# CHAPTER CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Instrument calibration</td>
<td>1010</td>
</tr>
<tr>
<td>6.1.1 Introduction</td>
<td>1010</td>
</tr>
<tr>
<td>6.1.2 Factors affecting calibration</td>
<td>1010</td>
</tr>
<tr>
<td>6.1.3 Pre-launch calibration</td>
<td>1011</td>
</tr>
<tr>
<td>6.1.4 On-board calibration</td>
<td>1011</td>
</tr>
<tr>
<td>6.1.5 Vicarious calibration</td>
<td>1012</td>
</tr>
<tr>
<td>6.1.6 Intercalibration by simultaneous observations</td>
<td>1012</td>
</tr>
<tr>
<td>6.1.7 Bias adjustment of long-term data records</td>
<td>1013</td>
</tr>
<tr>
<td>6.1.8 Using calibration information</td>
<td>1013</td>
</tr>
<tr>
<td>6.1.9 Traceability of space-based measurements</td>
<td>1014</td>
</tr>
<tr>
<td>6.2 Product validation</td>
<td>1014</td>
</tr>
<tr>
<td>6.2.1 Factors to be accounted for in validation</td>
<td>1014</td>
</tr>
<tr>
<td>6.2.2 Validation strategies</td>
<td>1015</td>
</tr>
<tr>
<td>6.2.3 Impact studies</td>
<td>1016</td>
</tr>
<tr>
<td>REFERENCES AND FURTHER READING</td>
<td>1017</td>
</tr>
</tbody>
</table>
CHAPTER 6. CALIBRATION AND VALIDATION

6.1 INSTRUMENT CALIBRATION

6.1.1 Introduction

Calibration is the process of quantitatively defining the satellite instrument response to known, controlled signal inputs. The calibration information is contained in a calibration formula or in calibration coefficients that are then used to convert the instrument output (measured in “counts”, or, previously, “analogue signals”) into physical units (for example, radiance values). Instrument calibration is critical for any higher-level data processing, especially for deriving quantitative products or when data from different instruments need to be merged (such as for composite imagery). For climate applications, the requirement for accurate calibration is particularly stringent since detection of small trends over long periods requires the ability to compare different instruments flown on different satellites at different times. Building homogeneous climate data records is contingent on very stable calibration and error characterization.

The following considerations apply to passive and active instruments alike.

Five calibration domains should be generally considered: radiometric, spectral, spatial, temporal, and polarization. A complete calibration record should include estimates of uncertainties in calibration parameters. Satellite instrument calibration should take into account all phases of an instrument’s lifetime: from design and pre-launch phases to post-launch and on-orbit operations.

The intercalibration of instruments against a common reference instrument allows for consistency among satellite measurements at a given point in time. By comparing model-simulated and observed satellite radiances in data assimilation schemes, major numerical weather prediction (NWP) centres can also help determine relative biases between instruments. Calibration to absolute standards is, nevertheless, necessary to allow traceability of errors and to detect any long-term drift over time unambiguously.

Calibration using well-characterized, stable Earth targets (called vicarious calibration) is a fallback when a satellite instrument cannot be directly traceable to an agreed reference standard, for example due to the absence of a reliable on-board calibration device. Data records from past instruments can be “recalibrated” retrospectively, if additional information on the state of these instruments becomes available, for example through comparison with reprocessed, well-known historical time series.

6.1.2 Factors affecting calibration

The response of an instrument to signal input, i.e. the relationship between the irradiance the instrument is exposed to and the numerical value assigned to the measurement (in physical units, for example, W m⁻²) depends on several elements, such as:

(a) The viewing geometry, shielding effects, stray light, and antenna pattern;

(b) Detector sensitivity and ageing;

(c) Filter optics, as well as the possible contamination and stability of the filter;

---

1 From the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation. (The terms defined in this Part differ in some instances from those defined in JCGM, 2012).

2 See, for example, Ohring (2007).

3 For guidance on reference standards, see, for example, Fox (2010).
(d) The temperature of all parts of the instrument, including the front-end optics, detector and back-end electronics (focal plane electronics, preamplifier, etc.);

(e) The signal-processing system (gain, analogue-to-digital converter, etc.).

All of these elements help to determine the spectral response function and the point spread function that characterize the instrument from a radiometric and geometric viewpoint, respectively. They must be modelled before launch and monitored in flight by a set of on-board internal measurements (the housekeeping system). The instrument model and the housekeeping system are useful for understanding the status of the instrument and its trend as well as for predicting and correcting biases. However, it is generally not possible to analytically describe the exact variation of the instrument response resulting from these factors. Reference measurements are mandatory to characterize the actual calibration.

6.1.3 Pre-launch calibration

The pre-launch calibration of an instrument is performed in the laboratory, by using accurately known radiation sources under controlled conditions. Simulating all possible instrument states and stress factors before launch is very important because it is the only way to accurately characterize and model the instrument before it is exposed to the harsh orbital environment. Housekeeping systems and instruments need to be robust enough to withstand physical stress incurred during the launch, commissioning and exploitation phases. Housekeeping data, in combination with post-launch calibration information, will then allow operators to infer the calibration status of the instrument in orbit and to resolve on-orbit anomalies.

6.1.4 On-board calibration

On-board calibration involves monitoring the instrument performance (and stability) while in orbit. It is performed using reference targets (such as black bodies in the infrared, solar diffusers, and lamp line sources in short wave) for passive instruments, or by internal calibration systems (such as gain monitors) for active instruments. Some heritage instruments have been in operation without adequate means of on-board calibration, such as the Advanced Very High Resolution Radiometer (AVHRR), which provides long-term observations in the visible and near-infrared regions. Other means of calibration (e.g. vicarious, intercalibration) need to be used for characterizing such instruments. The accuracy of in-flight instrument calibration is a function of the stability of the on-board calibration systems throughout the instrument’s lifetime. Therefore, the calibration itself must be regularly checked by intercalibration against highly accurate references.

In the case of infrared instruments, if the radiometer detectors are assumed to have linear response, the output voltage is given as:

\[ V = \alpha R + V_0 \]

where \( R \) is the input radiance, \( \alpha \) is the radiometer responsivity and \( V_0 \) is the system offset. Calibration consists of determining \( \alpha \) and \( V_0 \), which is accomplished by exposing the radiometer to at least two reference targets with significantly different brightness temperatures.

For infrared and microwave instruments, one reference target is deep space, at a temperature of 2.725 K. Direct viewing of deep space is not always possible for instruments on a satellite platform. For instance, pushbroom instruments constantly pointing to the Earth’s surface need to be equipped with a sub-reflector to supply the deep space view at intervals. A second target is usually a well-characterized source with temperature in the medium to upper dynamic range, often a black body, which is ideally traceable to the International System of Units (SI), i.e. to a radiance scale provided by a national metrology institute.
If the instrument response is not linear across the dynamic range, this needs to be accounted for in the pre-launch instrument characterization, for example by using a quadratic function, or through linearization in different parts of the dynamic range and the possible addition of a second black body kept at a different temperature.

For ultraviolet, visible and near-infrared instruments, on-board calibration is more challenging since it is affected by many factors. At the low-signal extreme, deep space is a useful reference, provided that disrupting effects (for example, reflections from other parts of the satellite) are avoided. At the high-signal end, an absolute source is generally replaced by solar diffusers that provide a relatively stable reference. The moon also may be used as a reference target, with the advantage that it can be viewed without an attenuator; however, it must be used in conjunction with an accurate model of the moon's brightness. Neither the solar diffuser nor the moon provides an absolute calibration. Another system often used is a bench of lamp line sources of well-controlled intensity. Spectrally-dependent polarization effects induced by the reflecting surfaces of the instrument optics also need to be taken into account.

Another problem with on-board calibration is that often the instrument structure does not allow illumination of the full primary optics with reference sources. For example, a spin-stabilized radiometer in geostationary Earth orbit uses an internal black body requiring a model of the contributions of the telescope and foreoptics to the background radiation. Often, the reference source only illuminates a fraction of the total instrument optics and, therefore, is more used for stability monitoring than for absolute calibration.

### 6.1.5 Vicarious calibration

On-board calibration can be complemented by stable ground targets used as references in a process termed vicarious calibration. The target needs to be well characterized in order to infer the emitted or reflected radiance towards space. The combined effects of the viewing geometry and, in short wave, the bidirectional reflectance distribution function of the surface and atmosphere must be taken into account. The radiative transfer through the atmosphere between the satellite and the ground reference source must be accurately known at the time of the satellite overpass. In a cloud-free case, the short-wave spectrum is particularly affected by aerosols, whereas the long-wave spectrum is particularly influenced by the presence of water vapour.

Vicarious calibration can involve different kinds of targets: polar ice fields as a black body for microwave radiometers; snow fields, sunglint, homogeneous desert areas, and deep convective cloud tops for the upper end of the visible dynamic range; cloud-free ocean surface for dark targets in the visible spectrum; cube-corner reflectors for synthetic aperture radars; the rainforest as a black body for radar scatterometers; etc. Calibration field sites equipped with in situ observations are used for the calibration of high spatial resolution space-based instruments. During initial payload commissioning or at regular intervals, aircraft overflights of a target area synchronous with the satellite overpass offer additional vicarious calibration data.

### 6.1.6 Intercalibration by simultaneous observations

The intercalibration of satellite instruments involves relating the measurements of one instrument to those of another. This is done for the dual purposes of:

(a) Providing vicarious calibration to instruments that have no or a defective internal calibration device (intercalibration being performed against a high-quality, well-calibrated instrument serving as a reference);

(b) Merging the data from several instruments to generate consistent time-series.

The intercalibration of instruments operated during the same period requires careful collocation wherein instrument outputs are compared when the instruments are viewing the same Earth scenes, at the same times and from the same viewing angles. As part of the International Satellite Cloud Climatology Project of the World Climate Research Programme, simultaneous...
observations from collocations between geostationary Earth orbit (GEO) imagers and a low Earth orbit reference imager have been performed on a monthly basis for almost 30 years as a means to normalize GEO satellite imagery. More recently, the Global Space-based Inter-calibration System (GSICS) has developed an operational methodology for such intercalibrations, specifically for simultaneous collocated observations. The methodology considers the trade-off between accurate spatial–temporal co-registration of the instruments and the frequency of such events, and takes into account the corrections to be applied for:

(a) Different viewing geometries (with regard to both the instrument scan angle and the solar position);

(b) Different atmospheric states in the line of sight, including aerosols and clouds;

(c) Different spectral response functions.

It should be noted that simultaneous observations between two Sun-synchronous satellites can only occur at the intersections of their orbital planes, which are always located at a given local solar time and at a given, generally high north or south, latitude.\(^\text{4}\)

### 6.1.7 Bias adjustment of long-term data records

An alternative approach for instrument intercalibration, which is less demanding in computation and applicable a posteriori to long data series, is to simply compare the statistical distribution of overlapping time-series of two satellite instrument data records without imposing individual matches of individual scenes. Using this approach, it is possible to identify the relative bias between the two data records. The observed bias is analysed so that the different conditions of observation (for example, different local solar times) are accounted for, leaving the remaining bias as the part that is actually due to the difference in instrument calibration. One successful example of this approach is the intercalibration of the nine microwave sounding units on board the early National Oceanic and Atmospheric Administration satellites, representing a 26-year record of global tropospheric temperatures.

### 6.1.8 Using calibration information

The type of calibration information available depends on the processing level and on the instrument considered. Each instrument has its own operating mode and calibration cycle, which includes regular measurements of calibration targets each time a certain number of observations are performed. For instance, the table in this section indicates the calibration cycles of the Advanced Microwave Sounding Unit – A (AMSU-A), the Microwave Humidity Sounder (MHS) and the High-resolution Infrared Sounder 4 (HIRS/4).

<table>
<thead>
<tr>
<th></th>
<th>AMSU-A</th>
<th>MHS</th>
<th>HIRS/4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Earth views</td>
<td>1 line of 30 pixels</td>
<td>1 line of 90 pixels</td>
<td>38 lines of 56 pixels</td>
</tr>
<tr>
<td>Number of warm target views</td>
<td>2 (~300 K)</td>
<td>4 (~273 K)</td>
<td>48 (~290 K)</td>
</tr>
<tr>
<td>Number of cold target views</td>
<td>2 (deep space ~2.73 K)</td>
<td>4 (deep space ~2.73 K)</td>
<td>56 (deep space ~2.73 K)</td>
</tr>
<tr>
<td>Overall duration of the cycle</td>
<td>8 s</td>
<td>8/3 s</td>
<td>256 s</td>
</tr>
</tbody>
</table>

\(^\text{4}\) For a 98° inclination, the crossing latitude is above 70° when the equatorial crossing times (ECT) of the two orbits differ by less than 8 h, and only drops significantly when the ECT difference increases towards 12 h.
An important step in the pre-processing from Level 0 to Level 1b data (see Part III, Chapter 2, 2.3.2.6) is to extract the calibration information in the form of warm/cold view counts and then to compute the resulting calibration coefficients in accordance with the calibration model (such as a linear or quadratic calibration function, or a lookup table) defined by the satellite operator for that particular instrument. This provides the operational calibration for that instrument.

For applications requiring high accuracy and consistency among different instrument data records, a correction can be applied on top of the operational calibration to take into account the latest results of the intercalibration activities. Such corrections are provided by GSICS. The corrected calibration coefficients may be included in the Level 1.b/Level 1.5 data formats as additional calibration information.

### 6.1.9 Traceability of space-based measurements

While intercalibration can ensure consistency between satellite instruments, it does not necessarily provide traceability to SI unless a reference instrument in orbit is SI-traceable. There are major challenges to achieving SI traceability in orbit as most sensors degrade physically during and after launch. Achieving SI traceability poses instrument design challenges and remains a research topic for all but a few measurement types.

The Climate Absolute Radiance and Refractivity Observatory (CLARREO) mission proposed by the United States National Research Council consists of a highly accurate infrared interferometer with a high-emissivity reference black body using multiple phase-change cells for SI-traceable thermometer calibration, an ultraviolet, visible and near-infrared spectrometer calibrated by Sun and moon views, a cryogenically cooled active cavity radiometer, and radio occultation measurements. This suite of instruments is intended to provide fully traceable measurements of the entire Earth-emitted and -reflected solar spectrum. Implementing and maintaining such a mission would provide an anchor point in support of the calibration and traceability of the whole fleet of operational radiometers.

For measurement traceability, one should take advantage of instruments that do not depend on radiometric calibration, such as radio occultation and Sun or star occultation sensors (see Part III, Chapter 2, 2.2.4.3 and 2.2.5.1).

### 6.2 PRODUCT VALIDATION

#### 6.2.1 Factors to be accounted for in validation

Validation is the process of assessing, by independent means, the quality of the data products derived from satellite instrument measurements. Product validation should be performed by product developers, downstream of instrument calibration, and should be documented in instrument-specific product validation plans. Guidelines for documenting product quality are provided in the Quality Assurance Framework for Earth Observation (QA4EO).

Geophysical products are generated from satellite data (often radiance measurements) by applying an algorithm that is either physically or empirically based. Comparing the retrieved products and their trends with in situ observations or model outputs is an important part of the process to assess and document the reliability of given retrieval algorithms and define their domain of applicability.

If a particular trend is detected, it may relate to the instrument’s performance; a careful analysis of the satellite instrument’s calibration and environmental data must be performed before any empirical correction can be applied.

---

5 From the CEOS Working Group on Calibration and Validation.
For many products, validation is a complex problem since the comparison between products derived from satellite measurements and independent reference products often from in situ measurements is subject to several errors: (i) an inherent satellite-derived product error, (ii) the error in the reference data, and (iii) the error introduced by the comparison methodology, often due to non-collocation in time and space. In general, different measurement techniques measure different things: a satellite observation usually refers to a relatively large area (the instantaneous field of view) and nearly-instantaneous measurements (within milliseconds); ground (in situ) measurements are generally very local and integrated over a relatively long time. Surface-based remote-sensing usually provides information representative of the atmospheric column. Comparison of the different types of measurements requires downscaling or upscaling methods that can introduce spatially or temporally dependent errors.

A validation assessment model can be used to improve comparisons by understanding and accounting for these differences and to better appreciate the advantages and disadvantages of different validation approaches. Validation campaigns run by satellite operators are usually accompanied by such an assessment model.

It should be noted that for certain satellite products, independent validation measurements may not exist, and validation can only be performed by evaluating the impact of the product when used in an application (for example, when assimilated in an NWP model).

### 6.2.2 Validation strategies

The validation of satellite-derived products should follow defined best-practice and variable-dependent protocols, such as those developed by the CEOS Working Group on Calibration and Validation. The validation of satellite-derived parameters and products can be carried out using the following sources:

(a) Surface-based in situ measurements;
(b) Surface-based remote-sensing measurements;
(c) Model comparison and assimilation;
(d) Other satellite-derived or blended products of a similar type.

To use such validation sources, it is essential that:

(a) Measurement errors be well known;
(b) Temporal and spatial sampling follow best-practice protocols;
(c) Sampling be representative of the typical application-dependent environment (e.g. climatic zones, marine regimes, atmospheric regions, land-cover types).

For example, to support validated generation of combined sea-surface temperature satellite products, the Group for High Resolution Sea Surface Temperature has developed a comprehensive validation strategy, which includes detailed descriptions of protocols, strategies to harmonize validation concepts for different satellite sensors contributing to sea-surface temperature measurements, needs for in situ (buoy) measurements as in situ data sources, and metrics to monitor product quality.

---

6.2.3 **Impact studies**

Experience shows that the results of direct validation are less significant for some satellite-derived products than for others. Since validation tests combine the effects of different error sources (satellite product, ground measurement, comparison method), the error due to the satellite product itself may be difficult to single out. For certain geophysical variables, ground measurements may be rather inaccurate. For others, the comparison method may depend too much on the observation environment.

One option for evaluating a product for a particular application is thus to assess its impact on the application skill. In this case, the evaluation reflects the quality of the product combined with the ability of the application to use it. For example, NWP models showed marginal impact for a couple of decades from atmospheric temperature–humidity soundings; this changed to a significant positive impact only when direct radiance assimilation was introduced. An opposite example is the assimilation of cloud-motion winds, which exhibited a strong positive impact at first, although the initial validation exercises were disappointing.

In summary, validation requires rigorous analysis of all error sources and of all steps in the comparison method. If the analysis shows that the error of the satellite product cannot be singled out, performing an impact study is a remaining validation mechanism.
**REFERENCES AND FURTHER READING**

