cloud condensate. We summarize the convection schemes used in each reanlysis in **Table 2.6**. In *Chapter 2E*, we briefly describe the two aspects, trigger functions and closure assumptions.

Gravity wave drag (GWD) parameterisations are used in reanalysis forecast models to represent the systematic effects of momentum deposition on the resolved flow by small-scale (*i.e.*, unresolved) gravity waves. As a relative fraction of the momentum budget the importance of GWD forcing generally increases with altitude, becoming a dominant contribution in the mesosphere (*Polavarapu et al.*, 2005), but effects can also be significant at lower altitudes, such as on the upper poleward flank of the tropospheric subtropical jet (*McFarlane*, 1987; *Palmer et al.*, 1986). GWD parameterisations are typically implemented in atmospheric models via separate schemes for orographic and non-orographic gravity waves. All reanalysis systems considered here include orographic GWD parameterisation, but only ERA-20C, ERA5, MERRA, MERRA-2, and CFSv2 include non-orographic GWD parameterizations (**Table 2.7**). *Chapter 2E* has some further discussions on orographic and non-orographic gravity waves.

Table 2.7: *Gravity wave drag parameterizations used in the forecast models of the reanalysis.*

Reanalysis System	Gravity Wave Drag Parameterization
ERA-40	Subgrid-scale orographic drag is parameterized using the scheme developed by *Lott and Miller* (1997). The representation of the orographic gravity wave source follows *Miller* (1989) and *Baines and Palmer* (1990), and accounts for three-dimensional variability in the amplitude and orientation of wave stress. Non-orographic gravity wave drag is represented as Rayleigh friction above the stratopause.
ERA-Interim	Same as ERA-40.
ERA-20C	Subgrid-scale orographic drag is parameterized similarly to ERA-40 and ERA-Interim, but with slight modifications that increase gravity wave activity. Non-orographic gravity wave drag is included using the parameterization proposed by *Scinocca* (2003); see also *Orr et al.* (2010).
ERA5	Same as ERA-20C, except with a latitudinal dependence of non-orographic launch flux.
JRA-25 / JCDAS	The orographic gravity wave drag parameterization consists of a long wave (wavelengths over 100 km) component and a short wave (wavelengths of ~10 km) component (*Iwasaki et al.*, 1989a, 1989b). Long waves are assumed to propagate upward and break mainly in the stratosphere, where they exert drag (*Palmer et al.*, 1986). Short waves are regarded as trapped and dissipating within the troposphere. Non-orographic gravity wave drag is not included.
JRA-55	Same as JRA-25.
MERRA	MERRA includes parameterizations that compute drag due to the breaking of orographic (McFarlane, 1987) and non-orographic (after *Garcia and Boville*, 1994) gravity waves.
MERRA-2	Similar to MERRA, but with an increased latitudinal profile of the gravity wave drag background source at tropical latitudes and increased intermittency (*Molod et al.*, 2015).
NCEP-NCAR R1	An orographic gravity wave drag scheme based on *Palmer et al.* (1986), *Pierrehumbert* (1987), and *Helfand et al.* (1987) is included in the forecast model. Non-orographic gravity wave drag is not included.
NCEP-DOE R2	Same as NCEP-NCAR R1.
CFSR / CFSv2	The orographic gravity wave drag parameterization is based on the scheme proposed by *Kim and Arakawa* (1995). Sub-grid scale mountain blocking is represented using the scheme developed by *Lott and Miller* (1997). Although non-orographic gravity wave drag is not considered in CFSR, a simple representation of non-orographic gravity wave drag is included in CFSv2 via the parameterization proposed by *Chun and Baik* (1998).
NOAA-CIRES 20CR v2	The orographic gravity wave drag parameterization is the same as in CFSR. Non-orographic gravity wave drag is not considered.

Table 2.8 briefly describes the implementations of horizontal and vertical diffusion in the atmospheric forecast models used by the reanalysis systems. All of the systems that use spectral dynamical cores on Gaussian or reduced Gaussian grids (see above) use implicit linear diffusion in spectral space, although the implementations vary from second-order (NCEP-NCAR R1, NCEP-DOE R2, and 20CR) to eighth-order (CFSR). MERRA and MERRA-2, which are built on finite volume dynamical cores, use slightly different implementations of explicit second-order diffusion. Representations of vertical diffusion in the free troposphere and above are based on first order *K*-type closures. One of the most notable differences among these parameterizations as implemented in the reanalysis systems is the presence or absence of a critical Richardson number, above which turbulent mixing no longer occurs (*Flannaghan and Fueglistaler*, 2014). See the extended *Chapter 2E* (and Figure 2.4) for additional information. Consideration of turbulence in the surface layer and ABL introduces a wider array of parameterizations for turbulent mixing, which are listed in **Table 2.8** but not introduced in detail. Differences in these parameterisations may influence surface exchanges of enthalpy and

Table 2.8: *Representations of vertical and horizontal diffusion in the forecast models used by reanalysis systems.*

Reanalysis System	Representations of Vertical and Horizontal Diffusion
ERA-40	Horizontal diffusion: Implicit linear fourth-order diffusion in spectral space. Vertical diffusion: Vertical diffusion in the free atmosphere and in the ABL under stable conditions is based on the revised Louis scheme (*Beljaars*, 1995; *Louis*, 1979;) for positive Richardson numbers and on Monin–Obukhov similarity for negative Richardson numbers. Vertical diffusion in the ABL under unstable conditions is based on the non-local scheme proposed by *Troen and Mahrt* (1986). Turbulent fluxes in the surface layer are calculated using bulk formulae based on Monin–Obukhov similarity.
ERA-Interim	Horizontal diffusion: Same as ERA-40. Vertical diffusion: Vertical diffusion in the free atmosphere and in the ABL under stable conditions is as in ERA-40. Vertical diffusion in the ABL under unstable conditions is based on an eddy-diffusivity mass-flux (EDMF) scheme (*Köhler et al.*, 2011). Turbulent fluxes in the surface layer are calculated using bulk formulae based on Monin–Obukhov similarity.
ERA-20C	Horizontal diffusion: Same as ERA-40. Vertical diffusion: Similar to ERA-Interim, but with vertical diffusion above the lower troposphere based on Monin–Obukhov similarity under all conditions (rather than the revised Louis scheme) and the inclusion of a simple empirical parameterization to represent unresolved vertical wind shear.
ERA5	Similar to ERA-20C, but with the empirical parameterization of unresolved vertical wind shear removed.
JRA-25 / JCDAS	Horizontal diffusion: Implicit linear fourth-order diffusion in spectral space. Vertical diffusion: Vertical diffusion of momentum, heat, and moisture are represented using the level 2 turbulence closure scheme developed by *Mellor and Yamada* (1974). Surface turbulent fluxes are calculated using bulk formulae based on Monin–Obukhov similarity.
JRA-55	Same as JRA-25.
MERRA	Horizontal diffusion: Explicit second-order horizontal divergence damping is included in the dynamical core. Vertical diffusion: Vertical diffusion in the free atmosphere and in the boundary layer under stable conditions is based on a local gradient Richardson number closure (*Louis et al.,* 1982), but a tuning parameter severely suppresses turbulent mixing at pressures less than ~900 hPa. Vertical diffusion in the boundary layer under unstable conditions is based on the non-local scheme proposed by *Lock et al.* (2000).
MERRA-2	Horizontal diffusion: Similar to MERRA, but with an additional second-order Smagorinsky divergence damping. Vertical diffusion: Similar to MERRA in most respects, with the addition of a Monin–Obukhov-type parameterization to represent turbulent fluxes across the surface layer (*Helfand and Schubert*, 1995). The tuning parameter that suppressed turbulent mixing at pressures less than ~900 hPa in MERRA has been removed, but diffusion coefficients are still usually very small in the free atmosphere.
NCEP-NCAR R1	Horizontal diffusion: Implicit linear second-order diffusion in spectral space. Horizontal diffusion along model σ layers led to the occurrence of spurious "spectral precipitation", particularly in mountainous areas at high latitudes. A special precipitation product was produced to correct this issue. Vertical diffusion: Local K diffusion (*Louis et al.*, 1982) is applied in both the ABL and the free atmosphere with a uniform background diffusion coefficient.
NCEP-DOE R2	Horizontal diffusion: Implicit linear second-order diffusion in spectral space. Issues with spectral precipitation caused by horizontal diffusion are greatly reduced relative to *R1*. Vertical diffusion: Local K diffusion (*Louis et al.*, 1982) is applied in the free atmosphere with a uniform background diffusion coefficient. Non-local diffusion is applied in the ABL (*Hong and Pan*, 1996).
CFSR / CFSv2	Horizontal diffusion: Implicit linear eighth-order diffusion in spectral space. Vertical diffusion: Local K diffusion (*Louis et al.*, 1982) is applied in the free atmosphere with a background diffusion coefficient that decreases exponentially with pressure. Non-local vertical diffusion is applied in the ABL (*Hong and Pan*, 1996).
NOAA-CIRES 20CR v2	Horizontal diffusion: Implicit linear second-order diffusion in spectral space. Vertical diffusion: Same as CFSR.

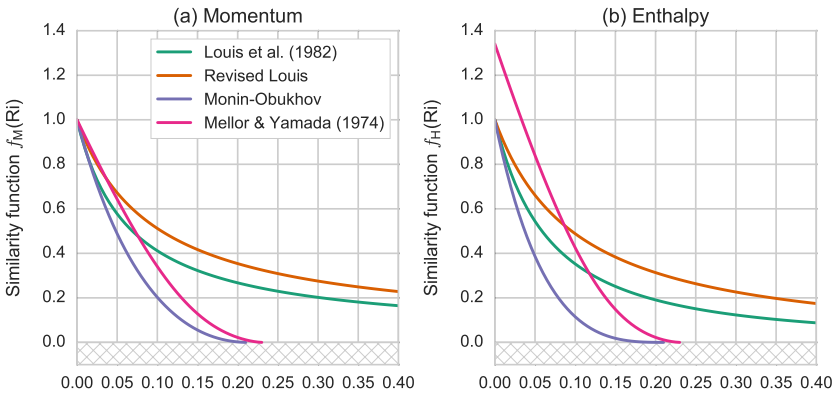


Figure 2.4: *Similarity functions for parameterized turbulent transfer of (a) momentum and (b) enthalpy (heat and moisture) as a function of the gradient Richardson number (Ri) based on four turbulence schemes used in the free troposphere by reanalysis systems. See Chapter 2E for details.*

momentum. Different treatments of surface roughness lengths over land and ocean can also influence energy and momentum fluxes into the atmosphere; these aspects are documented in Table 2.9 of *Chapter 2E* but are omitted here.

2.2.3 Boundary and other specified conditions

This section describes the boundary and other specified conditions that can be regarded as "externally supplied forcings" for each reanalysis system. These conditions comprise the elements of the reanalysis that are not taken from the forecast model or data assimilation but are used to produce the outputs. **Figure 2.5** shows three examples of how externally-specified boundary conditions may vary amongst reanalysis systems.

The factors that may be considered "external" vary somewhat among reanalyses because the forecast and assimilation components have provided a progressively larger fraction of the inputs (initial conditions) for the forecast model as reanalysis systems have developed. Ozone is a prime example. As discussed below, all of the

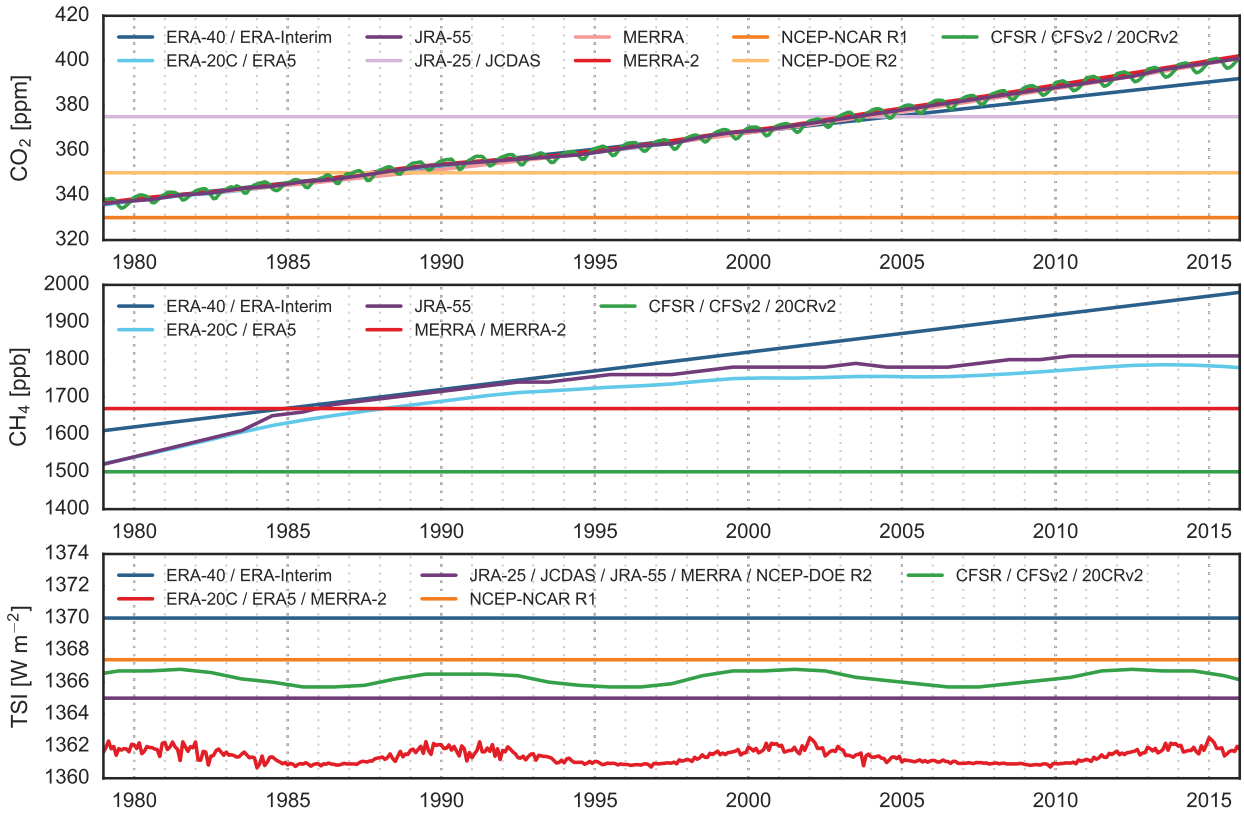


Figure 2.5: *Time series of boundary and specified conditions for CO_2 (top), CH_4 (center), and TSI (bottom) used by the reanalysis systems. The CH_4 climatology used in MERRA and MERRA-2 varies in both latitude and height; here a "tropospheric mean" value is calculated as a mass- and area-weighted integral between 1000 hPa and 288 hPa to facilitate comparison with the "well-mixed" values used by most other systems. ERA-20C and ERA5 also apply rescalings of annual mean values of both CO_2 and CH_4 that vary in latitude and height; here the base values are shown (note that the ERA-20C/ERA5 time series in panel a is obscured by those for JRA-55 and MERRA-2). Time series of TSI neglect seasonal variations due to the ellipticity of the Earth's orbit, as these variations are applied similarly (but not identically) across reanalysis systems. Additional information on CO_2 and CH_4 is provided in* ***Table 2.13****, and additional information on TSI is provided in* ***Table 2.14****. Reproduced from Fujiwara et al. (2017).*

[1] **Table 2.9** (titled as "Sources and representations of surface roughness in the reanalysis systems") is only shown in *Chapter 2E*.

reanalysis systems except for NCEP-NCAR R1, NCEP-DOE R2, and NOAA-CIRES 20CR; JRA-55 and ERA-40 prior to 1978; and ERA5 prior to April 1970) assimilate satellite ozone measurements. Some of these reanalysis systems (notably ERA-40, ERA-Interim, ERA-20C, and ERA5) use ozone climatologies rather than internally generated ozone fields for radiation calculations in the forecast model. MERRA-2 assimilates aerosol optical depths and uses internally generated aerosol fields for the radiation calculations, while other systems use

Table 2.10: *Treatment of sea surface temperature and sea ice.* [1]

Reanalysis System	Sea Surface Temperature and Sea Ice
ERA-40	Monthly data from the Met Office HadISST1 product was used before November 1981, replaced by weekly data from the NOAA–NCEP 2D-Var product from December 1981 through June 2001 and NOAA OISSTv2 from July 2001 through August 2002 (*Reynolds et al.*, 2002). A special sea ice analysis and a method of specifying SST in grid boxes with partial ice-cover were used. Interpolation was used to produce daily values.
ERA-Interim	Similar to ERA-40 but NCEP RTG sea surface temperatures were used from January 2002 through January 2009 and OSTIA (*Donlon et al.*, 2012) was used from February 2009 through August 2019.
ERA-20C	Daily gridded SST and sea ice are from HadISST version 2.1.0.0 (*Titchner and Rayner*, 2014) at 0.25° horizontal resolution. Daily fields are obtained via cubic interpolation from monthly analyses, with the temporal average of daily fields constrained to match the analysed monthly mean.
ERA5	Daily gridded SSTs are from HadISST version 2.1.0.0 between January 1949 and August 2007, and from OSTIA for September 2007 onwards. Sea ice cover is from HadISST version 2.0.0.0 from January 1950 through December 1978, from reprocessed OSI SAF fields between January 1979 and August 2007, and from operational OSI SAF estimates for September 2007 onwards. Data through August 2007 are at 0.25° horizontal resolution, while data from September 2007 to present are at 0.05° horizontal resolution. When necessary, daily fields are obtained from monthly analyses using the same procedure as ERA-20C.
JRA-25 / JCDAS	Daily COBE SSTs (*Ishii et al.*, 2005) were used. COBE SSTs are based on the ICOADS and Kobe data collections, and do not include satellite data. Daily sea ice distributions prepared for COBE are based on reports by *Walsh and Chapman* (2001) for the Northern Hemisphere and *Matsumoto et al.* (2006) for the Southern Hemisphere.
JRA-55	Daily COBE SSTs and sea ice distributions are used, with minor updates from those used for JRA-25/JCDAS. Southern Hemisphere sea ice coverage is based on a climatology before October 1978, and based on *Matsumoto et al.* (2006) after October 1978.
MERRA	Weekly NOAA OISST data at 1° resolution (*Reynolds and Smith*, 1994) are linearly interpolated in time to the model time steps.
MERRA-2	Monthly 1° gridded data (*Taylor et al.*, 2000) are used prior to 1982, daily 0.25° gridded data (*Reynolds et al.*, 2007) through March 2006, and daily 0.05° gridded data from OSTIA (*Donlon et al.*, 2012) from April 2006.
NCEP-NCAR R1	SSTs are taken from the Met Office Global Ice and Sea Surface Temperature (GISST) data set for 1981 and earlier, and from the NOAA OISST data set from 1982 to the present. Sea ice cover is from Navy/NOAA Joint Ice Center analyses before 1978, from SMMR observations for 1978 through 1987, and from SSM/I observations for 1988 through the present. Snow cover is from the NESDIS weekly snow cover analysis (Northern Hemisphere only) for September 1998 and earlier, and from the US Air Force global snow cover analysis from October 1998 through the present.
NCEP-DOE R2	SSTs and sea ice cover for January 1979 through 15 August 1999 are taken from data prepared for AMIP-II and provided by the PCMDI at Lawrence Livermore National Laboratory. SSTs and sea ice cover for 16 August 1999 through December 1999 are from monthly NOAA OISST and monthly NCEP operational sea ice analyses, interpolated to daily resolution. SSTs and sea ice cover for January 2000 to present are from daily NOAA OISST and NCEP operational sea ice analyses.
CFSR / CFSv2	The atmospheric model is coupled to the GFDL MOM version 4 ocean model and a two-layer sea ice model. AVHRR and AMSR satellite infrared observations of SST are assimilated in the SST analysis, along with in situ data from ships and buoys. The sea (and lake) ice concentration analysis products assimilate different observational data depending on the period, including microwave satellite observations when available. Temperatures at the atmosphere–ocean boundary are relaxed every six hours to separate SST analyses, including the 1° gridded HadISST1.1 from January 1979 through October 1981 and versions 1 and 2 of the 0.25° gridded OI analyses described by *Reynolds et al.* (2007) from November 1981. Further details of the coupling procedure and SST/sea ice analysis have been provided by *Saha et al.* (2010).
NOAA-CIRES 20CR v2	HadISST1.1 monthly mean SST and sea ice data are interpolated to daily resolution. Sea ice concentrations were accidentally mis-specified in coastal regions. This error results in warmer lower tropospheric temperatures in polar regions relative to ERA-40 and NCEP-NCAR R1 (*Compo et al.*, 2011). The error has been corrected in Version 2c of the reanalysis.

[1] Table 2.9: Sources and representations of surface roughness in the reanalysis systems is provided in *Chapter 2E*.

climatologies or neglect the role of aerosols altogether. CFSR is a coupled atmosphere–ocean–sea ice system, in which the SST and sea ice lower boundary conditions for the atmospheric model are generated by an ocean model (although temperatures at the boundary are relaxed every six hours to SST analyses similar to those used by other reanalysis systems). This section summarizes the treatment of SST, sea ice, ozone, aerosols, trace greenhouse gases (other than water vapour), and the solar cycle, with special notes where necessary. Dynamical variables, water vapour, and internally generated ozone (*i.e.*, variables that are often directly constrained by the set of assimilated observations) are discussed and evaluated in *Chapters 3 and 4* of this report.

Table 2.11: *Treatment of ozone. See also Chapter 4 of this report.*

Reanalysis System	Treatment of Ozone
ERA-40	TOMS and SBUV ozone retrievals were assimilated from 1978 onward. Ozone in the model is described using a linearization of the ozone continuity equation, including photochemical sources and sinks (*Cariolle and Déqué*, 1986; *Dethof and Hólm*, 2004). The model does not account for heterogeneous chemistry, but does include an empirical ozone destruction term to account for chemical loss in polar stratospheric clouds. Model-generated ozone is not used in the radiation calculations, which instead assume the climatological ozone distribution reported by *Fortuin and Langematz* (1995).
ERA-Interim	Ozone retrievals are assimilated from TOMS (1979-present), SBUV (1979-present), GOME (1996-2002), MIPAS (2003-2004), SCIAMACHY (2003-2008), MLS (2008-present), and OMI (2008-present). The ozone scheme is an updated version of that used in ERA-40 (*Dragani*, 2011; *Cariolle and Teyssèdre*, 2007). As in ERA-40, climatological ozone distributions from *Fortuin and Langematz* (1995) are used for radiation calculations.
ERA-20C	No ozone data are assimilated. The forecast model ozone parameterization is identical to that used in ERA-Interim. Model-generated ozone is not used in the radiation calculations, which instead use monthly three-dimensional ozone fields that evolve in time (*Cionni et al.*, 2011).
ERA5	The ozone scheme is the same as that used in ERA-Interim, but with substantial updates to the assimilated data. Reprocessed retrievals are assimilated from BUV (1970-1977), TOMS (1979-2003), SBUV v8.6 (1979-present), CCI MIPAS (2005-2012) and SCIAMACHY (2003-2012), Aura MLS v4.2 (2004-present) and OMI-DOAS (2004-present). ERA5 also assimilates IR ozone-sensitive radiance not used in ERA-Interim, and uses variational bias correction (see *Section 2.4.3.2*) during the ozone analysis. Analyzed ozone is not used in the radiation calculations, which instead use an in-house ozone climatology from CAMSiRA (*Flemming et al.*, 2017).
JRA-25 / JCDAS	Daily ozone distributions were prepared in advance using the MRI-CCM1 offline chemical transport model with output "nudged" to satellite retrievals of total ozone. These distributions were provided to the forecast model for use in radiation calculations.
JRA-55	For 1979 and later, the approach is similar to that used by JRA-25/JCDAS, but uses an updated chemical transfer model with 68 vertical levels rather than 45. For 1958-1978, a monthly mean climatology generated from the 1980-1984 ozone analyses was used. These distributions were provided to the forecast model for use in radiation calculations.
MERRA	Version 8 SBUV ozone retrievals have been assimilated from October 1978 onward. The ozone parameterization is based on an empirical relationship between ozone and prognostic odd-oxygen that varies with height and the diurnal cycle (*Rienecker et al.*, 2008). The parameterization uses zonally-symmetric monthly production and loss rates derived from a 2-dimensional model as described by *Stajner et al.* (2008), but without representation of heterogeneous chemistry in polar regions. The forecast model uses analyzed ozone data in radiation calculations.
MERRA-2	Version 8.6 SBUV retrievals have been assimilated in reanalyses between 1980 and 2004. Starting from October 2004, these data have been replaced by retrieved MLS profiles (version 2.2 through 31 May 2015; version 4.2 from 1 June 2015) and OMI observations of total ozone (*McCarty et al.*, 2016). Assimilation of MLS retrievals at 261 hPa was discontinued starting on 1 May 2016 (*Wargan et al.*, 2017). The ozone parameterization is the same as that used in MERRA. The forecast model uses analyzed ozone data in radiation calculations.
NCEP-NCAR R1	Seasonal ozone climatologies reported by *London* (1962) and *Hering et al.* (1965) are used in radiation calculations. No ozone analysis is produced.
NCEP-DOE R2	The zonal mean ozone climatology published by *Rosenfield et al.* (1987) is used in radiation calculations, but the latitudinal orientation was reversed north-to-south. Although this error may cause some problems in the stratosphere, *Kanamitsu et al.* (2002) report that the overall impact is minor. No ozone analysis is produced.
CFSR / CFSv2	Version 8 SBUV profiles and total ozone retrievals were assimilated without bias adjustment. Prognostic ozone is parameterized using concentration-dependent climatological production and destruction terms generated by a 2-dimensional chemistry model (*McCormack et al.*, 2006). The forecast model uses analyzed ozone data for radiation calculations. Late 20th century levels of CFCs are included implicitly in the gas phase chemistry and ozone climatology used in the prognostic ozone parameterization.
NOAA-CIRES 20CR v2	No ozone data are assimilated. The ozone model is the same as that used in CFSR.

Table 2.12: *Treatment of aerosols.*

Reanalysis System	Treatment of Aerosols
ERA-40	Aerosols have been included in the radiation calculations using prescribed climatological aerosol distributions (*Tanré et al.*, 1984). These distributions include annual mean geographical distributions for maritime, continental, urban and desert aerosol types, in addition to uniformly distributed tropospheric and stratospheric 'background' aerosol loading. No trends or temporal variations (such as variations due to volcanic eruptions) were included.
ERA-Interim	Aerosols are included in the radiation calculations using updated climatological distributions (*Tegen et al.*, 1997). The climatological annual cycles of tropospheric aerosols have been revised relative to those used by ERA-40, as have the optical thickness values for tropospheric and stratospheric background aerosols. There is no evolution of volcanic aerosols.
ERA-20C	The evolution of tropospheric aerosols is based on data prepared for CMIP5 (*an Vuuren et al.*, 2011; *Lamarque et al.*, 2010). Volcanic sulphates (*Sato et al.*, 1993) and ash (*Tanré et al.*, 1984) are also included in the stratosphere. A more detailed description of the aerosol fields used in ERA-20C and ERA-20CM has been provided by *Hersbach et al.* (2015).
ERA5	Same as ERA-20C.
JRA-25 / JCDAS	Aerosols are represented using two aerosol profiles, one over land and one over sea (*WMO*, 1986). Neither interannual nor seasonal variations are considered.
JRA-55	Similar to JRA-25, but with optical depths adjusted to a 2-dimensional monthly climatology (*JMA*, 2013). Interannual variations, such as those due to volcanic eruptions, are not considered.
MERRA	Aerosols are represented using a climatological aerosol distribution generated using the Goddard Chemistry, Aerosol, Radiation, and Transport (GOCART) model (*Colarco et al.*, 2010).
MERRA-2	Aerosol optical depths from AVHRR, MODIS, MISR and AERONET are assimilated into the GEOS-5 GAAS (*Buchard et al.*, 2015, 2017; *Randles et al.*, 2017). Volcanic aerosols are included. The forecast model uses analyzed aerosols in radiation calculations for the entire production period. Additional details have been provided by *Randles et al.* (2017).
NCEP-NCAR R1	No aerosols.
NCEP-DOE R2	No aerosols.
CFSR / CFSv2	Aerosols are represented using a seasonally varying climatological global distribution of aerosol vertical profiles on a 5° grid (*Koepke et al.*, 1997). Monthly zonal mean volcanic aerosols in four latitude bands (90 - 45°S, 45°S - equator, equator - 45°N, 45 - 90°N) are specified based on data reported by *Sato et al.* (1993).
NOAA-CIRES 20CR v2	Same as CFSR.

2.2.3.1 Sea surface temperature and sea ice

Table 2.10 summarizes the treatment of SST and sea ice distributions in the reanalysis systems, including the names of SST and sea-ice datasets, special calibration or preprocessing details (*e.g.*, bias corrections, interpolations), and details of how the datasets were produced.

2.2.3.2 Ozone

Table 2.11 briefly summarizes the treatment of ozone in the reanalysis systems (detailed intercomparisons are deferred to *Chapter 4*). Some reanalysis systems assimilate satellite ozone measurements (from 1978/1979, and in one case 1970, onward) to produce an ozone analysis product, while some systems do not. Moreover, some systems that produce an ozone analysis use a climatological ozone distribution (rather than the ozone analysis) for radiation calculations in the forecast model. These distinctions are made explicit in **Table 2.11**. None of the reanalysis systems considered here assimilate data from ozonesondes.

2.2.3.3 Aerosols

Table 2.12 summarizes the treatment of stratospheric and tropospheric aerosols in the reanalysis systems. Some reanalysis systems consider tropospheric aerosols over continents and over oceans separately in the radiation scheme. Some reanalysis systems (but not all) account for changes in stratospheric aerosols due to volcanic eruptions. One reanalysis (MERRA-2) assimilates aerosol optical depths and uses analyzed aerosols in radiation calculations.

2.2.3.4 Carbon dioxide and other radiatively active gases

Table 2.13 summarizes the treatment of carbon dioxide and other radiatively active gases (except for water

Table 2.13: *Treatment of carbon dioxide and other radiatively active gases.*

Reanalysis System	CO_2 and Reactive Trace Gases
ERA-40	CO_2, CH_4, N_2O, CFC-11, and CFC-12 are assumed to have globally uniform concentrations throughout the atmosphere. The concentrations of these gases were set to the observed 1990 values plus a linear trend as specified by *IPCC* (1996).
ERA-Interim	Same as ERA-40.
ERA-20C	CO_2, CH_4, N_2O, CFC-11, and CFC-12 are specified according to CMIP5-recommended values (*Meinshausen et al.*, 2011). The IPCC RCP3PD scenario is followed for 2006 - 2010. Greenhouse gases are not assumed to be globally uniform; rather, they are rescaled to match specified seasonal cycles and zonal mean vertical distributions (*Hersbach et al.*, 2015).
ERA5	Same as ERA-20C, with extension of RCP3PD after 2010.
JRA-25 / JCDAS	A constant, globally uniform CO_2 concentration of 375 ppmv was assumed. CH_4, N_2O, CFCs, and HCFCs were not considered.
JRA-55	Daily values of CO_2, CH_4, N_2O, CFC-11, CFC-12, and HCFC-22 are specified by interpolating from annual mean values. For CO_2, CH_4, and N_2O these annual mean values are valid on 1 July; for CFC-11, CFC-12, and HCFC-22 they are valid on 31 December. All species are assumed to be globally uniform, with sources that vary in time (*Kobayashi et al.*, 2015; their Table 7).
MERRA	CO_2 concentrations are assumed to be globally uniform and are specified according to historical observed values. CH_4, N_2O, CFCs, and HCFCs are specified according to steady state monthly climatologies from the Goddard two-dimensional chemistry transport model (*Rienecker et al.*, 2008). These monthly climatologies vary in both latitude and pressure, but do not contain interannual variability.
MERRA-2	Annual global mean CO_2 concentrations follow the IPCC RCP4.5 scenario and are assumed to be uniform throughout the atmosphere. CH_4, N_2O, CFCs, and HCFCs are specified as in MERRA.
NCEP-NCAR R1	A constant, globally uniform CO_2 concentration of 330 ppmv is assumed. CH_4, N_2O, CFCs, and HCFCs are not considered.
NCEP-DOE R2	Similar to *R1*, but with a constant, globally uniform CO_2 concentration of 350 ppmv.
CFSR / CFSv2	Monthly mean 15°×15° distributions of CO_2 concentrations derived from historical WMO Global Atmosphere Watch observations are used. Constant values of CH_4, N_2O, O_2, and four types of halocarbons are also included in the radiation calculations.
NOAA-CIRES 20CR v2	Similar to CFSR for 1956 and later. Estimates of semi-annual average global mean CO_2 concentrations based on ice core data are specified for the period before 1956. Values of CH_4, N_2O, O_2, and four types of halocarbons are constant throughout.

vapour) in the reanalysis systems (see also **Figure 2.5**). Notes on the treatment of water vapour are provided in *Section 2.4.4*.

2.2.3.5 Solar cycle

The solar cycle (*i.e.*, changes in TSI with a period of ~11 years) is an important driver of atmospheric variability. This variability is incorporated in reanalysis systems in a variety of ways, including specified solar radiation at the TOA (boundary condition) and/or observations of temperature or ozone (data assimilation). **Table 2.14** briefly briefly summarizes the extent to which interannual variations in TSI are represented in each reanalysis system (see also **Figure 2.5**).

2.2.4 Surface air and land surface treatments

Treatments of surface air and land surface properties present a number of challenges for reanalyses. For example, sharp gradients and other types of spatial heterogeneity in land cover are difficult to represent in global models, but have important influences on the magnitudes and variability of water and energy fluxes between the land surface and the atmosphere. More specific to reanalyses, the spatial region for which near-surface observations may be considered representative is reduced in coastal regions and regions of complex topography. Land surface properties, such as soil moisture and soil temperature, also evolve relatively slowly, especially at deeper layers. As a result, these variables are among the main targets of model spin-up. Discontinuities in the land surface state at stream transitions (*Section 2.5*) can propagate into the atmosphere.

Reanalyses use two main approaches for producing surface air analysis variables over land (**Table 2.15**). The first approach, which is taken by ERA-40, ERA-Interim, ERA5, JRA-25, and JRA-55, is to assimilate screen-level station observations (*i.e.*, temperatures and dewpoint temperatures at 2-m height) in separate two-dimensional OI analyses (*Section 2.3*) of surface air variables (*e.g.*, *Simmons et al.*, 2004). The main benefits include stronger constraints on surface meteorological conditions and their influences on the LSM (see below);

however, this approach can also generate inconsistencies between the upper air and surface fields in the analysis. None of the reanalysis systems use the results of OI surface air analyses to initialize subsequent forecasts, although these analyses can still indirectly affect subsequent forecasts via influences on the land surface state. The second approach, which is taken by all other reanalyses described in this document, omits screen-level station observations from the analysis. Surface air analysis variables over land are still affected by surface pressure and (in the case of full-input reanalyses) upper air measurements assimilated during the standard analysis cycle. This approach establishes weaker observational constraints on the evolution of surface air and land surface conditions in regions where the observational network is dense, but has the

Table 2.14: *Influence of solar cycle on the reanalysis systems.*

Reanalysis System	Influence of the solar cycle
ERA-40	The ~11-year solar cycle is not included in the TSI boundary condition, with the base irradiance assumed to be constant at 1370 W m^{-2}; however, variations in this value due to changes in the distance between the Earth and the Sun have been incorporated as prescribed by *Paltridge and Platt* (1976). A programming error artificially increased the effective TSI by about 2 W m^{-2} relative to the specified value. *Dee et al.* (2011) reported that the impact of this error is mainly expressed as a warming of approximately 1 K in the upper stratosphere; systematic errors in other regions are negligible. The effects of the solar cycle are included in the assimilated upper-air temperatures, but are not included in the ozone passed to the forecast model (see **Table 2.11**).
ERA-Interim	Same as ERA-40.
ERA-20C	ERA-20C uses TSI variations provided for CMIP5 historical simulations by the SPARC SOLARIS-HEPPA working group with the TIM scaling applied, which take values ranging from 1360.2 W m^{-2} to 1362.7 W m^{-2} between 1900 and 2008. These variations account for solar cycle changes through 2008 and repeat the final cycle (April 1996 - June 2008) thereafter. Seasonal variations due to the ellipticity of the Earth's orbit are also included.
ERA5	Same as ERA-20C.
JRA-25 / JCDAS	A constant base TSI of 1365 W m^{-2} was assumed, including seasonal effects due to the ellipticity of the Earth's orbit (*Spencer*, 1971). Interannual variations in incoming solar radiation were not included in the TSI boundary condition, but were included in assimilated temperature and ozone observations.
JRA-55	Same as JRA-25. Note that interannual variations in incoming solar radiation are included in assimilated temperature observations for the whole period, but only included in ozone observations for 1979 and later.
MERRA	MERRA assumes a constant base TSI of 1365 W m^{-2}. Seasonal variations due to the ellipticity of the Earth's orbit are included. Although interannual variations in incoming solar radiation were not included in the TSI boundary condition, these variations could influence the model state through assimilated temperature and ozone observations.
MERRA-2	MERRA-2 uses TIM-corrected TSI variations provided for CMIP5 historical simulations by the SPARC SOLARIS-HEPPA working group, which take values ranging from 1360.6 to 1362.5 W m^{-2} between 1980 and 2008. These variations account for solar cycle changes through 2008 and repeat the final cycle (April 1996 - June 2008) thereafter. Seasonal variations due to the ellipticity of the Earth's orbit are included.
NCEP-NCAR R1	R1 uses a constant TSI of 1367.4 W m^{-2}. The ~11-year solar cycle is not included in the TSI boundary condition, but variations due to changes in orbital geometry are accounted for. The effects of the solar cycle are included in the assimilated upper-air temperatures, but are not included in the ozone passed to the forecast model (see **Table 2.11**).
NCEP-DOE R2	Similar to R1, but with a constant TSI of 1365 W m^{-2}.
CFSR / CFSv2	Annual average variations in TSI were specified according to data prepared by *H. van den Dool (personal communication,* 2006), with values ranging from 1365.7 W m^{-2} to 1367.0 W m^{-2}. The solar cycle after 2006 is repeated forwards (*e.g.*, insolation for 2007 is the same as that for 1996, that for 2008 is the same as that for 1997, and so on). The effects of the solar cycle are included in assimilated temperature and ozone observations; however, the prognostic ozone parameterization does not otherwise account for variations in incoming solar radiation.
NOAA-CIRES 20CR v2	Annual average variations in TSI were specified according to data prepared by *H. van den Dool (personal communication*, 2006), with values ranging from 1365.7 W m^{-2} to 1367.0 W m^{-2}. The solar cycle before 1944 is repeated backwards (*e.g.*, insolation for 1943 equals that for 1954, that for 1942 equals that for 1953, and so on) and the solar cycle after 2006 is repeated forwards (as in CFSR). Upper-air observations were neither assimilated nor included. The prognostic ozone scheme does not account for variations in incoming solar radiation.

Table 2.15: *Information about land surface models and analyses of surface air variables (if applicable) in the reanalysis systems. Surface air station observations are assimilated in ERA-40, ERA-Interim, ERA5, JRA-25, and JRA-55 in analysis steps separate from the standard upper-air analysis cycles. Other reanalyses do not assimilate these data. Additional details are provided in Chapter 2E.*

Reanalysis System	Surface models and analyses of surface air variables
ERA-40	The surface air and land surface analyses are performed outside of the main atmospheric reanalysis. Six-hourly OI analyses of surface air temperature and dewpoint temperature at 2-m height are produced using station observations over land and the background state from the most recent atmospheric analysis. Empirical relationships between surface air fields and soil properties are then used to update soil temperature and soil moisture in a four-level land surface model (*van den Hurk et al.*, 2000).
ERA-Interim	Essentially the same as ERA-40. The additional global land surface reanalysis ERA-Interim/Land was conducted for 1979-2010 using a newer version of the land surface model (*Balsamo et al.*, 2015) with atmospheric forcing from ERA-Interim and precipitation from GPCP.
ERA-20C	Surface pressure and surface winds (over ocean) are the only variables directly constrained by the data assimilation; no land surface analysis is performed. The land surface scheme is based on a new version of the land surface model (*Balsamo et al.*, 2015) relative to that used in ERA-Interim.
ERA5	Similar to ERA-Interim, but with substantial updates to the land surface analysis (*de Rosnay et al.*, 2014) and a new formulation of the LSM that better represents subgrid-scale water bodies and coastlines. A separate global land surface reanalysis ERA5-Land is being conducted with atmospheric forcing from ERA5.
JRA-25 / JCDAS	Surface air temperature, winds, and relative humidity are based on univariate OI analyses that assimilate meteorological station observations. Observation departures are computed relative to the background state at the analysis time rather than at the observation time. Soil temperature and soil moisture on three levels are based on a modified version of the SiB model (*Sato et al.*, 1989; *Sellers et al.*, 1986) forced by atmospheric reanalysis fields applied every 6 h.
JRA-55	Surface air analyses differ from those in JRA-25 in two ways. First, comparisons between observations and the first-guess background state are evaluated at observation times rather than analysis times. Second, screen-level observations over islands are not used as they may not be appropriately representative of conditions at the scale of the model grid cell. Representation of the land surface state is similar to that in JRA-25, but atmospheric forcing is applied every 3 h instead of every 6 h.
MERRA	MERRA did not conduct separate surface air or land surface analyses. Screen-level temperature and humidity measurements over land are not assimilated, although surface air variables in both ANA and IAU products are affected by surface pressure and upper air measurements assimilated during the analysis cycle. Estimates of land surface properties represent the time-integrated effects of coupling between the LSM (*Koster et al.*, 2000) and surface conditions and fluxes generated by the atmospheric model during the IAU "corrector" segment (see *Section 2.3*). A separate land surface analysis (MERRA-Land) was conducted by replacing model-generated precipitation with pentad-resolution GPCP data and using an updated version of the LSM (*Reichle et al.*, 2011).
MERRA-2	Like MERRA, MERRA-2 does not conduct a land surface analysis; however, precipitation inputs to the LSM are primarily based on observations rather than model-generated values between 60°S and 60°N (*Reichle et al.*, 2017a). The reanalysis does not assimilate screen-level temperature or humidity measurements over land. Surface meteorological variables over land thus primarily reflect the net effects of assimilated surface pressures, model-generated surface fluxes (which are directly affected by precipitation corrections), and the upper-air assimilated state (which is not). The LSM features several adjustments relative to MERRA and MERRA-Land (*Reichle et al.*, 2017b).
NCEP-NCAR R1	The reanalysis does not assimilate screen-level temperature or humidity measurements over land, although surface air variables are affected by surface pressure and upper air measurements assimilated during the standard analysis cycle. The land surface analysis includes soil moisture and soil temperature on two layers. Rather than an assimilation, this analysis is constructed by driving the 2-layer OSU LSM (*Pan and Mahrt*, 1987; *Mahrt and Pan*, 1984) using analyses of snow cover (**Table 2.16**) and atmospheric reanalysis fields as forcings. Soil moisture and temperature are relaxed toward a specified climatology.
NCEP-DOE R2	Similar to *R1*, but with precipitation inputs to the LSM corrected for consistency with pentad-mean precipitation data from CMAP. Also, the relaxation of soil variables to climatological values used in *R1* was not used in *R2*.
CFSR / CFSv2	Similar to *R1* and *R2*, but using the 4-layer Noah LSM (*Ek et al.*, 2003). The precipitation forcing is a blended estimate combining pentad-mean CMAP data, the CPC daily-mean gauge-based analysis, and precipitation produced by the atmospheric model. The weights for the blending depend on location, especially latitude. Other forcing data are taken from the coupled atmosphere–ocean reanalysis. The LSM is fully coupled to the atmosphere throughout the diurnal cycle, but the land surface analysis is performed only once per day (at 00UTC) for better consistency with the temporal resolution of the precipitation forcing.
NOAA-CIRES 20CR v2	Surface pressure is the only variable assimilated by the system; no land surface analysis is performed. The model is coupled to the 4-layer Noah LSM (*Ek et al.*, 2003).

benefit of producing a more internally-consistent atmospheric state. Reanalyses using this second approach are mutually independent with respect to external analyses of surface air temperatures over land (e.g., *CRUTEM*; *Osborn and Jones, 2014*); reanalyses using the first approach are not.

Land surface state variables that are simulated by atmospheric reanalyses include soil moisture and soil temperature. Analyses of these variables are not directly affected by data assimilation, but are instead produced by LSMs forced entirely or primarily by the reanalysis atmospheric state. In addition to the different treatments of surface air variables discussed above, a key difference among reanalyses in this respect is the source of the precipitation forcing, which may be taken from the atmospheric model, from observations, or from a combination of the two. The complexity and implementation of the land surface models used by reanalyses also varies widely. These aspects are covered in more detail in *Chapter 2E*.

Table 2.16: *Treatment of snow in the reanalysis systems. Additional details are provided in Chapter 2E.*

Reanalysis System	Treatment of Snow
ERA-40	A snow analysis is performed outside of the main atmospheric reanalysis using Cressman interpolation with successive corrections. Assimilated observations include station observations of snow depth and gridded estimates of snow cover from satellites. Observations of snow depth are limited to Canada before 1966 and to Canada and the former Soviet Union between 1966 and 1976 (*Uppala et al.*, 2005). The snow depth analysis is relaxed toward a climatology when observations are unavailable.
ERA-Interim	Similar to ERA-40.
ERA-20C	Snow depth, albedo, temperature and density are simulated using the model described by *Dutra et al.* (2010).
ERA5	Similar to ERA-Interim, but using a two-dimensional OI analysis (*de Rosnay et al.*, 2015) as opposed to Cressman interpolation. The snow model has also been updated relative to that used by ERA-Interim (*Dutra et al.*, 2010), and the snow depth analysis is no longer relaxed toward a climatology when observations are unavailable.
JRA-25 / JCDAS	A separate OI snow depth analysis is performed once per day. The first-guess background state combines the land-surface analysis and gridded satellite observations. Weekly NOAA snow cover analyses are used in place of gridded satellite observations when the latter are unavailable. The analysis ingests in situ observations of snow depth from selected archives (*Onogi et al.*, 2007).
JRA-55	Some differences relative to JRA-25. The first-guess background state combines the land-surface analysis, gridded satellite observations, and climatological values over ice sheets. Climatological values are used in place of gridded satellite observations when the latter are unavailable. The analysis ingests in situ observations of snow depth from selected archives (*Kobayashi et al.*, 2015).
MERRA	The evolution of snow mass, depth, and heat content is simulated using a three-layer snow model (*Stieglitz et al.*, 2001). No snow analysis is produced.
MERRA-2	Similar to MERRA in most respects; however, a detailed representation of the surface properties of land ice sheets is introduced that includes the evolution of overlying snow layers (*Gelaro et al.*, 2017). No snow analysis is produced.
NCEP-NCAR R1	Snow is treated as a single layer of frozen water with a uniform density. Weekly snow cover analyses from the NSIDC are used for the NH between 1967 and September 1998, after which they are replaced with daily analyses. Snow cover analyses are not available in the SH or in the NH before 1967; climatologies are used instead. Weekly analyses are not interpolated in time, so snow variables change discontinuously every seven days. Model-simulated snow depths are ignored and replaced using an empirical function of model temperature. Several errors have been identified (*Kanamitsu et al.*, 2002; *Kistler et al.*, 2001). For example, the snow cover analysis mistakenly re-used 1973 data for the entire 1974 - 1994 period, and conversion of snow to water during melting was overestimated by three orders of magnitude.
NCEP-DOE R2	Snow is simulated as a single layer of frozen water with a uniform density via a budget equation that accounts for accumulation and melting. Weekly analyses of NH snow cover from the NSIDC are interpolated to daily resolution until September 1998, after which they are replaced with daily analyses. Snow cover analyses are not available in the SH, where model-generated values are used instead. The model-predicted evolution of snow depth is used when it is consistent with ingested snow cover. When this condition is not met, snow is either removed or added, with snow depth in the latter case determined via an empirical function of model temperature.
CFSR / CFSv2	Snow is simulated using a three-layer snow model (*Ek et al.*, 2003; *Koren et al.*, 1999). Simulated snow variables are evaluated and adjusted using external analyses of global snow depth and NH snow cover. These external analyses are not available for dates prior to February 1997, but are used to supplement and correct the snow depth analyses after this date. Model-estimated snow depths are only adjusted if they differ from the analysed depth by more than a factor of two, and are used as is when analysed values are unavailable. A prognostic snow layer is also included in the sea ice model.
NOAA-CIRES 20CR v2	Snow is simulated using a three-layer snow model (*Ek et al.*, 2003; *Koren et al.*, 1999).

Snow cover and its evolution have important impacts on climate (*e.g.*, *Cohen and Entekhabi*, 1999), including the stratospheric circulation and its coupling with the troposphere (*Cohen et al.*, 2014; *Allen and Zender*, 2010; *Orsolini and Kvamstø*, 2009). **Table 2.16** summarizes the models and analysis techniques used to represent snow in reanalyses. Several of the reanalyses produce analyses of snow cover and snow depth using station observations of snow depth. Gridded, observationally-based analyses of snow cover and/or depth may be assimilated as additional constraints, used to help constrain the background state prior to assimilating station observations, or applied (when available) as the primary determinant for the presence or absence of snow. Four of the reanalyses (ERA-20C, MERRA, MERRA-2, NOAA–CIRES 20CRv2) simulate the evolution of snow using snow models forced by the atmospheric reanalysis and the land surface state, with no adjustment based on observational data.

2.3 Assimilation Schemes

2.3.1 Basics of data assimilation

This section provides a brief overview of data assimilation concepts and methods as implemented in current reanalysis systems. More detailed summaries have been provided by *Krishnamurti and Bounoua* (1996), *Bouttier and Courtier* (1999), and *Kalnay* (2003), among others. In this context, an analysis is a best estimate of the true state of the atmosphere at a given time t. Reanalysis systems use objective analysis methods that employ mathematical optimization (data assimilation) techniques to combine model-generated forecasts and observed data, given constraints that are intended to preserve consistency. The results should be reproducible, internally consistent, and spatially continuous.

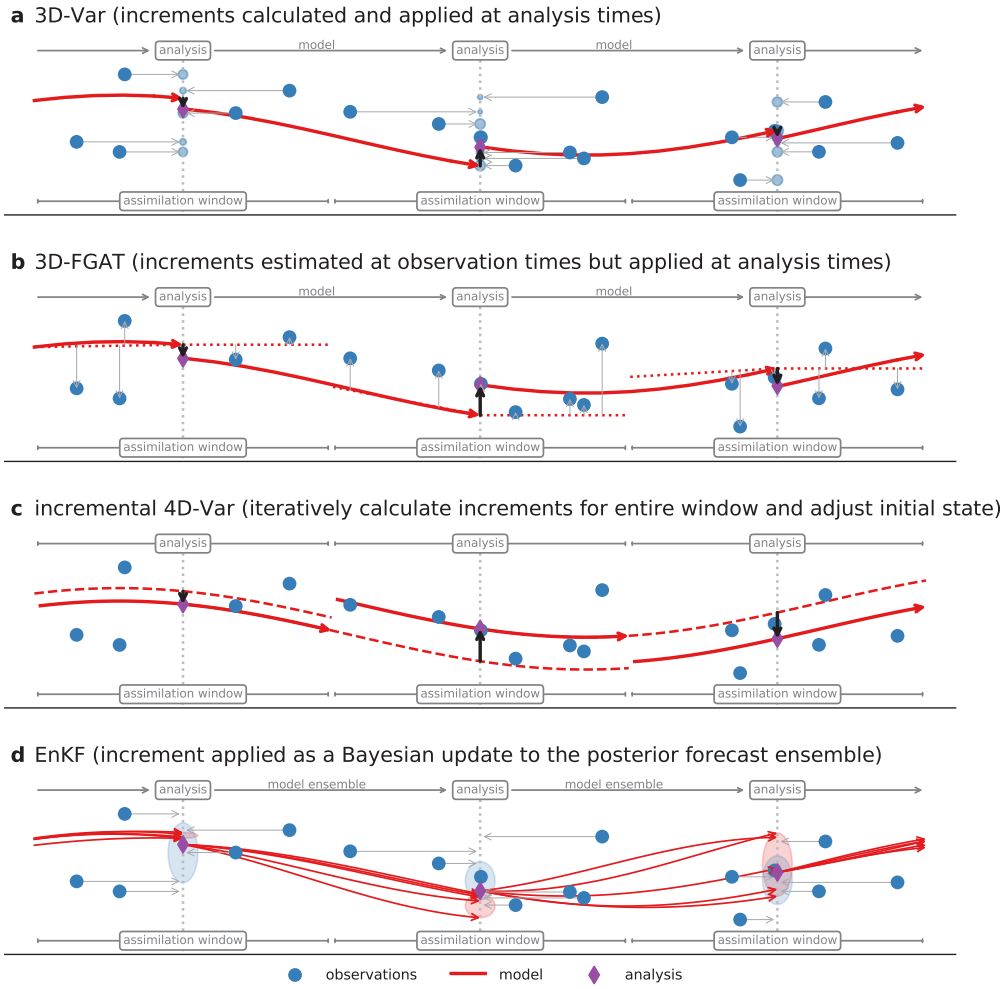


Figure 2.6: *Simplified schematic representations of four data assimilation strategies used by current reanalyses: (a) 3D-Var; (b) 3D-FGAT (here the 'semi-FGAT' approach used by NCEP–NCAR R1 and NCEP–DOE R2 is shown); (c) incremental 4D-Var; and (d) EnKF. Blue circles represent observations, red lines represent the model trajectory, and purple diamonds indicate the analysis. The dotted red lines in (b) represent linearly interpolated/extrapolated first guesses used to estimate increments at observation times. The dashed red lines in (c) represent the initial forecasts, prior to iterative adjustments. These illustrations are conceptual, and should not be taken as exact depictions of the much more complex strategies used by reanalysis systems. Updated from Fujiwara et al. (2017).*

Data ingested into an analysis system may include observations and variables from a first guess background state (such as a previous analysis or forecast). Analysis systems are constructed to be consistent with known or assumed physical properties (such as smoothness, hydrostatic balance, geostrophic or gradient-flow balance, or more complex non-linear balances). Both the observations and the background state include important information, and neither should be considered as 'truth': both the model and observations include errors and uncertainties. An analysis system must therefore adopt a consistent and objective strategy for minimizing the differences between the analysis and the (unknown) true state of the atmosphere. Such strategies are intended to reduce the extent to which errors and uncertainties in both observations and the first-guess background state influence the final analysis state. To this end, data assimilation algorithms often employ statistics to represent the range of potential uncertainties in the background state, observations, and any techniques used to convert between model and observational space (*i.e.*, observation operators), and ultimately aim to minimize these potential uncertainties.

An observation operator (also sometimes referred to as a "forward operator") is a function that converts information from the first guess background state space to the observation space, thus permitting direct comparisons between the model state and observed variables. Different types of observations require different types of observation operators. Key functions performed by observation operators include spatial interpolation from the model grid to observation locations and the transformation of model variables to observed quantities (*i.e.*, the estimation of satellite radiances via the application of a radiative transfer model to the first guess profile; see also **Table 2.19**). Errors in the observation operators constitute a portion of the observation errors considered by the data assimilation scheme.

The analysis methods used by current reanalysis systems include variational methods (3D-Var and 4D-Var) and the ensemble Kalman filter (EnKF). Variational methods (*e.g., Talagrand*, 2010) minimize an objective cost function that simultaneously penalizes differences between the analysis and observations and differences between the analysis and the model background state, with consideration of uncertainties in both the observations and the model. Implementations of variational data assimilation may be applied to derive optimal states at discrete times (3D-Var), or to identify optimal state trajectories within finite time windows (4D-Var). In EnKF (*e.g., Evensen*, 2009), an ensemble of forecasts is used to define a set of background states (the prior ensemble), which is then combined with observations and associated uncertainties to derive a set of analysis states that is consistent with the posterior distribution. The optimal analysis states are determined by applying a Kalman filter (*Kalman*, 1960) to this posterior ensemble (see also *Evensen and van Leeuwen*, 2000). If a single analysis state is required, this is typically derived by averaging the ensemble members, although this approach often leads to fields that are spatially smoother than any of the individual ensemble members, particularly in regions of sharp gradients. One of the key advantages of 3D-Var, 4D-Var, and EnKF methods relative to many earlier implementations of data assimilation is the ability to account for indirect and possibly nonlinear relationships between observed quantities and analysis variables. This ability permits the direct assimilation of satellite radiance data without an intermediate retrieval step (*Tsuyuki and Miyoshi*, 2007), and underpins many of the recent advances in reanalysis development.

Figure 2.6 shows simplified one-dimensional schematic representations of four data assimilation strategies used by current reanalysis systems (3D-Var, 3D-FGAT, 4D-Var, and EnKF). In the following discussion, we frequently refer to the analysis increment, which is defined as the adjustment applied to the first guess (forecast) background state following the assimilation of observational data (*i.e.*, the difference between the analysis state and the first guess background state). We also use the term observation increment, which refers to the difference between the observation and the background state after the observation operator is applied. This concept is also referred to in the literature as the observational 'innovation' (see detailed discussion by *Kalnay*, 2003). The analysis increment reflects the combined adjustment after evaluating and weighting (see also *Section 2.4.2*) all observation increments within an assimilation window, where the assimilation window is the time period containing observations that influence the analysis. The assimilation window used in reanalyses is typically between 6 and 12 hours long but can be as long as 24 hours. This window is often (but not always) centred at the analysis time. Core differences among the data assimilation strategies used in current reanalysis systems can be understood in terms of how the analysis increment is calculated and applied.

The 3D-Var method (**Figure 2.6a**) calculates and applies analysis increments only at discrete analysis times. Observation increments within the assimilation window may either be treated as though they were all at the analysis time (which approximates the average observation time) or weighted by when they occurred (so that observations collected closer to the analysis time have a stronger impact on the analysis increment). JRA-25 uses a 3D-Var method for data assimilation under the former assumption, in which all observations within the assimilation window are treated as valid at the analysis time. In practice, many 3D-Var systems estimate observation increments at observation times rather than analysis times (**Figure 2.6b**). This approach is referred to as 3D-FGAT ('first guess at the appropriate time'; *Lawless*, 2010). The implementation of 3D-FGAT in reanalysis systems varies. For example, *R1* and *R2* are 'semi-FGAT' systems in that observation increments are estimated relative to a linear interpolation between the initial and final states of the forecast before the analysis time and relative to a constant state after the analysis time (*i.e.*, these systems effectively use a pure 3D-Var approach for the portion of the assimilation window after the analysis time). The illustration provided in **Figure 2.6b** corresponds to this semi-FGAT approach. Other 3D-FGAT systems break each forecast into multiple piecewise segments of 30 minutes (ERA-40), one hour (CFSR), or three hours (MERRA and

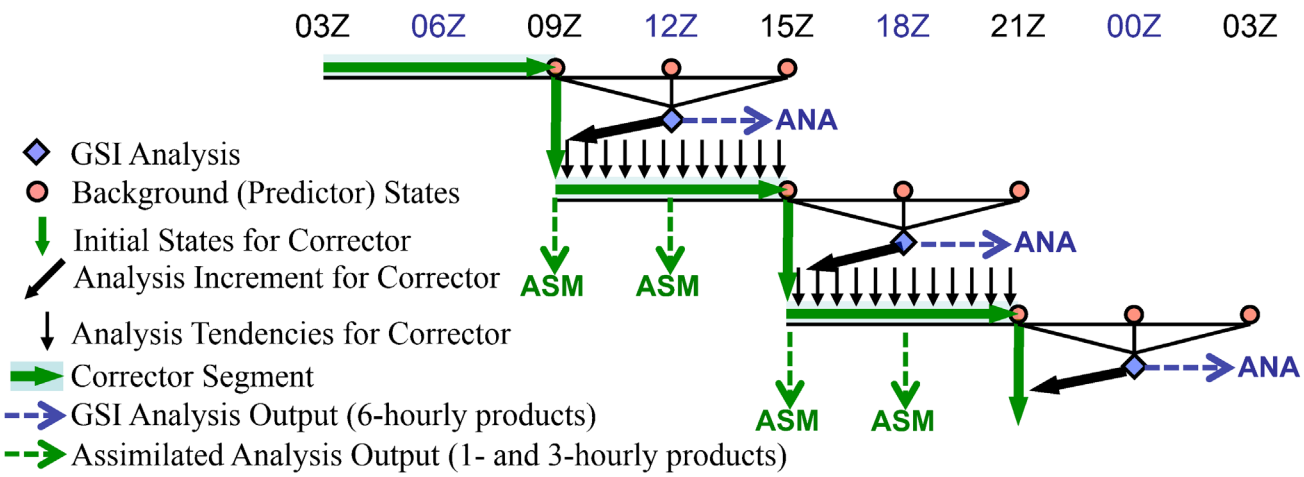


Figure 2.7: *A schematic illustration of the DAS procedure used to create ANA products and the IAU procedure used to create ASM products as implemented in MERRA and MERRA-2 (modified from Rienecker et al., 2011). See text for details.*

MERRA-2) in length. Observation increments are calculated by interpolating to observation times within each piecewise segment and then used to estimate analysis tendencies for each piecewise segment. These analysis tendencies are then combined to construct the full analysis increment.

MERRA and MERRA-2 include an additional step relative to other 3D-FGAT systems, and generate two separate sets of reanalysis products (designated 'ANA' for the analyzed state and 'IAU' for the incremental analysis update state) using an iterative predictor–corrector approach (*Rienecker et al.*, 2011). The ANA products are analogous to the analyses produced by other 3D-FGAT systems, and are generated by using the data assimilation scheme to adjust the background state produced by a 12-h 'predictor' forecast (from 9 h before the analysis time to 3 h after). The IAU products (also referred to as 'ASM') have no analogue among other 3D-FGAT reanalyses. These latter products are generated by conducting a 6-h 'corrector' forecast centred on the analysis time and using the IAU procedure (*Bloom*, 1996) to apply the previously calculated analysis increment gradually at each model time step rather than abruptly at the analysis time. The corrector forecast is then extended 6 h to generate the next predictor state. This iterative predictor–corrector procedure is illustrated in **Figure 2.7**. Note that the IAU state has only seen half of the analysis increment by the original analysis time, so that differences between the IAU and ANA states correspond to approximately half of the analysis increment. Moreover, the inclusion of the analysis increment as an additional tendency term may alter the physical tendency terms produced by the atmospheric model. For example, diabatic temperature tendencies produced by MERRA and MERRA-2 are archived during the corrector step rather than the predictor step. This arrangement applies to all tendency terms (moisture, momentum, ozone, *etc.*) and introduces a conceptual difference relative to the tendencies produced by other reanalyses (which are archived prior to the analysis during the initial forecast step), though it is important to emphasize that the analysis tendency is needed to close the budget in either case. For MERRA and MERRA-2, ANA products represent the closest match to assimilated observations, while the IAU products provide a more complete and consistent suite of atmospheric variables and tendency terms with reduced wind and tracer imbalances relative to the 3D-FGAT analyzed state (see also **Table 2.18** and associated discussion). IAU products should be used for transport simulations and other applications for which internal consistency is a priority (see also technical note on appropriate use of MERRA-2 products at **https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/docs/ANAvsASM.pdf**). MERRA and MERRA-2 analysis increments for temperature, winds, water vapor, and ozone are included in a subset of the data products provided by these systems.

Unlike 3D-Var and 3D-FGAT, which attempt to optimize the fit between assimilated observations and the atmospheric state at discrete analysis times, 4D-Var (**Figure 2.6c**) attempts to optimize the fit between assimilated observations and the time-varying forecast trajectory within the full assimilation window (*e.g.*, *Park and Županski*, 2003). 4D-Var makes more complete use of observations collected between analysis times than 3D-Var or 3D-FGAT, and has been shown to substantially improve the resulting analysis (*Talagrand*, 2010). However, the computational resources required to run a 4D-Var analysis are much greater than the computational resources required to run a 3D-Var or 3D-FGAT analysis, and the full implementation of 4D-Var is impractical for atmospheric reanalyses. Current reanalysis systems using 4D-Var (such as ERA-Interim, ERA-20C, ERA5, and JRA-55) therefore apply the simplified 'incremental 4D-Var' approach described by *Courtier et al.* (1994). Under this approach, the model state at the beginning of the assimilation window is iteratively adjusted to obtain progressively better fits between the assimilated observations and the forecast trajectory. This iterative adjustment process propagates information both forward and backward in time, which benefits the analysis but requires the derivation and maintenance of an adjoint model. The latter is a difficult and time-consuming process, and is a significant impediment to the implementation of 4D-Var. Incremental 4D-Var is tractable (unlike full 4D-Var), but it is still computationally expensive, and is therefore usually implemented in

two nested loops for computational efficiency. Analysis increments are first tested and refined in an inner loop using the tangent linear model (and its adjoint) with reduced resolution and simplified physics. This approach takes advantage of the fact that the cost function for the tangent linear model is perfectly quadratic, thus permitting the use of efficient optimization algorithms designed especially for quadratic functions. The final analysis increments are then applied in an outer loop with full resolution and full physics after the inner loop converges.

Most implementations of variational methods in reanalysis systems are based on single deterministic forecasts. By contrast, EnKF (**Figure 2.6d**) uses an ensemble approach to evaluate and apply analysis increments, thus generating an ensemble of analysis states at each analysis time. Major advantages of the ensemble Kalman filter technique include ease of implementation (unlike 4D-Var, EnKF does not require an adjoint model) and the generation of useful estimates of analysis uncertainties, which are difficult to obtain when using variational techniques with single forecasts (ERA5 uses 4D-Var in a reduced-resolution 'ensemble of data assimilations', in part to address this issue). Although the assimilation of satellite radiances presents some unique challenges in EnKF (P*olavarapu and Pulido*, 2017 ; *Campbell et al.*, 2010), recent work provides approaches to overcome this problem (*Lei et al.*, 2018). *Whitaker et al.* (2009) found that 4D-Var and EnKF perform comparably well in the case of a reanalysis that assimilates only surface pressure observations, and that both 4D-Var and EnKF give more accurate results than 3D-Var in this case. 20CR uses an EnKF method for data assimilation.

As discussed in *Section 2.2.4*, some reanalyses use simpler methods (such as OI or Cressman interpolation) for certain types of data assimilation, especially analyses of screen-level meteorological variables or snow depth. In Cressman interpolation (*Cressman*, 1959), the analysis is iteratively 'corrected' toward the set of observed values, with weighted observation increments that reduce with distance according to a specified window function. The radius of influence defined by this window function is typically reduced on successive iterations so that the closest observations have the largest influence on the final analysis. OI (*Gandin*, 1963) is formulated as a multiple linear regression problem in which both the observations and the background state are assumed to be unbiased, with known random errors. Standard OI is a special case of two of the methods discussed above, and can be functionally equivalent to both 3D-Var (assuming linear observation operators and Gaussian errors) and to the Kalman filter (assuming constant background error covariance). Although the assumptions involved in Cressman interpolation and OI are rarely satisfied, they offer a flexibility in application that can be valuable for estimating analysis increments in variables with highly heterogeneous spatial distributions (such as surface air temperature).

Additional details regarding these methods, including relative advantages and disadvantages, have been discussed and summarized by *Park and Županski* (2003), L*orenc and Rawlins* (2005), *Kalnay et al.* (2007a; 2007b), *Gustafsson* (2007), and *Buehner et al.* (2010a; 2010b), among others.

The assimilation of observational data can introduce spurious artefacts into reanalyses of the state and variability of the upper troposphere, stratosphere, and mesosphere. For example, data assimilation can act to smooth sharp vertical gradients in the vicinity of the tropopause. The potential importance of this effect is illustrated by abrupt changes in vertical stratification near the tropopause at the beginning of the satellite era in *R1* (*Birner et al.*, 2006). Changes in data sources and availability can also lead to biases and artificial oscillations in temperature in various regions of the stratosphere, particularly in the polar and upper stratosphere where observations are sparse (*Lawrence et al.*, 2015; *Simmons et al.*, 2014; *Uppala et al.*, 2005; *Randel et al.*, 2004). Information and errors introduced by the input data and data assimilation system propagate upwards through the middle atmosphere in both resolved waves and parameterized gravity wave drag (*Polavarapu and Pulido*, 2017). The effects of this propagation are often but not always undesirable. The abrupt application of analysis increments can generate spurious gravity waves in systems that use intermittent data assimilation techniques (*Schoeberl et al.*, 2003), including most implementations of 3D-Var, 3D-FGAT, and EnKF, and may also generate instabilities that artificially enhance mixing and transport in the subtropical lower stratosphere (*Tan et al.*, 2004). Although most reanalysis systems use techniques to reduce these effects (see **Table 2.18** in the following section), reanalyses of the stratosphere and mesosphere are nonetheless quite sensitive to the details of the data assimilation scheme and input data at lower altitudes.

2.3.2 Data assimilation in reanalysis systems

Table 2.17 summarizes the schemes used for atmospheric data assimilation in the reanalysis systems, which include variations on the 3D-Var, 3D-FGAT, 4D-Var, and EnKF techniques.

As noted above, the application of analysis increments can generate spurious instabilities in the atmospheric state, particularly when these increments are applied intermittently (as in 3D-Var). Several methods have been developed to mitigate these effects, including nonlinear normal mode initialization techniques and the application of digital filters. Nonlinear normal mode initialization (*Daley*, 1981; *Machenhauer*, 1977) limits the impacts of spurious instabilities by reducing or eliminating the tendencies associated with all "fast-mode" disturbances (*i.e.*, gravity waves) in the vertical and horizontal domains. By contrast, digital filter initialization (*Lynch*, 1993) aims to reduce or eliminate high-frequency noise in the temporal domain. Both approaches can be applied as strong constraints (in which all potentially undesirable modes are eliminated) or as weak constraints (in which potentially undesirable modes are penalized rather than eliminated entirely).

Table 2.17: *List of assimilation schemes used for atmospheric analyses.*

Reanalysis System	Assimilation Schemes
ERA-40	3D-FGAT with a 9-h forecast step ending three hours after the analysis time and a 6-h assimilation window centred on the analysis time. Analysis tendencies are calculated in 30-minute windows and then combined to construct the analysis increment.
ERA-Interim	Incremental 4D-Var atmospheric analysis with 12-h assimilation windows extending from 03 UTC to 15 UTC and from 15 UTC to 03 UTC. Analysis increments are calculated on coarser grids that approach the model resolution over successive iterations.
ERA-20C	Incremental 4D-Var analysis with 24-h assimilation windows extending from 09 UTC to 09 UTC. Like earlier ECMWF reanalyses, assumed background error covariances are invariant in time; however, a scaling is applied for consistency with time-varying background errors produced by an earlier 10-member ensemble pilot reanalysis that also assimilated only surface observations (*Poli et al.*, 2013, 2016).
ERA5	Similar to ERA-Interim, but with assimilation windows extending from 09 UTC to 21 UTC and from 21 UTC to 09 UTC. A 10-member 'ensemble of data assimilations' is conducted on a coarser grid, providing more robust estimates of analysis uncertainties and background error covariances.
JRA-25 / JCDAS	3D-Var (not 3D-FGAT) with 6-h forecast steps. Observations from 3 hours before the analysis to 3 hours afterwards are considered.
JRA-55	Incremental 4D-Var with a 9-h forecast step that extends 3 h past the analysis time and a 6-h assimilation window centred on the analysis time. Analysis increments are calculated on a coarser T106/F80 inner grid (rather than the TL319/N160 outer grid used in the forecast model) to limit computational expense.
MERRA	3D-FGAT using the gridpoint statistical interpolation (GSI; *Wu et al.*, 2002; *Kleist et al.*, 2009) scheme with incremental analysis update (IAU; *Bloom et al.*, 1996) and 6-h assimilation windows centred on each analysis time. The IAU procedure (illustrated in **Figure 2.7**) is summarized in the text.
MERRA-2	GSI with IAU as in MERRA, but with updated background error specifications. A global constraint is imposed on the analysis increment of total water (*Takacs et al.*, 2015).
NCEP-NCAR R1	Spectral statistical interpolation (SSI; Parrish and Derber, 1992) in a 3D-Var 'semi-FGAT' configuration (see text) with a 6-hour assimilation window centred on each analysis time. For times before the analysis time, first guesses are based on linear interpolation between the initial and final model states. For times after the analysis time, first guesses are estimated as the first guess at the analysis time.
NCEP-DOE R2	Same as NCEP-NCAR R1.
CFSR / CFSv2	GSI with 9-h forecasts (from 6 h before to 3 h after each analysis time) and 6-h assimilation windows (centred on each analysis time). The implementation of GSI in CFSR is a form of 3D-FGAT with hourly first guesses.
NOAA-CIRES 20CR v2	Ensemble Kalman filter (EnKF) with a 6-h window centred on each analysis time. Observations from 3 hours before the analysis to 3 hours afterwards are used. The EnKF implementation in 20CR uses a window that straddles the analysis time, and is therefore technically an Ensemble Kalman Smoother (*Compo et al.*, 2011).

Table 2.18: *Initialization procedures used to mitigate assimilation-driven instabilities.*

Reanalysis System	Initialization procedure
ERA-40	Nonlinear normal mode initialization
ERA-Interim	Weak constraint digital filter
ERA-20C	Weak constraint digital filter
ERA5	Weak constraint digital filter
JRA-25 / JCDAS	Nonlinear normal mode initialization
JRA-55	None
MERRA	IAU
MERRA-2	IAU
NCEP-NCAR R1	None
NCEP-DOE R2	None
CFSR / CFSv2	6-h digital filter (*Lynch and Huang*, 1992)
NOAA-CIRES 20CR v2	None

Table 2.19: *List of radiative transfer schemes used for assimilating satellite radiances.*

Reanalysis System	Radiative transfer scheme used for assimilating satellite radiances
ERA-40	RTTOV-5 is used for assimilating satellite radiances.
ERA-Interim	RTTOV-7 is used for assimilating satellite radiances.
ERA-20C	Satellite radiances are not assimilated (see also **Table 2.21**).
ERA5	RTTOV-11 is used for assimilating satellite radiances. Note that where ERA-40 and ERA-Interim only assimilated clear-sky radiances (see also **Table 2.23**), ERA5 assimilates all-sky radiances from certain sensors.
JRA-25 / JCDAS	RTTOV-6 is used for assimilating TOVS radiances and RTTOV-7 is used for assimilating ATOVS radiances.
JRA-55	RTTOV-9 is used for assimilating satellite radiances.
MERRA	The GLATOVS radiative transfer model is used for assimilating SSU radiances; the CRTM is used for assimilating all other satellite radiances.
MERRA-2	All radiances are assimilated using version 2.1.3 of the CRTM.
NCEP-NCAR R1	Satellite radiances are not assimilated (see also **Table 2.21**).
NCEP-DOE R2	Satellite radiances are not assimilated (see also **Table 2.21**).
CFSR / CFSv2	CFSR uses the CRTM developed at NOAA/NESDIS and the JCSDA for assimilating satellite radiances.
NOAA-CIRES 20CR v2	Satellite radiances are not assimilated (see also **Table 2.21**).

Certain data assimilation techniques also aim to reduce the impacts of spurious instabilities and/or eliminate the need for initialization techniques. For example, one of the benefits of the SSI analysis technique (*Parrish and Derber*, 1992) developed at NCEP and used in R1 and R2 was that it imposed a global balance constraint on the analysis that eliminated the need for nonlinear normal mode initialization (*Kalnay et al.*, 1996). It should be noted, however, that balance constraints and filters (particularly those applied as strong constraints) may eliminate real information along with spurious noise. The loss of this information can have particularly detrimental effects in the middle atmosphere, where gravity waves that propagate upward from lower levels play important roles in the dynamics (*Polavarapu and Polido*, 2017). The application of IAU, as in MERRA and MERRA-2, can help to eliminate spurious instabilities without affecting other "fast-mode" disturbances in the model atmosphere. The use of IAU has been shown to improve the representation of the mesosphere in data assimilation systems (*e.g.*, *Sankey et al.*, 2007).

The assimilation of observed satellite radiances by a reanalysis system requires the use of a radiative transfer scheme. This scheme typically differs from that used in the forecast model (**Table 2.4**). **Table 2.19** lists the radiative transfer schemes used by each reanalysis system for assimilating satellite radiances.

2.4 Observational Data

2.4.1 Summary of basic information

This section provides information on key observational data assimilated in the reanalysis systems. Reanalysis systems assimilate observational data from a variety of sources. These sources are often grouped into two main categories: conventional data (*e.g.* surface records, radiosonde profiles, and aircraft measurements) and satellite data (*e.g.* microwave and infrared radiances, atmospheric motion vectors inferred from satellite imagery, and various retrieved quantities).

The densities and distributions of both types of observational data have changed considerably over time. **Figure 2.8** shows examples of the spatial distributions of observations assimilated by JRA-55 in the 1980s (00 UTC, 22 September 1983), while **Figure 2.9** shows examples of the spatial distributions of observations assimilated by the same reanalysis system in the 2010s (00 UTC, 23 June 2010). These two sets of examples are representative of the distribution and number of observations assimilated in most recent reanalysis systems (with the notable exception of ERA-20C and 20CR, which do not assimilate upper-air observations). **Figures 2.10** through **2.13** summarize the availability of different

Table 2.20: *List of codes/acronyms for selected observations assimilated by reanalysis systems.*

SYNOP (conventional)	Surface meteorological observation reported by manned and automated weather stations.
SHIP (conventional)	Surface meteorological observations reported by ships.
BUOY (conventional)	Surface meteorological observations reported by buoys.
PAOBS (conventional)	Surface pressure bogus data for the southern hemisphere. This was a product of human analysts in the Australian Bureau of Meteorology who estimated sea level pressure based on satellite imagery, conventional data and temporal continuity. Production and distribution of PAOBS ceased in mid-August 2010.
AMV (satellite)	Atmospheric motion vectors derived by tracing the movement of individual cloud or water vapour features in successive images from geostationary and polar-orbiting satellites.

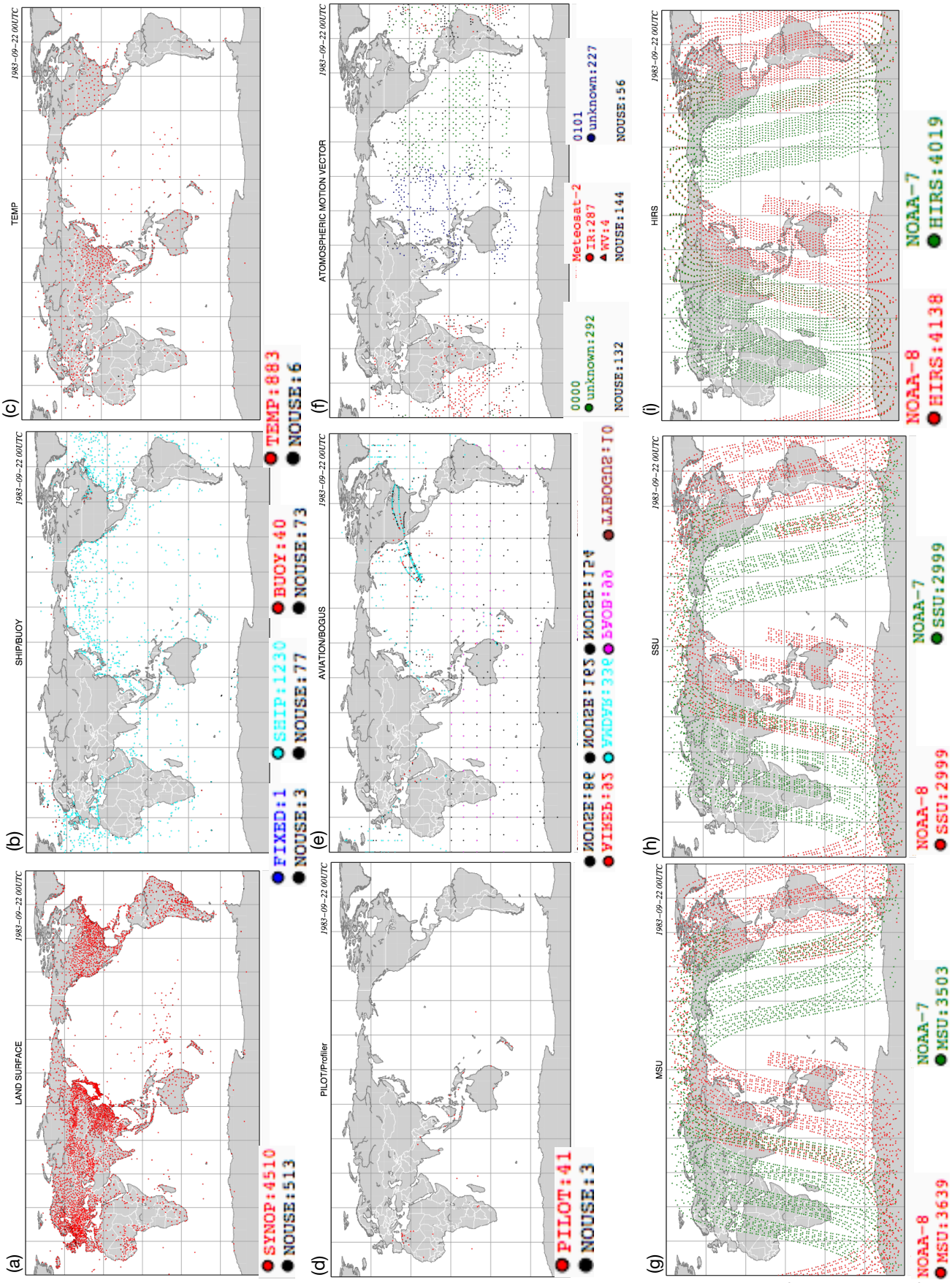
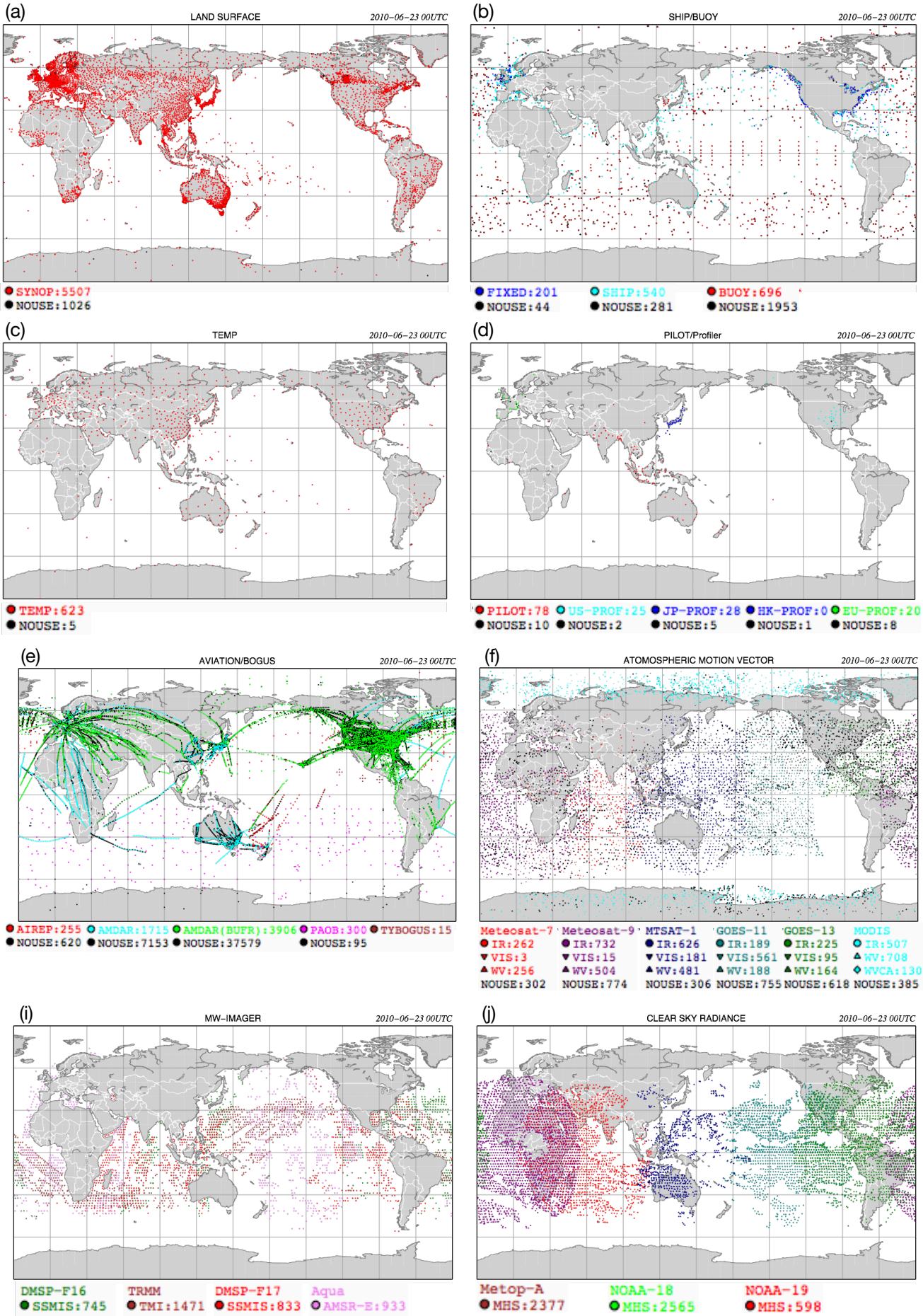


Figure 2.8: *Observations assimilated by JRA-55 at 00UTC 22 September 1983 (±3 hours): (a) land surface data, (b) surface meteorological data reported by ships and buoys, (c) radiosonde profiles, (d) pilot balloons, (e) aircraft, PAOBS, and tropical cyclone wind retrievals, and (f) atmospheric motion vectors from METEOSAT, GMS, and GOES satellites, (g) Microwave temperature sounder radiances from NOAA satellites, (h) stratospheric temperature sounder radiances from NOAA satellites, and (i) infrared sounder radiances (sensitive to temperature and moisture) from NOAA satellites.*


(a)
LAND SURFACE
2010-06-23 00UTC
SYNOP:5507
NOUSE:1026
(b)
SHIP/BUOY
2010-06-23 00UTC
FIXED:201
NOUSE:44
SHIP:540
NOUSE:281
BUOY:696
NOUSE:1953
(c)
TEMP
2010-06-23 00UTC
TEMP:623
NOUSE:5
(d)
PILOT/Profiler
2010-06-23 00UTC
PILOT:78
NOUSE:10
US-PROF:25
NOUSE:2
JP-PROF:28
NOUSE:5
HK-PROF:0
NOUSE:1
EU-PROF:20
NOUSE:8
(e)
AVIATION/BOGUS
2010-06-23 00UTC
AIREP:255
NOUSE:620
AMDAR:1715
NOUSE:7153
AMDAR(BUFR):3906
NOUSE:37579
PAOB:300
NOUSE:95
TYBOGUS:15
(f)
ATOMOSPHERIC MOTION VECTOR
2010-06-23 00UTC
Meteosat-7
IR:262
VIS:3
WV:256
NOUSE:302
Meteosat-9
IR:732
VIS:15
WV:504
NOUSE:774
MTSAT-1
IR:626
VIS:181
WV:481
NOUSE:306
GOES-11
IR:189
VIS:561
WV:188
NOUSE:755
GOES-13
IR:225
VIS:95
WV:164
NOUSE:618
MODIS
IR:507
WV:708
WVCA:130
NOUSE:385
(i)
MW-IMAGER
2010-06-23 00UTC
DMSP-F16
SSMIS:745
TRMM
TMI:1471
DMSP-F17
SSMIS:833
Aqua
AMSR-E:933
(j)
CLEAR SKY RADIANCE
2010-06-23 00UTC
Metop-A
MHS:2377
NOAA-18
MHS:2565
NOAA-19
MHS:598

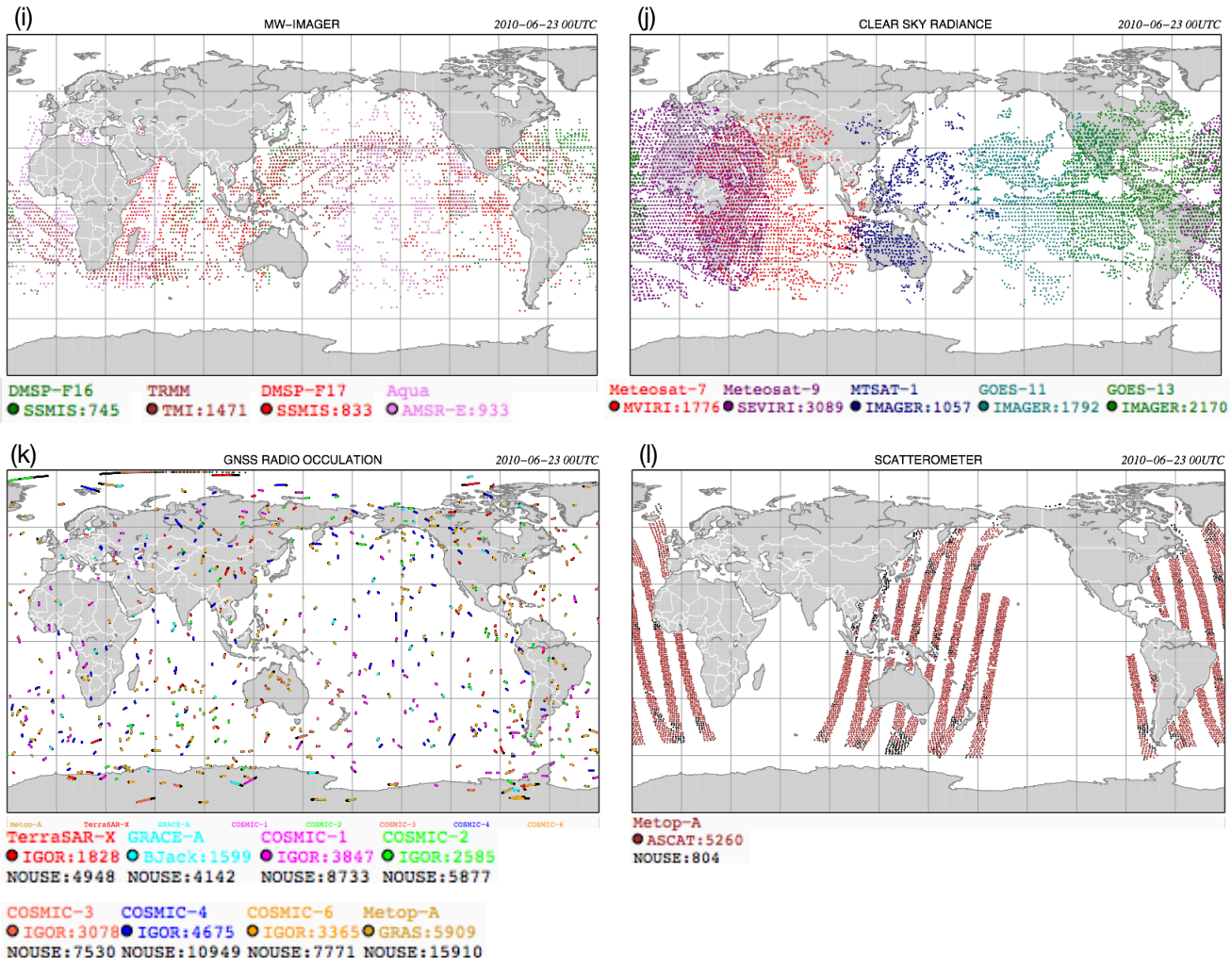


Figure 2.9: *Observations assimilated by JRA-55 at 00UTC 23 June 2010 (±3 hours): (a) land surface data, (b) surface meteorological data reported by ships and buoys, (c) radiosonde profiles, (d) pilot balloons and wind profilers, (e) aircraft, PAOBS, and tropical cyclone wind retrievals, and (f) atmospheric motion vectors from the METEOSAT, MTSAT, GOES, Aqua, and Terra satellites (g) microwave temperature sounder radiances from the NOAA, MetOp, and Aqua satellites, (h) microwave humidity sounder radiances from NOAA and MetOp satellites, (i) microwave imager radiances (sensitive to moisture) from the DMSP, TRMM, and Aqua satellites, (j) clear-sky radiances from METEOSAT, MTSAT, and GOES satellites, (k) GNSS-RO refractive index data (sensitive to temperature and moisture) from the COSMIC, GRACE, MetOp, and TerraSAR-X satellites, and (l) ocean surface winds from MetOp (ASCAT scatterometer).*

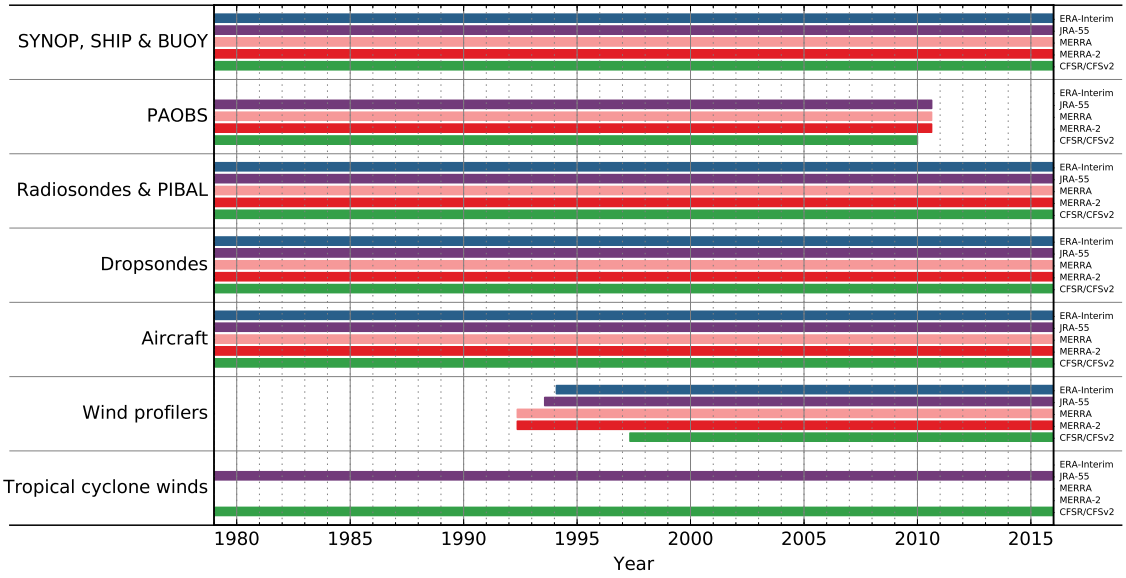


Figure 2.10: *Availability of conventional observations assimilated by ERA-Interim (blue), JRA-55 (purple), MERRA (dark red), MERRA-2 (light red), and CFSR (green) reanalysis systems as a function of time. See* ***Table 2.20*** *and Appendix B for acronym definitions. Reproduced from Fujiwara et al. (2017).*

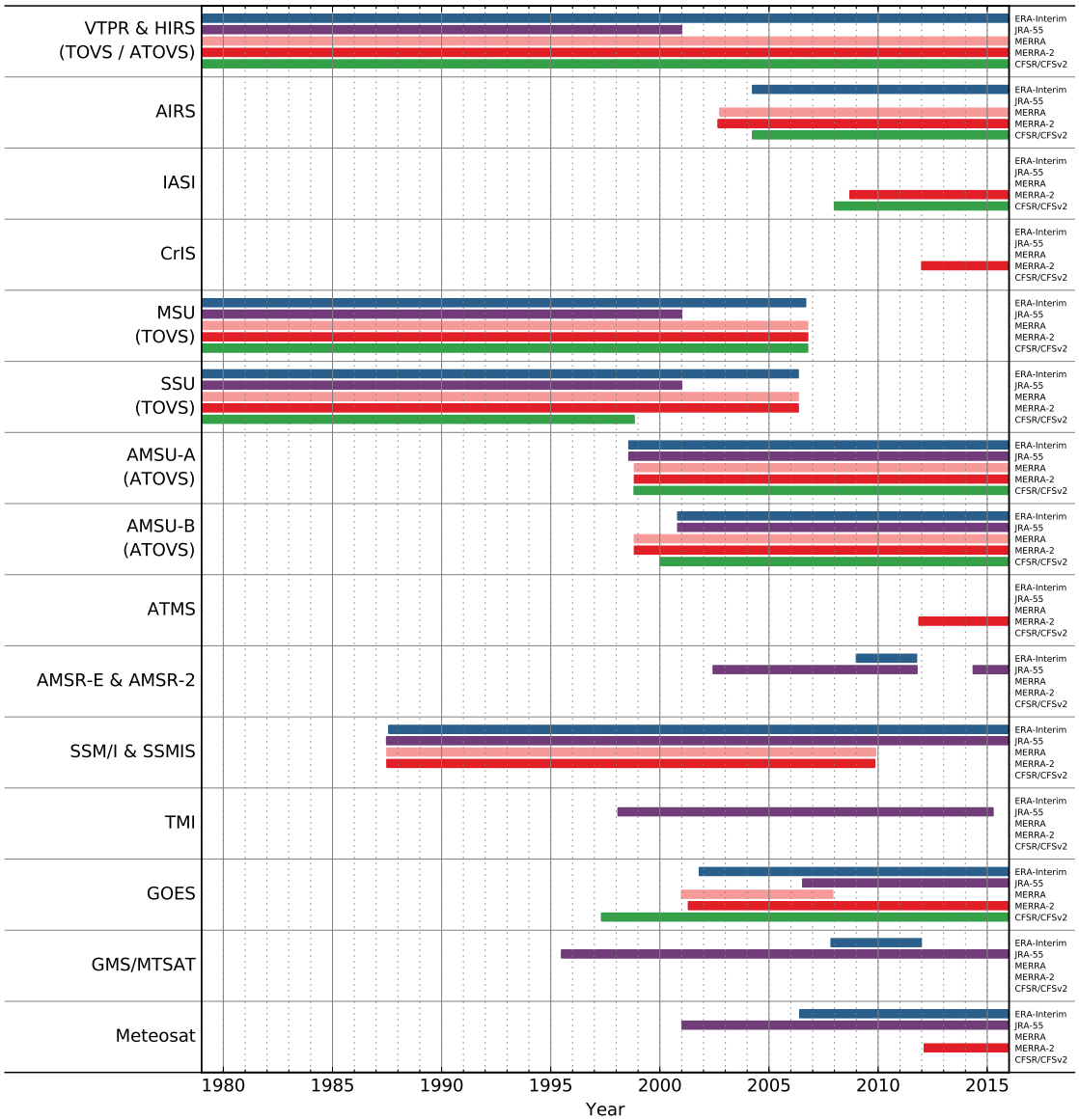


Figure 2.11: *As in* ***Figure 2.10****, but for satellite radiances assimilated by the reanalysis systems. Reproduced from Fujiwara et al. (2017).*

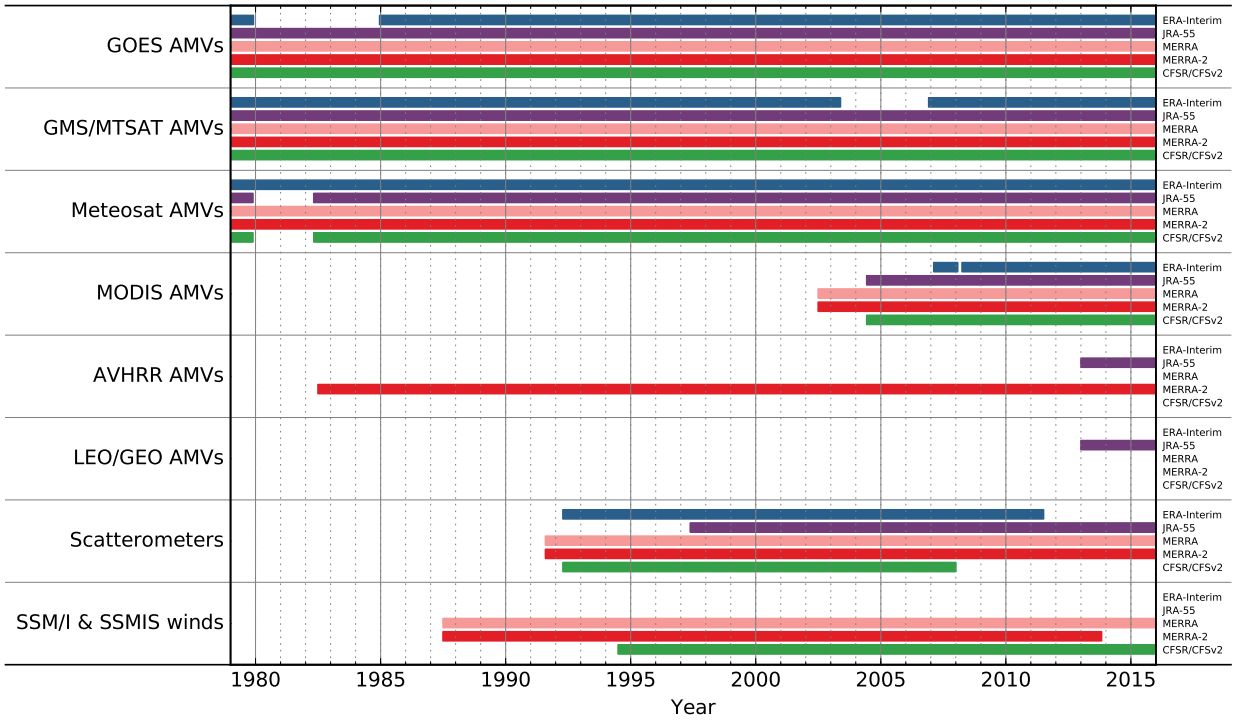


Figure 2.12: *As in* ***Figure 2.10****, but for AMVs and ocean surface wind products derived from satellites and assimilated by the reanalysis systems. Reproduced from Fujiwara et al. (2017).*

types of observations assimilated in five of the most recent reanalysis systems as a function of time. **Figure 2.14** provides a more detailed look at how the availability of radiances observed by certain instruments changes as satellites are launched and retired. Common codes and terminology for assimilated observations are listed in **Table 2.20**.

Several key details are apparent in **Figures 2.8** through **2.14**. First, conventional in-situ data (such as surface, radiosonde, and aircraft data) are unevenly distributed in space. Second, satellite data (microwave and infrared sounder data, air motion vector data from geostationary and polar satellites, *etc.*) are often more evenly distributed but still inhomogeneous in space. Third, none of these datasets are continuous and homogeneous in time. For example, microwave and infrared sounders (*i.e.*, the TOVS suite) were introduced in 1979, while advanced sounders (*i.e.*, the ATOVS suite) were introduced in 1998. Such changes in the availability of observational data for assimilation have strong impacts on the quality of the reanalysis datasets that assimilate them, so that discontinuities in reanalysis data should be carefully evaluated and checked for coincidence with changes in the input observations. The quality of a given type of measurement is also not necessarily uniform in time; for example, virtually all radiosonde sites have adopted different instrument packages over time (see *Section 2.4.2.1*), while TOVS and ATOVS data were collected using several different sounders on several different satellites with availability that changed over time (see **Figure 2.14** and *Section 2.4.2.2*). Finally, **Figures 2.10** through **2.13** show that, although modern reanalysis systems assimilate observations from many common sources, different reanalysis

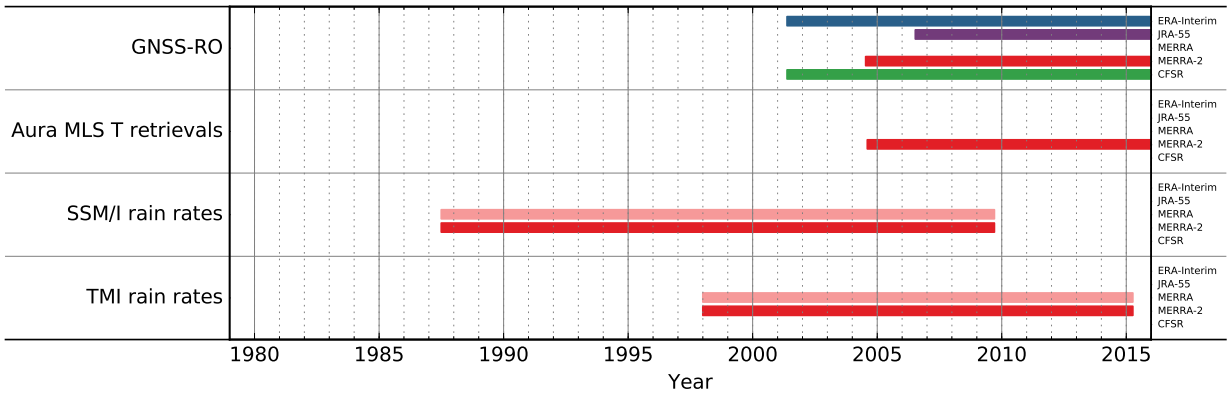


Figure 2.13: *As in* ***Figure 2.10****, but for other types of satellite observations assimilated by the reanalysis systems. Timelines of satellite retrievals of total column ozone and ozone profiles assimilated by the reanalysis systems are provided in Chapter 4 of this report (****Figures 4.1*** *and* ***4.2****). Timelines of GNSS-RO observations assimilated from different sets of sensors are provided in* ***Figure 2.17****. Reproduced from Fujiwara et al. (2017).*

systems assimilate different subsets of the available observations. Such discrepancies are particularly pronounced for certain categories of satellite observations and, like differences in the underlying forecast models, are an important potential source of inter-reanalysis differences.

Timelines of conventional data assimilated by reanalyses are quite consistent among modern full input reanalyses (**Figure 2.10**), as well as the conventional input JRA-55C (not shown). All of the reanalysis systems discussed in this chapter assimilate records of surface pressure from manned and automated weather stations, ships, and buoys, while all but 20CR assimilate at least some records of surface winds over oceans. All but ERA-Interim, ERA5, ERA-20C, 20CR, and JRA-55C assimilated synthetic surface pressure data for the Southern Hemisphere (PAOBS) through at least 2009. PAOBS are subjective analyses of surface pressure produced by the Australian BOM based on available observations and temporal continuity, which are used to compensate for the scarcity of direct observations in the Southern Hemisphere. The influence of these data in reanalysis systems has waned in recent years, as the availability of direct observations covering the Southern Hemisphere has expanded. All of the full input reanalyses and JRA-55C assimilate upper-air observations made by radiosondes, dropsondes, and wind profilers. JRA-25, JRA-55, and JRA-55C assimilate wind speed profiles in tropical cyclones, while 20CR assimilates records of tropical cyclone central pressures. CFSR uses the NCEP tropical storm relocation package (*Liu et al.*, 1999) to relocate tropical storm vortices to observed locations. ERA5 assimilates PAOBS before 1979 to improve its representation of tropical cyclones during the pre-satellite era. ERA-40, ERA-Interim, MERRA, MERRA-2, NCEP-NCAR R1 and NCEP-DOE R2 have no special treatment for tropical cyclones.

Timelines of satellite data assimilated by current reanalysis systems are more varied (**Figures 2.11** through **2.13**; see also **Figure 2.17** and **Figures 4.1** and **4.2**), but still include many commonalities. The core satellite data assimilated by most reanalyses are microwave and infrared radiances from a variety of instruments. All of the full input reanalyses (including NCEP-NCAR R1 and NCEP-DOE R2) also assimilate atmospheric motion vector (AMV) data derived from geostationary and polar-orbiting satellite imagery. Many of the more recent systems assimilate GNSS-RO data, while MERRA-2 assimilates

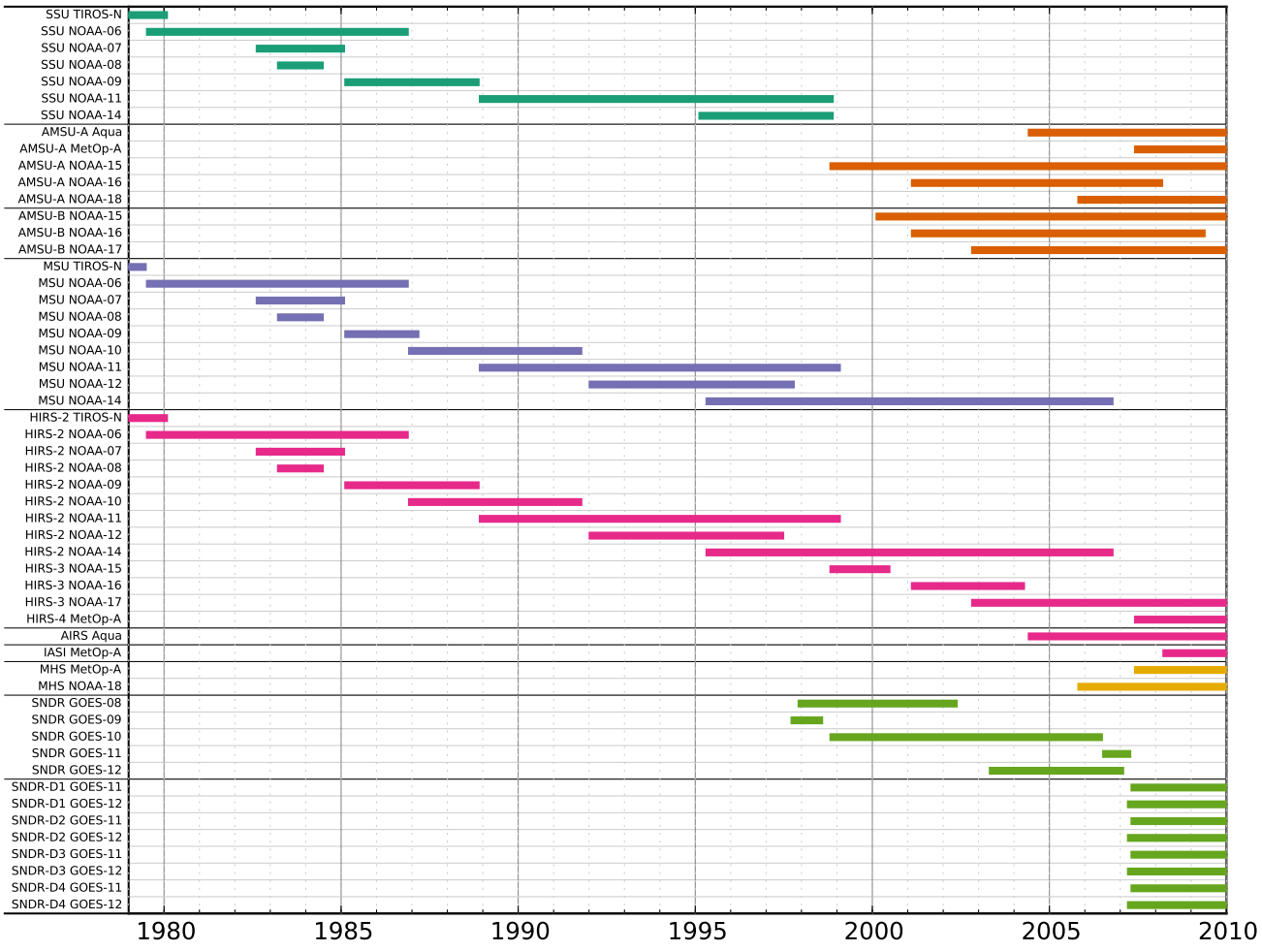


Figure 2.14: *Usage of satellite instruments with radiances assimilated by CFSR as a function of time. Adapted from Saha et al. (2010). Original © American Meteorological Society. Used with permission.*

Table 2.21: *Special features regarding observational data assimilated in each reanalysis system (see also* ***Figures 2.10*** *through* ***2.13*** *for five recent full input reanalyses).*

Reanalysis System	Special features of assimilated observational data
ERA-40	SSM/I total column water vapor and surface wind retrievals were assimilated. Neither GNSS-RO data nor AIRS radiances were assimilated (ERA-40 effectively predates these data types). No special treatment for tropical cyclones was included.
ERA-Interim	GNSS-RO bending angles and AIRS radiances are assimilated. Unlike ERA-40, SSM/I radiances are assimilated directly (in place of TCWV and surface wind retrievals). No special treatment for tropical cyclones was included.
ERA-20C	ERA-20C assimilated surface pressure observations from ISPD (*Cram et al.,* 2015) and surface pressure and surface wind observations from ICOADS (*Woodruff et al.*, 2011). Reports that appear in both the IPSD and ICOADS databases were taken from ICOADS, with the IPSD report discarded. Tropical cyclone best track data were assimilated, but with relatively large rejection rates during quality control (*Poli et al.*, 2016).
ERA5	GNSS-RO bending angles are assimilated. AIRS radiances are assimilated, as are hyperspectral radiances observed by IASI and CrIS, microwave soundings from ATMS, and infrared and microwave radiances from several sounding instruments on the Chinese FY-3 series of meteorological satellites. Radiances from several microwave imagers are assimilated directly, including SSM/I and SSMIS, TMI, and GMI, as well as visible and infrared radiances from AHI. Variational bias corrections have been added for ozone, aircraft measurements, and surface pressure. PAOBS are assimilated to improve the representation of tropical cyclones in the pre-satellite era.
JRA-25 / JCDAS	Total column water vapor retrievals from SSM/I and AMSR-E were assimilated, as were wind profile retrievals in tropical cyclones. SSM/I surface winds, GNSS-RO data, and AIRS radiances were not assimilated.
JRA-55	GNSS-RO refractivity data are assimilated, as are wind profile retrievals in tropical cyclones. Clear-sky radiances from selected channels of microwave imagers such as SSM/I, TMI, and AMSR-E are assimilated over ocean (*Kobayashi et al.*, 2015). Neither SSM/I surface winds nor hyperspectral radiances were assimilated. Variational bias corrections have been added for non-blacklisted satellite radiances.
MERRA	AIRS radiances were assimilated, as were rain rates from SSM/I and TMI. SSM/I radiances were assimilated through late 2009, and surface winds were assimilated throughout. GNSS-RO data were not assimilated and no special treatment for tropical cyclones was included.
MERRA-2	GNSS-RO bending angles are assimilated up to 30 km. AIRS radiances are assimilated, as are hyperspectral radiances observed by IASI, CrIS and ATMS. MLS temperature retrievals are assimilated above 5 hPa (version 3.3 through 31 May 2015; version 4.2 from 1 June 2015). A new adaptive bias correction scheme is applied to aircraft observations (see also *Section 2.4.2.3*). Assimilated aerosol optical depths are also bias-corrected. Rain rates from SSM/I and TMI and satellite observations of AOD are assimilated, as are SSM/I surface wind retrievals. SSM/I radiances were assimilated through late 2009. No special treatment for tropical cyclones was included.
NCEP-NCAR R1	Temperature retrievals from microwave and infrared sounders are assimilated, rather than radiances. The horizontal and vertical resolutions of temperature retrievals are downgraded to reduce the weight given to satellite data in recent analyses. Satellite moisture retrievals and SSM/I surface winds are not assimilated.
NCEP-DOE R2	Same as NCEP-NCAR R1.
CFSR / CFSv2	GNSS-RO bending angles and radiances from AIRS and IASI are assimilated. SSM/I radiances are not assimilated, but surface wind retrievals are. The NCEP tropical storm relocation package is applied to relocate tropical storm vortices to observed locations.
NOAA-CIRES 20CR v2	Only observations of surface pressure, sea level pressure, and tropical cyclone central pressure were assimilated. No upper-air or satellite data were assimilated.

temperature retrievals from Aura MLS at pressures 5 hPa and less. Timelines of satellite ozone retrievals assimilated by reanalyses are discussed in Chapter 4 of this report (**Figures 4.1** and **4.2**).

Table 2.21 lists special features of each reanalysis system regarding observational data assimilated. Note that NCEP-NCAR R1 and NCEP-DOE R2 assimilated temperature retrievals from microwave and infrared sounders (*e.g.*, *Reale*, 2001), while the other reanalysis systems (except for surface-input reanalyses) assimilated radiance observations directly. Some systems use bias correction procedures. These are described in *Section 2.4.3*.

2.4.2 Quality control procedures

The observations assimilated by reanalyses are subjected to rigorous quality control procedures that are intended to prevent the introduction of errors into the analysis. Key steps in the quality control algorithm for each reanalysis are listed in **Table 2.22**. Common quality control procedures are briefly described in the following paragraphs (see also *Kalnay, 2003*).

A typical first step in quality control is a preliminary screening. This step eliminates observations with incomplete or duplicate data records, as well as observations that have previously been 'blacklisted' by either the data provider or the reanalysis center.

Many data assimilation systems include automated procedures that try to correct incomplete data records to reduce the number of observations that are eliminated at this stage. The preliminary screening is typically followed by tests to identify and exclude data with physically unreasonable values. The latter may take several different forms. The simplest, the 'gross check', involves comparison against climatological values. Observations are excluded from the analysis if the gross check indicates that they differ from the expected value by more than a specified threshold amount. This type of test may be supplemented (or superseded) by comparison to other reasonable expected values, such as the average of other nearby observations (*i.e.*, a 'buddy check') or the forecast background

Table 2.22: *Standard quality control procedures applied in the reanalysis systems.*

Reanalysis System	Quality control procedure
ERA-40	• Preliminary screening and exclusion of incomplete, duplicate, and blacklisted data • Thinning of selected observation types • Check that the departure from the first-guess is below a threshold that depends on expected error statistics • Variational quality control applied during the analysis step
ERA-Interim	Similar to ERA-40, but with updated thresholds.
ERA-20C	• Preliminary screening and exclusion of incomplete, duplicate, and blacklisted data • In the case of duplicates, precedence is given to ICOADS over ISPD • Wind observations over land and near coasts are excluded • Data are excluded if more than three constant values are reported within a five-day window • Background check eliminates data with departures large (more than seven times expected) relative to the combined error variance from the pilot ensemble • Variational quality control applied during the analysis step
ERA5	Similar to ERA-Interim, but with updated thresholds and additional information from the reduced-resolution ensemble of data assimilations.
JRA-25 / JCDAS	• Preliminary screening and exclusion of incomplete, duplicate, and blacklisted data • Gross check against climatology for most observation types, with thresholds determined using the "dynamic" method proposed by *Onogi* (1998) • Track checks against expected locations for ships, buoys, and aircraft • Complex quality control for radiosondes • Data thinning is applied to AMVs and some TOVS radiances to make the data distribution more uniform
JRA-55	Similar to JRA-25, but thresholds have been reviewed and updated (*Sakamoto*, 2009)
MERRA	• Preliminary screening and exclusion of incomplete, duplicate, and blacklisted data • Check that the departure from the first-guess background state is below a threshold that depends on observation type • Data thinning is applied to all radiance data
MERRA-2	Similar to MERRA, but with revised thresholds for departures from the first-guess background state.
NCEP-NCAR R1	• Complex quality control, including a hydrostatic check and correction • Data exclusion for unrealistic values, duplicate records, ship measurements over land, and blacklisted data • Thinning of selected observation types • Aircraft rejected during certain phases of flight • Background and buddy checks to eliminate observations with large departures • Quality control based on observations within ±24 hours rather than only the assimilation window • Horizontal and vertical thinning of satellite temperature retrievals to reduce the impact of resolution improvements over time
NCEP-DOE R2	Similar to NCEP-NCAR R1.
CFSR / CFSv2	• Complex quality control, including a hydrostatic check and correction • Data exclusion for unrealistic values, duplicate records, ship measurements over land, and blacklisted data • Thinning of selected observation types • Aircraft rejected during certain phases of flight • Variational quality control penalizes observations based on magnitude of departure from the preliminary analysis
NOAA-CIRES 20CR v2	• Pressure observations reduced to sea level and subjected to a gross check against the plausible range 880 to 1060 hPa • Background check eliminates data with departures large (more than three times expected) relative to the combined error variance • Buddy check against nearby observations; can override the results of the background check • Data thinning eliminates observations with weak impacts on the analysis; has the added effect of capping assimilated observations at near mid-20th century levels • Correction of systematic biases (recalibrated every 60 days)

state itself. These comparisons may also be combined, for instance by performing a simple OI analysis using nearby observations (except for the observation being evaluated) and then checking for consistency between the observation and the result of the OI analysis. One benefit of this kind of approach is that it can applied iteratively, rescuing data that might have been excluded by comparison to the initial background state or eliminating data that passed the initial checks but is too far from the OI analysis. In addition to expected values, observations may be checked for consistency with expected balance criteria. For example, height measurements might be compared against heights calculated from virtual temperature measurements via the hypsometric equation. Complex quality control refers to the common practice of applying these checks in combination, and then using an algorithm to decide whether each observation should be included or excluded.

The quality control procedures described above are used to pre-select observational data for use in the analysis. Many 3D-Var and 4D-Var data assimilation systems use variational quality control (*Anderson and Järvinen*, 1999), in which observations that are far from the expected value are penalized in the analysis rather than eliminated entirely. This means that observations that fail to meet the desired criteria have less impact on the analysis, but may still be influential, especially in regions where observations are sparse. Data pre-selection and variational quality control are not mutually exclusive. For example, ERA-Interim conducts a preliminary screening for incomplete, duplicate, and blacklisted data records before starting the incremental 4D-Var assimilation. The initial iterations of the assimilation (see *Section 2.3*) are then conducted without variational quality control, so that all observations that meet the pre-selection criteria are weighted equally. Variational quality control is then turned on for the later iterations of the assimilation to limit the impacts of outlier observations on the final analysis state.

In addition to consistency checks, data may be thinned to reduce redundancy in regions where many observations are available. This procedure can have several benefits, including identifying previously undetected duplicates and reserving an independent set of observations for validating the analysis (*Compo et al.*, 2011). Quality control criteria are also intimately connected to bias correction procedures. Bias corrections may be applied to certain observations either before or during the analysis step to keep otherwise good observations with known biases from being excluded from the analysis. Some typical bias correction procedures for radiosonde, satellite, and aircraft measurements are described in the following section.

2.4.3 Summary of key upper air observations and known issues

This section discusses a selection of upper air observational data that are assimilated in one or more of the reanalysis systems and are key for SPARC sciences. Radiosondes provide high vertical resolution profiles of temperature, horizontal wind, and humidity worldwide; however, most radiosonde stations are located in the Northern Hemisphere at middle and high latitudes over land (**Figure 2.15**). The typical vertical coverage of radiosonde data extends from the surface up to ~30 hPa for temperature and wind and from the surface up to 300 ~ 200 hPa for humidity. Operational satellite radiance measurements provide constraints for temperature and moisture with more homogeneous spatial coverage, but at the cost of coarse vertical resolution (*e.g.*, **Figure 2.16**). Moreover, the majority of these measurements were not available before 1978, and no radiance data have been assimilated prior to late 1972 in these reanalyses. Both observing systems have known biases, as well as jumps and drifts in the time series that may cause the quality of reanalysis products to change over time. Bias corrections prior to and/or within the assimilation step are therefore essential for creating more reliable reanalysis products (see below for examples). In addition to radiosonde and satellite data, atmospheric motion vector (AMV) data created from geostationary and polar-orbiter satellite images and wind and temperature observations collected by aircraft are influential in the upper troposphere and lower stratosphere.

2.4.3.1 Radiosonde data

The main source of systematic errors in radiosonde temperature measurements is the effect of solar radiative heating and (to a lesser extent) infrared cooling on the temperature sensor (*Nash et al.*, 2011). This issue, which is sometimes called the 'radiation error', can cause particularly pronounced warm biases in raw daytime stratospheric measurements. These biases may be corrected onsite in the ground data receiving system before reporting, and further corrections may be applied at each reanalysis centre before assimilation. The major issue with radiosonde humidity measurements is that the sensor response is too slow at cold temperatures (*Nash et al.*, 2011). Recent advances in radiosonde instrumentation are beginning to improve this issue, particularly in the upper troposphere; however, radiosonde observations of humidity at pressures less than 300 hPa are typically not assimilated by reanalysis systems. Other issues include frequent (and often undocumented) changes in radiosonde instrumentation and observing methods at radiosonde stations, which may cause jumps in the time series of temperature and relative humidity. Several 'homogenization' activities for radiosonde temperature data exist to support climate monitoring and trend analyses (see, *e.g.*, *Seidel et al.*, 2009). Although some of these activities have been effectively independent of reanalysis activities, cooperation between the two groups has increased substantially in recent years. Particularly notable is the production of RAOBCORE (*Haimberger et al.*, 2008, 2012), which was conducted with reanalysis applications in mind. One or more versions of RAOBCORE are used in ERA-Interim (v1.3), MERRA and MERRA-2 (v1.4 through 2005), and JRA-55 (v1.4 through 2005; v1.5 thereafter). ERA5 uses the RICH dataset (v1.5.1)

rather than RAOBCORE. Further efforts on data rescue, reprocessing, homogenization, and uncertainty evaluation by the broader research community are likely to be an essential part of the next generation of reanalyses (*e.g., ACRE (Allan et al., 2011), and GRUAN (Bodeker et al., 2016)*).

The following example describes a 'homogenization' (or bias correction) of radiosonde temperature measurements for assimilation in a reanalysis system:

i. Radiosonde temperatures are corrected for estimated biases from 1980 onwards;
ii. Stations are separated into groups representing different countries or regions (because stations within the same country often use the same type of radiosonde from the same manufacturer);
iii. Mean differences between background forecasts and observations are accumulated for each group of stations;
iv. The mean error for all groups is subtracted from the bias computed for each group to provide a correction for radiation effects;

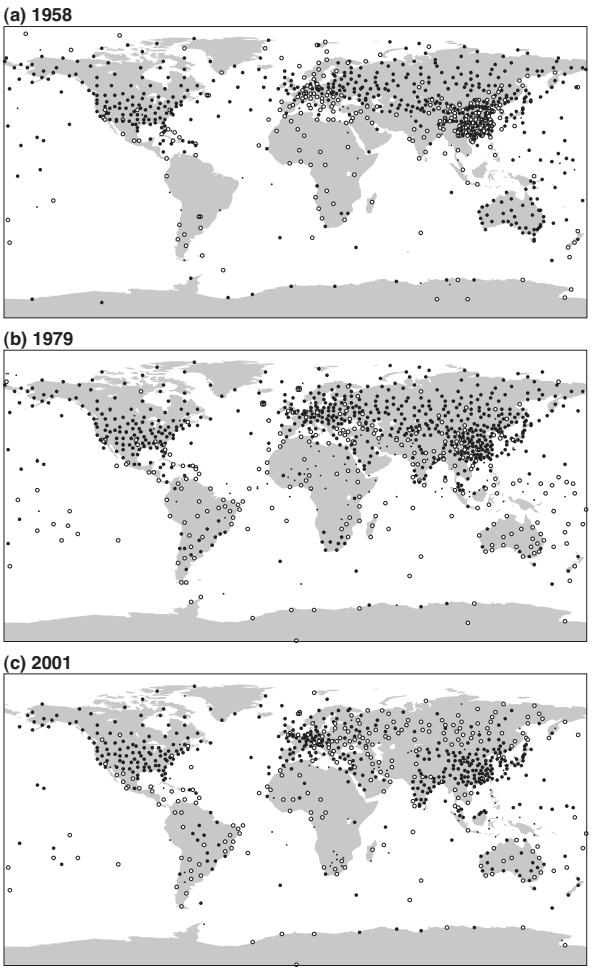


Figure 2.15: *Frequency of radiosonde reports assimilated by ERA-40 during (a) 1958, (b) 1979, and (c) 2001. Solid circles denote stations reporting three times every 2 days on average, open circles denote stations reporting at least once every 2 days, and small dots denote stations reporting at least once per week (from Uppala et al., 2005). ©Royal Meteorological Society. Used with permission.*

This approach corrected for many daily and seasonal variations of the biases but did not account for variations in annual mean biases. Radiosonde temperature measurements homogenized using this approach were assimilated in both ERA-40 and JRA-25 (*Onogi et al.*, 2007; *Uppala et al.*, 2005; *Andrae et al.*, 2004). The homogenizations applied to produce the RAOBCORE temperatures assimilated by many later reanalyses (including ERA-Interim, JRA-55, MERRA, and MERRA-2, as discussed above) have been conducted using updated versions of this procedure. Although radiosonde humidity measurements are also known to suffer from biases, current reanalysis systems do not include schemes to correct for biases in radiosonde humidities.

Major quality control criteria for radiosonde profiles (and other conventional data) include checks for completeness, physical and climatological consistency, and duplicate reports (*Section 2.4.2*). Data may also be filtered using locally compiled blacklists or blacklists acquired from other data providers and reanalysis centres. Further information on the quality control criteria applied by different reanalysis is available in the text and supporting material of the publications listed in **Table 2.1**.

Radiosonde and other upper-air in situ data are also often shared among different reanalysis centres. For example, *Rienecker et al.* (2011) listed the sources for historical radiosonde, dropsonde, and PIBAL data used by MERRA as:

i. NCEP–NCAR: Office Note 20, Office Note 29, NMC/NCEP/GTS ingest;
ii. ECMWF: ECMWF/FGGE, ECMWF/MARS/GTS ingest;
iii. JMA: Japan Meteorological Agency GTS ingest;
iv. NCAR: International archives from Argentina, Australia, Brazil, Canada, China, Dominica, France, India, Japan, NCDC, New Zealand, Russia, Singapore, South Africa, United Kingdom Research sets: PermShips, RemoteSites, Ptarmigan, Scherhaug, LIE, GATE and BAS;
v. NCDC: U.S. military and academic sources, including TD52, TD53, TD54, TD90, USCNTRL, USAF, U.S. Navy, CCARDS and MIT.

These data sources overlap substantially with those used in ERA-40 and ERA-Interim (*Tavolato and Isaksen*, 2011; *Uppala et al.*, 2005, their Appendix B), JRA-25 and JRA-55 (*Kobayashi et al.*, 2015, their Table A1; *Onogi et al.*, 2007, their section 2.1a), MERRA-2 (*McCarty et al.*, 2016), NCEP-NCAR R1 (*Kalnay et al.*, 1996, their Section 3a), and CFSR (*Saha et al.*, 2010, their section "Conventional observing systems in the CFSR"); however, individual reanalyses may supplement standard data sets with data from unique sources. A detailed intercomparison of the conventional data used in each reanalysis is beyond the

scope of this chapter; however, we note that at least four of the reanalyses (ERA-40, ERA-Interim, JRA-25, and JRA-55) use the ERA-40 ingest as a starting point, and that the ERA-40 ingest has much in common with the conventional data archives used by NCEP (*R1*, *R2*, and CFSR) and the NASA GMAO (MERRA and MERRA-2). More recent updates in data holdings at ECMWF, JMA, GMAO, and NCEP rely heavily on near-real-time data gathered from the WMO GTS, which also contributes to the use of a largely (but not completely) common set of conventional data among reanalysis systems.

2.4.3.2 Satellite data

Reanalysis systems assimilate data from several different types of satellite instruments, most notably the microwave and infrared sounders in the TOVS suite (1979–2006 on several satellites) and the ATOVS suite (1998–present on several satellites). The TOVS suite included the Stratospheric Sounding Unit (SSU), the Microwave Sounding Unit (MSU), and the High-resolution Infrared Sounder-2 (HIRS/2). The ATOVS suite includes the Advanced MSU-A (AMSU-A) and HIRS/3 (updated to HIRS/4 starting with NOAA-18). NCEP-NCAR R1 and NCEP-DOE R2 assimilate temperature retrievals from these instruments (see, *e.g.*, *Reale*, 2001). All of the other full input reanalyses described in this chapter assimilate microwave and infrared radiances from the TOVS and ATOVS suites. ERA-Interim, ERA5, MERRA, MERRA-2, and CFSR also assimilate radiances from AIRS, the first hyperspectral infrared sounder with data assimilated in reanalyses (2002–present). ERA5, MERRA-2, and CFSR assimilate hyperspectral infrared radiances from IASI (2008-present), while ERA5 and MERRA-2 also assimilate radiances from the hyperspectral infrared sounder CrIS and the most recent generation of microwave sounder ATMS (late 2011-present). ERA-Interim, ERA5, JRA-55, MERRA-2, and CFSR assimilate data from GNSS-RO instruments (CHAMP: 2001–2008; FORMOSAT-3/COSMIC: 2006–present; MetOp-A: 2008–present; and several other recent missions), in the form of bending angles or refractivity at the tangent point rather than temperature or water vapour retrievals.

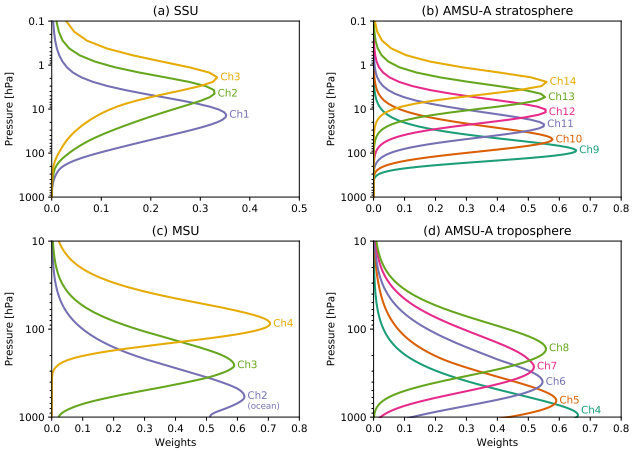


Figure 2.16: *Vertical weighting functions of radiance measurements for (a) SSU (1979–2005) channels 1–3, (b) AMSU-A (1998–present) stratospheric temperature channels 9–14, (c) MSU (1979–2006) channels 2–4, and (d) AMSU-A tropospheric temperature channels 4–8. Weighting functions are for nadir or near-nadir scan positions and have been normalized as described by Zou and Qian (2016).*

Satellite sounding instruments often have several channels with different vertical weighting functions (see, *e.g.*, **Figure 2.16**). Even when using the same satellite instrument, different reanalysis systems may assimilate data from different sets of channels. Bias corrections and quality control criteria for satellite radiances may also vary by channel. **Table 2.23** lists details of satellite data usage for five of the full input reanalysis systems considered in this chapter.

Radiances observed by the SSU instruments, which covered the period 1979–2005, represent an important archive of stratospheric temperatures (*e.g.*, *Nash*

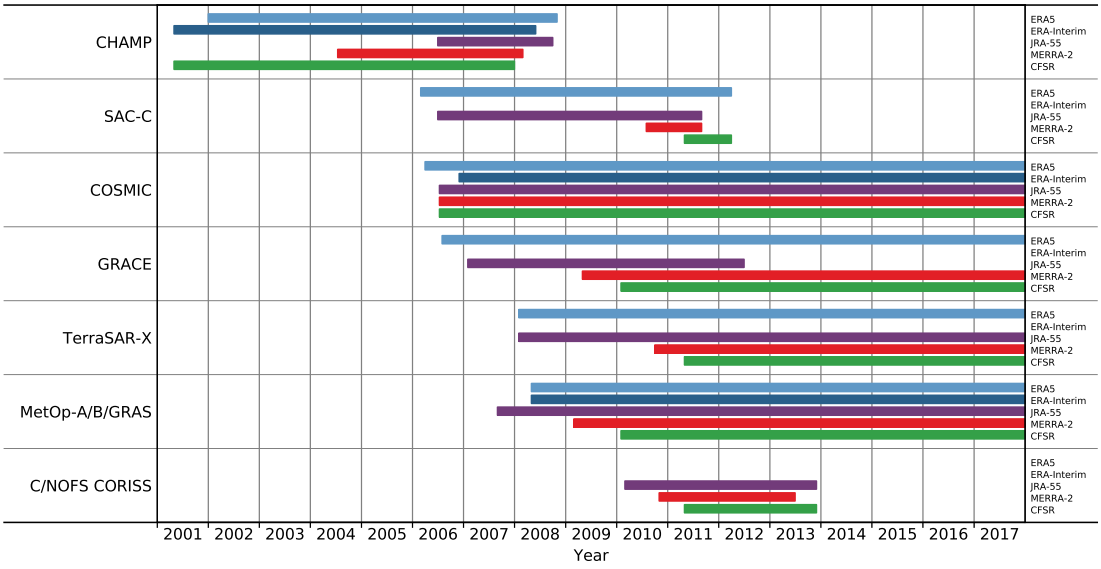


Figure 2.17: *Assimilation of GNSS-RO observations from different campaigns by five recent reanalyses.*

Table 2.23: *Overview of satellite data usage in five of the most recent full input reanalysis systems. Adapted and updated from* ***http://reanalyses.org/observations/satellite-1****. Refer to the website for source information and the latest version of this table (including information for JRA-25/JCDAS). See Appendix for acronym definitions.*

Instrument (observable)	CFSR / CFSv2	MERRA	MERRA-2	JRA-55	ERA-Interim
MSU (radiances)	Channels 1,2,3,4 **Notes**: NESDIS SNO corrected calibration coefficients applied (NOAA-10 to -14) **Exclusions**: • More restrictive QC in tropics and over high terrain • Window test ch. 2	Channels 1,2,3,4 **Notes**: NESDIS SNO corrected calibration coefficients applied **Exclusions**: • Snow, ice, mixed surfaces for ch. 1 – 2	Channels 2,3,4 **Notes**: NESDIS SNO corrected calibration coefficients applied **Exclusions**: • Restrictive QC over snow, ice and mixed surfaces • Observation errors inflated over non-water surfaces	Channels 2,3,4 **Exclusions**: • Land or rain for ch. 2 • Land for ch. 3	Channels 2,3,4 **Exclusions**: • Land or rain for ch. 2 • Land for ch. 3
AMSU-A (radiances)	Channels 1 – 13, 15 **Exclusions**: • Estimated cloud liquid water large for ch. 1 – 5, 15 • Scattering index large for ch. 1 – 6, 15 • Ch. 4 gross check large for ch. 1 – 5, 15 • Ch. 6 gross check large for ch. 1 – 6, 15 • High terrain for ch. 1 – 5, 15 • Fit to emissivity or surface temp large for ch. 1 – 5, 15	Channels 1 – 15 **Exclusions**: • Snow, ice, mixed surfaces for ch. 1– 6, 15 • No offset bias correct for ch. 14	Channels 4 – 14 **Exclusions**: • Restrictive QC • Observation errors inflated for ch. 4 – 6 over non-water surfaces	Channels 4 – 14 **Exclusions**: • Sea ice or land for ch. 4 – 5 • High terrain for ch. 6 – 7 • Rain for ch. 4 – 8	Chanels 5 – 14 **Exclusions**: • High terrain for ch. 5 – 6 • Rain for ch. 5 – 7 • No offset bias correct for ch. 14
AMSU-B / MHS (radiances)	Channels 1 – 5 **Exclusions**: • Scattering index too large • Channel 1 fit too large • Any channel failing gross check • High terrain	Channels 1 – 5 **Exclusions**: • Snow, ice, mixed surfaces for ch. 1, 2, 5	Channels 1 – 5 **Exclusions**: • Restrictive gross check • Observation errors inflated for all channels over non-water surfaces	Channels 3 – 5 **Exclusions**: • Land, sea-ice, rain	Channels 3–5 **Exclusions**: • Sea ice, rain, high terrain for ch. 3 – 4 • Land for ch. 5
SSM/I (radiances)		Channels 1 – 7 **Exclusions**: • Land	Channels 1 – 7 **Exclusions**: • All non-water surfaces	Channels 1,3,4,6 **Exclusions**: • Land, rain	Channels 1 – 7 **Exclusions**: • Land, rain
HIRS (radiances)	Channels 2 – 15 **Exclusions**: • Over water wavenumbers > 2400 during day • High terrain • Above model top • Channels without signal over clouds • Surface sensing channels with large difference.	Channels 2 – 15 **Exclusions**: • Land for ch. 5 – 8	Channels 2 – 12 **Exclusions**: • Surface-sensitive channels • Observation errors inflated over non-water surfaces	Ch. 2 – 7,11,12,14,15 **Exclusions**: • Land for ch. 4-7, 11,14,15 • High terrain for ch. 12 • Clouds for ch. 3 and above	Ch. 2 – 7,11,12,14,15 **Exclusions**: • Clouds, land for ch. 4 – 7,11,14,15 • High terrain for ch. 12
SSU (radiances)	Channels 1 – 3 **Notes**: • All channels bias-corrected.	Channels 1 – 3 **Notes**: • No offset bias correction for ch. 3	Channels 1 – 3 **Notes**: • Only ch. 1–2 after onset of NOAA-15 AMSU-A (1 Nov 1998) • No offset bias correction for ch. 3	Channels 1 – 3	Channels 1 – 3 **Notes**: • No offset bias correction for ch. 3
GEO (radiances)	GOES sounder **Notes**: • 5°×5° 1993–2007 • 1°×1° 2007–present	GOES sounder	GOES, Meteosat (after early 2012)	GOES, METEOSAT, GMS, MTSAT imagers	GOES, METEOSAT, MTSAT imagers
SSM/I (retrievals)	• Surface wind speed over oceans	• Surface wind speed ov. oceans • Rain rate	• Surface wind speed over oceans • Rain rate	• Snow cover	• Total column water vapor (rainy areas over oceans)
Imager (upper-air winds)	GOES, METEOSAT, GMS, MTSAT, MODIS	GOES, METEOSAT, GMS, MTSAT, MODIS	GOES, METEOSAT, GMS, MTSAT, MODIS	GOES, METEOSAT, GMS, MTSAT, MODIS	GOES, METEOSAT, GMS, MTSAT, MODIS
Scatterometer (winds over ocean surface)	ERS, Quikscat, ASCAT	ERS, Quikscat	ERS, Quikscat, ASCAT	ERS, Quikscat, ASCAT	ERS, Quikscat
Ozone sensors (retrievals)	SBUV V8 retrievals	SBUV V8 retrievals	SBUV V8 retrievals, OMI, MLS (v2.2 through 31 May 2015, switching to v4.2 from 1 June 2015; 261 hPa switched off from 1 May 2016)	TOMS, OMI (nudging)	TOMS, SBUV, GOME, MIPAS, SCIAMACHY, MLS, OMI
Other notable elements	• AIRS • IASI • GNSS-RO • AMSR-E • Reprocessed ERS • Reprocessed GMS • AMSU-B (NOAA-15 only)	• TMI rain rate • AIRS • NOAA-15 AMSU-B	• TMI rain rate • AIRS • IASI • CrIS • GNSS-RO • NOAA-15 AMSUB • ATMS • SEVIRI • MLS temperature retrievals (v3.3 through 31 May 2015, switching to v4.2 from 1 June 2015) above 5 hPa • AOD from MISR, MODIS, AVHRR and AERONET	• Reprocessed winds from GMS, GOES-9, MTSAT (revised) and METEOSAT • Reprocessed radiances from GMS, GOES-9, MTSAT • TMI (NASA) • AMSR-E (JAXA) • GNSS-RO • SSM/I-S • VTPR • Exclude HIRS from NOAA-15 and later	• GNSS-RO • AIRS • SSM/I-S • AMSR-E • HIRS NOAA-18

and Saunders, 2015; *Zou et al.*, 2014; *Wang et al.*, 2012) and serve as a useful illustration of the types of issues that may be encountered in assimilating satellite data. The SSU was a pressure-modulated radiometer with an on-board CO_2 cell for spectral filtering at 15 µm. The calibration of SSU radiances is affected by the following known issues:

i. Space-view anomalies due to electrical interference;
ii. CO2 gas leakage and cell pressure changes;
iii. Changes in atmospheric CO2 concentrations;
iv. Satellite orbital drift and diurnal sampling biases;
v. Short overlap periods between successive instruments.

Raw radiance data from SSU include drifts and jumps in the time series due to these issues (*e.g.*, **Figure 2.18**), which must be accounted for in the data assimilation system. Drifts and jumps of this type are not unique to SSU, and other long-term satellite radiance archives are also affected by issues specific to individual instruments. For example, *Simmons et al.* (2014; their Figure 13) have shown that estimated biases for certain MSU, HIRS, and AMSU-A channels can be of similar orders of magnitude to those for SSU, while trends in atmospheric CO_2 concentrations also cause long-term drifts in estimated biases for HIRS, AIRS, and IASI radiances unless accounted for in the observation operator. Biases in radiances observed by MSU and AMSU-A can be attributed mainly to inaccurate calibration offsets and non-linearity (*Zou et al.*, 2006).

Post-launch inter-satellite calibration (or "homogenization") efforts by the satellite remote sensing community, such as the WMO GSICS (*Goldberg et al.*, 2011) have substantially reduced inter-satellite differences in some cases, including MSU (*Zou et al.*, 2006), AMSU-A (*Zou and Wang*, 2011), and SSU (*Zou et al.*, 2014). In practice, this type of inter-satellite calibration is usually performed by reanalysis systems internally via bias correction terms applied during the data assimilation step. It is therefore not strictly necessary for satellite data to be homogenized prior to its assimilation in a reanalysis system, although it is beneficial to assimilate data with biases as small as possible.

The use of externally homogenized data has been found to improve some aspects of recent reanalyses. For example, homogenized MSU data (NESDIS SNO corrected calibration coefficients; *Zou et al.*, 2006) assimilated by CFSR, MERRA and MERRA-2 (**Table 2.23**) have been found to improve temporal consistency in bias correction patterns (*Rienecker et al.*, 2011), and may have helped MERRA to produce a more realistic stratospheric temperature response following the eruption of Mount Pinatubo (*Simmons et al.*, 2014). In situations where conventional data are unavailable or insufficient to provide a reference for satellite bias correction, such as SSU in the middle and upper stratosphere, homogenized radiance data may be even more effective in eliminating artificial drifts and jumps in the analysis state. Homogenized satellite radiance time series only represent a relatively small fraction of the satellite data ingested by current reanalysis systems (several of which do not assimilate homogenized data at all); however, the availability of homogenized satellite radiance time series is increasing and these data are likely to become more influential in future reanalysis efforts.

Bias corrections for assimilated satellite data often vary by satellite platform and/or reanalysis system. Although bias corrections are intended to limit the impacts of changing satellite biases within the reanalysis, these impacts may still manifest as spurious trends or discontinuities in the time series of temperature and other reanalysis variables. In older reanalyses that assimilated satellite radiances, such as ERA-40 and JRA-25, bias corrections were often (but not always) based on a fixed regression that spanned the lifetime of the instrument (*Sakamoto and Christy*, 2009; *Onogi et al.*, 2007; *Uppala et al.*, 2005). This approach, which occasionally required the reanalysis to be interrupted for manual retuning of bias correction terms, has been replaced by adaptive (or variational) bias correction schemes in recent reanalysis systems. Adaptive bias corrections for satellite radiances are based on differences between observed radiances and expected radiances calculated from model-generated background states. Some early implementations of adaptive bias corrections, such as that applied to TOVS data in JRA-25, left the reanalysis vulnerable to jumps and drifts inherited from the assimilated radiances (*Sakamoto and Christy*, 2009). These problems are addressed in most recent reanalysis systems by defining observational "anchors" that are regarded as unbiased and are therefore allowed to contribute directly to the background state (*Dee*, 2005). A key example is the use of homogenized radiosonde data to anchor bias corrections for satellite radiances (*e.g., Auligné et al.*, 2007). Versions of this approach have been implemented in ERA-Interim, ERA5, JRA-55, MERRA, and MERRA-2. GNSS-RO observations are also useful for anchoring bias corrections (*e.g.*, *Poli et al.*, 2010), and are used in this capacity in ERA-Interim, ERA5, JRA-55, and MERRA-2; however, GNSS-RO data are only available after May 2001, and in useful numbers only from 2006. The approach to bias correction taken by CFSR and CFSv2

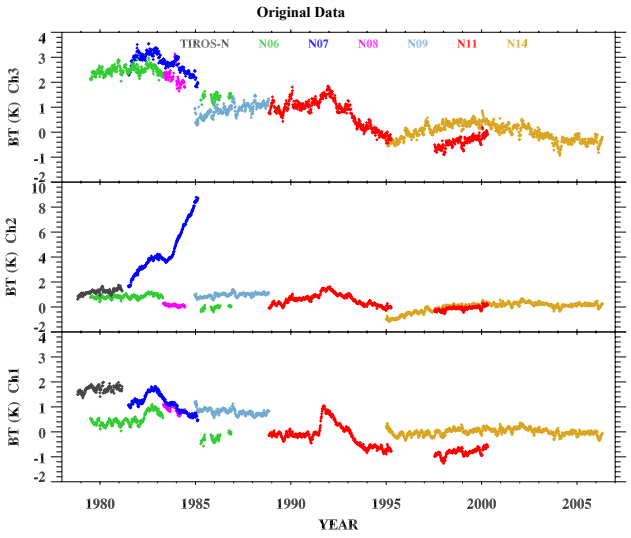


Figure 2.18: *Global mean pentad brightness temperature anomalies based on raw SSU radiances from different satellites. Anomalies are calculated relative to the 1995–2005 mean NOAA-14 annual cycle (from Wang et al., 2012). ©American Meteorological Society. Used with permission.*

(*Saha et al.*, 2010; *Derber and Wu*, 1998) differs from that taken by other systems in that anchor observations are not used. Instead, initial bias corrections are determined for each new satellite instrument via a three-month spin-up assimilation and then allowed to evolve slowly. The effects of satellite-specific drifts and jumps are kept small by assigning very low weights to the most recent biases between the observed and expected radiances, and by accounting for known historical variations in satellite performance as catalogued by multiple research centres. One byproduct of this procedure is an oscillating warm bias in CFSR in the upper stratosphere (see *Chapter 3* of this report). This bias, which is intrinsic to the forecast model, largely disappears when a new execution stream is introduced, only to slowly return as the model bias is imprinted on the observational bias correction terms.

A further example of the type of temporal discontinuities that can result from changes in satellite instrumentation is the cold bias (~2 K) in middle stratospheric temperature in JRA-25 between 1979 and 1998 (*Onogi et al.*, 2007). This feature resulted from a known cold bias in the radiative transfer model used by JRA-25. The SSU had only three channels sensitive to stratospheric temperature (too few to correct the model bias). The AMSU-A instruments, first launched in 1998, have more channels (*i.e.*, higher vertical resolution) in the stratosphere (see also **Figure 2.16**). Assimilation of the higher-resolution AMSU-A radiances effectively corrected the model bias. The JRA-55 system uses an improved radiative transfer model, and produces more realistic stratospheric temperatures during 1979 – 1998 (*Kobayashi et al.*, 2015; *Ebita et al.*, 2011).

A final illustrative example concerns temperatures in the upper stratosphere. MERRA shows artificial annual cycles in the upper stratosphere (*Rienecker et al.*, 2011; their Figure 16), which probably arise because the forward radiative transfer model used to assimilate SSU radiances did not consider variations in atmospheric CO_2. These issues have been corrected in MERRA-2, which uses version 2.1.3 of the CRTM to assimilate SSU radiances (**Table 2.19**). Several reanalyses also show jumps in upper stratospheric temperature in or around 1998 (the sign varies by vertical level and reanalysis) due to the introduction of AMSU-A, which includes channels that peak higher in the stratosphere. See *Chapter 3* of this report for further details and additional examples.

2.4.3.3 Aircraft data

Measurements made by aircraft, such as the AMDAR data collection, are influential inputs in many atmospheric analyses and reanalyses (*Petersen*, 2016). Horizontal wind data from aircraft are assimilated in all of the reanalysis systems but ERA-20C and 20CR, while temperature data from aircraft are assimilated in all of the reanalysis systems except for ERA-20C, JRA-55, JRA-25, and 20CR. In principle, aircraft data were assimilated from the outset by ERA- 40 (September 1957; *Uppala et al.*, 2005), JRA-55 (January 1958; *Kobayashi et al.*, 2015), and NCEP-NCAR R1 (January 1958; *Kalnay et al.*, 1996; see also *Moninger et al.*, 2003), although many of the data from these early years do not meet the necessary standards for assimilation. The volume of aircraft data suitable for assimilation increased substantially after January 1973 (*Uppala et al.*, 2005; *Kobayashi et al.*, 2015).

Aircraft temperature data have been reported to have a warm bias with respect to radiosonde observations (*Ballish and Kumar*, 2008). This type of discrepancy among ingested data sources can have important impacts on the analysis. For example, *Rienecker et al.* (2011) and *Simmons et al.* (2014) have shown that an increase in the magnitude of the temperature bias at 300 hPa in MERRA with respect to radiosondes in the middle to late 1990s coincides with a large increase in the number of aircraft observations assimilated by the system. Moreover, they conclude that differences in temperature trends at 200 hPa between MERRA and ERA-Interim reflect the different impacts of aircraft temperatures in these two reanalysis systems. MERRA-2 applies adaptive bias corrections to AMDAR observations that may help to reduce the uncertainties associated with assimilating these data (*McCarty et al.*, 2016): after each analysis step the updated bias is estimated as a weighted running mean of the aircraft observation increments from preceding analysis times. These adaptive bias corrections are calculated and applied for each aircraft tail number in the database separately.

2.4.4 Water vapour

The assimilation of radiosonde and satellite observations of humidity fields is problematic in the upper troposphere and above, where water vapour mixing ratios are very low and measurement uncertainties are relatively large. The impact of saturation means that humidity probability density functions are often highly non-Gaussian (*Ingleby et al.*, 2013). These issues are particularly pronounced near the tropopause, where sharp temperature gradients complicate the calculation and application of bias corrections for humidity variables during the assimilation step. Reanalysis systems therefore often do not assimilate observations of water vapour provided by radiosondes and/or microwave and infrared sounders (mostly in the form of radiances; see *Section 2.4.2.2*) above a specified upper bound, which is typically between ~300 hPa and ~100 hPa. In regions of the atmosphere that lie above this upper bound (*i.e.,* the uppermost troposphere and stratosphere), the water vapour field is typically determined by the forecast model alone. In this case, water vapour in the stratosphere is determined mainly by transport from below, turbulent mixing, and dehydration in the vicinity of the tropical cold point tropopause (*e.g.*, *Gettelman et al.*, 2010). **Table 2.24** provides brief descriptions of special treatments and caveats affecting reanalysis estimates of water vapour in the upper troposphere and stratosphere. A more detailed discussion and assessment of reanalysis estimates of water vapour is provided in *Chapter 4* of this report.

2.5 Execution streams

2.5.1 What is an 'execution stream'?

The production of reanalyses often must be completed under strict deadlines determined by external factors. To meet these deadlines, most reanalyses have been executed in two or more distinct 'streams', which are then combined. Discontinuities in the time series of some analyzed variables may occur when streams are joined. These potential discontinuities should be considered (along with the changes in assimilated observations described in *Section 2.4*) when reanalysis variables are used for assessments of climate variability and/or trends.

2.5.2 Summary of stream execution

Table 2.25 and **Figure 2.19** briefly summarize the streams used for generating each set of reanalysis products. Refer to the reference papers listed in **Table 2.1** for the procedures used to transition between streams in creating the final data product, as different reanalysis systems may use different approaches. Certain periods have been reprocessed to correct errors in the input data. The reprocessed periods and associated potential discontinuities listed in **Table 2.25** and shown in F**igure 2.19** may be incomplete, and are also likely to change subsequent to the publication of this report. Users are therefore recommended to contact the reanalysis centres directly if they encounter unexplained shifts or jumps in reanalysis products.

Table 2.24: *Notes on treatment of water vapour in the upper troposphere and stratosphere. Additional information is provided in Chapter 2E.*

Reanalysis System	Special treatments and caveats affecting reanalysis estimates of water vapour
ERA-40	No adjustments due to data assimilation are applied in the stratosphere (above the diagnosed tropopause. Methane oxidation is included via a simple parameterization in the stratosphere.
ERA-Interim	The ERA-Interim system contains a parameterization that allows supersaturation with respect to ice in the cloud-free portions of grid cells with temperatures less than 250 K. As in ERA-40, no adjustments due to data assimilation are applied in the stratosphere, and methane oxidation is included via a simple parameterization.
ERA-20C	ERA-20C does not assimilate any water vapour observations. Supersaturation with respect to ice is permitted in cloud-free portions of grid cells with temperatures less than 250 K, and methane oxidation is included via a simple parameterization in the stratosphere.
ERA5	Similar to ERA-Interim, but the parameterization of supersaturation with respect to ice in cloud-free portions of grid cells has been extended to all temperatures less than 273 K (as opposed to only temperatures less than 250 K as in ERA-Interim) and a more consistent treatment of potentially negative values in the stratosphere has been added.
JRA-25 / JCDAS	Observations of humidity are not assimilated and analyses of moisture variables are not provided at pressures less than 100 hPa. Vertical correlations of humidity background errors are set to zero at pressures less than 50 hPa to prevent spurious analysis increments above this level. No moisture source due to methane oxidation is applied to water vapour in the stratosphere. The radiation scheme assumes a constant volume mixing ratio of 2.5 ppmv in the stratosphere.
JRA-55	Analyses of moisture variables are not provided at pressures less than 100 hPa in the pressure-level analysis (anl_p), although analyses of moisture variables are provided for all model levels in the model-level analysis (anl_mdl). Observations of humidity are not assimilated at pressures less than 100 hPa, and vertical correlations of humidity background errors are set to zero at pressures less than 5 hPa to prevent spurious analysis increments above this level. No moisture source due to methane oxidation is applied to water vapour in the stratosphere. The radiation scheme uses climatological annual mean mixing ratios observed by HALOE and UARS MLS during 1991 – 1997 (without seasonal variations) in the stratosphere.
MERRA	The MERRA system tightly constrains stratospheric water vapour to a specified profile, which is based on zonal mean climatologies from HALOE and Aura MLS (*Rienecker et al.*, 2011; *Jiang et al.*, 2010). Water vapour does not undergo physically meaningful variations at pressures less than ~50 hPa.
MERRA-2	Essentially the same as MERRA.
NCEP-NCAR R1	Analyses of moisture variables are not provided at pressures less than 300 hPa. Satellite humidity retrievals are not assimilated.
NCEP-DOE R2	Satellite humidity retrievals are not assimilated.
CFSR / CFSv2	Although there is no upper limit to assimilated GNSS-RO data, radiosonde humidities are only assimilated at pressures 250 hPa and greater. Moisture variables are provided in the stratosphere, but dehydration processes in the tropopause layer may yield negative values, which are replaced by very small positive values for the radiation calculations, but are not replaced in the analysis. Methane oxidation is not included.
NOAA-CIRES 20CR v2	Moisture variables are provided in the stratosphere, but dehydration processes in the tropopause layer may yield negative values, which are artificially replaced by very small positive values for the radiation calculations, but are not replaced in the output fields. Methane oxidation is not included.

Table 2.25: *Information on the execution streams for each reanalysis system.*

Reanalysis System	Execution sreams
ERA-40	ERA-40 was planned for execution in three streams covering 1989–2002, 1957–1972, and 1972–1988. In practice, a small number of parallel-running sub-streams bridging gaps between the main streams had to be run in order to meet the production deadline.
ERA-Interim	ERA-Interim was carried out in two main streams, the first from 1989 to present and the second from 1979 to 1988. The period of the first stream covering January 1989 to August 1993 was rerun to include from the outset all changes made on the fly in the original production for this period; these changes were also included in the second main production stream. The second stream was extended to the end of 1989 to check consistency during the overlap period (see also discussion by *Simmons et al.*, 2014).
ERA-20C	The reanalysis consists of 22 streams, all but the last of which are six years in length. The first stream starts on 1 January 1899 and extends through 31 December 1904. Each subsequent stream starts on 1 January in years ending in 4 or 9 and ends on 31 December of the next year ending in 4 or 9. The final stream starts on 1 January 2004 and extends seven years through the end of the reanalysis. The first year of each stream is discarded from the final product.
ERA5	ERA5 comprises one high-resolution (31-km) analysis (HRES) and a 10-member reduced-resolution (62-km) ensemble of data assimilations (EDA). Seven production streams were run between 1979 and the present for the EDA, and additional shorter streams were run for the HRES to resolve, where practicable, issues encountered in the original production streams. Details are given in Table 3 of *Hersbach et al.* (2020). A further four streams have been run to provide analyses from 1950 to 1978. In addition to these streams for ERA5 core production, a rerun covering the period 2000–2006 has been conducted and is now publicly available under the name ERA5.1. This rerun offers improved representations of temperature and humidity in the stratosphere but differs little from ERA5 in the lower and middle troposphere.
JRA-25 / JCDAS	JRA-25 was conducted in two main streams: the first covers January 1979–December 1990, and the second covers January 1991–January 2014. Note also the transition from JRA-25 (conducted jointly by JMA and CRIEPI) to JCDAS (conducted by JMA only) in January 2005. The execution of JCDAS was conducted entirely in real time. Two periods (January 1994–December 1999 and January 2000–January 2002) were recalculated and replaced to fix problems with data quality; these two periods may be considered as separate sub-streams in addition to the two main streams.
JRA-55	JRA-55 has been executed in two streams. Stream A covers January 1958 through August 1980, while stream B covers September 1980 through the present. Three periods have also been reprocessed after errors were identified: January to June 1958, December 1974 to August 1980 and June 1987 to September 1992 (see also *Kobayashi et al.*, 2015; their Figure 7). JRA-55C has been executed in three streams: Stream A covers 1 November 1972 through 31 August 1980, Stream B covers 1 September 1980 through 31 August 2005, and Stream C covers 1 September 2005 through 31 December 2012. JRA-55AMIP has been executed in one continuous stream.
MERRA	MERRA was executed in three streams. Stream 1 covers January 1979–December 1992, stream 2 covers January 1993–December 2000, and stream 3 covers January 2001–present. Each stream was spun up in two stages: a 2-year analysis at 2°×2.5° followed by a 1-year analysis on the native MERRA grid (see **Table 2.2**). The production version of stream 2 (after spin-up) overlaps with the final four years of stream 1 (January 1989–December 1992), while the production version of stream 3 overlaps with the final three years of stream 2 (January 1998–December 2000).
MERRA-2	MERRA-2 was executed in four streams covering January 1980–December 1991, January 1992–December 2000, January 2001–December 2010, and January 2011–present. Each stream was spun up for one year on the full MERRA-2 system.
NCEP-NCAR R1	NCEP-NCAR R1 was run in three streams. The first stream, which produced data covering 1982–present, was started in December 1978. The second stream, covering 1958–1981 (post-IGY), was started second. For the third and final stream, which covers 1948–1957 (pre-IGY), the analyses were conducted at 03Z, 09Z, 15Z and 21Z (rather than 00Z, 06Z, 12Z and 18Z). There may be additional discontinuities involving updates. For example, the original analyses may have been affected by a problem with the sea ice boundary condition. A second simulation with an improved sea ice boundary condition may be run for a few months, and then replace the original analyses. Transitions between the original product and these "patches" may cause discontinuities.
NCEP-DOE R2	NCEP-DOE R2 was executed in one continuous stream; however, like NCEP-NCAR R1, there may be discontinuities involving updates.
CFSR / CFSv2	CFSR was produced by running six simultaneous streams covering the following periods: • Stream 1: 1 December 1978 to 31 December 1986 • Stream 2: 1 November 1985 to 31 December 1989 • Stream 5: 1 January 1989 to 31 December 1994 • Stream 6: 1 January 1994 to 31 March 1999 • Stream 3: 1 April 1998 to 31 March 2005 • Stream 4: 1 April 2004 to 31 December 2009 A full 1-year overlap between the streams was used to address spinup issues concerning the deep ocean, the upper stratosphere and the deep soil. The entire CFSR thus covers 31 years (1979–2009) plus five overlap years. Each earlier stream is used to its end, so that the switch to the next stream occurs at the end of the overlap period. A separate one-year stream was run for 2010, after which the analysis system was updated to CFSv2 (with an increase in horizontal resolution from T382 to T574). For most applications, CFSR can be extended through the present using output from CFSv2.
NOAA-CIRES 20CR v2	20CR v2 was executed in 28 streams. With some exceptions, each stream typically produced five years of data with 14 months of spinup. The following text gives the data coverage provided by each stream (the streams are numbered sequentially), with the spin-up start year provided in parentheses: 1871–1875 (1869), 1876–1880 (1874), 1881–1885 (1879), 1886–1890 (1884), 1891–1895 (1889), 1896–1900 (1894), 1901–1905 (1899), 1906–1910 (1904), 1911–1915 (1909), 1916–1920 (1914), 1921–1925 (1919), 1926–1930 (1924), 1931–1935 (1929), 1936–1940 (1934), 1941–1945 (1939), 1946–1951 (1944), 1952–1955 (1949), 1956–1960 (1954), 1961–1965 (1959), 1966–1970 (1964), 1971–1975 (1969), 1976–1980 (1974), 1981–1985 (1979), 1986–1990 (1984), 1991–1995 (1989), 1996–2000 (1994), and 2001–2012 (1999). The spin-up start date for each stream was 00 UTC 1 November, the production start date was 00 UTC 1 January, and the production end date was 21 UTC 31 December.

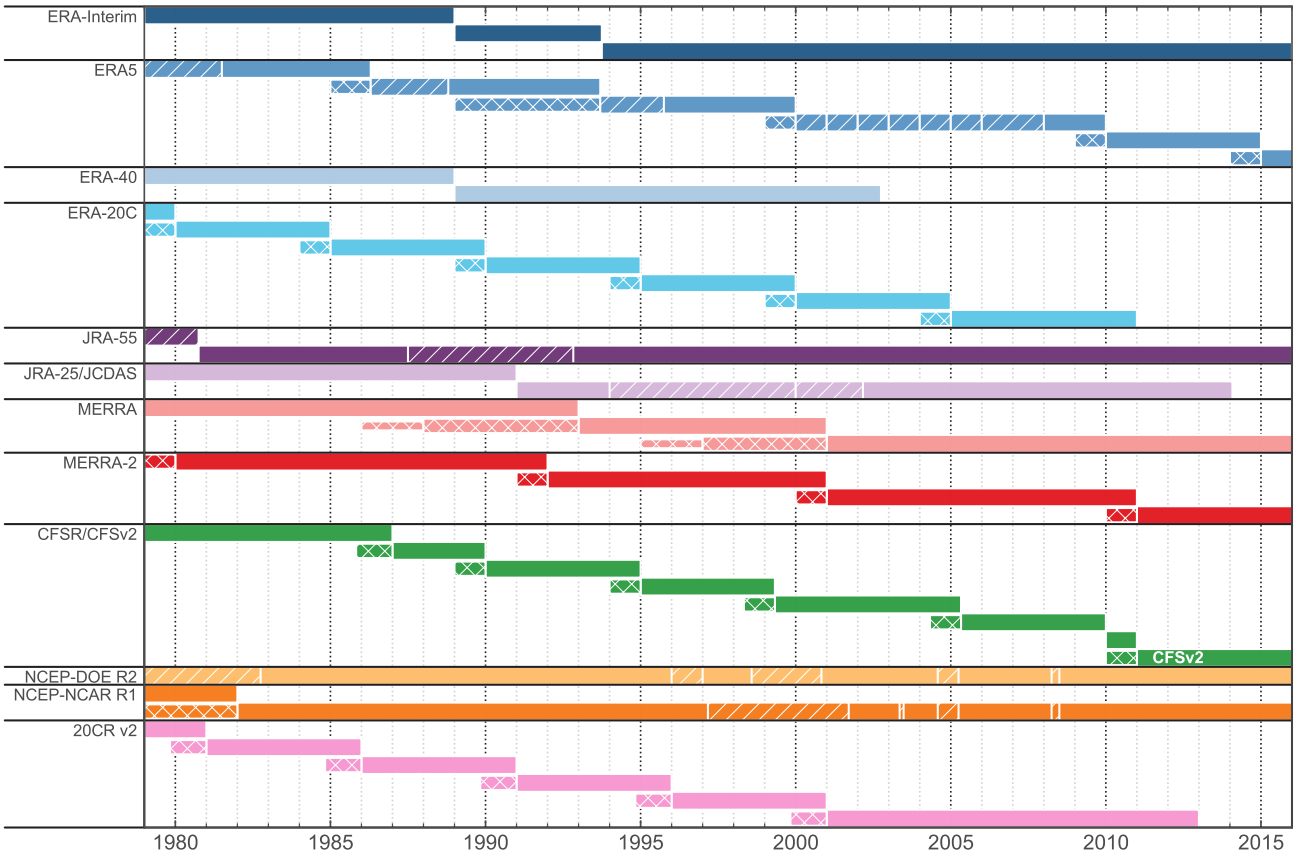


Figure 2.19: *Summary of the execution streams of the reanalyses for the period 1979 – 2016. Hatching indicates known reprocessed 'patches' or 'repair runs'. The narrowest cross-hatched segments indicate known spin-up periods, while the medium-narrow cross-hatched segments indicate overlap periods. See also* ***Table 2.25****. Reproduced from Fujiwara et al. (2017).*

2.6 Archived data

The original data at model resolution and model levels (**Table 2.2**) are converted by each reanalysis centre to data on regular horizontal grids (sometimes at multiple resolutions) and on pressure levels (see *Appendix A*) for public release. The converted data (and sometimes the original data) can typically be obtained via the reanalysis centre websites (see the S-RIP website for links). Some other institutes or projects, such as the NCAR Research Data Archive (RDA), have also constructed public archives of one or more of the reanalysis datasets. Such institutes may have used independent conversions for the data grid, levels, and/or units. Pre-processed data sets have also been produced for the S-RIP activity, including zonal-mean data sets containing dynamical (*Martineau*, 2017) and diabatic (*Wright*, 2017) diagnostics on pressure levels. These pre-processed data are stored in the S-RIP archive at CEDA (**http://data.ceda.ac.uk/badc/srip/**), together with detailed documentation (see also *Martineau et al.*, 2018). Additional data produced for S-RIP include supplementary data files for this chapter (many also provided as a supplement to *Fujiwara et al.*, 2017) and common grid files containing basic variables (*Davis*, 2020). CFSR/CFSv2 products on model levels have also been converted to netCDF format for S-RIP using the High-Resolution Initial Conditions binary files and forecast files archived by NOAA NCEI (**https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/climate-forecast-system-version2-cfsv2**). Data users of these or any other public release of reanalysis or reanalysis-based products should always read the documentation for that release carefully.

It is particularly important to check unit information, as different reanalysis centres or public archives may use different units for the same variable. For example, temperature may be provided in units of °C or K. Some centres provide geopotential height in meters (or "gpm"), while others provide geopotential in $m^2\,s^{-2}$. For water vapour, specific humidity (not volume mixing ratio) is provided in most cases, in units of either $kg\,kg^{-1}$ or $g\,kg^{-1}$. Some reanalyses do not provide vertical pressure velocity (ω, in $Pa\,s^{-1}$) and/or specific humidity data in the stratosphere. Ozone is provided as mass mixing ratio (not volume mixing ratio) in most cases, in units of either $kg\,kg^{-1}$ or $mg\,kg^{-1}$ (*i.e.*, ppmm). Care is also recommended when using precipitation or other 'flux' data, because the integration time period may not be explicitly documented in the data file. Precipitation data may also be divided into multiple categories (such as anvil, convective, and large-scale), the exact definitions of which vary by reanalysis.

Monthly mean products may also differ across different reanalyses, and even for different variables from the same reanalysis, owing to differences in the sampling times or intervals (hourly, 3-hourly, or 6-hourly; instantaneous or time-average). Such differences can be especially impactful for variables with distinct diurnal or sub-diurnal signals (*e.g.*, land-sea breezes in the boundary layer and atmospheric tides in the upper stratosphere and above). In the following, we describe the exact definitions of monthly means for major variable groups in recent reanalyses.

- ERA-Interim and ERA5 divide variables into "instantaneous", "forecast", and "accumulated" products. For ERA-Interim, monthly means of instantaneous products are calculated from 6-hourly data valid at 00, 06, 12, and 18 UTC throughout the month. For ERA5, monthly means of instantaneous products are calculated from hourly fields valid from 00 through 23 UTC throughout the month. Monthly means of accumulated products account for all forecast time steps, although in some cases (*e.g.*, temperature and moisture tendencies) these products are not provided and must be calculated by the user. In cases where ECMWF does provide a monthly mean, partial time steps have been accounted for so that only time steps within the specified month have been included in the average.

- For JRA-55, monthly means of upper-air winds, temperature, geopotential height, and other core analysis fields are calculated from instantaneous analyses at 00, 06, 12, and 18 UTC. Distinctions between instantaneous and time-averaged forecast diagnostics also apply for JRA-55. Instantaneous forecast products are output either every three hours (for two-dimensional fields) or every six hours (for three-dimensional fields), and the monthly means reflect this sampling. Time-averaged fields are designated by the fcst_phy collections, with monthly means representing all time steps.

- For MERRA and MERRA-2, monthly means for each product are calculated by averaging the corresponding instantaneous or time-averaged data. For example, inst3_3d_asm_Np (3-hourly instantaneous data) gets averaged over a month to produce instM_3d_asm_Np and inst6_3d_asm_Np (6-hourly instantaneous data) gets averaged over a month to produce instM_3d_asm_Np. By contrast, the tavg files contain fields averaged from all the (15-min) model time steps within a given time window. Like inst3_3d_asm, these fields are forecast model outputs from the IAU "corrector" step as described in section 2.3 above.

- For CFSR/CFSv2, monthly mean analysis fields are calculated from instantaneous values at 00, 06, 12, and 18 UTC. Monthly means of most forecast variables are also calculated from instantaneous outputs. Only radiation, precipitation, and other 'flux'-type variables are aggregated from averages over the forecast step. These distinctions are directly embedded in the metadata of original GRIB2 files for CFSR/CFSv2 (e.g., 'anl' for analysis variables, '6 hour fcst' for instantaneous forecast variables, and '0-6 hour ave fcst' for time-average forecast variables).

The file formats for archived data may include GRIB, GRIB2, NetCDF, and HDF. Grid boundaries and orientations, such as the starting point for longitude (0 °E or 180 °W), the order of latitudes (from the North Pole or from the South Pole), and the vertical orientation (from the surface or from the TOA) may also vary by reanalysis and/or data source.

After interpolation to pressure levels, most reanalyses (with the exception of MERRA and MERRA-2) provide data below the surface (*e.g.*, at 1000 hPa over the continents). These data are calculated via vertical extrapolation, and are provided for two reasons. First, they enable the use of a complete field when plotting or taking derivatives, and second, they allow data users to visualize variability over the whole globe (including features over mountains) using data from a single pressure surface. The extrapolation procedure may differ by variable and/or reanalysis system. Users of data in the lower part of the troposphere should be aware of this feature, particularly in regions of complex topography.

Acknowledgements

The materials contained in the tables and figures have been compiled from a variety of sources, for which we acknowledge the contributions of Santha Akella, Peter Bechtold, Michael Bosilovich, Dick Dee, John Derber, Ron Gelaro, Yun-Tai Hou, Robert Kistler, Daryl Kleist, Shinya Kobayashi, Shrinivas Moorthi, Eric Nielsen, Yvan Orsolini, Saroja Polavarapu, Paul Poli, Bill Putman, Suranjana Saha, Jack Woollen, Fanglin Yang, and Valery Yudin. Yoshio Kawatani prepared an early draft of the text on calculation of monthly means. We thank Wenhao Dong, Suqin Duan, and Jacob Smith for providing comments on advance drafts of this document, and an anonymous reviewer for detailed and insightful comments on the first submitted draft. For C.-Z. Zou, the views, opinions, and findings contained in this report are those of the authors and should not be construed as an official NOAA or U.S. Government position, policy, or decision.

Figures 2.1, **2.5, 2.6, 2.10, 2.11, 2.12, 2.13**, and **2.19** are reproduced from *Fujiwara et al.* (2017). These reproductions are made under a creative commons attribution 3.0 license **https://creativecommons.org/licenses/by/3.0/**.

References

Allan, P., *et al.*, 2011: The international Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative. *Bull. Am. Meteor. Soc.*, **92**, 1421 - 1425, doi: 10.1175/2011BAMS3218.1.

Allen, R.J., and C.S. Zender, 2011: Forcing of the Arctic Oscillation by Eurasian snow cover. *J. Climate*, **24**, 6528 - 6539, doi: 10.1175/2011JCLI4157.1.

Andersson, E. and H. Järvinen, 1999: Variational quality control. *Q. J. R. Meteorol. Soc.*, **125**, 697 - 722, doi: 10.1002/qj.49712555416.

Andrae, U., N. Sokka, and K. Onogi, 2004: The radiosonde temperature bias corrections used in ERA-40. *ERA-40 Project Report Series*, No. **15**, ECMWF, Shinfield Park, Reading, UK, 34 pp. [available at **https://www.ecmwf.int/en/elibrary/7773-radiosonde-temperature-bias-corrections-used-era-40**; accessed August 2018].

Arakawa, A., 2004: The cumulus parameterization problem: past, present, and future. *J. Climate*, **17**, 2493 - 2525, doi: 10.1175/1520-0442(2004)017<2493:RATCPP>2.0.CO;2.

Arakawa, A., and W.H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment, part I. *J. Atmos. Sci.*, **31**, 674 - 701, doi: 10.1175/1520-0469(1974)031<0674:IOACCE>2.0.CO;2.

Auligné, T., A.P. McNally, and D.P. Dee, 2007: Adaptive bias correction for satellite data in a numerical weather prediction system. *Q. J. R. Meteorol. Soc.*, **133**, 631 - 642, doi: 10.1002/qj.56.

Bacmeister, J.T., M.J. Suarez, and F.R. Robertson, 2006: Rain re-evaporation, boundary layer–convection interactions, and Pacific rainfall patterns in an AGCM. *J. Atmos. Sci.*, **63**, 3383 - 3403, doi: 10.1175/JAS3791.1.

Baines, P.G., and T.N. Palmer, 1990: Rationale for a new physically-based parametrization of subgrid-scale orographic effects. *Tech. Memo*. No. **169**, ECMWF, Shinfield Park, Reading, UK, 11 pp. [available at **https://www.ecmwf.int/en/elibrary/7875-rationale-new-physically-based-parametrization-subgrid-scale-orographic-effects**; accessed August 2018].

Ballish, B.A., and V.K. Kumar, 2008: Systematic differences in aircraft and radiosonde temperatures: Implications for NWP and climate studies. *Bull. Am. Meteor. Soc.*, **89**, 1689 - 1708, doi: 10.1175/2008BAMS2332.1.

Balsamo, G., *et al.*, 2015: ERA-Interim/Land: a global land surface reanalysis data set. *Hydrol. Earth Syst. Sci.*, **19**, 389 - 407, doi: 10.5194/hess-19-389-2015.

Bechtold, P., *et al.*, 2014: Representing equilibrium and nonequilibrium convection in large-scale models. *J. Atmos. Sci.*, **71**, 734 - 753, doi:10.1175/JAS-D-13-0163.1.

Beljaars, A.C.M., 1995: The impact of some aspects of the boundary layer scheme in the ECMWF model. In *Proc. of ECMWF Seminar on Parameterization of Sub-grid Scale Physical Processes*, pp. 125 - 161, Shinfield Park, Reading, UK, 5 - 9 September 1994. [available at **https://www.ecmwf.int/en/elibrary/8035-impact-some-aspects-boundary-layer-scheme-ecmwf-model**; accessed August 2018].

Bengtsson, L., M. Kanamitsu, P. Kållberg, and S. Uppala, 1982: FGGE research activities at ECMWF. *Bull. Am. Meteor. Soc.*, **63**, 277 - 303, doi: 10.1175/1520-0477-63.3.277.

Birner, T., D. Sankey, and T.G. Shepherd, 2006: The tropopause inversion layer in models and analyses. *Geophys. Res. Lett.*, **33**, L14804, doi: 10.1029/2006GL026549.

Bloom, S., L. Takacs, A. DaSilva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256 - 1271, doi: 10.1175/1520-0493(1996)124<1256:DAUIAU>2.0.CO;2.

Bodeker, G.E., *et al.*, 2016: Reference upper-air observations for climate: From concept to reality. *Bull. Am. Meteor. Soc.*, **97**, 123 - 135, doi: 10.1175/BAMS-D-14-00072.1.

Boussetta, S., *et al.*, 2013: Impact of a satellite-derived leaf area index monthly climatology in a global numerical weather prediction model. *Int. J. Remote Sens.*, **34**, 3520 - 3542, doi: 10.1080/01431161.2012.716543.

Bouttier, F., and P. Courtier, 1999: Data assimilation concepts and methods. *Meteorological Training Course Lecture Series*, ECMWF, Shinfield Park, Reading, UK. [available at **https://www.ecmwf.int/en/elibrary/16928-data-assimilation-concepts-and-methods**; accessed August 2018].

Briegleb, B.P., 1992: Delta-Eddington approximation for solar radiation in the NCAR community climate model. *J. Geophys. Res. Atmos.*, **97**, 7603 - 7612, doi:10.1029/92JD00291.

Buchard, V., *et al.*, 2015: Using the OMI aerosol index and absorption aerosol optical depth to evaluate the NASA MERRA Aerosol Reanalysis. *Atmos. Chem. Phys.*, **15**, 5743 - 5760, doi:10.5194/acp-15-5743-2015.

Buchard, V., *et al.*, 2017: The MERRA-2 Aerosol Reanalysis, 1980–onward, Part II: Evaluation and Case Studies. *J. Climate*, **30**, 6851 - 6872, doi: 10.1175/JCLI-D-16-0613.1.

Buehner, M., *et al.*, 2010a: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: description and single-observation experiments. *Mon. Wea. Rev.*, **138**, 1550 - 1566, doi: 10.1175/2009MWR3157.1.

Buehner, M., *et al.*, 2010b: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: one-month experiments with real observations. *Mon. Wea. Rev.*, **138**, 1567 - 1586, doi: 10.1175/2009MWR3158.1.

Campbell, W.F., C.H. Bishop, and D. Hodyss, 2010: Vertical covariance localization for satellite radiances in ensemble Kalman filters. *Mon. Wea. Rev.*, **138**, 282 - 290, doi: 10.1175/2009MWR3017.1.

Cariolle, D. and M. Déqué, 1986: Southern hemisphere medium-scale waves and total ozone disturbances in a spectral general circulation model. *J. Geophys. Res. Atmos.*, **91**, 10825 - 10846, doi: 10.1029/JD091iD10p10825.

Cariolle, D. and H. Teyssèdre, 2007: A revised linear ozone photochemistry parameterization for use in transport and general circulation models: multi-annual simulations. *Atmos. Chem. Phys.*, **7**, 2183 - 2196, doi: 10.5194/acp-7-2183-2007.

Charnock, H., 1955: Wind stress on a water surface. *Q. J. R. Meteorol. Soc.*, **81**, 639 - 640, doi: 10.1002/qj.49708135027.

Chou, M.-D., and K.-T. Lee, 1996: Parameterizations for the absorption of solar radiation by water vapor and ozone. *J. Atmos. Sci.*, **53**, 1203 - 1208, doi: 10.1175/1520-0469(1996)053<1203:PFTAOS>2.0.CO;2.

Chou, M.-D., and M.J. Suarez, 1999: A solar radiation parameterization for atmospheric studies. *NASA Tech. Rep. Series on Global Modeling and Data Assimilation*, Vol. **15**, NASA/TM-1999-104606, Greenbelt, Maryland, 40 pp. [available at **https://ntrs.nasa.gov/search.jsp?R=19990060930**; accessed August 2018].

Chou, M.-D., M.J. Suarez, X.-Z. Liang, and M.M.-H. Yan, 2001: A thermal infrared radiation parameterization for atmospheric studies. *NASA Tech. Rep. Series on Global Modeling and Data Assimilation*, Vol. **19**, NASA/TM-2001-104606, Greenbelt, Maryland, 56 pp. [available at **https://ntrs.nasa.gov/search.jsp?R=20010072848**; accessed August 2018].

Chun, H.-Y, and J.-J. Baik, 1998: Momentum flux by thermally induced internal gravity waves and its approximation for large-scale models. *J. Atmos. Sci.*, **55**, 3299 - 3310, doi: 10.1175/1520-0469(1998)055<3299:MFBTII>2.0.CO;2.

Cionni, I., *et al.*, 2011: Ozone database in support of CMIP5 simulations: results and corresponding radiative forcing. *Atmos. Chem. Phys.*, **11**, 11267 - 11292, doi: 10.5194/acp-11-11267-2011.

Clough, S.A., *et al.*, 2005: Atmospheric radiative transfer modeling: A summary of the AER codes. *J. Quant. Spectrosc. Radiat. Transfer*, **91**, 233 - 244, doi: 10.1016/j.jqsrt.2004.05.058.

Cohen, J., and D. Entekhabi, 1999: Eurasian snow cover variability and Northern Hemisphere climate variability. *Geophys. Res. Lett.*, **26**, 345 - 348, doi: 10.1029/1998GL900321.

Cohen, J., *et al.*, 2014: Linking Siberian snow cover to precursors of stratospheric variability. *J. Climate*, **27**, 5422 - 5432, doi: 10.1175/JCLI-D-13-00779.1.

Colarco, P.R., A. da Silva, M. Chin, and T. Diehl, 2010: Global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth. *J. Geophys. Res. Atmos.*, **115**, D14207, doi: 10.1029/2009JD012820.

Compo, G.P., *et al.*, 2011: The twentieth century reanalysis project. *Q. J. R. Meteorol. Soc.*, **137**, 1 - 28, doi: 10.1002/qj.776.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.*, **120**, 1367 - 1387, doi: 10.1002/qj.49712051912.

Cram, T.A., *et al.*, 2015: The International Surface Pressure Databank version 2. *Geosci. Data J.*, **2**, 31 - 46, doi: 10.1002/gdj3.25.

Cressman, G.P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367 - 374, doi: 10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2.

Daley, R., 1981: Normal mode initialization. *Rev. Geophys.*, **19**, 450 - 468, doi: 10.1029/RG019i003p00450.

Davis, S., 2020: SPARC Reanalysis Intercomparison Project (S-RIP) reanalysis common grid files. doi: 10.5281/zenodo.3754752.

de Rosnay, P., *et al.*, 2014: Initialisation of land surface variables for numerical weather prediction. *Surv. Geophys.*, **35**, 607 - 621, doi: 10.1007/s10712-012-9207-x.

de Rosnay, P., L. Isaksen, and L. Dahoui, 2015: Snow data assimilation at ECMWF. *ECMWF Newsletter* No. **143**, Spring 2015. [available at **https://www.ecmwf.int/en/elibrary/17328-snow-data-assimilation-ecmwf**; accessed September 2018].

Dee, D.P., 2005: Bias and data assimilation. *Q. J. R. Meteorol. Soc.*, **131**, 3323 - 3343, doi: 10.1256/qj.05.137.

Dee, D.P., *et al.*, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**, 553 - 597, doi: 10.1002/qj.828.

Derber, J.C., and W.-S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287 - 2299, doi: 10.1175/1520-0493(1998)126<2287:TUOTCC>2.0.CO;2.

Dethof, A., and E. V. Hólm, 2004: Ozone assimilation in the ERA-40 reanalysis project. *Q. J. R. Meteorol. Soc.*, **130**, 2851 - 2872, doi: 10.1256/qj.03.196.

Donlon, C., *et al.*, 2007: The global ocean data assimilation experiment high-resolution sea surface temperature pilot project. *Bull. Am. Meteor. Soc.*, **88**, 1197 - 1213, doi: 10.1175/BAMS-88-8-1197.

Dorman, J.L., and P.J. Sellers, 1989: A global climatology of albedo, roughness length and stomatal resistance for atmospheric general circulation models as represented by the Simple Biosphere model (SiB). *J. Appl. Meteorol.*, **28**, 833 - 855, doi: 10.1175/1520-0450(1989)028<0833:AGCOAR>2.0.CO;2.

Dragani, R., 2011: On the quality of the ERA-Interim ozone reanalyses: Comparisons with satellite data. *Q. J. R. Meteorol. Soc.*, **137**, 1312 - 1326, doi: 10.1002/qj.821.

Dutra, E., et al., 2010: An improved snow scheme for the ECMWF land surface model: Description and offline validation. *J. Hydrometeorol.*, **11**, 899 - 916, doi: 10.1175/2010JHM1249.1.

Ebita, A., *et al.*, 2011: The Japanese 55-year reanalysis "JRA-55": An interim report. *SOLA*, **7**, 149 - 152, doi: 10.2151/sola.2011-038.

Ek, M.B., *et al.*, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res. Atmos.*, **108**, 8851, doi: 10.1029/2002JD003296.

Evensen, G., and P.J. van Leeuwen, 2000: An Ensemble Kalman Smoother for nonlinear dynamics. *Mon. Wea. Rev.*, **128**, 1852 - 1867, doi: 10.1175/1520-0493(2000)128<1852:AEKSFN>2.0.CO;2.

Evensen, G., 2009: Data Assimilation: The Ensemble Kalman Filter. *Springer*, Berlin, 307 pp., doi: 10.1007/978-3-642-03711-5.

Fels, S., and M.D. Schwarzkopf, 1975: The simplified exchange approximation: A new method for radiative transfer calculations. *J. Atmos. Sci.*, **32**, 1475 - 1488, doi: 10.1175/1520-0469(1975)032<1475:TSEAAN>2.0.CO;2.

Flannaghan, T.J., and S. Fueglistaler, 2014: Vertical mixing and the temperature and wind structure of the tropical tropopause layer. *J. Atmos. Sci.*, **71**, 1609 - 1622, doi: 10.1175/JAS-D-13-0321.1.

Flemming, J., *et al.*, 2017: The CAMS interim Reanalysis of carbon monoxide, ozone and aerosol for 2003 - 2015, *Atmos. Chem. Phys.*, **17**, 1945 - 1983, doi: 10.5194/acp-17-1945-2017.

Fortuin, J.P., and U. Langematz, 1995: An update on the current ozone climatology and on concurrent ozone and temperature trends. *Proc. SPIE*, **2311**, 207 - 216, doi: 10.1117/12.198578.

Fouquart, Y., and B. Bonnel, 1980: Computation of solar heating of the Earth's atmosphere: a new parameterization. *Beitr. Phys.*, **53**, 35 - 62.

Freidenreich, S.M., and V. Ramaswamy, 1999: A new multiple-band solar radiative parameterization for general circulation models. *J. Geophys. Res. Atmos.*, **104**, 31389 - 31409, doi: 10.1029/1999JD900456.

Fujiwara, M., *et al.*, 2017: Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems. *Atmos. Chem. Phys.*, **17**, 1417 - 1452, doi: 10.5194/acp-17-1417-2017.

Gandin, L.S., 1963: Objective analysis of meteorological fields. *Gidrometerologischeskoe Izdatelstvo*, Leningrad. [English translation by Israeli Program for Scientific Translations, Jerusalem, 1965].

Garcia, R.R. and B.A. Boville, 1994: Downward control of the mean meridional circulation and temperature distribution of the polar winter stratosphere. *J. Atmos. Sci.*, **51**, 2238 - 2245, doi: 10.1175/1520-0469(1994)051<2238:COTMMC>2.0.CO;2.

Gelaro, R., *et al.*, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Climate*, **30**, 5419 - 5454, doi: 10.1175/JCLI-D-16-0758.1.

Gettelman, A., *et al.*, 2010: Multimodel assessment of the upper troposphere and lower stratosphere: Tropics and global trends. *J. Geophys. Res. Atmos.*, **115**, D00M08, doi: 10.1029/2009JD013638.

Gibson, J.K., *et al.*, 1997: ERA description, *ERA-15 Rep. Series*, No. **1**, ECMWF, Shinfield Park, Reading, UK, 72 pp. [available at **https://www.ecmwf.int/en/elibrary/9584-era-description**; accessed August 2018].

Goldberg, M., *et al.*, 2011: The Global Space-based Inter-Calibration System. *Bull. Am. Meteor. Soc.*, **92**, 467 - 475, doi: 10.1175/2010BAMS2967.1.

Goody, R.M., 1952: A statistical model for water-vapour absorption. *Quart. J. R. Meteor. Soc.*, **78**, 165 - 169, doi: 10.1002/qj.49707833604.

Gustafsson, N., 2007: Discussion on '4D-Var or EnKF?', *Tellus A*, **59**, 774 - 777, doi: 10.1111/j.1600-0870.2007.00262.x.

Haimberger, L., C. Tavolato, and S. Sperka, 2008: Towards elimination of the warm bias in historic radiosonde temperature records - some new results from a comprehensive intercomparison of upper air data. *J. Climate*, **21**, 4587 - 4606, doi: 10.1175/2008JCLI1929.1.

Haimberger, L., C. Tavolato, and S. Sperka, 2012: Homogenization of the global radiosonde temperature dataset through combined comparison with reanalysis background series and neighboring stations. *J. Climate*, **25**, 8108 - 8131, doi: 10.1175/JCLI-D-11-00668.1.

Harada, Y., *et al.*, 2016: The JRA-55 Reanalysis: representation of atmospheric circulation and climate variability. *J. Meteor. Soc. Jpn.*, **94**, 269 - 302, doi:1 0.2151/jmsj.2016-015.

Helfand, H.M., *et al.*, 1987: The effect of a gravity wave drag parameterization scheme on GLA fourth-order GCM forecasts. Collection of papers presented at the *WMO/IUGG NWP Symposium*, Tokyo, 4 - 8 August 1986, 729 - 742. [available at **https://ntrs.nasa.gov/search.jsp?R=19880050777**; accessed August 2018].

Helfand, H.M. and S.D. Schubert, 1995: Climatology of the simulated Great Plains low-level jet and its contribution to the continental moisture budget of the United States. *J. Climate*, **8**, 784 - 806, doi: 10.1175/1520-0442(1995)008<0784:COTSGP>2.0.CO;2.

Hering, W.S., and T.R. Borden, Jr., 1965: Mean distributions of ozone density over North America, 1963–1964. T*ech. Rep. AFCRL-65-913*, Air Force Cambridge Research Laboratories.

Hersbach, H., *et al.*, 2015: ERA-20CM: a twentieth-century atmospheric model ensemble. *Q. J. R. Meteorol. Soc.*, **141**, 2350 - 2375, doi: 10.1002/qj.2528.

Hersbach, H., 2020: The ERA5 Global Reanalysis, *Q. J. R. Meteorol. Soc.*, **146**, 1999 - 2049, doi: 10.1002/qj.3803.

Holton, J.R., and W.M. Wehrbein, 1980: A numerical model of the zonal mean circulation of the middle atmosphere. *Pure Appl. Geophys.*, **118**, 284 - 306, doi: 10.1007/BF01596455.

Hong, S.-Y. and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.*, **124**, 2322 - 2339, doi: 10.1175/1520-0493(1996)124<2322:NBLVDI>2.0.CO;2.

Hortal, M., and A.J. Simmons, 1991: Use of reduced Gaussian grids in spectral models. *Mon. Wea. Rev.*, **119**, 1057 - 1074, doi: 10.1175/1520-0493(1991)119<1057:UORGGI>2.0.CO;2.

Iacono, M.J., 2008: Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models. *J. Geophys. Res. Atmos.*, **113**, D13103, doi:10.1029/2008JD009944.

Ingleby, N.B., 2013: Improved variational analyses using a nonlinear humidity control variable. *Q. J. R. Meteorol. Soc.*, **139**, 1875 - 1887, doi: 10/1002/qj.2073.

IPCC, 1996: Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the *Second Assessment Report of the Intergovernmental Panel on Climate Change*, Houghton, J.T., L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg, and K. Maskell (eds.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 572 pp.

Ishii, M., A. Shouji, S. Sugimoto, and T. Matsumoto, 2005: Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the KOBE collection. *Int. J. Climatol.*, **25**, 865 - 879, doi: 10.1002/joc.1169.

Iwasaki, T., S. Yamada, and K. Tada, 1989a: A parameterization scheme of orographic gravity wave drag with the different vertical partitionings, Part I: Impact on medium range forecasts. *J. Meteor. Soc. Jpn.*, **67**, 11 - 27, doi: 10.2151/jmsj1965.67.1_11.

Iwasaki, T., S. Yamada, and K. Tada, 1989b: A parameterization scheme of orographic gravity wave drag with the different vertical partitionings, Part II: Zonally averaged budget analyses based on transformed Eulerian-mean method. *J. Meteor. Soc. Jpn.*, **67**, 29 - 41, doi: 10.2151/jmsj1965.67.1_29.

Janssen, P.A.E.M., *et al.*, 2001: Impact and feedback of ocean waves on the atmosphere. *Tech. Memo.* No. **341**, ECMWF, Shinfield Park, Reading, UK, 34 pp. [available at **https://www.ecmwf.int/en/elibrary/10225-impact-and-feedback-ocean-waves-atmosphere**; accessed August 2018].

Jiang, J. H., *et al.*, 2010: Five year (2004–2009) observations of upper tropospheric water vapor and cloud ice from MLS and comparisons with GEOS-5 analyses. *J. Geophys. Res. Atmos.*, **115**, D15103, doi: 10.1029/2009JD013256.

JMA, 2007: Outline of the operational numerical weather prediction at the Japan Meteorological Agency. *WMO Technical Progress Report on the Global Data-processing and Forecasting System (GDPFS) and Numerical Weather Prediction (NWP) Research*, JMA, Japan. [available at **http://www.jma.go.jp/jma/jma-eng/jma-center/nwp/outline2007-nwp/index.htm**; accessed August 2018].

JMA, 2013: Outline of the operational numerical weather prediction at the Japan Meteorological Agency. *WMO Technical Progress Report on the Global Data-processing and Forecasting System (GDPFS) and Numerical Weather Prediction (NWP) Research*, JMA, Japan. [available at **http://www.jma.go.jp/jma/jma-eng/jma-center/nwp/outline2013-nwp/index.htm**; accessed August 2018].

Kållberg, P., A. Simmons, S. Uppala, and M. Fuentes, 2007: The ERA-40 Archive. *ECMWF ERA-40 Project Report Series*, No. **17**, 32 pp. (originally published in September 2004, revised in October 2007).

Kalman, R.E., 1960: A new approach to linear filtering and prediction problems. *J. Basic Eng.*, **82**, 35 - 45, doi: 10.1115/1.3662552.

Kalnay, E., *et al.*, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteor. Soc.*, **77**(3), 437 - 471, doi: 10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kalnay, E., 2003: Atmospheric Modeling, Data Assimilation, and Predictability. *Cambridge University Press*, Cambridge, 341 pp.

Kalnay, E., *et al.*, 2007a: 4-D-Var or ensemble Kalman filter?, *Tellus A*, **59**, 758 - 773, doi: 10.1111/j.1600-0870.2007.00261.x.

Kalnay, E., *et al.*, 2007b: Response to the discussion on "4-D-Var or EnKF?" by Nils Gustafsson, *Tellus A*, **59**, 778 - 780, doi: 10.1111/j.1600-0870.2007.00263.x.

Kanamitsu, M., *et al.*, 2002: NCEP-DOE AMIP-II reanalysis (R-2). *Bull. Am. Meteor. Soc.*, **83**(11), 1631 - 1643, doi: 10.1175/BAMS-83-11-1631.

Kawai, H. and T. Inoue, 2006: A simple parameterization scheme for subtropical marine stratocumulus. *SOLA*, **2**, 17 - 20, doi: 10.2151/sola.2006-005.

Kim, Y.-J., and A. Arakawa, 1995: Improvement of orographic gravity wave parameterization using a mesoscale gravity wave model. *J. Atmos. Sci.*, **83**, 1875 - 1902, doi: 10.1175/1520-0469(1995)052<1875:IOOGWP>2.0.CO;2.

Kistler, R., *et al.*, 2001: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Am. Meteor. Soc.*, **82**(2), 247 - 267, doi: 10.1175/1520-0477(2001)082<0247:TNNYRM>2.3.CO;2.

Kleist, D.T., *et al.*, 2009: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wea. Forecasting*, **24**, 1691 - 1705, doi: 10.1175/2009WAF2222201.1.

Kobayashi, C., *et al.*, 2014: Preliminary results of the JRA-55C, an atmospheric reanalysis assimilating conventional observations only. *SOLA*, **10**, 78 - 82, doi: 10.2151/sola.2014-016.

Kobayashi, S., *et al.*, 2015: The JRA-55 reanalysis: General specifications and basic characteristics. *J. Meteorol. Soc. Jpn.*, **93**, 5 - 48, doi: 10.2151/jmsj.2015-001.

Koepke, P., M. Hess, I. Schult, and E.P. Shettle, 1997: Global Aerosol Data Set. *Report* No. **243**, Max-Planck-Institut für Meteorologie, Hamburg, ISSN 0937-1060, 44 pp. [available at **http://opac.userweb.mwn.de/radaer/gads.html**; accessed September 2018].

Köhler, M., M. Ahlgrimm, and A. Beljaars, 2011: Unified treatment of dry convective and stratocumulus-topped boundary layers in the ECMWF model. *Q. J. R. Meteorol. Soc.*, **137**, 43 - 57, doi: 10.1002/qj.713.

Komen, G.J., *et al.*, 1994: Dynamics and Modelling of Ocean Waves. *Cambridge University Press*, Cambridge, 532 pp.

Kondo, J., 1975: Air–sea bulk transfer coefficients in diabatic conditions. *Bound.-Lay. Meteorol.*, **9**, 91 - 112, doi: 10.1007/BF00232256.

Koren, V., *et al.*, 1999: A parameterization of snowpack and frozen ground intended for NCEP weather and climate models. *J. Geophys. Res. Atmos.*, **104**, 19,569 - 19,585, doi: 10.1029/1999JD900232.

Koster, R.D., *et al.*, 2000: A catchment-based approach to modeling land surface processes in a general circulation model. Part 1: Model structure. *J. Geophys. Res. Atmos.*, **105**, 24,809 - 24,822, doi: 10.1029/2000JD900327.

Krishnamurti, T.N., and L. Bounoua, 1996: An Introduction to Numerical Weather Prediction Techniques. *CRC Press*, Florida, USA, 293 pp.

Lacis, A.A., and J.E. Hansen, 1974: A parameterization for the absorption of solar radiation in the Earth's atmosphere. *J. Atmos. Sci.*, **31**, 118 - 133, doi: 10.1175/1520-0469(1974)031<0118:APFTAO>2.0.CO;2.

Laloyaux, P., *et al.*, 2016: A coupled data assimilation system for climate reanalyses. *Q. J. R. Meteorol. Soc.*, **142**, 65 - 78, doi: 10.1002/qj.2629.

Lamarque, J.-F., *et al.*, 2010: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application. *Atmos. Chem. Phys.*, *10*, 7017 - 7039, doi: 10.5194/acp-10-7017-2010.

Large, W.G., and S. Pond, 1981: Open ocean momentum flux measurements in moderate to strong winds,.*J. Phys. Oceanogr.*, **11**, 324 - 336, doi: 10.1175/1520-0485(1981)011<0324:OOMFMI>2.0.CO;2.

Lawless, A.S., 2010: A note on the analysis error associated with 3D-FGAT. *Q. J. R. Meteorol. Soc.*, **136**, 1094 - 1098, doi: 10.1002/qj.619.

Lawrence, Z.D., *et al.*, 2015: Comparisons of polar processing diagnostics from 34 years of the ERA-Interim and MERRA reanalyses. *Atmos. Chem. Phys.*, **15**, 3873 - 3892, doi: 10.5194/acp-15-3873-2015.

Lei, L., J.S. Whitaker, and C. Bishop, 2018: Improving assimilation of radiance observations by implementing model space localization in an ensemble Kalman filter. J. *Adv. Model. Earth Syst.*, **10**, 3221 - 3232, doi: 10.1029/2018MS001468.

Lin, S.-J., 2004: A "vertically Lagrangian" finite-volume dynamical core for global models. *Mon. Wea. Rev.*, **132**, 2293 - 2307, doi: 10.1175/1520-0493(2004)132<2293:AVLFDC>2.0.CO;2.

Lindzen, R.S., E.S. Batten, and J.W. Kim, 1968: Oscillations in atmospheres with tops. *Mon. Wea. Rev.*, **96**, 133 - 140, doi: 10.1175/1520-0493(1968)096<0133:OIAWT>2.0.CO;2.

Liu, Q., *et al.*, 1999: Improvements in hurricane initialization and forecasting at NCEP with global and regional (GFDL) models. *NOAA Tech. Proc. Bull.*, No. **472**, National Weather Service, Office of Meteorology, Silver Spring, MD, 1 - 7. [available at **http://www.nws.noaa.gov/om/tpb/472.htm**; accessed August 2017].

Lock, A.P., *et al.*, 2000: A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Mon. Wea. Rev.*, **128**, 3187 - 3199, doi: 10.1175/1520-0493(2000)128<3187:ANBLMS>2.0.CO;2.

London, J., 1962: Mesospheric dynamics, part III. The distribution of total ozone in the Northern Hemisphere. Final Report, Dept. of Meteorology and Oceanography, New York University, 99 pp.

Lorenc, A.C. and Rawlins, F., 2005: Why does 4D-Var beat 3D-Var? *Q. J. R. Meteorol. Soc.*, **131**, 3247 - 3257, doi: 10.1256/qj.05.85.

Lott, F., and M.J. Miller, 1997: A new subgrid-scale orographic drag parameterization: Its formulation and testing. *Q. J. R. Meteorol. Soc.*, **123**, 101 - 127, doi: 10.1002/qj.49712353704.

Louis, J.F., 1979: Parameteric model of vertical eddy fluxes in the atmosphere. *Bound.-Lay. Meteorol.*, **17**, 187 - 202, doi: 10.1007/BF00117978.

Louis, J.F., M. Tiedtke, and J.-F. Geleyn, 1982: A short history of the operational PBL parameterization at ECMWF. In *Proc. ECMWF Workshop on Boundary Layer Parameterization*, pp. 59 - 80, Reading, 25 - 27 November 1981 [available at **https://www.ecmwf.int/en/elibrary/10845-short-history-pbl-parameterization-ecmwf**; accessed 9 August 2017].

Lynch, P. and X. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, **120**, 1019 - 1034, doi: 10.1175/1520-0493(1992)120<1019:IOTHMU>2.0.CO;2.

Lynch, P., 1993: Digital filters for numerical weather prediction. *HIRLAM Technical Reports*, No. **10**, 55 pp. [available at **http://hirlam.org/index.php/publications-54/hirlam-technical-reports-a/doc_download/1563-hirlam-technical-report-no-10-digital-filters-for-numerical-weather-prediction**; accessed 7 October 2017].

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model with applications to normal mode initialization. *Contrib. Atmos. Phys.*, **50**, 253 - 271.

Machenhauer, B., 1979: The Spectral Method, in: Numerical Methods used in Atmospheric Models, Vol. II. *GARP Publication Series*, **17**, World Meteorological Organization, Geneva, 121 - 275.

Mahfouf, J.F., *et al.*, 1995: The land surface scheme ISBA within the Météo-France climate model ARPEGE. *J. Climate*, **8**, 2039 - 2057, doi: 10.1175/1520-0442(1995)008<2039:TLSSIW>2.0.CO;2.

Mahrt, L., and H. Pan, 1984: A two-layer model of soil hydrology. *Bound.-Lay. Meteorol.*, **29**, 1 - 20, doi: 10.1007/BF00119116.

Martineau, P., 2017: S-RIP: Zonal-mean dynamical variables of global atmospheric reanalyses on pressure levels [data set]. *Centre for Environmental Data Analysis (CEDA)*, doi: 10.5285/b241a7f536a244749662360bd7839312.

Martineau, P., J.S. Wright, N. Zhu, and M. Fujiwara, 2018: Zonal-mean data set of global atmospheric reanalyses on pressure levels. *Earth Syst. Sci. Data,* **10***,* 1925 - 1941, doi: 10.5194/essd-10-1925-2018.

Matsumoto, T., M. Ishii, Y. Fukuda, and S. Hirahara, 2006: Sea ice data derived from microwave radiometer for climate monitoring. *Proceedings of the 14th Conference on Satellite Meteorology and Oceanography*, Atlanta, USA, P2.21.

McCarty, W., *et al.*, 2016: MERRA-2 input observations: Summary and assessment. *NASA Tech. Rep. Series on Global Modeling and Data Assimilation*, Vol. **46**, NASA/TM-2016-104606, Greenbelt, Maryland. [available at **https://ntrs.nasa.gov/search.jsp?R=20160014544**; accessed 9 August 2017]

McCormack, J.P., S.D. Eckermann, D.E. Siskind, and T. McGee, 2006: CHEM2D-OPP: A new linearized gas-phase ozone photochemistry parameterization for high-altitude NWP and climate models. *Atmos. Chem. Phys.*, **6**, 4943 - 4972, doi: 10.5194/acp-6-4943-2006.

McFarlane, N.A., 1987: The effect of orographically excited gravity-wave drag on the circulation of the lower stratosphere and troposphere. *J. Atmos. Sci.*, **44**, 1775 - 1800, doi: 10.1175/1520-0469(1987)044<1775:TEOOEG>2.0.CO;2.

Meinshausen, M., *et al.*, 2011: The RCP greenhouse gas concentrations and their extension from 1765 to 2300. *Clim. Change*, **109**, 213 - 241, doi: 10.1007/s10584-011-0156-z.

Mellor, G.L. and T. Yamada, 1974: A hierarchy of turbulence closure models for planetary boundary layers. *J. Atmos. Sci.*, **31**, 1791 - 1806, doi: 10.1175/1520-0469(1974)031<1791:AHOTCM>2.0.CO;2.

Miller, M.J., T.N. Palmer, and R. Swinbank, 1989: Parameterization and influence of subgrid-scale orography in general circulation and numerical weather prediction models. *Meteorol. Atmos. Phys.*, **40**, 84 - 109, doi:10.1007/BF01027469.

Mlawer, E.J., *et al.*,1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the long wave. *J. Geophys. Res. Atmos.*, **102**, 16663 - 16682, doi: 10.1029/97JD00237.

Molod, A., 2012: Constraints on the profiles of total water PDF in AGCMs from AIRS and a high-resolution model. *J. Climate*, **25**, 8341 - 8352, doi: 10.1175/JCLI-D-11-00412.1.

Molod, A., L. Takacs, M. Suarez, and J. Bacmeister, 2015: Development of the GEOS-5 atmospheric general circulation model: Evolution from MERRA to MERRA-2. *Geosci. Model Dev.*, **8**, 1339 - 1356, doi: 10.5194/gmd-8-1339-2015.

Moninger, W.R., R.D. Mamrosh, and P.M. Pauley, 2003: Automated meteorological reports from commercial aircraft. *Bull. Am. Meteor. Soc.*, **84**, 203 - 216, doi: 10.1175/BAMS-84-2-203.

Moorthi, S., and M.J. Suarez, 1992: Relaxed Arakawa-Schubert. A parameterization of moist convection for general-circulation models. *Mon. Wea. Rev.*, **120**, 978 - 1002, doi: 10.1175/1520-0493(1992)120<0978:RASAPO>2.0.CO;2.

Moorthi, S., H.-L. Pan, and P. Caplan, 2001: Changes to the 2001 NCEP operational MRF/AVN global analysis/forecast system. *Technical Procedures Bulletin* **484**, NOAA NWS, Silver Spring, MD, 14 pp. [available at **https://rda.ucar.edu/datasets/ds093.0/docs/484.pdf**; accessed August 2018].

Moorthi, S., R. Sun, H. Xiao, and C.R. Mechoso, 2010: Low-cloud simulation in the Southeast Pacific in the NCEP GFS: Role of vertical mixing and shallow convection. *NCEP Office Note* **463**, 28 pp. [available at **http://www.emc.ncep.noaa.gov/officenotes/FullTOC.html**; accessed August 2018].

Morcrette, J-J., *et al.*, 2008: Impact of a new radiation package, McRad, in the ECMWF Integrated Forecasting System. *Mon. Wea. Rev.*, **136**, 4773 - 4798, doi: 10.1175/2008MWR2363.1.

Murai, S., S. Yabu, and H. Kitagawa, 2005: Development of a new radiation scheme for the global atmospheric NWP model. *21st Conference on Weather Analysis and Forecasting and 17th Conference on Numerical Weather Prediction*, American Meteorological Society, Washington, D. C. [5pp., available at **https://ams.confex.com/ams/pdfpapers/94316.pdf**; accessed November 2019].

Nash, J., T. Oakley, H. Vömel, and W. Li, 2011: WMO intercomparison of high quality radiosonde systems, Yangjiang, China, 12 July–3 August 2010. *Instruments and Observing Methods Report* No. **107**, WMO/TD-No. 1580, 238 pp. [available at **https://library.wmo.int/index.php?lvl=notice_display&id=15531**; accessed 9 August 2017].

Nash, J., and R. Saunders, 2015: A review of Stratospheric Sounding Unit radiance observations for climate trends and reanalyses. *Q. J. R. Meteorol. Soc.*, **141**, 2103 - 2113, doi: 10.1002/qj.2505.

Onogi, K., 1998: A data quality control method using forecasted horizontal gradient and tendency in a NWP system: Dynamic QC. *J. Meteorol. Soc. Japan*, **76**, 497 - 516, doi: 10.2151/jmsj1965.76.4_497.

Onogi, K., *et al.*, 2007: The JRA-25 reanalysis. *J. Meteorol. Soc. Jpn.*, **85**(3), 369 - 432, doi: 10.2151/jmsj.85.369.

Orr, A., *et al.*, 2010: Improved middle atmosphere climate and forecasts in the ECMWF model through a nonorographic gravity wave drag parameterization. *J. Climate*, **23**, 5905 - 5926, doi: 10.1175/2010JCLI3490.1.

Orsolini, Y.J., and N. Kvamstø, 2009: The role of the Eurasian snow cover upon the wintertime circulation: decadal simulations forced with satellite observations. *J. Geophys. Res. Atmos.*, **114**, D19108, doi: 10.1029/2009JD012253.

Osborn, T.J., and P.D. Jones, 2014: The CRUTEM4 land-surface air temperature data set: construction, previous versions and dissemination via Google Earth. *Earth Syst. Sci. Data*, **6**, 61 - 68, doi: 10.5194/essd-6-61-2014.

Palmer, T.N., G.J. Schutts, and R. Swinbank, 1986: Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parameterization. *Q. J. R. Meteorol. Soc.*, **112**, 1001 - 1039, doi: 10.1002/qj.49711247406.

Paltridge, G.W., and C.M.R. Platt, 1976: Radiative Processes in Meteorology and Climatology. *Elsevier*, New York, 336 pp.

Pan, H.-L., and L. Mahrt, 1987: Interaction between soil hydrology and boundary-layer development. *Bound.-Lay. Meteorol.*, **38**, 185 - 202, doi: 10.1007/BF00121563.

Pan, H.-L., and W. Wu, 1995: Implementing a mass flux convective parameterization package for the NMC Medium-Range Forecast model. *NMC Office Note*, No. **409**, 43 pp. [available at **http://www.ncep.noaa.gov/officenotes/NOAA-NPM-NCEPON-0005/01408A42.pdf**; accessed August 2018].

Park, S.K., and D. Županski 2003: Four-dimensional variational data assimilation for mesoscale and storm-scale applications. *Meteorol. Atmos. Phys.*, **82**, 173 - 208, doi: 10.1007/s00703-001-0586-7.

Parrish, D.F., and J.C. Derber, 1992: The National Meteorological Center's Spectral Statistical-Interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747 - 1763, doi: 10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

Petersen, R.A., 2016: On the impact and benefits of AMDAR observations in operational forecasting - Part I: A review of the impact of automated aircraft wind and temperature reports. *Bull. Am. Meteor. Soc.*, **97** 585 - 602, doi: 10.1175/BAMS-D-14-00055.1.

Pierrehumbert, R.T., 1986: An essay on the parameterization of orographic gravity wave drag. *Seminar/Workshop 1986 Observation, Theory and Modelling of Orographic Effects*, ECMWF, Shinfield Park, Reading, UK, 251 - 282. [available at **https://www.ecmwf.int/en/elibrary/11673-essay-parameterization-orographic-gravity-wave-drag**; accessed August 2018].

Polavarapu, S. , T.G. Shepherd, Y. Rochon, and S. Ren, 2005: Some challenges of middle atmosphere data assimilation. *Q. J. R. Meteorol. Soc.*, **131**, 3513 - 3527, doi: 10.1256/qj.05.87.

Polavarapu, S.M. and M. Pulido, 2017: Stratospheric and mesospheric data assimilation: The role of middle atmospheric dynamics. In: *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*, vol. **3**, edited by: S. K. Park and L. Xu, Springer-Verlag, Berlin, pp. 429 - 454, doi: 10.1007/978-3-319-43415-5_19.

Poli, P., S.B. Healy, and D.P. Dee, 2010: Assimilation of Global Positioning System radio occultation data in the ECMWF ERA–Interim reanalysis. *Q. J. R. Meteorol. Soc.*, **136**, 1972 - 1990, doi: 10.1002/qj.722.

Poli, P., *et al.*, 2013: The data assimilation system and initial performance of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20C), ERA Report Series, Vol. 14. [available at **https://www.ecmwf.int/en/elibrary/11699-data-assimilation-system-and-initial-performance-evaluation-ecmwf-pilot-reanalysis**; accessed August 2018].

Poli, P., *et al.*, 2016: ERA-20C: An atmospheric reanalysis of the 20th century, *J. Climate*, **29**, 4083 - 4097, doi: 10.1175/JCLI-D-15-0556.1.

Putman, W. and S.-J. Lin, 2007: Finite volume transport on various cubed sphere grids. *J. Comput. Phys.*, **227**, 55 - 78, doi: 10.1016/j.jcp.2007.07.022.

Randel, W., *et al.*, 2004: The SPARC intercomparison of middle-atmosphere climatologies, *J. Climate*, **17**, 986 - 1003, doi: 10.1175/1520-0442(2004)017<0986:TSIOMC>2.0.CO;2.

Randles, C.A., *et al.*, 2017: The MERRA-2 Aerosol Reanalysis, 1980–onward, Part I: System Description and Data Assimilation Evaluation. *J. Climate*, **30**, 6823 - 6850, doi: 10.1175/JCLI-D-16-0609.1.

Reale, A.L., 2001: NOAA operational sounding products from Advanced-TOVS polar orbiting environmental satellites. *NOAA Technical Report NESDIS* **102**, 59 pp. [available at **https://repository.library.noaa.gov/view/noaa/1123**; accessed August 2018].

Reichle, R.H., *et al.*, 2011: Assessment and enhancement of MERRA land surface hydrology estimates. *J. Climate*, **24**, 6322 - 6338, doi: 10.1175/JCLI-D-10-05033.1.

Reichle, R.H., *et al.*, 2017a: Land surface precipitation in MERRA-2. *J. Climate*, **30**, 1643 - 1664, doi: 10.1175/JCLI-D-16-0570.1.

Reichle, R.H., *et al.*, 2017b: Assessment of MERRA-2 land surface hydrology estimates. *J. Climate*, **30**, 2937 - 2960, doi: 10.1175/JCLI-D-16-0720.1.

Reynolds, R.W., and T.M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 1609 - 1625, doi: 10.1175/1520-0442(1994)007<0929:IGSSTA>2.0.CO;2.

Reynolds, R.W., *et al.*, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate*, **15**, 1609 - 1625, doi: 10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2.

Reynolds, R.W., *et al.*, 2007: Daily high-resolution blended analyses for sea surface temperature. *J. Climate*, **20**, 5473 - 5496, doi: 10.1175/2007JCLI1824.1.

Riehl, H., and J.S. Malkus, 1958: On the heat balance in the equatorial trough zone. *Geophysica*, **6**, 503 - 538.

Rienecker, M.M., *et al.*, 2008: The GEOS-5 Data Assimilation System—Documentation of versions 5.0.1 and 5.1.0, and 5.2.0. *NASA Tech. Rep. Series on Global Modeling and Data Assimilation*, Vol. **27**, NASA/TM-2008-104606, Greenbelt, Maryland, 92 pp. [available at **https://ntrs.nasa.gov/search.jsp?R=20120011955**; accessed August 2018].

Rienecker, M.M., *et al.*, 2011: MERRA: NASA's modern-era retrospective analysis for research and applications. *J. Climate*, **24**, 3624 - 3648, doi: 10.1175/JCLI-D-11-00015.1.

Rosenfield, J.E., M.R. Schoeberl, and M.A. Geller, 1987: A computation of the stratospheric diabatic circulation using an accurate radiative transfer model. *J. Atmos. Sci.*, **44**, 859 - 876, doi: 10.1175/1520-0469(1987)044<0859:ACOTSD>2.0.CO;2.

Saha, S., *et al.*, 2010: The NCEP climate forecast system reanalysis. *Bull. Am. Meteor. Soc.*, **91**(8), 1015 - 1057, doi: 10.1175/2010BAMS3001.1.

Saha, S., *et al.*, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185 - 2208, doi: 10.1175/JCLI-D-12-00823.1.

Sakamoto, M., 2009: An update to the quality control thresholds of the conventional observing system for global data assimilation. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **39**, 33 - 34.

Sakamoto, M. and J.R. Christy, 2009: The influences of TOVS radiance assimilation on temperature and moisture tendencies in JRA-25 and ERA-40. *J. Atmos. Ocean. Tech.*, **26**, 1435 - 1455, doi:10.1175/2009JTECHA1193.1.

Sankey, D., *et al.*, 2007: Impact of data assimilation filtering methods on the mesosphere. *J. Geophys. Res. Atmos.*, **112**, D24104, doi: 10.1029/2007JD008885.

Sato N., *et al.*, 1989: Effects of implementing the simple biosphere model (SiB) in a general circulation model. *J. Atmos. Sci.*, **46**, 2757 - 2782, doi: 10.1175/1520-0469(1989)046<2757:EOITSB>2.0.CO;2.

Sato, M., J.E. Hansen, M.P. McCormick, and J.B. Pollack, 1993: Stratospheric aerosol optical depths. 1850 - 1990, *J. Geophys. Res. Atmos.*, **98**, 22987 - 22994, doi: 10.1029/93JD02553.

Schoeberl, M.R., A.R. Douglass, Z.Zhu, and S. Pawson, 2003: A comparison of the lower stratospheric age spectra derived from a general circulation model and two data assimilation systems. *J. Geophys. Res. Atmos.*, **108**, 4113, doi: 10.1029/2002JD002652.

Schubert, S.D., J. Pfaendtner, and R. Rood, 1993: An assimilated data set for Earth science applications. *Bull. Am. Meteor. Soc.*, **74**, 2331 - 2342, doi: 10.1175/1520-0477(1993)074<2331:AADFES>2.0.CO;2.

Schubert, S.D., *et al.*, 1995: A multiyear assimilation with the GEOS-1 system: Overview and results. *NASA Tech. Rep. Series on Global Modeling and Data Assimilation*, Vol. **6**, NASA/TM-104606, Greenbelt, Maryland, 201 pp. [available at **https://ntrs.nasa.gov/search.jsp?R=19950020003**; accessed August 2018].

Schwarzkopf, M.D., and S. Fels, 1991: The simplified exchange method revisited: An accurate, rapid method for computation of infrared cooling rates and fluxes. *J. Geophys Res. Atmos.*, **96**, 9075 - 9096, doi: 10.1029/89JD01598.

Scinocca, J.F., 2003: An accurate spectral nonorographic gravity wave drag parameterization for general circulation models. *J. Atmos. Sci.*, **60**, 667 - 682, doi: 10.1175/1520-0469(2003)060<0667:AASNGW>2.0.CO;2.

Seidel, D.J., et al., 2009: Reference upper-air observations for climate: Rationale, progress, and plans. *Bull. Am. Meteor. Soc.*, **90**, 361 - 369, doi: 10.1175/2008BAMS2540.1.

Sellers, P.J., Y. Mintz, Y.C. Sud, and A. Dalcher, 1986: A simple biosphere model (SiB) for use within general circulation models. *J. Atmos. Sci.*, **43**, 505 - 531, doi: 10.1175/1520-0469(1986)043<0505:ASBMFU>2.0.CO;2.

Shepherd, T.G., K. Semeniuk, and J.N. Koshyk, 1996: Sponge layer feedbacks in middle-atmosphere models. *J. Geophys. Res. Atmos.*, **101**, 23447 - 23464, doi: 10.1029/96JD01994.

Shepherd, T.G. and T.A. Shaw, 2004: The angular momentum constraint on climate sensitivity and downward influence in the middle atmosphere. *J. Atmos. Sci.*, **61**, 2899 - 2908, doi: 10.1175/JAS-3295.1.

Simmons, A.J., and D.M. Burridge, 1981: An energy and angular-momentum conserving finite-difference scheme and hybrid vertical coordinates. *Mon. Wea. Rev.*, **109**, 758 - 766, doi: 10.1175/1520-0493(1981)109<0758:AEAAMC>2.0.CO;2.

Simmons, A.J., *et al.*, 2004: Comparison of trends and low-frequency variability in CRU, ERA-40, and NCEP/NCAR analyses of surface air temperature. *J. Geophys. Res. Atmos.*, **109**, D24115, doi: 10.1029/2004JD005306.

Simmons, A.J., *et al.*, 2014: Estimating low-frequency variability and trends in atmospheric temperature using ERA-Interim. *Q. J. R. Meteorol. Soc.*, **140**, 329 - 353. doi: 10.1002/qj.2317.

Slivinski, L.C., *et al.*, 2019: Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. *Q. J. R. Meteorol. Soc.,* **145**, 2876 - 2908, doi: 10.1002/qj.3598.

Smith, R.N.B., 1990: A scheme for predicting layer clouds and their water content in a general circulation model. *Q. J. R. Meteorol. Soc.*, **116**, 435 - 460. doi: 10.1002/qj.49711649210.

Spencer, J.W., 1971: Fourier series representation of the position of the sun. *Search*, **2**, 172 [available at **http://www.mail-archive.com/sundial@uni-koeln.de/msg01050.html**; accessed August 2018].

Stajner, I., *et al.*, 2008: Assimilated ozone from EOS-Aura: Evaluation of the tropopause region and tropospheric columns. *J. Geophys. Res.*, **113**, D16S32, doi: 10.1029/2007JD008863.

Stieglitz, M., A. Ducharne, R. Koster, and M. Suarez, 2001: The impact of detailed snow physics on the simulation of snow cover and subsurface thermodynamics at continental scales. *J. Hydrometeorol.*, **2**, 228 - 242, doi: 10.1175/1525-7541(2001)002<0228:TIODSP>2.0.CO;2.

Sundqvist, H., 1978: A parameterization scheme for non-convective condensation including prediction of cloud water content. *Q. J. R. Meteorol. Soc.*, **104**, 677 - 690, doi: 10.1002/qj.49710444110.

Takacs, L.L., M. Suarez, and R. Todling, 2015: Maintaining atmospheric mass and water balance within reanalysis. *NASA Tech. Rep. Series on Global Modeling and Data Assimilation*, Vol. **37**, NASA/TM-2014-104606, Greenbelt, Maryland, 46 pp. [available at **https://ntrs.nasa.gov/search.jsp?R=20150007977**; accessed August 2018].

Talagrand, O., 2010: Variational assimilation. In: *Data Assimilation: Making Sense of Observations*, edited by: W. Lahoz, B. Khattatov, and R. Menard, Springer-Verlag, Berlin, pp. 41 - 67, doi: 10.1007/978-3-540-74703-1_3.

Tan, W.W., M.A. Geller, S. Pawson, and A. da Silva, 2004: A case study of excessive subtropical transport in the stratosphere of a data assimilation system. *J. Geophys. Res. Atmos.*, **109**, D11102, doi: 10.1029/2003JD004057.

Tanré, D., J.-F. Geleyn, and J.M. Slingo, 1984: First results of the introduction of an advanced aerosol-radiation interaction in the ECMWF low resolution global model. In: *Aerosols and their Climatic Effects*, H. E. Gerber and A. Deepak (eds.), A. Deepak Publishing: Hampton, VA, USA, pp. 133 - 177.

Tavolato, C., and L. Isaksen, 2011: Data usage and quality control for ERA-40, ERA-Interim, and the operational ECMWF data assimilation system. *ERA Report Series*, No. **7**, ECMWF, Reading, UK, 44 pp. [available at **https://www.ecmwf.int/en/elibrary/12573-data-usage-and-quality-control-era-40-era-interim-and-operational-ecmwf-data**; accessed August 2018].

Taylor, K.E., D. Williamson, and F. Zwiers, 2000: The sea surface temperature and sea ice concentration boundary conditions for AMIP II simulations. *Program for Climate Model Diagnosis and Intercomparison (PCMDI) Report* **60**, Lawrence Livermore National Laboratory. [available at **https://pcmdi.llnl.gov/mips/amip/**; accessed August 2018].

Tegen, I., *et al.*, 1997: Contribution of different aerosol species to the global aerosol extinction optical thickness: Estimates from model results. *J. Geophys. Res. Atmos.*, **102**, 23895 - 23915, doi: 10.1029/97JD01864.

Tiedtke, M., 1983: The sensitivity of the time-mean large-scale flow to cumulus convection in the ECMWF model. *ECMWF Workshop on Convection in Large-Scale Models, Reading*, UK, ECMWF, 297 - 316. [available at **https://www.ecmwf.int/en/elibrary/12733-sensitivity-time-mean-large-scale-flow-cumulus-convection-ecmwf-model**; accessed 9 August 2017].

Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea. Rev.*, **117**, 1779 - 1800, doi: 10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2.

Tiedtke, M., 1993: Representation of clouds in large-scale models. *Mon. Wea. Rev.*, **121**, 3040 - 3061, doi: 10.1175/1520-0493(1993)121<3040:ROCILS>2.0.CO;2.

Titchner, H.A. and N.A. Rayner, 2014: The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concentrations. *J. Geophys. Res. Atmos.*, **119**, 2864 - 2889, doi: 10.1002/2013JD020316.

Tompkins, A.M., K. Gierens, and G. Rädel, 2007: Ice supersaturation in the ECMWF integrated forecast system. *Q. J. R. Meteorol. Soc.*, **133**, 53 - 63, doi: 10.1002/qj.14.

Troen, I. and Mahrt, L., 1986: A simple model of the atmospheric boundary layer; sensitivity to surface evaporation. *Bound.-Lay. Meteorol.*, **37**, 129 - 148, doi: 10.1007/BF00122760.

Tsuyuki, T., and T. Miyoshi, 2007: Recent progress of data assimilation methods in meteorology, *J. Meteorol. Soc. Jpn.*, **85B**, 331 - 361, doi: 10.2151/jmsj.85B.331.

Uppala, S.M., *et al.*, 2005: The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.*, **131**(612), 2961 - 3012, doi:10.1256/qj.04.176.

USGS 2000: Global land cover characteristics data base, version 2.0. [Available at **https://lta.cr.usgs.gov/glcc/globdoc2_0**; accessed August 2018].

van den Hurk, B.J.J., P. Viterbo, A.C.M. Beljaars, and A. Betts, 2000: Offline validation of the ERA40 surface scheme. *Tech. Memo*. **295**, ECMWF, Shinfield Park, Reading, UK, 43 pp. [available at **https://www.ecmwf.int/en/elibrary/12900-offline-validation-era40-surface-scheme**; accessed August 2018].

van Vuuren D.P., *et al.*, 2011: The representative concentration pathways: An overview. *Clim. Change*, **109**, 5 - 31, doi: 10.1007/s10584-011-0148-z.

Walsh, J.E., and W.L. Chapman, 2001: 20th-century sea-ice variations from observational data. *Ann. Glaciol.*, **33**, 444 - 448, doi: 10.3189/172756401781818671.

Wang, L., C.-Z. Zou, and H. Qian, 2012: Construction of stratospheric temperature data records from Stratospheric Sounding Units. *J. Climate*, **25**, 2931 - 2946, doi: 10.1175/JCLI-D-11-00350.1.

Wang, W. and C.-Z. Zou, 2014: AMSU-A-only atmospheric temperature data records from the lower troposphere to the top of the stratosphere. *J. Atmos. Ocean. Tech.*, **31**, 808 - 825, doi: 10.1175/JTECH-D-13-00134.1.

Wargan, K., *et al.*, 2017: Evaluation of the ozone fields in NASA's MERRA-2 reanalysis. *J. Climate*, **30**, 2961 - 2988, doi: 10.1175/JCLI-D-16-0699.1.

Whitaker, J.S., G.P. Compo, and J.-N. Thépaut, 2009: A comparison of variational and ensemble-based data assimilation systems for reanalysis of sparse observations. *Mon. Wea. Rev.*, **137**, 1991 - 1999, doi:10.1175/2008MWR2781.1.

WMO, 1986: A preliminary cloudless standard atmosphere for radiation computation. *World Climate Programme (WCP)*, **112**, 53 pp. [Available at **https://library.wmo.int/index.php?lvl=notice_display&id=11668**; accessed April 2021].

Woodruff, S.D., et al., 2011: ICOADS release 2.5: Extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.*, **31**, 951 - 967, doi: 10.1002/ joc.2103.

Wright, J., 2017: S-RIP: Zonal-mean heating rates of global atmospheric reanalyses on pressure levels [data set]. *Centre for Environmental Data Analysis (CEDA)*, doi: 10.5285/70146c789eda4296a3c3ab6706931d56.

Wu, W., R.J. Purser, and D.F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905 - 2916, doi: 10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2.

Xie, S., and M. Zhang, 2000: Impact of the convection triggering function on single-column model simulations, *J. Geophys. Res. Atmos.*, **105**, 14983 - 14996, doi: 10.1029/2000JD900170.

Xu, K.-M., and S.K. Krueger, 1991: Evaluation of cloudiness parameterizations using a cumulus ensemble model. *Mon. Wea. Rev.*, **119**, 342 - 367, doi: 10.1175/1520-0493(1991)119<0342:EOCPUA>2.0.CO;2.

Xu, K.-M., and D.A. Randall, 1996: A semiempirical cloudiness parameterization for use in climate models. *J. Atmos. Sci.*, **53**, 3084 - 3102, doi: 10.1175/1520-0469(1996)053<3084:ASCPFU>2.0.CO;2.

Zhao, Q.Y., and F.H. Carr, 1997: A prognostic cloud scheme for operational NWP models. *Mon. Wea. Rev.*, **125**, 1931 - 1953, doi: 10.1175/1520-0493(1997)125<1931:APCSFO>2.0.CO;2.

Zou, C.-Z., et al., 2006: Recalibration of microwave sounding unit for climate studies using simultaneous nadir overpasses. *J. Geophys. Res.*, **111**, D19114, doi: 10.1029/2005JD006798.

Zou, C.-Z., and W. Wang, 2011: Intersatellite calibration of AMSU-A observations for weather and climate applications. *J. Geophys. Res.*, **116**, D23113, doi: 10.1029/2011JD016205.

Zou, C.-Z., et al., 2014: Recalibration and merging of SSU observations for stratospheric temperature trend studies. *J. Geophys. Res.*, **119**, 13180 - 13205, doi: 10.1002/2014JD021603.

Zou, C.-Z., and H. Qian, 2016: Stratospheric temperature climate data record from merged SSU and AMSU-A observations. *J. Atmos. Ocean Tech.*, **33**, 1967 - 1984, doi: 10.1175/JTECH-D-16-0018.1.

Appendix A: Vertical levels of the models

A2.1 ERA-40 and ERA-Interim

ERA-40 and ERA-Interim both use hybrid sigma–pressure (hybrid σ–p) vertical coordinates (*Simmons and Burridge*, 1981), which are also sometimes referred to as eta (η) vertical coordinates (see also **http://rda.ucar.edu/datasets/ds627.0/docs/Eta_coordinate/**). Both systems use the same vertical resolution with 61 levels (*Källberg et al.*, 2007). The pressure on each level is calculated as $p_k = A_k + B_k \times p_{srf}$, where p_{srf} is surface pressure. The following table provides example pressures at layer interfaces (k–1/2) and layer midpoints (k) for a surface pressure of 1013.25 hPa, from TOA to surface. Pressures at layer midpoints are defined as the average of pressures at layer interfaces. Pressure levels in brackets are used for ERA-Interim products but not for ERA-40 products.

Model Levels					Pressure Levels
k	$A_{k-1/2}$ (hPa)	$B_{k-1/2}$	$p_{k-1/2}$ (hPa)	p_k (hPa)	p (hPa)
1	0.00	0.00000	0.00	0.10	
2	0.20	0.00000	0.20	0.29	
3	0.38	0.00000	0.38	0.51	
4	0.64	0.00000	0.64	0.80	
5	0.96	0.00000	0.96	1.15	1
6	1.34	0.00000	1.34	1.58	
7	1.81	0.00000	1.81	2.08	2
8	2.35	0.00000	2.35	2.67	3
9	2.98	0.00000	2.98	3.36	
10	3.74	0.00000	3.74	4.19	
11	4.65	0.00000	4.65	5.20	5
12	5.76	0.00000	5.76	6.44	7
13	7.13	0.00000	7.13	7.96	
14	8.84	0.00000	8.84	9.89	10
15	10.95	0.00000	10.95	12.26	
16	13.56	0.00000	13.56	15.19	
17	16.81	0.00000	16.81	18.81	20
18	20.82	0.00000	20.82	23.31	
19	25.80	0.00000	25.80	28.88	30
20	31.96	0.00000	31.96	35.78	
21	39.60	0.00000	39.60	44.33	
22	49.07	0.00000	49.07	54.62	50
23	60.18	0.00000	60.18	66.62	70
24	73.07	0.00000	73.07	80.40	
25	87.65	0.00008	87.73	95.98	100
26	103.76	0.00046	104.23	113.42	
27	120.77	0.00182	122.61	132.76	(125)
28	137.75	0.00508	142.90	154.00	150
29	153.80	0.01114	165.09	177.12	(175)
30	168.19	0.02068	189.15	202.09	200

Model Levels					Pressure Levels
k	$A_{k-1/2}$ (hPa)	$B_{k-1/2}$	$p_{k-1/2}$ (hPa)	p_k (hPa)	p (hPa)
31	180.45	0.03412	215.03	228.84	(225)
32	190.28	0.05169	242.65	257.36	250
33	197.55	0.07353	272.06	287.64	300
34	202.22	0.09967	303.22	319.63	
35	204.30	0.13002	336.04	353.23	(350)
36	203.84	0.16438	370.41	388.27	400
37	200.97	0.20248	406.13	424.57	
38	195.84	0.24393	443.01	461.90	(450)
39	188.65	0.28832	480.79	500.00	500
40	179.61	0.33515	519.21	538.591	(550)
41	168.99	0.38389	557.97	577.38	
42	157.06	0.43396	596.78	616.04	600
43	144.11	0.48477	635.31	654.27	(650)
44	130.43	0.53571	673.24	691.75	700
45	116.33	0.58617	710.26	728.16	
46	102.10	0.63555	746.06	763.20	(750), 775
47	88.02	0.68327	780.35	796.59	(800)
48	74.38	0.72879	812.83	828.05	(825)
49	61.44	0.77160	843.26	857.34	850
50	49.42	0.81125	871.42	884.27	(875)
51	38.51	0.84737	897.11	908.65	(900)
52	28.88	0.87966	920.19	930.37	925
53	20.64	0.90788	940.55	949.35	(950)
54	13.86	0.93194	958.15	965.57	
55	8.55	0.95182	972.99	979.06	(975)
56	4.67	0.96765	985.14	989.95	
57	2.10	0.97966	994.75	998.39	1000
58	0.66	0.98827	1002.02	1004.64	
59	0.07	0.99402	1007.26	1009.06	
60	0.00	0.99763	1010.85	1012.05	
	0.00	1.00000	1013.25		

A2.2 ERA-20C

ERA-20C uses hybrid sigma–pressure (hybrid σ–p) vertical coordinates (*Simmons and Burridge*, 1981) with 91 levels. The pressure on each level is calculated as $p_k = A_k + B_k \times p_{srf}$, where p_{srf} is surface pressure. The following table provides example pressures at layer interfaces (k–1/2) and layer midpoints (k) for a surface pressure of 1013.25 hPa, from TOA to surface. Pressures at layer midpoints are defined as the average of pressures at layer interfaces.

Model Levels					Pressure Levels
k	$A_{k-1/2}$ (hPa)	$B_{k-1/2}$	$p_{k-1/2}$ (hPa)	p_k (hPa)	p (hPa)
1	0	0	0	0.01	
2	0.02	0	0.02	0.03	
3	0.03980832	0	0.04	0.06	
4	0.07387186	0	0.07	0.10	
5	0.12908319	0	0.13	0.17	
6	0.21413612	0	0.21	0.28	
7	0.33952858	0	0.34	0.43	
8	0.51746601	0	0.52	0.64	
9	0.76167656	0	0.76	0.92	1
10	1.08715561	0	1.09	1.30	
11	1.50986023	0	1.51	1.78	2
12	2.04637451	0	2.05	2.38	
13	2.71356506	0	2.71	3.12	3
14	3.52824493	0	3.53	4.02	
15	4.50685791	0	4.51	5.09	5
16	5.66519226	0	5.67	6.34	7
17	7.01813354	0	7.02	7.80	
18	8.57945801	0	8.58	9.47	10
19	10.36166504	0	10.36	11.37	
20	12.37585449	0	12.38	13.50	
21	14.6316394	0	14.63	15.88	
22	17.13709595	0	17.14	18.52	
23	19.8987439	0	19.90	21.41	20
24	22.92155518	0	22.92	24.57	
25	26.20898438	0	26.21	27.99	
26	29.76302246	0	29.76	31.67	30
27	33.58425781	0	33.58	35.63	
28	37.67196045	0	37.67	39.85	
29	42.02416504	0	42.02	44.33	
30	46.63776367	0	46.64	49.07	50
31	51.50859863	0	51.51	54.07	
32	56.6315625	0	56.63	59.31	
33	61.99839355	0	62.00	64.80	
34	67.59727051	0	67.60	70.51	70
35	73.41469727	0	73.41	76.43	
36	79.4292627	0.000014	79.44	82.57	
37	85.64624023	0.000055	85.70	88.96	
38	92.08305664	0.000131	92.22	95.62	
39	98.73560547	0.000279	99.02	102.58	100
40	105.5888184	0.000548	106.14	109.89	
41	112.6248438	0.001	113.64	117.59	
42	119.8266211	0.001701	121.55	125.75	125
43	127.1389746	0.002765	129.94	134.40	
44	134.5322559	0.004267	138.86	143.59	
45	141.9200977	0.006322	148.33	153.35	150
46	149.2268555	0.009035	158.38	163.72	
47	156.3805371	0.012508	169.05	174.72	175
48	163.2956055	0.01686	180.38	186.38	
49	169.9062305	0.022189	192.39	198.76	200
50	176.1328125	0.02861	205.12	211.87	
51	181.910293	0.036227	218.62	225.77	225
52	187.1696875	0.045146	232.91	240.48	
53	191.8454492	0.055474	248.05	256.07	250
54	195.8751367	0.067316	264.08	272.56	
55	199.1979688	0.080777	281.05	290.02	
56	201.7539453	0.095964	298.99	308.48	300
57	203.4891602	0.112979	317.97	327.99	
58	204.341582	0.131935	338.02	348.62	350
59	204.2621875	0.152934	359.22	370.42	
60	203.1901172	0.176091	381.61	393.44	400
61	201.0703125	0.20152	405.26	417.73	
62	197.8535742	0.229315	430.21	443.34	450
63	193.4877539	0.259554	456.48	470.17	
64	187.9882227	0.291993	483.85	497.96	500
65	181.4129688	0.326329	512.07	526.46	
66	173.855957	0.362203	540.86	555.40	550
67	165.4458594	0.399205	569.94	584.49	
68	156.3356641	0.436906	599.03	613.50	600
69	146.6564551	0.475016	627.97	642.29	650
70	136.5321973	0.51328	656.61	670.73	
71	126.0838379	0.551458	684.85	698.70	700
72	115.4316699	0.589317	712.56	726.07	
73	104.7131055	0.626559	739.57	752.67	750
74	94.05222656	0.662934	765.77	778.40	775
75	83.5625293	0.698224	791.04	803.16	800
76	73.35164551	0.732224	815.28	826.81	825
77	63.53920898	0.764679	838.35	849.25	850
78	54.22802734	0.795385	860.15	870.38	875
79	45.5021582	0.824185	880.61	890.13	
80	37.43464355	0.85095	899.66	908.44	900
81	30.10146973	0.875518	917.22	925.22	925
82	23.56202637	0.897767	933.22	940.44	
83	17.84854614	0.917651	947.66	954.09	950
84	12.97656128	0.935157	960.52	966.17	
85	8.95193542	0.950274	971.82	976.67	975
86	5.76314148	0.963007	981.53	985.63	
87	3.36772369	0.973466	989.73	993.30	
88	1.62043427	0.982238	996.87	999.84	1000
89	0.54208336	0.989153	1002.80	1005.12	
90	0.06575628	0.994204	1007.44	1009.15	
91	0.0000316	0.99763	1010.85	1012.05	
	0	1	1013.25		

A2.3 ERA5

ERA5 uses hybrid sigma–pressure (hybrid σ–p) vertical coordinates (*Simmons and Burridge*, 1981) with 137 levels. The pressure on each level is calculated as $p_k = A_k + B_k \times p_{srf}$, where p_{srf} is surface pressure. The following table provides example pressures at layer interfaces (k–1/2) and layer midpoints (k) for a surface pressure of 1013.25 hPa, from TOA to surface. Pressures at layer midpoints are defined as the average of pressures at layer interfaces.

Model Levels					Pressure Levels
k	$A_{k-1/2}$ (hPa)	$B_{k-1/2}$	$p_{k-1/2}$ (hPa)	p_k (hPa)	p (hPa)
1	0	0	0.00	0.01	
2	0.02000365	0	0.02	0.03	
3	0.03102241	0	0.03	0.04	
4	0.04666084	0	0.05	0.06	
5	0.06827977	0	0.07	0.08	
6	0.09746966	0	0.10	0.12	
7	0.13605424	0	0.14	0.16	
8	0.18608931	0	0.19	0.22	
9	0.24985718	0	0.25	0.29	
10	0.3298571	0	0.33	0.38	
11	0.42879242	0	0.43	0.49	
12	0.54955463	0	0.55	0.62	
13	0.69520576	0	0.70	0.78	
14	0.86895882	0	0.87	0.97	1
15	1.07415741	0	1.07	1.19	
16	1.31425507	0	1.31	1.45	
17	1.59279404	0	1.59	1.75	
18	1.91338562	0	1.91	2.10	2
19	2.27968948	0	2.28	2.49	
20	2.69539581	0	2.70	2.93	3
21	3.16420746	0	3.16	3.43	
22	3.68982361	0	3.69	3.98	
23	4.27592499	0	4.28	4.60	
24	4.92616028	0	4.93	5.29	5
25	5.64413452	0	5.64	6.04	
26	6.43339905	0	6.43	6.87	7
27	7.29744141	0	7.30	7.77	
28	8.23967834	0	8.24	8.75	
29	9.2634491	0	9.26	9.82	10
30	10.37201172	0	10.37	10.97	
31	11.56853638	0	11.57	12.21	
32	12.85610352	0	12.86	13.55	
33	14.23770142	0	14.24	14.98	
34	15.71622925	0	15.72	16.51	
35	17.29448975	0	17.29	18.13	
36	18.97519287	0	18.98	19.87	20
37	20.76095947	0	20.76	21.71	
38	22.65431641	0	22.65	23.66	
39	24.65770508	0	24.66	25.72	
40	26.77348145	0	26.77	27.89	
41	29.00391357	0	29.00	30.18	30
42	31.35119385	0	31.35	32.58	
43	33.81743652	0	33.82	35.11	
44	36.40468262	0	36.40	37.76	
45	39.11490479	0	39.11	40.53	
46	41.94930664	0	41.95	43.43	
47	44.90817383	0	44.91	46.45	
48	47.99149414	0	47.99	49.60	50
49	51.1989502	0	51.20	52.86	
50	54.52990723	0	54.53	56.26	
51	57.98344727	0	57.98	59.77	
52	61.56074219	0	61.56	63.42	
53	65.26946777	0	65.27	67.19	
54	69.11870605	0	69.12	71.12	70
55	73.11869141	0	73.12	75.20	
56	77.27412109	0.000007	77.28	79.45	
57	81.59354004	0.000024	81.62	83.88	
58	86.08525391	0.000059	86.15	88.51	
59	90.76400391	0.000112	90.88	93.35	
60	95.62682617	0.000199	95.83	98.42	
61	100.6597852	0.00034	101.00	103.71	100
62	105.8463184	0.000562	106.42	109.24	
63	111.1666211	0.00089	112.07	115.02	
64	116.6006738	0.001353	117.97	121.05	
65	122.1154785	0.001992	124.13	127.35	125
66	127.6687305	0.002857	130.56	133.92	
67	133.2466895	0.003971	137.27	140.77	
68	138.8133106	0.005378	144.26	147.91	150
69	144.3213965	0.007133	151.55	155.34	
70	149.7561523	0.009261	159.14	163.09	
71	155.0825684	0.011806	167.04	171.16	175
72	160.2611523	0.014816	175.27	179.55	
73	165.2732227	0.018318	183.83	188.29	
74	170.0878906	0.022355	192.74	197.37	200
75	174.6761328	0.026964	202.00	206.81	
76	179.0162109	0.032176	211.62	216.62	
77	183.0843359	0.038026	221.61	226.80	225
78	186.8571875	0.044548	232.00	237.38	
79	190.3128906	0.051773	242.77	248.36	250
80	193.4351172	0.059728	253.95	259.75	
81	196.2004297	0.068448	265.56	271.57	
82	198.5939063	0.077958	277.58	283.82	
83	200.5993164	0.088286	290.06	296.52	300
84	202.1966406	0.099462	302.98	309.67	
85	203.3786328	0.111505	316.36	323.29	
86	204.1230859	0.124448	330.22	337.39	
87	204.4207813	0.138313	344.57	351.99	350
88	204.2571875	0.153125	359.41	367.09	
89	203.6181641	0.16891	374.77	382.71	
90	202.4951172	0.185689	390.64	398.85	400
91	200.8708594	0.203491	407.06	415.54	
92	198.7402539	0.222333	424.02	432.78	
93	196.0857227	0.242244	441.54	450.59	450
94	192.9022656	0.263242	459.63	468.97	
95	189.1746094	0.285354	478.31	487.95	
96	184.8970703	0.308598	497.58	507.50	500
97	180.0692578	0.332939	517.42	527.57	
98	174.7183984	0.358254	537.72	548.03	550
99	168.886875	0.384363	558.34	568.77	
100	162.6204688	0.411125	579.19	589.68	600
101	155.9669531	0.438391	600.17	610.66	
102	148.9845313	0.466003	621.16	631.62	
103	141.7332422	0.4938	642.08	652.44	650
104	134.2776953	0.521619	662.81	673.03	
105	126.6825781	0.549301	683.26	693.30	700
106	119.0133984	0.576692	703.35	713.16	
107	111.3330469	0.603648	722.98	732.53	
108	103.7017578	0.630036	742.09	751.34	750
109	96.17515625	0.655736	760.60	769.53	775
110	88.80453125	0.680643	778.47	787.05	
111	81.63375	0.704669	795.64	803.86	800
112	74.7034375	0.727739	812.08	819.93	825
113	68.04421875	0.749797	827.78	835.24	
114	61.6853125	0.770798	842.70	849.77	850
115	55.64382813	0.790717	856.84	863.52	
116	49.93796875	0.809536	870.20	876.50	875
117	44.57375	0.827256	882.79	888.71	
118	39.55960938	0.843881	894.62	900.17	900
119	34.89234375	0.859432	905.71	910.90	
120	30.57265625	0.873929	916.08	920.92	925
121	26.59140625	0.887408	925.76	930.26	
122	22.94242188	0.8999	934.77	938.95	
123	19.615	0.911448	943.14	947.02	
124	16.59476563	0.922096	950.91	954.51	950
125	13.87546875	0.931881	958.10	961.43	
126	11.4325	0.94086	964.76	967.83	
127	9.26507813	0.949064	970.90	973.74	975
128	7.34992188	0.95655	976.57	979.19	
129	5.680625	0.963352	981.80	984.20	
130	4.24414063	0.969513	986.60	988.81	
131	3.02476563	0.975078	991.02	993.05	
132	2.02484375	0.980072	995.08	996.95	
133	1.22101563	0.984542	998.81	1000.52	1000
134	0.6278125	0.9885	1002.23	1003.79	
135	0.22835938	0.991984	1005.36	1006.79	
136	0.03757813	0.995003	1008.22	1009.54	
137	0	0.99763	1010.85	1012.05	
	0	1	1013.25		

A2.4 JRA-25/JCDAS

JRA-25 used a hybrid sigma–pressure (hybrid σ–p) vertical coordinate after *Simmons and Burridge* (1981). The pressure on each level is calculated as $p_k = A_k + B_k \times p_{srf}$, where p_{srf} is surface pressure. The following table provides example pressures at layer interfaces (k–1/2) and layer midpoints (k) for a surface pressure of 1013.25 hPa, from TOA to surface. Pressures at layer midpoints are defined as the average of pressures at layer interfaces.

Model Levels					Pressure Levels
k	$A_{k-1/2}$ (hPa)	$B_{k-1/2}$	$p_{k-1/2}$ (hPa)	p_k (hPa)	p (hPa)
1	0.000000	0.000000	0.00	0.40	0.4
2	0.800000	0.000000	0.80	1.13	1
3	1.460000	0.000000	1.46	2.01	2
4	2.560000	0.000000	2.56	3.45	3
5	4.330000	0.000000	4.33	5.72	5
6	7.100000	0.000000	7.10	9.15	7
7	11.200000	0.000000	11.20	14.10	10
8	17.000000	0.000000	17.00	21.00	20
9	25.000000	0.000000	25.00	30.15	30
10	35.299999	0.000000	35.30	41.70	
11	48.099998	0.000000	48.10	55.55	50
12	62.634430	0.000366	63.01	71.53	70
13	76.105057	0.003895	80.05	89.60	
14	88.363998	0.010636	99.14	109.71	100
15	98.876595	0.021123	120.28	131.88	
16	107.299492	0.035701	143.47	156.10	150
17	113.447090	0.054553	168.72	182.38	
18	117.259979	0.077740	196.03	210.71	200
19	118.777374	0.105223	225.39	241.10	250
20	118.113609	0.136886	256.81	273.55	

Model Levels					Pressure Levels
k	$A_{k-1/2}$ (hPa)	$B_{k-1/2}$	$p_{k-1/2}$ (hPa)	p_k (hPa)	p (hPa)
21	115.438545	0.172561	290.29	308.05	300
22	110.961449	0.212039	325.81	344.09	
23	105.094887	0.253905	362.36	381.16	
24	98.151306	0.297849	399.95	419.76	400
25	90.192863	0.344807	439.57	460.40	
26	81.437820	0.394562	481.23	502.57	500
27	72.323532	0.445676	523.90	545.75	
28	63.056015	0.497944	567.60	589.95	600
29	53.811684	0.551188	612.30	635.16	
30	44.741348	0.605259	658.02	680.87	700
31	36.158020	0.658842	703.73	726.58	
32	28.130577	0.711869	749.43	771.77	
33	20.862747	0.763137	794.11	815.43	
34	14.485500	0.811514	836.75	856.55	850
35	9.064261	0.855936	876.34	894.10	
36	4.611954	0.895388	911.86	932.15	925
37	1.105610	0.938894	952.44	960.05	
38	0.000000	0.955000	967.65	977.79	
39	0.000000	0.975000	987.92	995.52	1000
40	0.000000	0.990000	1003.12	1008.18	
	0.000000	1.000000	1013.25		

A2.5 JRA-55

JRA-55 uses a hybrid sigma–pressure (hybrid σ–p) vertical coordinate after *Simmons and Burridge*, (1981). The pressure on each level is calculated as $p_k = A_k + B_k \times p_{srf}$, where p_{srf} is surface pressure. The following table provides example pressures at layer interfaces (k–1/2) and layer midpoints (k) for a surface pressure of 1013.25 hPa, from TOA to surface. Pressures at layer midpoints are defined as the average of pressures at layer interfaces.

Model Levels					Pressure Levels
k	$A_{k-1/2}$ (hPa)	$B_{k-1/2}$	$p_{k-1/2}$ (hPa)	p_k (hPa)	p (hPa)
1	0.000000	0.000000	0.00	0.10	
2	0.200000	0.000000	0.20	0.30	
3	0.390000	0.000000	0.39	0.52	
4	0.650000	0.000000	0.65	0.81	
5	0.970000	0.000000	0.97	1.17	1
6	1.360000	0.000000	1.36	1.59	
7	1.820000	0.000000	1.82	2.10	2
8	2.370000	0.000000	2.37	2.69	
9	3.010000	0.000000	3.01	3.39	3
10	3.770000	0.000000	3.77	4.23	
11	4.690000	0.000000	4.69	5.25	5
12	5.810000	0.000000	5.81	6.51	7
13	7.200000	0.000000	7.20	8.07	
14	8.930000	0.000000	8.93	9.99	10
15	11.050000	0.000000	11.05	12.38	
16	13.700000	0.000000	13.70	15.35	
17	17.000000	0.000000	17.00	19.03	20
18	21.050000	0.000000	21.05	23.58	
19	26.100000	0.000000	26.10	29.20	30
20	32.300000	0.000000	32.30	36.15	
21	40.000000	0.000000	40.00	44.75	
22	49.500000	0.000000	49.50	55.25	50
23	60.886730	0.000113	61.00	67.77	70
24	72.015690	0.002484	74.53	81.81	
25	82.262449	0.006738	89.09	97.13	100
26	91.672470	0.013328	105.18	114.24	
27	100.146151	0.022854	123.30	133.39	125
28	107.299494	0.035701	143.47	154.58	150
29	112.854041	0.052146	165.69	177.82	175
30	116.633554	0.072366	189.96	203.12	200

Model Levels					Pressure Levels
k	$A_{k-1/2}$ (hPa)	$B_{k-1/2}$	$p_{k-1/2}$ (hPa)	p_k (hPa)	p (hPa)
31	118.554343	0.096446	216.28	230.46	225
32	118.612531	0.124387	244.65	259.35	250
33	116.953716	0.155046	274.05	289.78	300
34	113.696478	0.189304	305.51	321.75	
35	109.126384	0.225874	337.99	355.26	350
36	103.294362	0.265706	372.52	390.30	400
37	96.561819	0.307438	408.07	426.36	
38	89.140822	0.350859	444.65	463.45	450
39	81.221598	0.395778	482.24	501.55	500
40	72.974699	0.442025	520.86	540.16	550
41	64.767182	0.488233	559.47	578.77	
42	56.718242	0.534282	598.08	617.38	600
43	48.918808	0.580081	636.69	655.48	650
44	41.629564	0.62437	674.27	693.06	700
45	34.688715	0.668311	711.85	729.63	750
46	28.474848	0.709525	747.40	764.16	775
47	22.948417	0.748052	780.91	797.16	800
48	17.909074	0.785091	813.40	828.63	825
49	13.4768	0.819523	843.86	858.07	850
50	9.597972	0.851402	872.28	884.97	875
51	6.346027	0.879654	897.66	908.82	900
52	3.649041	0.904351	919.98	930.13	925
53	1.33051	0.926669	940.28	949.41	950
54	0	0.946	958.53	965.63	
55	0	0.96	972.72	978.80	975
56	0	0.972	984.88	989.95	
57	0	0.982	995.01	998.56	1000
58	0	0.989	1002.10	1004.64	
59	0	0.994	1007.17	1008.69	
60	0	0.997	1010.21	1011.73	
	0	1	1013.25		

A2.6 MERRA and MERRA-2

MERRA and MERRA-2 use identical hybrid sigma–pressure (hybrid σ–p) vertical coordinates after *Simmons and Burridge*, (1981). The pressure on each level is calculated as $p_k = A_k + B_k \times p_{srf}$, where p_{srf} is surface pressure. The following table provides example pressures at layer interfaces (k–1/2) and layer midpoints (k) for a surface pressure of 1013.25 hPa, from TOA to surface. Pressures at layer midpoints are defined as the average of pressures at layer interfaces. NASA GMAO is transitioning away from this vertical grid and recommends that data users use the three-dimensional pressure fields provided with MERRA and MERRA-2 instead.

Model Levels					Pressure Levels
k	$A_{k-1/2}$ (hPa)	$B_{k-1/2}$	$p_{k-1/2}$ (hPa)	p_k (hPa)	p (hPa)
1	0.0100	0	0.01	0.015	
2	0.0200	0	0.02	0.026	
3	0.0327	0	0.03	0.040	
4	0.0476	0	0.05	0.057	
5	0.0660	0	0.07	0.078	
6	0.0893	0	0.09	0.105	0.1
7	0.1197	0	0.12	0.140	
8	0.1595	0	0.16	0.185	
9	0.2113	0	0.21	0.245	
10	0.2785	0	0.28	0.322	0.3
11	0.3650	0	0.37	0.420	0.4
12	0.4758	0	0.48	0.546	0.5
13	0.6168	0	0.62	0.706	0.7
14	0.7951	0	0.80	0.907	1
15	1.0194	0	1.02	1.160	
16	1.3005	0	1.30	1.476	
17	1.6508	0	1.65	1.868	2
18	2.0850	0	2.08	2.353	
19	2.6202	0	2.62	2.948	3
20	3.2764	0	3.28	3.677	4
21	4.0766	0	4.08	4.562	5
22	5.0468	0	5.05	5.632	
23	6.2168	0	6.22	6.918	7
24	7.6198	0	7.62	8.456	
25	9.2929	0	9.29	10.285	10
26	11.2769	0	11.28	12.460	
27	13.6434	0	13.64	15.050	
28	16.4571	0	16.46	18.124	
29	19.7916	0	19.79	21.761	20
30	23.7304	0	23.73	26.049	
31	28.3678	0	28.37	31.089	30
32	33.8100	0	33.81	36.993	40
33	40.1754	0	40.18	43.910	
34	47.6439	0	47.64	52.016	50
35	56.3879	0	56.39	61.496	
36	66.6034	0	66.60	72.558	70

Model Levels					Pressure Levels
k	$A_{k-1/2}$ (hPa)	$B_{k-1/2}$	$p_{k-1/2}$ (hPa)	p_k (hPa)	p (hPa)
37	78.5123	0	78.51	85.439	
38	92.3657	0	92.37	100.514	100
39	108.6630	0	108.66	118.250	
40	127.8370	0	127.84	139.115	150
41	150.3930	0	150.39	163.662	
42	176.9300	0	176.93	192.587	200
43	201.1920	0.006960	208.24	226.745	
44	216.8650	0.028010	245.25	267.087	250
45	224.3630	0.063720	288.93	313.966	300
46	223.8980	0.113602	339.01	358.038	350
47	218.7760	0.156224	377.07	396.112	400
48	212.1500	0.200350	415.15	434.212	450
49	203.2590	0.246741	453.27	472.335	
50	193.0970	0.294403	491.40	510.475	500
51	181.6190	0.343381	529.55	548.628	550
52	169.6090	0.392891	567.71	586.793	600
53	156.2600	0.443740	605.88	624.966	
54	142.9100	0.494590	644.05	663.146	650
55	128.6960	0.546304	682.24	694.969	700
56	118.9590	0.581041	707.70	720.429	725
57	109.1820	0.615818	733.16	745.890	750
58	99.3652	0.650635	758.62	771.355	775
59	89.0999	0.685900	784.09	796.822	800
60	78.8342	0.721166	809.56	819.742	825
61	70.6220	0.749378	829.93	837.570	
62	64.3626	0.770637	845.21	852.852	850
63	58.0532	0.791947	860.49	868.135	875
64	51.6961	0.813304	875.78	883.418	
65	45.3390	0.834661	891.06	898.701	900
66	38.9820	0.856018	906.34	913.984	
67	32.5708	0.877429	921.63	929.268	925
68	26.0920	0.898908	936.91	944.553	950
69	19.6131	0.920387	952.20	959.837	
70	13.1348	0.941865	967.48	975.122	975
71	6.5938	0.963406	982.76	990.408	
72	0.0480	0.984952	998.05	1005.650	1000
	0	1	1013.25		

A2.7 NCEP-NCAR R1 and NCEP-DOE R2

NCEP-NCAR R1 and NCEP-DOE R2 use a sigma vertical coordinate. The pressure on each level is calculated as $p_k = \sigma_k \times p_{srf}$, where p_{srf} is surface pressure. The following table provides example pressures at each level for a surface pressure of 1013.25 hPa, from TOA to surface.

Model Levels			Pressure Levels
k	σ_k	p_k (hPa)	p (hPa)
1	0.00273	2.77	3
2	0.01006	10.19	10
3	0.01834	18.58	20
4	0.02875	29.13	30
5	0.04179	42.34	
6	0.05805	58.82	50
7	0.07815	79.19	70
8	0.10278	104.14	100
9	0.13261	134.37	
10	0.16823	170.46	150

Model Levels			Pressure Levels
k	σ_k	p_k (hPa)	p (hPa)
11	0.21006	212.84	200
12	0.25823	261.65	250
13	0.31248	316.62	300
14	0.37205	376.98	400
15	0.43568	441.45	
16	0.50168	508.33	500
17	0.56809	575.62	
18	0.63290	641.29	
19	0.69426	703.46	700
20	0.75076	760.71	

Model Levels			Pressure Levels
k	σ_k	p_k (hPa)	p (hPa)
21	0.80142	812.04	
22	0.84579	857.00	850
23	0.88384	895.55	
24	0.91592	928.06	925
25	0.94255	955.04	
26	0.96437	977.15	
27	0.98208	995.09	
28	0.99500	1008.18	1000
	1.00000	1013.25	

A2.8 CFSR

CFSR uses a hybrid sigma–pressure (hybrid σ–p) vertical coordinates after *Simmons and Burridge* (1981). The pressure on each level is calculated as $p_k = A_k + B_k \times p_{srf}$, where p_{srf} is surface pressure. The following table provides example pressures at layer interfaces (k–1/2) and layer midpoints (k) for a surface pressure of 1013.25 hPa, from TOA to surface. Pressures at layer midpoints are defined as the average of pressures at layer interfaces.

Model Levels					Pressure Levels
k	$A_{k-1/2}$ (hPa)	$B_{k-1/2}$	$p_{k-1/2}$ (hPa)	p_k (hPa)	p (hPa)
1	0.00000	0.000000	0.00	0.32	
2	0.64247	0.000000	0.64	1.01	1
3	1.37790	0.000000	1.38	1.80	2
4	2.21958	0.000000	2.22	2.70	3
5	3.18266	0.000000	3.18	3.73	
6	4.28434	0.000000	4.28	4.91	5
7	5.54424	0.000000	5.54	6.26	
8	6.98457	0.000000	6.98	7.81	7
9	8.63058	0.000000	8.63	9.57	10
10	10.51080	0.000000	10.51	11.58	
11	12.65752	0.000000	12.66	13.88	
12	15.10711	0.000000	15.11	16.50	
13	17.90051	0.000000	17.90	19.49	20
14	21.08366	0.000000	21.08	22.90	
15	24.70788	0.000000	24.71	26.77	
16	28.83038	0.000000	28.83	31.17	30
17	33.51460	0.000000	33.51	36.17	
18	38.83052	0.000000	38.83	41.84	
19	44.85493	0.000000	44.85	48.26	50
20	51.67146	0.000000	51.67	55.52	
21	59.37050	0.000000	59.37	63.71	
22	68.04874	0.000000	68.05	72.93	70
23	77.77150	0.000037	77.81	83.29	
24	88.32537	0.000431	88.76	94.89	100
25	99.36614	0.001636	101.02	107.87	
26	110.54853	0.004107	114.71	122.32	125
27	121.52937	0.008294	129.93	138.37	
28	131.97065	0.014637	146.80	156.11	150
29	141.54316	0.023556	165.41	175.63	175
30	149.93074	0.035442	185.84	197.00	200
31	156.83489	0.050647	208.15	220.26	225
32	161.97967	0.069475	232.37	245.44	250
33	165.11736	0.092167	258.51	272.50	
34	166.11603	0.118812	286.50	301.39	300
35	165.03144	0.149269	316.28	331.99	
36	161.97315	0.183296	347.70	364.14	350
37	157.08893	0.220570	380.58	397.64	400
38	150.56342	0.260685	414.70	432.25	
39	142.61435	0.303164	449.80	467.68	450
40	133.48671	0.347468	485.56	503.61	500
41	123.44490	0.393018	521.67	539.73	550
42	112.76348	0.439211	557.79	575.69	
43	101.71712	0.485443	593.59	611.17	600
44	90.57051	0.531135	628.74	645.84	650
45	79.56908	0.575747	662.94	679.44	
46	68.93117	0.618800	695.93	711.70	700
47	58.84206	0.659887	727.47	742.43	750
48	49.45029	0.698683	757.39	771.47	775
49	40.86614	0.734945	785.55	798.70	800
50	33.16217	0.768515	811.86	824.07	825
51	26.37553	0.799310	836.28	847.53	850
52	20.51150	0.827319	858.79	869.11	875
53	15.54789	0.852591	879.44	888.85	
54	11.43988	0.875224	898.26	906.80	900
55	8.12489	0.895355	915.34	923.06	925
56	5.52720	0.913151	930.78	937.72	
57	3.56223	0.928797	944.67	950.89	950
58	2.14015	0.942491	957.12	962.68	
59	1.16899	0.954434	968.25	973.21	975
60	0.55712	0.964828	978.17	982.58	
61	0.21516	0.973868	986.99	990.90	
62	0.05741	0.981742	994.81	998.27	1000
63	0.00575	0.988627	1001.73	1004.79	
64	0.00000	0.994671	1007.85	1010.55	
	0.00000	1.000000	1013.25		

A2.9 20CR

The 20CR uses a hybrid sigma–pressure (hybrid σ–p) vertical coordinates after *Simmons and Burridge*, (1981). The pressure on each level is calculated as $p_k = A_k + B_k \times p_{srf}$, where p_{srf} is surface pressure. The following table provides example pressures at layer interfaces (k–1/2) and layer midpoints (k) for a surface pressure of 1013.25 hPa, from TOA to surface. Pressures at layer midpoints are defined as the average of pressures at layer interfaces.

Model Levels					Pressure Levels
k	$A_{k-1/2}$ (hPa)	$B_{k-1/2}$	$p_{k-1/2}$ (hPa)	p_k (hPa)	p (hPa)
1	0.00000	0.000000	0.00	2.83	
2	5.66898	0.000000	5.67	9.29	10
3	12.90533	0.000000	12.91	17.51	20
4	22.10979	0.000000	22.11	27.94	30
5	33.76516	0.000000	33.77	41.10	
6	48.44036	0.000000	48.44	57.61	50
7	66.78608	0.000000	66.79	78.15	70
8	89.13767	0.000379	89.52	103.47	100
9	113.43654	0.003933	117.42	134.33	150
10	136.71427	0.014326	151.23	171.39	
11	156.13564	0.034950	191.55	215.13	200
12	169.12130	0.068675	238.71	265.66	250
13	173.64658	0.117418	292.62	322.64	300,350
14	169.59994	0.180667	352.66	385.13	400
15	158.12926	0.256084	417.61	451.65	450
16	140.89535	0.340293	485.70	520.25	500,550
17	119.91428	0.429195	554.80	588.72	600
18	97.31807	0.518457	622.64	654.89	650
19	75.08532	0.604055	687.14	716.87	700
20	54.81144	0.682747	746.60	773.25	750
21	37.57142	0.752347	799.89	823.16	800
22	23.89205	0.811785	846.43	866.32	850
23	13.81526	0.860975	886.20	902.86	900
24	7.01453	0.900581	919.53	933.27	
25	2.92577	0.931750	947.02	958.21	950
26	0.86457	0.955872	969.40	978.42	
27	0.11635	0.974402	987.43	994.63	1000
28	0.00009	0.988726	1001.83	1007.54	
	0.00000	1.000000	1013.25		

Major abbreviations and terms

20CR	20th Century Reanalysis (v2 for version 2, v2c for version 2c, and v3 for version 3)
2D-Var	2-dimensional variational assimilation scheme
3D-Var	3-dimensional variational assimilation scheme
3D-FGAT	3-dimensional variational assimilation scheme with FGAT
4D-Var	4-dimensional variational assimilation scheme
ABL	atmospheric boundary layer
ACARS	Aircraft Communications Addressing and Reporting System
ACRE	Atmospheric Circulation Reconstructions over the Earth
AER	Atmospheric and Environmental Research
AERONET	Aerosol Robotic Network
AGCM	atmospheric general circulation model
AHI	Advanced Himawari Imager
AIRS	Atmospheric Infrared Sounder
AMDAR	Aircraft Meteorological Data Relay
AMIP	Atmospheric Model Intercomparison Project
AMSR	Advanced Microwave Scanning Radiometer
AMSR-E	Advanced Microwave Scanning Radiometer for EOS
AMSU	Advanced Microwave Sounding Unit
AMV	atmospheric motion vectors
ANA	"analyzed" state produced prior to IAU for MERRA and MERRA-2
AOD	aerosol optical depth
Aqua	a satellite in NASA's Earth Observing System (EOS) A-Train constellation
ASCAT	Advanced Scatterometer
ASM	"assimilated" state produced by IAU for MERRA and MERRA-2
ATMS	Advanced Technology Microwave Sounder
ATOVS	Advanced TIROS Operational Vertical Sounder
Aura	a satellite in NASA's Earth Observing System (EOS) A-Train constellation
AVHRR	Advanced Very High Resolution Radiometer
BAS	British Antarctic Survey
BOM	Bureau of Meteorology (Australia)
BUOY	Surface meteorological observation report from buoys
CAMSiRA	Copernicus Atmosphere Monitoring Service Interim Reanalysis
CCARDS	Comprehensive Aerological Reference Dataset, Core Subset
CCI	Climate Change Initiative (ESA)
CEDA	Centre for Environmental Data Analysis
CERA	a coupled atmosphere–ocean data assimilation system developed by ECMWF
CFC	chlorofluorocarbon
CFS	Climate Forecast System developed at NCEP
CFSR	Climate Forecast System Reanalysis
CFSv2	Climate Forecast System version 2
CHAMP	CHAllenging Minisatellite Payload
CIRES	Cooperative Institute for Research in Environmental Sciences (NOAA and University of Colorado Boulder)
CMA	China Meteorological Administration
CMAP	CPC Merged Analysis of Precipitation
CMIP5	Coupled Model Intercomparison Project Phase 5

CNSA	China National Space Administration
COBE	Centenial in-situ Observation-Based Estimates of variability of SST and marine meteorological variables
COSMIC	Constellation Observing System for Meteorology, Ionosphere, and Climate
CPC	Climate Prediction Center (NOAA)
CRIEPI	Central Research Institute of Electric Power Industry
CrIS	Cross-track Infrared Sounder
CRTM	Community Radiative Transfer Model
CRUTEM	Climatic Research Unit Air Temperature Anomalies
CTM	chemical transport model
DAO	Data Assimilation Office (NASA); now GMAO
DAS	data assimilation system
DMSP	Defense Meteorological Satellite Program
DOE	Department of Energy
ECMWF	European Centre for Medium-Range Weather Forecasts
EDA	the 10-member "ensemble of data assimilations" produced for ERA5
EMC	Ensemble Modeling Center
EnKF	Ensemble Kalman Filter assimilation scheme
EOS	Earth Observing System of the NASA
ERA-15	ECMWF 15-year reanalysis
ERA-20C	ECMWF 20th century reanalysis
ERA-40	ECMWF 40-year reanalysis
ERA-CLIM	European Reanalysis of Global Climate Observations
ERA-Interim	ECMWF interim reanalysis
ERA5	the fifth major global reanalysis produced by ECMWF
ERA5L	a land surface reanalysis with atmospheric forcing from ERA5
ERS	European Remote Sensing satellite
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FGAT	first guess at appropriate time
FGGE	First GARP (Global Atmospheric Research Program) Global Experiment
FORMOSAT	The name given to the Republic of China Satellite (ROCSat) following a public naming competition.
FY-3	FengYun-3 (a series of polar-orbiting satellites launched by the CMA and CNSA)
GAAS	Goddard Aerosol Assimilation System
GARP	Global Atmospheric Research Program
GATE	GARP (Global Atmospheric Research Program) Atlantic Tropical Experiment
GAW	Global Atmosphere Watch
GCM	general circulation model
GEO	geostationary satellites
GEOS	Goddard Earth Observing System Model of the NASA
GFDL	Geophysical Fluid Dynamics Laboratory of the NOAA
GFS	Global Forecast System of the NCEP
GISST	UKMO Global Ice and Sea Surface Temperature dataset
GLATOVS	Goddard Laboratory for Atmospheres TOVS (radiative transfer model)
GLCC	Global Land Cover Characteristics data base
GLDAS	Global Land Data Assimilation System
GMI	GPM Microwave Imager
GMS	Geostationary Meteorological Satellite
GNSS-RO	Global Navigation Satellite System Radio Occultation (see also GPS-RO)
GOCART	Goddard Chemistry, Aerosol, Radiation, and Transport model
GODAS	NCEP Global Ocean Data Assimilation System
GOES	Geostationary Operational Environmental Satellite

GOME	Global Ozone Monitoring Experiment
GPCP	Global Precipitation Climatology Project
GPM	Global Precipitation Measurement mission
GPS-RO	Global Positioning System Radio Occultation (see also GNSS-RO)
GRACE	Gravity Recovery and Climate Experiment
GRIB	General Regularly-distributed Information in Binary form (a file format)
GRIB2	GRIB, Version 2 (a file format)
GRUAN	Global Climate Observing System Reference Upper Air Network
GSI	Gridpoint Statistical Interpolation assimilation scheme
GSICS	Global Space-based Inter-calibration System
GSM	Global Spectral Model of the JMA
GTS	Global Telecommunication System
GWD	gravity wave drag
HadISST	UKMO Hadley Centre Sea Ice and SST dataset
HALOE	Halogen Occultation Experiment
HCFC	hydrochlorofluorocarbon
HDF	Hierarchical Data Format (a file format)
HIRS	High-resolution Infrared Radiation Sounder
HRES	the high-resolution analysis produced for ERA5
IASI	Infrared Atmospheric Sounding Interferometer
IAU	Incremental Analysis Update procedure (or products resulting from that procedure)
ICOADS	International Comprehensive Ocean-Atmosphere Data Set
IFS	Integrated Forecast System of the ECMWF
IGY	International Geophysical Year (July 1957–December 1958)
IPCC	Intergovernmental Panel on Climate Change
IR	infrared
ISPD	International Surface Pressure Databank
JAXA	Japan Aerospace Exploration Agency
JCDAS	JMA Climate Data Assimilation System
JCSDA	Joint Center for Satellite Data Assimilation
JMA	Japan Meteorological Agency
JRA-25	Japanese 25-year Reanalysis
JRA-55	Japanese 55-year Reanalysis
JRA-55AMIP	Japanese 55-year Reanalysis based on AMIP-type simulations
JRA-55C	Japanese 55-year Reanalysis assimilating Conventional observations only
LAI	leaf area index
LCL	lifting condensation level
LEO/GEO	Low Earth Orbit / Geostationary
LIE	Line Islands Experiment
LSM	land surface model
MARS	Meteorological Archival and Retrieval System of the ECMWF
McICA	Monte Carlo Independent Column Approximation
MERRA	Modern Era Retrospective-Analysis for Research and Applications
MERRA-2	Modern Era Retrospective-Analysis for Research and Applications, Version 2
Met Office	see UKMO
METEOSAT	geostationary meteorological satellites operated by EUMETSAT
MetOp	A series of three polar orbiting meteorological satellites operated by the EUMETSAT
MHS	Microwave Humidity Sounder
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MISR	Multiangle Imaging Spectroradiometer
MIT	Massachusetts Institute of Technology

MLS	Microwave Limb Sounder
MODIS	MODerate resolution Imaging Spectroradiometer
MOM	Modular Ocean Model
MRF	Medium Range Forecast Version of the NCEP Global Forecast System
MRI-CCM1	Meteorological Research Institute (JMA) Chemistry Climate Model, version 1
MSU	Microwave Sounding Unit
MTSAT	Multi-functional Transport Satellite
MW	microwave (sounders)
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCDC	National Climatic Data Center of the NOAA
NCEI	National Centers for Environmental Information (NOAA)
NCEP	National Centers for Environmental Prediction (NOAA)
NCEP-DOE R-2	Reanalysis 2 of the NCEP and DOE
NCEP-NCAR R-1	Reanalysis 1 of the NCEP and NCAR
NESDIS	National Environmental Satellite, Data, and Information Service
NetCDF	Network Common Data Form (a file format)
NH	Northern Hemisphere
NIST	National Institute of Standards and Technology
NMC	National Meteorological Center
NOAA	National Oceanic and Atmospheric Administration
NOAA-CIRES 20C	20th Century Reanalysis of the NOAA and CIRES (see also 20CR)
NSIDC	National Snow and Ice Data Center
OI	optimal interpolation
OISST	NOAA Optimum Interpolation Sea Surface Temperature (v2 for version 2)
OMI	Ozone Monitoring Instrument
OSTIA	Operational Sea Surface Temperature and Sea-Ice Analysis
OSU LSM	Oregon State University LSM
PAOBS	Bogus surface pressure data for the Southern Hemisphere produced by the Australian Bureau of Meteorology
PCMDI	Program of Climate Model Diagnosis and Intercomparison
PDF	probability distribution function
PIBAL	Pilot Balloon
QBO	Quasi-Biennial Oscillation
QC	quality control
QuikSCAT	Quick Scatterometer
R1	see NCEP–NCAR R1
R2	see NCEP–DOE R2
RAOBCORE	Radiosonde Observation Correction using Reanalyses
RCP	representative concentration pathway (IPCC)
RDA	Research Data Archive (NCAR)
RH	relative humidity
RICH	Radiosonde Innovation Composite Homogenization
RO	radio occultation
RRTM	Rapid Radiative Transfer Model developed by AER
RRTM-G	Rapid Radiative Transfer Model for application to GCMs developed by AER
RTG	NCEP Real-Time Global sea surface temperature
RTTOV	Radiative Transfer for TOVS
S-RIP	SPARC Reanslysis Intercomparison Project
SBUV	Solar Backscatter Ultraviolet Radiometer
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY

SEVIRI	Spinning Enhanced Visible and Infrared Imager (EUMETSAT)
SH	Southern Hemisphere
SHIP	Surface meteorological observation report from ships
SiB	Simple Biosphere model
SIC	sea ice concentration
SMMR	Scanning Multichannel Microwave Radiometer
SNDR	Sounder (for radiance measurements by the GOES 8 to 12)
SNO	Simultaneous Nadir Overpass method
SOLARIS-HEPPA	Solar Influences for SPARC–High Energy Partical Precipitation in the Atmosphere
SPARC	Stratosphere-troposphere Processes And their Role in Climate
SSI	Spectral Statistical Interpolation (an assimilation scheme)
SSM/I or SSMI	Special Sensor Microwave Imager
SSMIS	Special Sensor Microwave Imager Sounder
SST	sea surface temperature
SSU	Stratospheric Sounding Unit
SYNOP	Surface meteorological observation report from manned and automated weather stations
TCWV	total column water vapour
TD	tape deck ("TD" is a name of a rawinsonde dataset. For example, TD54 is a dataset of mandatory level data from rawinsondes during 1946-1972 prepared by the USAF. See **http://rda.ucar.edu/docs/papers-scanned/pdf/rj0187.pdf** (accessed 29 May 2015).
Terra	a satellite in NASA's Earth Observing System (EOS).
TerraSAR-X	a German satellite with a phased array Synthetic Aperture Radar (SAR) antenna at the X-band wavelength
TIM	Total Irradiance Monitor
TIROS	Television Infrared Observation Satellite
TIROS-N	Television InfraRed Operational Satellite - Next-generation
TMI	TRMM Microwave Imager
TOA	top of atmosphere
TOMS	Total Ozone Mapping Spectrometer
TOVS	TIROS Operational Vertical Sounder
TRMM	Tropical Rainfall Measuring Mission
TSI	total solar irradiance
UARS	Upper Atmosphere Research Satellite
UKMO	United Kingdom Meteorological Office (or Met Office)
USAF	U.S. Air Force
USCNTRL	U.S. controlled oceanweather stations
USGS	U.S. Geological Survey
UTC	Universal Coordinated Time
VTPR	Vertical Temperature Profile Radiometer
WMO	World Meteorological Organization