Radio Occultation Modeling Experiment (ROMEX) Framework

Final version 3 August 2023 from original by Martin McHugh, modified by Hui Shao, Richard Anthes, Ben Ruston, and Christian Marquardt, with input from ROMEX members. Updates and other ROMEX news available to ROMEX Working Group at https://drive.google.com/drive/folders/16f8wWWjbFPFx5eLwOY2baZ4MKdRZvY6z

1. Overview and Purpose

The international Radio Occultation (RO) community has recently and collectively proposed to undertake a collaborative effort to explore the impact of a large number of RO observations on numerical weather prediction (NWP). This effort was first proposed by Richard Anthes in May 2022 in response to questions from by NOAA for input on future RO needs The resulting discussion led to a proposal for the Radio Occultation Modeling Experiment (ROMEX), which was endorsed by the IROWG in September, 2022 (IROWG-9).

The experiment seeks to answer some of the more pressing technical and programmatic questions facing the community and help inform the near- and long-term strategies for RO missions and acquisitions by NOAA, EUMETSAT, and other CGMS partners. Most important among these questions is to quantify the benefit of increasing the quantity of RO observations. These will be addressed by performing data assimilation experiments with real measurements. The purpose of this framework document is to describe the tasks and efforts needed to conduct the experiment and outline a proposed plan, schedule, and responsibilities for successfully executing the project.

ROMEX is envisioned to consist of at least two three-month periods during which all available RO data are collected, processed, archived, and made available to the global community free of charge for research and testing. Although the primary purpose is to test the impact of varying numbers of RO observations on NWP, the three-months of RO observations will be a rich data set for research on many atmospheric phenomena. The first ROMEX period (ROMEX-1) will be September through November 2022, which contains a number of tropical cyclones that can be studied. Although ROMEX-1 data will be restricted to neutral atmosphere data, if successful a broadened ROMEX that will also contain space weather data is also envisioned.

2. Background and Motivation

Historically RO has played an increasingly important role in atmospheric and environmental remote sensing. The direct application of RO observations to weather prediction began with the GPS/MET mission in 1995. In the decades that followed, there have been a plethora of dedicated RO satellite missions and hosted instruments, whose data are used for weather prediction and other applications. Examples include CHAMP, GRACE/GRACE-FO, SAC-C/D, TerraSAR-X, TanDEM-X, COSMIC-1/2, METOP-A/B/C, FengYun-3C/D/E, C/NOFS, KOMPSAT-5, and Paz. In 2020 Spire provided some free data to ECMWF for testing, and these additional RO data helped compensate for the loss of aircraft data during the pandemic. In 2021, commercial vendors began supplying operational RO data to NOAA and EUMETSAT for charge, and now contribute a significant portion of the total number of RO observations being used and shared by the international community.

The primary, but by no means only, application of RO is in NWP. RO is sensitive to temperature, pressure and specific humidity within the troposphere and lower stratosphere. This, and the highly calibrated nature of the measurement technique, makes RO well suited for NWP applications. Raw RO measurements are processed into quasi-vertical atmospheric profiles, and these profiles now form a significant part of the NWP portfolio, as illustrated in Figure 1. Other important applications include climate studies and reanalysis, as well as ionospheric observations for space weather applications, where RO provides much needed near-real-time measurements of the ionospheric state, including total electron content and scintillation events.

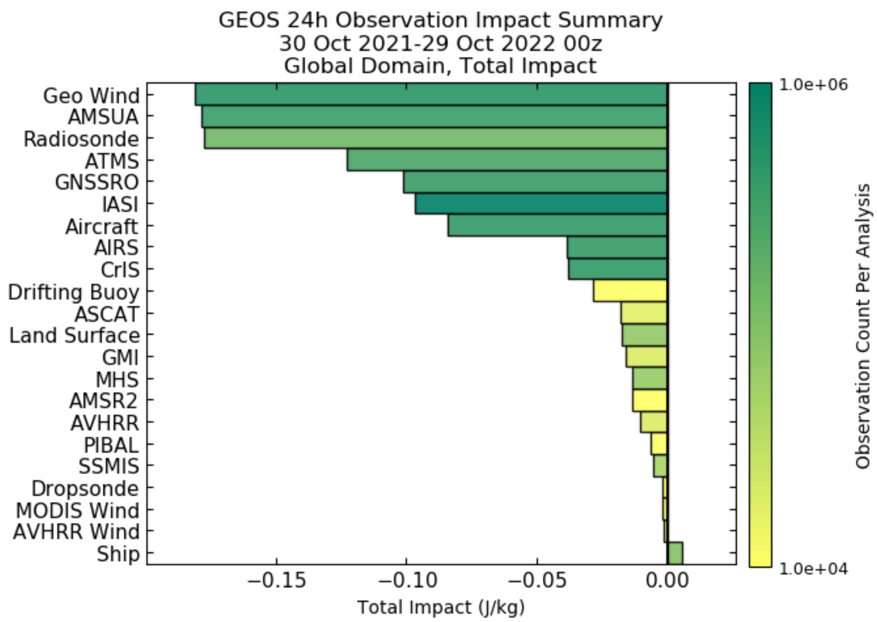

Figure 1. Example of impact results of various observation types, from NASA GEOS-5. This depicts a global measure of 24-h forecast error that combines errors in wind, temperature, specific humidity, and surface pressure with respect to the verifying analysis from the surface to 1 hPa in terms of moist total energy (J/kg). Observation impact is taken to be the difference in this error measure between 24-h forecasts initialized from the analysis and corresponding background state, where this difference is due entirely to the assimilation of the observations. (Courtesy NASA GMAO: https://gmao.gsfc.nasa.gov/forecasts/systems/fp/obs_impact)

The current strong impact of RO is in part due to the increase in the number of observations available for assimilation. Several studies (e.g., Harnish and Healy (2013), Privé et al. (2022)) used simulations that indicated the beneficial impact to NWP will increase steadily with additional RO observations up to at least 125 thousand profiles per day. Such studies prompted the International RO Working Group (IROWG) in 2015 to recommend acquiring a baseline of at least 20,000 RO profiles per day with full global and local-time coverage for NWP. Lonitz et al. (2021) showed empirically with RO profiles collected by COSMIC-2, Spire, and others that large positive impacts were attained on nearly all forecast metrics for quantities up to at least 10,000 profiles/day. Others have suggested that 30,000 profiles per day may be closer to a “knee” in the benefit vs number curve. The related cost-benefit curve will depend on the price per profile.

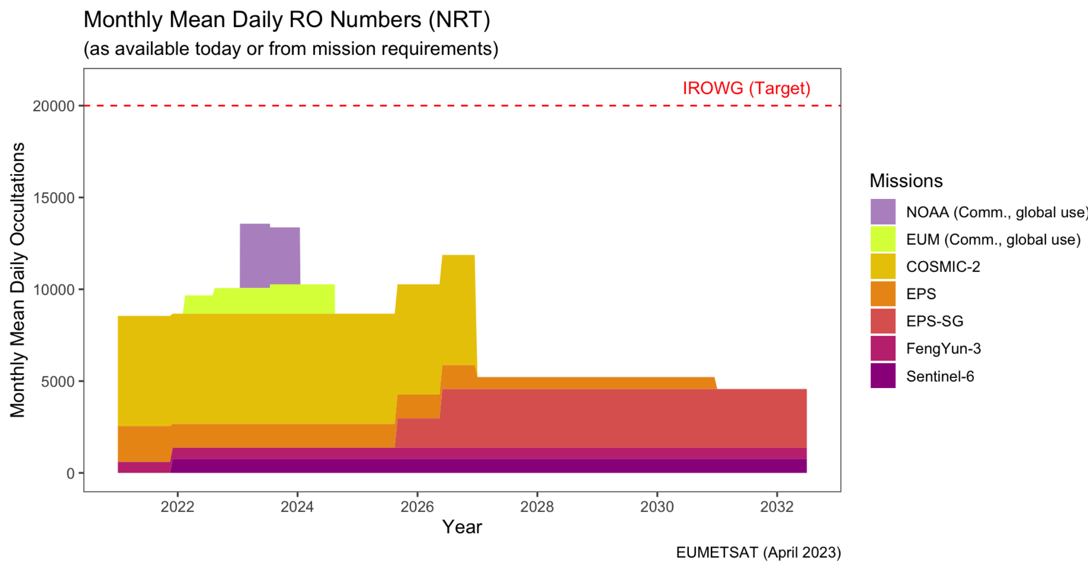

Figure 2. Projected daily number of RO occultations (C. Marquardt, EUMETSAT). This is based on the requirements for the listed set of ongoing and known upcoming missions and their planned lifetimes. Commercial data sources are not projected forward beyond 2024.

Currently NOAA assimilates about 12,000 RO profiles per day into its Global Forecast System, with approximately half coming from commercial sources. The acquisition of commercial RO data began with pilot studies in 2016 and progressed to a fully operational data program in 2021. EUMETSAT has also begun purchasing and distributing commercial RO data to supplement the government-sponsored data sources with about 10,000 profiles per day. Figure 2 shows the near real time monthly mean daily RO numbers available as of April 2023 or from mission requirements according to EUMETSAT. The numbers plotted here are only for open-accessed data. Therefore, even though there is a significant amount of commercial data available for purchase and future acquisitions are now being planned, the number acquired from commercial vendors now is limited. Figure 2 shows a clear shortfall from the recommended 20,000 profile baseline. This situation will become even worse in the next 10 years, since COSMIC-2 and Metop-B are expected to lose their data and the CGMS partners have no plans yet for government-sponsored RO missions (Fig. 3). Meanwhile, the commercial supplemental data continues to pose uncertainty to the future in terms of open data access and long-term stability.

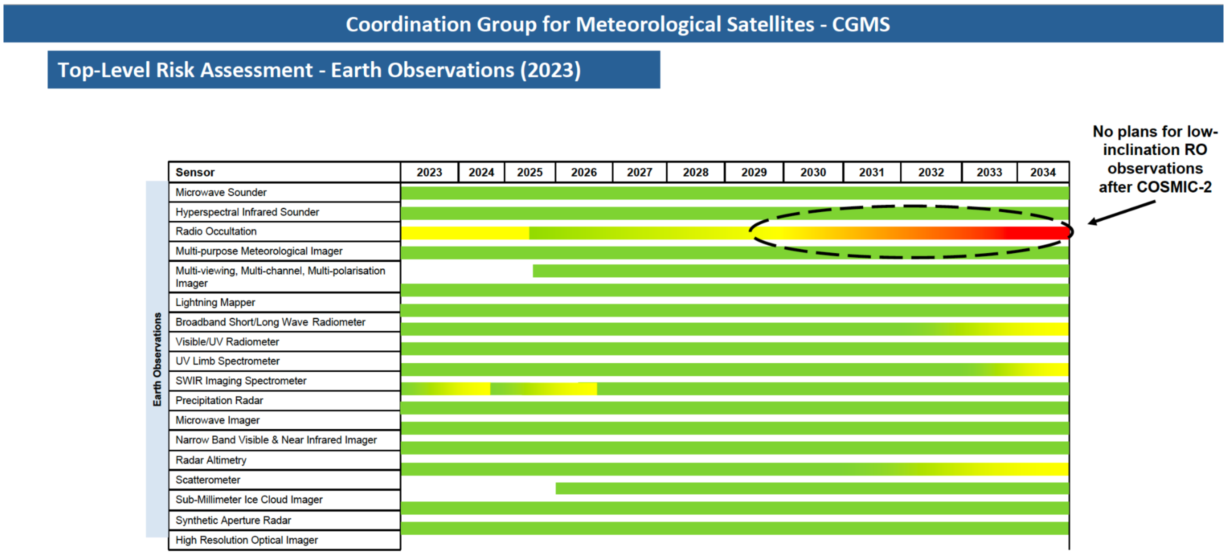

Figure 3: Top level risk assessment of Earth Observations. The risk is highest for radio occultation observations. (A. Mehta, CGMS WGIII report, 2023)

3. The RO Modeling Experiment (ROMEX)

The Radio Occultation Modeling Experiment (ROMEX), will assimilate historical RO profiles into NWP systems to quantify the impact of large quantities of profiles. Previous studies examining the impact of large quantities of RO data have had to rely on simulated data and theoretical estimates of the impact. A rough estimate shows that there are potentially 30 thousand profiles per day available for examination, and this will constitute the primary goal of the experiment. The ROMEX concept was presented to NOAA, EUMETSAT, and other international partners at the IROWG-9 meeting in September 2022. The IROWG community has gained approval from their respective institutions to perform data assimilation experiments with the additional RO measurements over the designated time period.

3.1. Questions to be addressed by ROMEX

High level questions about RO data fall into the following general categories.

1) Data Quantity

How many RO profiles per day are needed? Is there a “knee” or asymptote in the benefit vs. number curve using real data? With roughly 30 thousand profiles per day being collected today, the experiment can measure improvements up to that daily number.

2) Geographic / Temporal Sampling

How should RO observations be distributed around the globe and in local time? Should the tropics or other specific latitude regions receive relatively more focus? How important is local-time sampling to NWP and climate applications?

3) Data Quality

Are there systematic differences in the GNSS-RO sources and their processing algorithms? Issues include the value of higher SNR, latency, differences between GNSS sources and differences between processing algorithms. Beyond these lie questions about how to best exploit the various quality aspects of the data such as penetration depth. Finally, the issue of how to quantify these quality aspects into data requirements is an important topic, especially for commercial purchases.

4) Make vs. Buy

How much can CGMS partners rely on commercial services to provide operational data? Issues include total cost per observation, control of measurements, restrictions on data, robustness of system and long-term stability, level of transparency, and value of additional capabilities (e.g., space weather data and neutral atmosphere observations combined).

While a scientific study such as ROMEX cannot fully answer high level policy questions, it can identify the salient points and provide the technical underpinning to the decision makers. Note that the IROWG recommends that RO be regarded as a "core" data type per WMO Resolution 1, and that WMO and CGMS coordinate on any purchases of commercial RO data. The IROWG also strongly recommends that all RO data be purchased with a world-use license (all data freely available to everyone with no restrictions on use). The ROMEX results will provide additional scientific foundation for these recommendations and help the community formulate more specific recommendations moving forward.

In order to provide actionable answers in a reasonable timeframe, it is suggested that this initial ROMEX effort, designated as ROMEX-1, only focus on questions from the first and maybe a few more categories, potentially follow-on experiments can address the remainder. At the conclusion of ROMEX-1, a follow-on ROMEXs can be initiated to address remaining and emerging questions at that time.

3.2 ROMEX Project Plan:

3.2.1 First Steps

Prior to kicking off the project in earnest, some upfront planning is necessary to map out the scope, roles, activities, and overall schedule. The following steps (Fig. 4) should be taken more or less sequentially. A more detailed plan can be developed after these steps are complete.

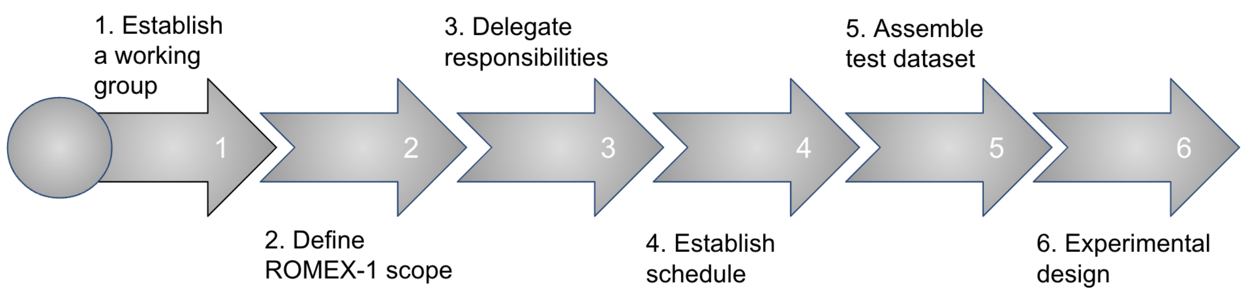

Figure 4. Workflow for the initial stage of ROMEX Project

1) ROMEX Steering Committee (RSC)

A ROMEX Steering Committee (RSC) has been established. Questions or requests for information can be addressed to any of the RSC members.

The members are:

Richard Anthes	anthes@ucar.edu
Christian Marquardt	christian.marquardt@eumetsat.int
Benjamin Ruston (Chair)	benr@ucar.edu
Hui Shao	huishao@ucar.edu

2) ROMEX Working Group (RWG) Representatives from all the major stakeholder institutions are represented. This group includes representatives with expertise spanning meteorology, space weather, RO technology, and climate science, as well as the potential providers of RO observations. The RWG will collectively decide on recommended actions to be taken throughout ROMEX-1. (These will be recommendations as opposed to directions, as the members are participating on a voluntary basis.)

3) ROMEX-1 Scope

The list of questions collected above constitutes just a sampling of the many research questions surrounding RO, but the focus should remain on answering the most pressing issues—those that impact the policy issues faced by the major institutions responsible for RO data acquisitions. To that end, a subset of the questions listed above should be defined as the scope of ROMEX-1, and this scope delineation should be the first output of the RSC.

4) Responsibilities

Because of the magnitude of the processing required for such an endeavor, there are only a few facilities with the capabilities to perform such a computationally intense experiment. Ways of segmenting or dividing the experiment should be discussed, and a plan for which centers will perform the data assimilations determined.

Analysis will be performed by individuals from the respective organizations listed above, and any others that may be included. While each may have their predilections for what studies they would like to conduct, the RSC and RWG should ensure that the studies conducted are sufficient and comprehensive enough to answer the primary questions defined in the scope of ROMEX-1.

5) Schedule

The overall ROMEX-1 project should be completed within two years to provide the guidance needed by NOAA, EUMETSAT and other CGMS organizations in planning their future observation architectures. Such a timeframe will accommodate the time needed for disseminating data and performing experiments. On the other hand, a time frame is also needed to provide prompt scientific evidence and basis for the IROWG community and CGMS partners to plan and coordinate satellite missions. Figures 2 and 3 clearly show the need for this. For example, with an end-of-life of COSMIC-2 and one of the EPS satellites as early as 2027, planning for follow-on missions and or acquisitions must proceed by 2024 at the latest. Accordingly, a timetable was proposed as shown in Appendix 1, with the final report expected during the next IROWG workshop scheduled 12-18 September 2024 in Boulder, Colorado.

6) Assemble Test Dataset

The number and length of the ROMEX test periods should be decided by the RSC and RWG. The first three-month period (ROMEX-1) will be September-November 2022, which overlaps the Northern Hemisphere tropical cyclone season. A second period during the Northern Hemisphere winter will be considered if the first period is successful. Constraining the number of intensive data periods will be the HPC resources available to conduct various data assimilation experiments. The data will need to be available to all scientists and centers worldwide without restriction. The data assembling process will follow the terms and conditions described in the Terms and Agreements section.

7) Experimental design

Once these steps are completed, the work on specific analyses can begin. At the forefront will be the major data assimilation and impact assessments, but there will likely be many additional studies and analyses that will kick off as soon as the data are available. The RWG meetings will be used to solicit, organize, and track the analyses underway to ensure the best coverage of the major questions enumerated above.

3.2.2 ROMEX Terms and Agreement

EUMETSAT and ROMSAF are taking the lead to collect, process (if necessary), store, and redistribute the ROMEX data. Therefore an agreement with the data providers will be signed and managed by EUMETSAT. Consequently, the rest of the ROMEX community will be required to sign a data license with EUMETSAT for the terms and conditions of using ROMEX data. Once EUMETSAT performs security checks, reformats and processes (if needed) the data, the observations (in BUFR format) will be made available to the ROMEX community through ROMSAF. Additional data processing centers may also receive the lower level data from EUMETSAT under mutual agreement. A draft data workflow is depicted in Figure 5.

ROMEX Framework

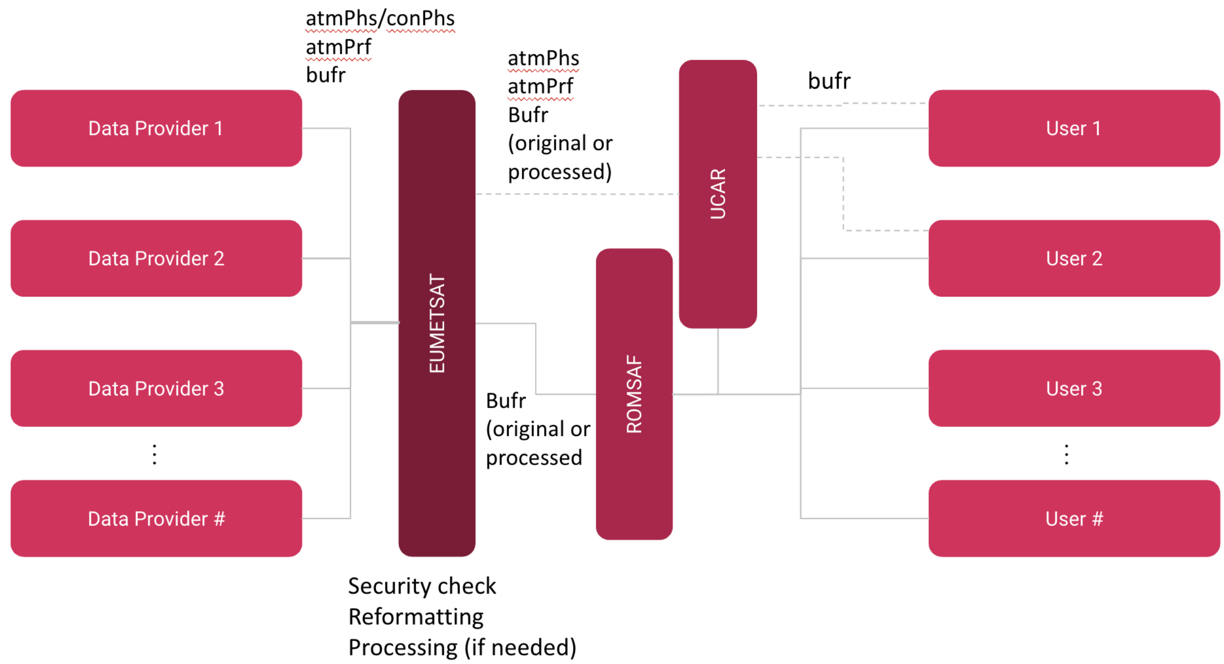

Figure 5. Workflow for ROMEX data handling and sharing.

Appendix 1:

Meeting schedule and timeline

Date (specific dates will be assigned once it is determined)	Task
29 November 2022	Kick-off meeting: introduction to ROMEX. Sign up sheet for ROMEX participants
19-20 December 2022	Meeting. Set up subgroup discussions: Processing and Experiment Design
22-23 February 2023	Meeting. Status update
January - May 2023	Data agreement and license document. Meetings with individual data providers
February - June 2023	Data collection and assemble
May 2023	RWG Meeting. Status update. Experimental Design
June 2023	Data starts flowing to EUMETSAT. RWG meeting.
3 August 2023	ROMEX white paper
September 2023	Initial data assessment begins
October 2023	All data available.
December 2023	RWG meeting to discuss initial results and identify any issues
April 2024	Workshop in Darmstadt (EUMETSAT)
12-18 September 2024	IROWG-10 in Boulder
December 2024	Comprehensive report of ROMEX-1 results

Appendix 2: Experimental Design Proposal

Baseline/Control: Align with operational configuration. All observing systems at normal usage levels and methodology.

- Remove all RO data not commonly used at all centers (e.g. GNOS; NOAA or EUMETSAT commercial purchase data)

ROMEX experiment: same as above.

- Add all available ROMEX data

Table 1 Estimated ROMEX Data Volume (updated 3 August 2023)			
Mission	RO/day	Control	Data Source
Metop B, C (GRAS)	1,200	y	EUM
COSMIC-2	6,000	y	UCAR
Spire	17,000	n	UCAR
FY3-C, D, E (GNOS)	2,100	n	CMA
PlanetiQ	3,300	n	PlanetIQ (commercial)
Yunyao	6,200	n	Yunyao (commercial)
Binhu (GNOS-mini)	100	n	Binhu (commercial)
KOMPSAT-5	300	y	UCAR
PAZ	200	y	UCAR
TerraSAR-X	100	y	UCAR
TanDEM-X	100	y	UCAR
Sentinel-6	800	y	EUM
Sum Control	**8700**	**y**	
Sum Supplemental	**28,700**	**n**	

ROMEX Grand total	**37,400**	**n**	

Experiment Impact and Verification

The control and ROMEX experiment will provide metrics for intercomparison.

Forecasts statistics bias and RMS statistics for air temperature, moisture, and wind at a series of standard pressure levels. These forecasts will range from 5-10 days.

Accumulated bias and RMS statistics on the observation fits particularly to radiosonde, and microwave and infrared sounders.

Additional diagnostics are encouraged. Systems which can readily provide an analysis of ensemble spread and/or forecast sensitivity to observation are welcome to include these results.