The forthcoming Upper-Air Instruments Intercomparaison (UAII2021) is organized under the auspices of the World Meteorological Organization (WMO) with the purpose to improve the quality of upper air observing systems and to develop the knowledge and expertise of national meteorological services (NMHS). The participating radiosonde systems (RS) are representative for those currently employed in the global observational network. A novelty with regard to previous RS intercomparison campaigns is to adopt principles and practices well established in the GCOS Reference Upper Air Network (GRUAN). This includes using GRUAN reference data products (GDPs) in the comparison of the RS. A distinguishing feature of a GDP is that the data are traceable to SI units. The data measured by a GDP are fully characterized in terms of their vertically-resolved uncertainty. This means that for each data point measured by the GDP at altitude z the obtained value V is represented as V(z) ± ΔV(z). Another novelty is the use of open source software for data analysis and visualization. The Data Visualization and Analysis Software (DVAS) is currently being developed and, as a basic feature, will use GDPs to establish the reference against which the performances of the participating RS are evaluated. This ensures a fair and transparent approach in the intercomparison of the RS. The DVAS uses a statistical combination of the available GDPs (two GDPs being the minimum number required) to act as a working standard (WS). Each data point measured by a GDP at the altitude z is assessed for consistency with the data points measured by the other GDPs and, if consistent, they are retained for calculation of the WS. In other words, the values V(z) ± ΔV(z) from the GDPs are evaluated and added together to yield the combined values for the measured parameter and its uncertainty. The evaluation of the values
$V(z) \pm \Delta V(z)$ consists of a consistency test, i.e. the values $V(z) \pm \Delta V(z)$ should lie within the uncertainties of the other GDPs. The details of the consistency test, including the case when the test fails, will be provided in the presentation of the DVAS.

In addition to the evaluation of the RS performances in terms of their mean bias and variability, each RS is assessed for its ability to be fit for purpose with respect to different application areas based on the Observing Systems Capability Analysis and Review Tool (OSCAR table of requirements, https://www.wmo-sat.info/oscar/). For example, a RS can be affected by a too-large bias/variability with respect to the reference WS for applications in the area of high-resolution numerical modelling, but can be fit for the purposes of global modelling.