



STARFIRE-2: Can We Detect the Global Redshifted 21 cm Signal from the Cosmic Dawn in Earth Orbit?

Yogen Pranesh^{1,2}, Mayuri Sathyanarayana Rao² , and Saurabh Singh²

¹Department of Physics, University of Rome Tor Vergata, 00133 Rome, Italy

²Raman Research Institute, C. V. Raman Avenue, Sadashivanagar, Bangalore 560080, India

Received 2025 December 9; revised 2026 April 28; accepted 2026 May 19; published 2026 June 24

Abstract

Detecting the redshifted global 21 cm signal from the cosmic dawn remains a major challenge due to strong terrestrial radio-frequency interference (RFI), particularly dominated by frequency modulation (FM) transmissions in the 88–110 MHz range. While observations from the radio-quiet lunar farside are ideal, Earth orbit offers an intermediate and simpler alternative that may mitigate several limitations of ground-based experiments. We assess the feasibility of detecting the global 21 cm signal from Earth orbit by quantifying FM-based RFI at different altitudes and orbital configurations. We present Simulation of Terrestrial Radio Frequency Interference in oRbits around Earth—2 (STARFIRE-2), an algorithm that estimates FM transmitter-based RFI intercepted by radiometers in orbit. The model constructs a global FM transmitter database and compensates for incomplete data using statistical methods. Using PRATUSH as the reference experiment, we simulate a range of orbital scenarios to identify configurations that minimize RFI and optimize sensitivity for global 21 cm detection. The algorithm can also be adapted for other experiments. Simulations indicate that conducting such an experiment from Earth orbit is feasible for a thermal noise limited instrument placed in a low Earth, near-polar orbit. Mock sky observations further demonstrate that most theoretically plausible cosmic dawn 21 cm signals can be recovered with high confidence under these optimized orbital conditions.

Unified Astronomy Thesaurus concepts: [Cosmology \(343\)](#); [Radio astronomy \(1338\)](#); [Space telescopes \(1547\)](#); [Reionization \(1383\)](#)

1. Introduction

It has long been predicted that the astrophysics and cosmology of the dark ages, cosmic dawn (CD), and epoch of reionization (EoR) can be studied using the redshifted 21 cm signal from hydrogen (J. R. Pritchard & A. Loeb 2012), either as a monopole global signal or as multipoint correlation functions such as power spectrum and beyond. Throughout this paper, we use the terms *monopole*, *sky-averaged signal*, and *global signal* interchangeably. Since then, several ground-based experiments have searched for the faint global signal (G. Bernardi et al. 2015; S. Singh et al. 2017, 2018a, 2018b, 2022; J. Bowman et al. 2018; L. Philip et al. 2018; D. C. Price et al. 2018; E. de Lera Acedo et al. 2022); still a confident detection remains elusive. Ground-based experiments aiming to detect the redshifted global 21 cm signal from CD/EoR encounter numerous challenges. For instance, chromatic effects induced in the measured sky spectrum from ionospheric refraction, reflection, and absorption can limit foreground separation and hence signal detection (H. K. Vedantham et al. 2014; A. Datta et al. 2016; E. Shen et al. 2021). Also, antenna properties depend critically on the operating environment, including but not limited to ground-plane effects and the dielectric constant of the near field, which vary in ways that are complex and hence difficult to model (R. F. Bradley et al. 2019; M. Spinelli et al. 2022; Y. Agrawal et al. 2024). Finally, terrestrial radio-frequency interference (RFI) presents a significant challenge for CD/EoR

experiments (A. R. Offringa et al. 2015; M. J. Wilensky et al. 2019). This issue is particularly problematic in the frequency modulation (FM) band (88–110 MHz), which has significant overlap with the frequency range of CD observations. Several experiments have been proposed to measure the low-frequency radio sky spectrum in the pristine radio environment of the lunar farside (J. O. Burns et al. 2012, 2019; X. Chen et al. 2020). Probing Reionization of the Universe using Signal from Hydrogen (PRATUSH) is one such proposed experiment from India that aims to detect the CD/EoR signal from the lunar farside (M. Sathyanarayana Rao et al. 2023).

Before heading to the Moon, one can study the feasibility of conducting this experiment in free space, far away from the surface of the Earth, to alleviate challenges posed by the ionosphere and medium/environment-induced antenna chromaticity. An experiment in Earth orbit as a precursor to a lunar orbiter provides an intermediate option with lower logistic and economic complexity, with more opportunities for launch and technological readiness and sensitivity. Whether such an experiment can generate competitive results depends critically on the RFI environment in Earth orbit, thus motivating this work. Since an experiment in Earth orbit will have a field of view (FOV) of Earth proportional to its altitude, a space-based radio telescope will inevitably capture a larger number of RFI sources within its beam than perhaps one observing at a radio-quiet location on Earth. However, it will certainly be devoid of the other two challenges faced by ground-based experiments, namely ionospheric effects and antenna coupling to the surrounding medium. In a previous paper we presented the Simulation of Terrestrial Radio Frequency Interference in oRbits around Earth (STARFIRE), an algorithm for estimating FM transmitter-based RFI at



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

different Earth orbits (S. Ghosh et al. 2023). While the algorithm proved effective, the user-provided FM transmitter information in the database used in STARFIRE was sparse and hence the final results were of limited application. Recognizing the need for more complete data, we propose a method for building a statistically complete FM transmitter database and conducting forecast studies for PRATUSH in Earth orbit, henceforth referred to as PRATUSH-I. We take this effort one step further, to quantify the feasibility of conducting a CD/EoR experiment in Earth orbit by examining the signal detection prospects predicted by standard models.

In Section 2, we build a complete FM transmitter database and account for incomplete information by statistically modeling various parameters and generating synthetic data, followed by Section 3, where we conduct various orbit simulations to estimate the amount of RFI intercepted by PRATUSH-I throughout its orbital lifetime. The radio-quiet zones in the orbit would thereby dictate the feasibility of conducting such an experiment in Earth orbit. Section 4 presents simulations of sky observations by PRATUSH at different points in its orbit path and quantify the sensitivity of the radiometer to various standard CD models. We discuss sources of uncertainty in Section 5 and present conclusions in Section 6.

2. Building a Complete FM Transmitter Database

The STARFIRE algorithm uses as its input an FM transmitter database that includes each FM transmitter’s geodetic coordinates, its frequency of transmission, and its effective isotropic radiated power (EIRP). Folding in instrument properties such as antenna beam, frequency span, and resolution, it generates an RFI spectral cube for user-defined satellite altitudes. The algorithm uses the HEALPix pixel scheme (K. M. Gorski et al. 2005) in Python using the “Healpy” library (A. Zonca et al. 2019). Accurate RFI estimation for a receiver in Earth orbit using STARFIRE critically depends on the completeness of the input RFI source database, which in this case is the FM transmitter database. Such a database has been populated using FMLIST,³ a publicly available worldwide FM transmitter database that sources its data from government websites and HAM radio operators. This open-access country-wise FM transmitter information was processed to match the specifications required by the STARFIRE algorithm and added to the database.

2.1. Incomplete Datasets for Select Countries

FM transmitter information on the FMLIST interface is displayed in ascending order of frequency of transmission. One can access information for up to the first 5000 transmitters per country in the free version of FMLIST. The interface also displays the total number of FM transmitters in the country. Thus transmitter datasets of countries with more than 5000 transmitters (see Table 1) are incomplete in the free version of FMLIST. This can result in underestimating the RFI by the STARFIRE algorithm. We build a statistical model of FM transmitters to keep the database flexible and future-proof (such as FM transmitters coming up or going down in the future). This statistical approach not only accounts for the

Table 1

Number of Missing Transmitters for Countries with Incomplete Databases

Country	Number of Missing Transmitters
Russia	10,508
China	4333
Argentina	2092
Brazil	4201
Spain	1340
France	4406

remaining data but also adds scalability to the STARFIRE input database, providing flexibility to replace FMLIST with other transmitter datasets if needed. Here we note that using a statistical model to account for missing transmitters instead of exact transmitter data does not alter the output of the STARFIRE algorithm significantly. It is unlikely to convert a radio-quiet location or channel into a radio-loud one (or vice versa), because any additional statistically generated transmitters would primarily appear in countries that already exceed 5000 registered transmitters—regions that are inherently high RFI to begin with. To account for the “missing” transmitters, we populate synthetic transmitters by generating the location, frequency, and EIRP parameters per transmitter using the statistical methods detailed as follows. We note that though the USA has more than 5000 FM transmitters, we have full data available from the Federal Communications Commission (FCC).⁴ This dataset was used as a reference dataset for testing the effectiveness of our statistical methods used to build a complete database.

2.1.1. Modeling FM Transmitter Location Parameter

The area of the geometric footprint of a satellite is given by

$$A = 2\pi R_e^2 \left(1 - \frac{R_e}{R_e + h} \right), \quad (1)$$

where R_e is the radius of the Earth and h is the altitude of the satellite above the mean sea level. For a satellite in low Earth orbit (LEO; ~ 410 km altitude), the geometric footprint, or FOV, is expected to cover 15 million square kilometers, making its footprint the smallest among all Earth orbits (LEOs, medium Earth orbits, and geostationary Earth orbits). This sets the requirement of the finest spatial resolution of the pixel when placing synthetic transmitters in the statistical model. As a result, the accuracy of synthetic transmitter locations, when compared to their real counterparts, is not highly constrained.

Since the location parameter of a transmitter is bivariate (latitude and longitude), a normalized bivariate probability distribution function was created using a two-dimensional histogram of transmitter latitude and longitude using the `histogramdd` function of the NumPy Python library (C. R. Harris et al. 2020). We assume that the 2D distribution, created out of 5000 points, is a good representation of the full distribution of the transmitters in the country. A transmitter location corresponds to a coordinate pair within the geopolitical boundary of the country it is located in. Hence sampling from the spatial distribution must be done using ordered pairs that also follow the geometry of the country corresponding to

³ <https://www.FMLIST.org>

⁴ <https://www.fcc.gov/media/radio/fm-query>

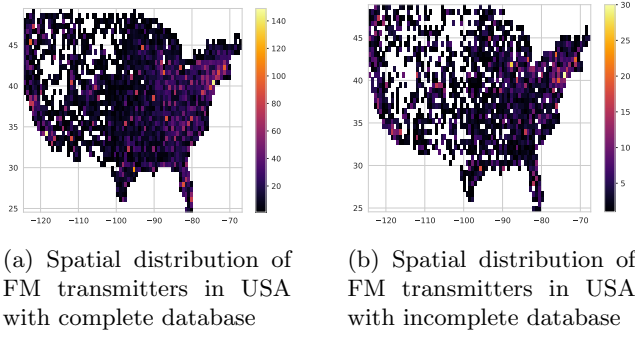


Figure 1. Spatial transmitter distribution comparison.

its geopolitical boundary. A HEALPix pixel located within a country represents an ordered pair, as it corresponds to a specific longitude and latitude value.

The accuracy of generating locations of synthetic FM transmitters using the prior spatial probability distribution functions strongly depends on the bin size. Since the bin width dictates the angular resolution of sampling for the probability distribution, a large bin width would not capture the intrinsic spatial variation in the distribution. In contrast, a very small bin width would tend to overfit the distribution by capturing random fluctuations (K. H. Knuth 2006). The optimal bin width was chosen by minimizing the “cost function” across various selections of bin widths, as directed by H. Shimazaki & S. Shinomoto (2007). The cost function is defined as

$$C(\Delta) = \frac{2k - v}{\Delta^2}, \quad (2)$$

where k and v are the mean and standard deviation of the number of variables per bin, and Δ is the selected bin width for the histogram. It is feasible to generate synthetic transmitter locations using the incomplete dataset transmitters’ spatial distribution as a 2D probability distribution function.

One can generate the (latitude, longitude) ordered pairs of coordinates within a country at a chosen angular resolution by obtaining the “shapefile” of the country. A shapefile is a geospatial vector data format mainly used in geographic information systems (GIS; K. Xia & C. Wei 2008). Such files are useful for obtaining information about country geometry, described using polygons and multipolygons. A shapefile of each country with an incomplete FM transmitter database was obtained from GADM.⁵ Since HEALPix works with raster data, the shapefile was converted into a high-resolution raster using the “rasterize” function in the Quantum Geographic Information System (QGIS; <https://www.qgis.org>) and converted into a HEALPix map, thereby obtaining all the HEALPix pixels within that country of selected N_{Side} resolution. The 2D spatial distribution of the incomplete dataset was then projected onto the HEALPix maps using a standard multidimensional interpolation scheme.

Synthetic transmitter locations were then generated by randomly sampling pixel positions within each country according to its normalized spatial probability distribution. The number of sampled points corresponded to the transmitter deficit for that country.

The statistical model was tested for the incomplete USA dataset against the already obtained complete USA dataset as

Table 2

Performance of the Model for Different N_{Side} HEALPix Maps

N_{Side}	Accuracy	Precision	Recall
32	0.83	0.93	0.86
64	0.80	0.66	0.66
128	0.80	0.48	0.37
256	0.87	0.19	0.28
512	0.94	0.06	0.15

Note. We indicate the bold values to be optimal, which corresponds to N_{side} 64 in the table.

mentioned in Section 2.1 (see Figure 1). The synthetic transmitter HEALPix map and the original transmitter HEALPix map were binarized—meaning any pixel having at least one FM transmitter would have a value of 1, while any pixel having no FM transmitters would have a value of 0. Performance of the model was measured using accuracy, precision, and recall (see Equation (3); S. Gagnon-Hartman et al. 2021),

$$\begin{aligned} \text{Accuracy} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \\ \text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\ \text{Recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \end{aligned} \quad (3)$$

where TP, TN, FP, and FN stand for true positive, true negative, false positive, and false negative, respectively. The performance of the model with different sampling resolutions were tested by varying the N_{Side} of the maps that were being sampled.

As previously mentioned, a low sampling resolution would result in underfitting, and a high sampling resolution would result in overfitting, forcing synthetic transmitters to be placed at the same location as the sample. Hence an optimal sampling resolution would be a value of N_{Side} that lies between these two extrema and has the best performance metrics. Looking at the performance of the model as seen in Table 2, a sampling resolution of N_{Side} 32 (and subsequently lower N_{Side} values) seems to have the best overall performance. However, this is attributed to the fact that a low-resolution binarized FM transmitter map will mostly have pixels with the value 1. This leads to a drastic increase in the number of TPs thus falsely returning a very high performance score (see Equation (3)), even if the synthetic transmitter locations did not reflect real locations with accuracy. Thus N_{Side} 64 was selected to be the optimal resolution for sampling, as it had the best verifiable performance metrics while having an intermediate spatial resolution.

2.1.2. Modeling FM Transmitter Frequency Parameter

The frequency of the synthetic transmitter will be higher than the last recorded frequency in the incomplete transmitter dataset, as FMLIST presents data in ascending order of frequency. Therefore, there were no existing spatially dependent distribution functions to sample for the missing frequencies. Hence frequency parameter was assumed to be spatially independent, thus making it a univariate distribution. We use the PRATUSH window of 55–110 MHz to sample synthetic transmitter frequency values.

⁵ <https://GADM.org>

New frequency values were generated using “reverse-interpolation” sampling. Reverse-interpolation sampling is done by creating a normalized cumulative distribution function (CDF) for the univariate distribution. New samples can be drawn from the normalized CDF by reverse interpolating a uniform distribution ranging from 0 to 1 with the CDF to obtain the corresponding values of the variable. We use the Freedman–Diaconis rule to bin frequency values (D. Freedman & P. Diaconis 1981) given by Equation (4),

$$h = \frac{2 \cdot IQR(x)}{\sqrt[3]{n}}, \quad (4)$$

where h is the bin width, $IQR(x)$ is the interquartile range of the data, and n is the number of data points. The frequencies were randomly sampled and assigned to each synthetic transmitter independent of their location parameter.

2.1.3. Modeling FM Transmitter EIRP Parameter

The spatial dependence of FM transmitter EIRP can be measured using the “Local Moran’s I test” for spatial autocorrelation (P. A. P. Moran 1950) given by Equation (5),

$$I_i = \frac{(x_i - \bar{x})}{S^2} \sum_{j=1}^n w_{ij}(x_j - \bar{x}), \quad (5)$$

where I_i is the local Moran’s I value at observation i , w_{ij} is the spatial weighting factor between observation i and nearest j neighbors, x_i is the value of the transmitter EIRP, x_j is the transmitter EIRP value at location j , $\bar{x}$ is the average transmitter EIRP in that country, and S^2 is the global transmitter EIRP variance given by

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N}. \quad (6)$$

Since the local Moran’s I is an inferential statistic, we compute the statistical significance of the obtained local Moran’s I value with hypothesis testing. We form a null hypothesis stating that the transmitter EIRP values are *not* spatially correlated at i th observation with j neighbors. We then infer the spatial dependence of transmitter EIRP distribution with the obtained Z-score given by Equation (7),

$$Z_i = \frac{I_i - \mathbb{E}[I_i]}{\sqrt{\text{Var}(I_i)}}, \quad (7)$$

where Z_i is the Z-score at observation i ; I_i is the local Moran’s I value obtained at observation i ; and $\mathbb{E}[I_i]$ is the expected local Moran’s I value at i , given by Equation (8):

$$\mathbb{E}[I_i] = -\frac{\sum_{j=1}^n w_{ij}}{n-1}. \quad (8)$$

Here n is the number of transmitters in the incomplete dataset and $\text{Var}(I_i)$ is the variance of the local Moran’s I under the null hypothesis and is given by

$$\text{Var}(I_i) = \mathbb{E}[I_i^2] - \mathbb{E}[I_i]^2. \quad (9)$$

The local Moran’s I was computed for each country with an incomplete dataset for the transmitter EIRP’s spatial autocorrelation locally, along with the statistical significance of the value in QGIS. The nature of spatial autocorrelation of the transmitter EIRP can be analyzed by categorizing the local Moran’s I value

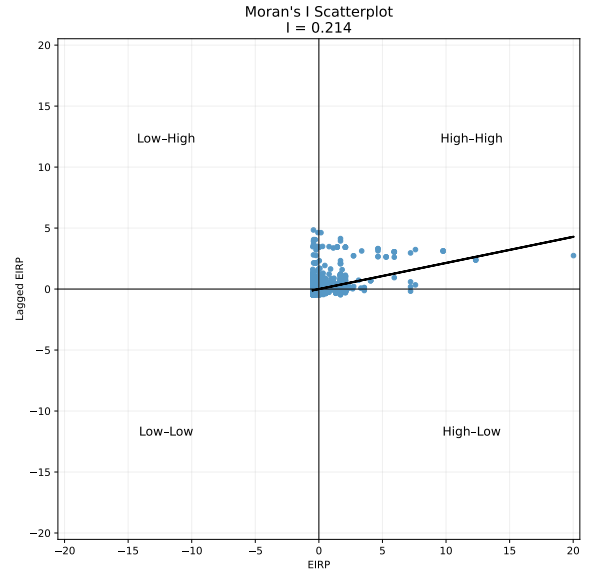


Figure 2. Local Moran’s I scatterplot of FM transmission EIRP for Russia.

into four different quadrants visualized in a Moran’s I scatterplot, as shown in Figure 2. It presents the spatial autocorrelation by plotting the standardized average transmitter EIRP value of “nearest j neighbors,” referred to as “lagged EIRP” against the standardized transmitter EIRP value. A straight-line fit of global Moran’s I index is also shown. The first quadrant denotes high-high autocorrelation (EIRP at i location is high, average EIRP of nearest j neighbors are also high, and so on), the second quadrant denotes high-low autocorrelation, the third quadrant denotes low-low autocorrelation, and the fourth quadrant denotes low-high autocorrelation. The first and third quadrants represent clustering, and the second and fourth quadrants represent dispersion. Most of the transmitter EIRP values are clustered toward the first quadrant, implying that a majority of the transmitter EIRP values have high-high spatial autocorrelation. This was found to be the case for all countries with incomplete datasets.

For each new transmitter that was populated, first the transmitter location and frequency were assigned following the procedures described in Sections 2.1.1 and 2.1.2. Then a mapping was created between the HEALPix pixel of the transmitter and the quadrant where its corresponding local Moran’s I value was situated. New EIRP values were assigned to the synthetic transmitters by reverse-interpolation sampling with EIRP values sampled from the corresponding quartiles of the transmitter EIRP CDF. For example, if a synthetic transmitter was located in a pixel that had its local Moran’s I value in the first quadrant (which corresponds to high-high autocorrelation), the synthetic EIRP value was sampled from the fourth quartile of the transmitter EIRP CDF (see Figure 3). If the synthetic transmitter lied in a region where there was no spatial autocorrelation (statistically insignificant local Moran’s I value, meaning it would lie at the center of the local Moran’s I scatterplot), the synthetic transmitter EIRP was assigned by drawing random samples from the entire CDF of transmitter EIRP.

2.2. The Complete FM Transmitter Database

We henceforth call the STARFIRE algorithm with the newly completed database STARFIRE-2. The RFI spectra seen

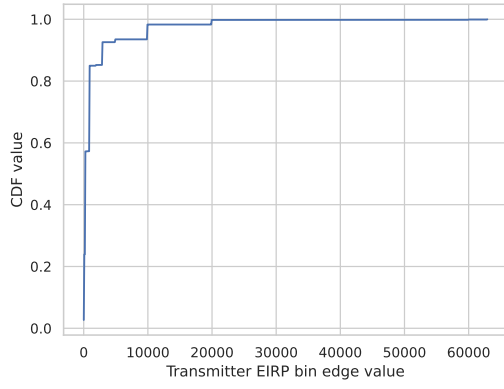


Figure 3. Cumulative function distribution of the EIRP of FM transmitters for an incomplete dataset.

by PRATUSH over Hanle (India) with different transmitter radiation patterns with the newly completed STARFIRE database is shown in Figure 4. Changes in the RFI spectral cube after running the newly completed FM transmitter database are shown in Figure 5. The older version of STARFIRE did not predict any RFI over this region due to incompleteness in its database. The complete database as a HEALPix map of $N_{\text{Side}} 128$ is seen in Figure 6. The algorithm was also run using a $\cos^2(\theta)$ radiation pattern for the FM transmitters, with θ being the elevation angle of the satellite with respect to the FM transmitter. This could be a more realistic assumption for radiation pattern used by FM transmitters to maximize the range of signal propagation on Earth. As seen in Figure 4, isotropic radiation patterns for FM transmitters have higher RFI levels in comparison to horizon-directed beams.

3. Orbit Simulations

The database was completed using the methods mentioned in the previous section. This complete database was run through the STARFIRE-2 algorithm with the following specifications:

1. Frequency range of receiver. 55–110 MHz
2. Frequency resolution of the receiver. 244 KHz

To estimate the RFI in Earth orbit, we propagate the satellite position over Earth using orbital parameters of interest and compute the expected RFI as the satellite moves over pixels corresponding to the “ground track”—the path that the satellite traces on the surface of the Earth directly beneath while in orbit (see Figure 7). For each pixel in the ground track, we estimate the corresponding RFI spectrum from the appropriate slicing of the STARFIRE-2 spectral cube.

Orbit simulations were done by using a “two-line element set” (TLE) of different satellite orbits. TLE is a data format used to list orbital elements of a satellite at a given epoch. Orbits are propagated with the TLE information using the simplified general perturbation model SGP4 (D. Vallado & P. Cefola 2012). The orbit-propagation algorithm uses the “Skyfield” Python library (B. Rhodes 2019) and returns the ground-track coordinates as a function of time.

The quality of an orbit can be assessed using the “figure of merit” (FOM), a metric introduced by S. Ghosh et al. (2023). The FOM quantifies the number of frequency channels free from any RFI, with all RFI-contaminated spectral channels flagged irrespective of the interference power in those

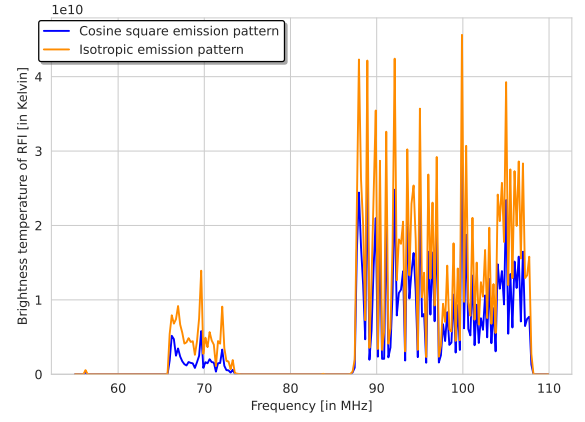
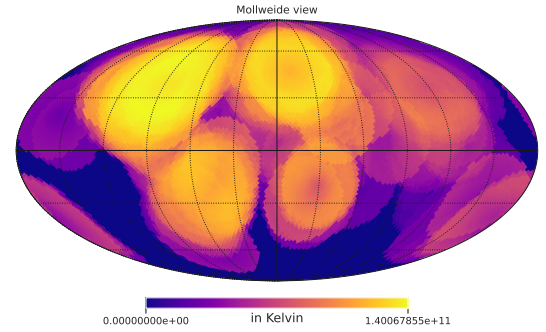
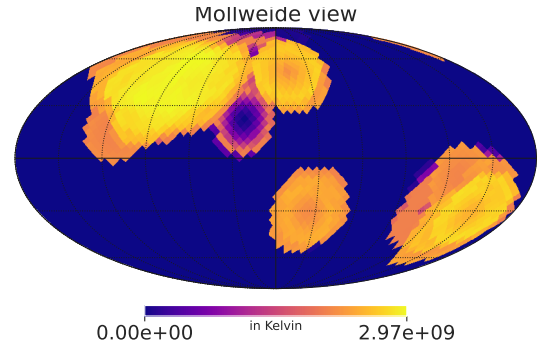


Figure 4. RFI spectrum over India with the new STARFIRE-2 database, with two radiation patterns used by the FM transmitters. The isotropic radiation pattern has higher RFI brightness temperature, thus making it a conservative assumption for PRATUSH.



(a) Beam averaged pixel-wise RFI power at 400 km altitude and frequency 107.22 MHz with new STARFIRE-2 database.



(b) Beam averaged pixel-wise RFI power at 400 km altitude and frequency 107.22 MHz with original STARFIRE database.

Figure 5. Comparison of beam-averaged pixel-wise RFI power at 400 km altitude for different STARFIRE database versions.

channels. Each FOM bin contains four frequency channels in the case of PRATUSH-I. Based on the FOM, regions where the satellite experiences minimal or no RFI are designated as low-RFI zones. Specifically, any pixel with an FOM value of 54 or higher (with 55 being the maximum possible FOM) is classified as a low-RFI zone. To evaluate different orbital configurations, the total time the satellite spends within these zones is calculated. An optimal orbit is therefore one that maximizes the time spent in high-FOM regions or,

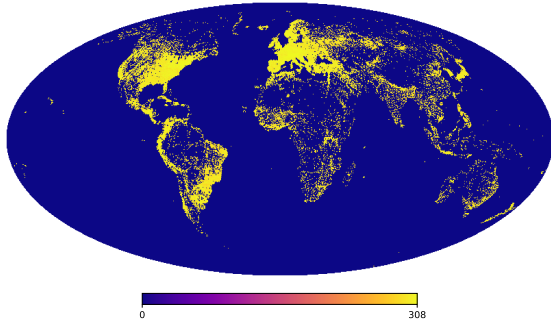


Figure 6. N_{Side} 128 map of number of transmitters per pixel with a complete database.

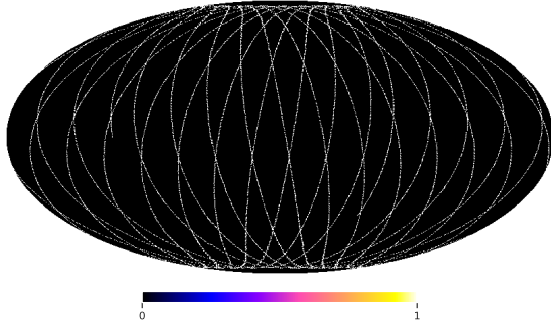


Figure 7. Ground track of 1 day of orbit propagation for a polar orbit at 705 km altitude.

equivalently, in zones characterized by low levels of RFI contamination.

We choose 1 yr as the orbit-propagation period. However, orbit-propagation errors using SGP4 models grow quadratically with propagation time for LEOs (R. Wang et al. 2009). A feasible method to simulate an orbit would be to obtain archival TLE information from missions that have been active for at least 1 yr, since historical TLE files for LEO satellites have a propagation error of ± 0.25 km for 1 day of orbit propagation (X. Xiao-li & X. Yong-qing 2019). This is well within the minimum resolution of the HEALPix maps (~ 205 km at the equator) used for RFI analysis. Therefore, in order to maintain an accurate ground track throughout the span of orbit propagation, TLE data for select satellites (see Section 3.2 for details) were obtained from the NORAD archive⁶ such that the TLE files would be renewed every day throughout the span of 1 yr of orbit propagation.

The time resolution for recording the ground-track coordinates of the satellite in orbit was set to 8.640 s. Finally, the total time spent by the satellite at each FOM bin was calculated for only regions and times when the satellite would not be sunlit since PRATUSH-I would conduct science observations when the Sun is below the horizon (nighttime) to avoid intercepting incoming solar radio flux.

3.1. South Atlantic Anomaly

There exists a region over Earth where the magnetic field strength is the weakest. This is due to the offset of the Earth’s magnetic dipole from the Earth’s center (T. Asikainen & K. Mursula 2008). Due to the offset of the two axes (rotational and magnetic), the Van Allen radiation belt containing high-

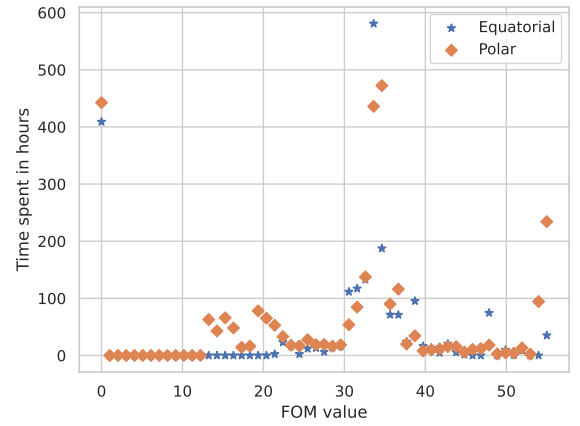


Figure 8. Time spent by the satellite at each FOM bin for an orbit propagation of 1 yr.

energy particles can penetrate deeper into the atmosphere (altitudes of 200 km from the Earth surface) in the South Atlantic Anomaly (SAA) region. This effect causes satellites in LEO to receive large doses of radiation in orbit (J. Heirtzler 2002) and potentially leads to “single event upsets,” which cause electronics to malfunction after interacting with a high-energy particle (F. Wang & V. D. Agrawal 2008). The SAA region spans from 90°W to 40°E longitude and 0°S to 50°S latitude (K. Affi Nasuddin et al. 2019).

We regard this as an unfavorable region for conducting observations. We account for this by setting the FOM value to 0 over this region, thereby down-weighting any pixels that lie in the SAA, which might otherwise have a high FOM.

3.2. Optimal Orbit Selection

Orbits of three satellites were simulated as representative to evaluate the time spent by the satellite in low-RFI zones—two of which were in LEO and a third one in geosynchronous orbit (GEO). A polar LEO was simulated using archival TLE data of “Landsat-7” (S. N. Goward et al. 2001), and an equatorial LEO was simulated using archival TLE data of “AstroSat” (K. P. Singh et al. 2014). “NSS-9” data was used to simulate GEO.

As seen in Figure 8, the polar orbit outperforms the equatorial orbit for the most amount of time spent over low-RFI zones for an orbit-propagation period of 1 year (234.582 hr a year). This is attributed to the fact that the satellite in polar orbit spends more time in radio-quiet zones, including Antarctica and the Pacific, South Atlantic, and Indian Oceans. In contrast, a satellite in equatorial orbit would mostly spend its time over RFI-contaminated zones. (It spends 35 hr in low-RFI zones per year.) GEOs have 0 hr of low-RFI zone time for a year of orbit propagation since the FOV of the satellite covers a hemisphere of the Earth, thereby containing multiple RFI sources in its beam at any location over the Earth. As mentioned in Section 2.2, the orbit simulations were also run setting the radiation pattern of the FM transmitters to a $\cos^2(\theta)$ pattern, with θ being the elevation angle. The time PRATUSH spends at radio-quiet zones changed to 253.2 hr per year in polar orbit, and remained unchanged for the equatorial orbit. With these results, we choose to set the radiation pattern of the FM transmitters to be isotropic to have conservative estimates for our sensitivity estimates in the next section of this paper.

⁶ <https://celestrak.org/NORAD/archives/>

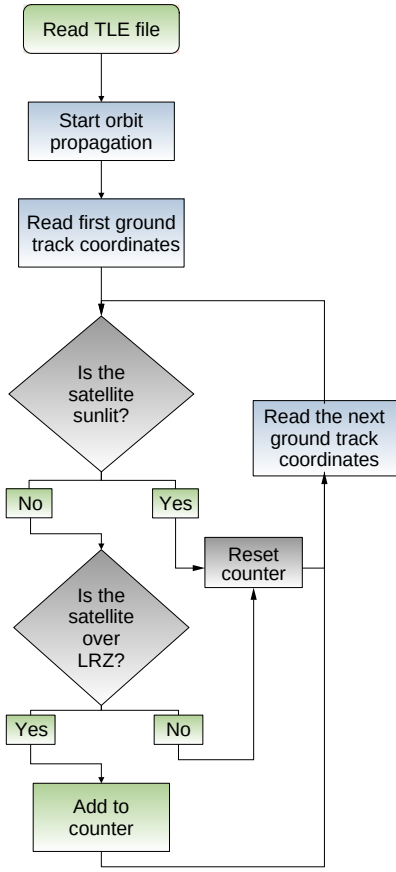


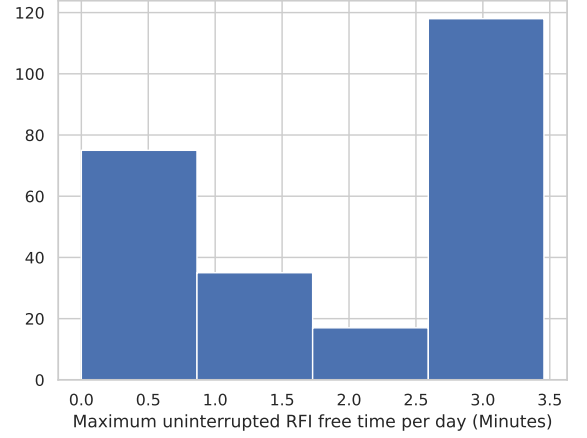
Figure 9. Uninterrupted low-RFI counter algorithm, where LRZ stands for “low-RFI zone.”

We also investigate the transit periods over low-RFI zones that the satellite has for different orbits per day. The orbit simulation algorithm records uninterrupted low-RFI time. The methodology is visualized in Figure 9. Figures 10(a) and (b) show the distribution of maximum uninterrupted low-RFI time per day in minutes for a satellite in polar and equatorial orbit, respectively. As expected, a satellite in polar orbit performs better in this parameter, as well as in contrast to an equatorial orbit.

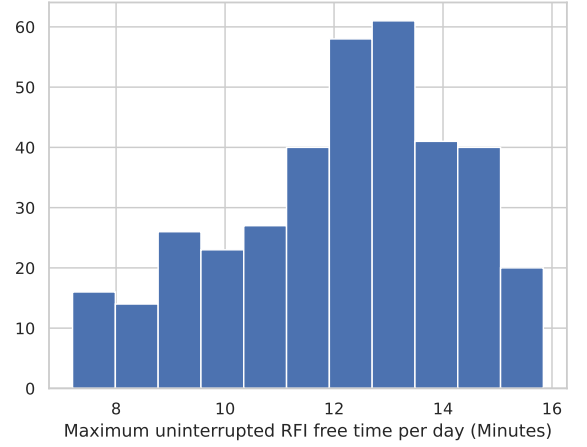
4. Mock Sky Observations

The sensitivity of PRATUSH-I to various CD/EoR models can be estimated by including mock observations of the radio sky included in the STARFIRE-2 algorithm with the complete database. This was done by obtaining the beam-weighted sky for the PRATUSH-I antenna given by Equation (10),

$$\begin{aligned}
 T_A(\nu) = & (1 - |S_{11}(\nu)|^2) \\
 & \times \left[\frac{\int_0^{2\pi} \int_{-\pi/2}^0 T_{\text{RFI}}(\theta, \phi, \nu) B(\theta, \phi, \nu) d\Omega}{\int_0^{2\pi} \int_{-\pi/2}^0 B(\theta, \phi, \nu) d\Omega} \right. \\
 & + \frac{\int_0^{2\pi} \int_0^{\pi/2} T_{\text{Sky}}(\theta, \phi, \nu) B(\theta, \phi, \nu) d\Omega}{\int_0^{2\pi} \int_0^{\pi/2} B(\theta, \phi, \nu) d\Omega} \\
 & \left. + T_{\text{Earth}}(\nu) \right], \quad (10)
 \end{aligned}$$



(a) Maximum Uninterrupted low RFI zone time per day in equatorial orbit



(b) Maximum Uninterrupted low RFI zone time per day in polar orbit

Figure 10. Maximum uninterrupted low-RFI zone time per day for equatorial and polar orbits.

where T_A is the antenna temperature, $S_{11}(\nu)$ is the reflection coefficient of the PRATUSH antenna, T_{RFI} is the RFI temperature that PRATUSH intercepts from the lower half of the azimuthal plane, T_{Sky} is the foreground brightness temperature from GMOSS maps (M. Sathyanarayana Rao et al. 2017), B is the beam pattern of the PRATUSH antenna, and T_{Earth} is the effective brightness temperature of the Earth approximated as a blackbody of temperature 300 K (E. G. Njoku & E. K. Smith 1985).

In addition to calculating the elevation angle an FM transmitter subtends with respect to the satellite, one requires the azimuthal angle subtended by the transmitter with respect to the subsatellite point. The azimuth angle of a ground point with respect to the subsatellite point is defined as

$$\text{Azimuth} = \arctan 2(\mathbf{v} \cdot \hat{\mathbf{e}}, \mathbf{v} \cdot \hat{\mathbf{n}}), \quad (11)$$

where $\mathbf{v} = \mathbf{r}_g - \mathbf{r}_{\text{ssp}}$ is the line-of-sight vector from the subsatellite point to the ground point, and $\hat{\mathbf{e}}, \hat{\mathbf{n}}$ are the local east and north unit vectors in the east–north–up (ENU) coordinate system (G. Seeber 2003).

4.1. Sensitivity to CD/EoR Models

We examine the sensitivity of the radiometer for different CD models at various points in the orbit corresponding to varying levels of FM-seeded RFI contamination. We compare three scenarios, a region over Earth with zero channels lost to RFI (72°23′15″S 37°30′00″E over Antarctica), a region with four channels lost/flagged (87°4′33″S 45°0′0″W over Antarctica), and a region with very high RFI levels (48°8′28″N 106°4′17″E) leading to high channel loss (153 channels lost out of 226 channels). A threshold based flagging routine was employed to flag channels with very high brightness temperature. From GMOSS, we can expect a Galactic contribution of ~ 6000 K at 55 MHz with the beam maxima pointing toward the Galactic center. Hence we round up the threshold for flagging any channel that contains more than 10,000 K as a conservative estimate. We do not explore more complex RFI flagging routines beyond threshold flagging, as this is beyond the scope of this paper. Sensitivity to various CD/EoR models at different regions over the Earth is shown in Figure 11. Foregrounds are expected to spectrally smooth in contrast to the global signal and can be modeled by a maximally smooth function (M. S. Rao et al. 2017). Foregrounds are separated by subtracting the best-fit maximally smooth function and minimizing the residual rms using the “Basinhopping” algorithm (D. J. Wales & J. P. K. Doye 1997). Various CD/EoR global-signal models (A. Cohen et al. 2017) were injected into the beam-weighted sky and were separated from the foregrounds using the maximally smooth function. Confidence in detecting each model was calculated by comparing the increase in the residual rms for the simulated sky spectrum with the 21 cm signal injected and that for a mock sky with no signal injected, which is a null hypothesis case. The noise floor for conducting the mock sky observations was determined by the radiometer equation

$$\Delta T = \frac{T_{\text{sys}}}{\sqrt{B \tau}}, \quad (12)$$

where ΔT represents the radiometer noise temperature, T_{sys} represents the system temperature, B is the channel width of the instrument, and τ is the integration time, which was obtained by calculating the total RFI free time for a year of orbit propagation in the polar orbit (J. J. Condon & S. M. Ransom 2016).

As seen in Table 3, having a small number of channels lost to RFI does not have an impact on confidence in the detection of various CD/EoR models. In contrast, high channel loss leads to a majority of the CD/EoR models not being detected with high confidence. With these results, we can conclude once again that any orbit that spends the most amount of time in low-RFI zones would be deemed optimal—the polar LEO in this case.

5. Discussion

5.1. PRATUSH at L2 Point of the Sun–Earth System

One can investigate the feasibility of conducting such an experiment much farther from the Earth to see if RFI power falls below the amplitude of the 21 cm signal. One possibility is the Sun–Earth L2 point, where many astronomy telescopes have flown in a Lissajous orbit (M. Hechler & J. Cobos 2003; M. Clampin 2008). For simplicity of the feasibility study, we

place PRATUSH at the L2 point, rather than in orbit around L2, in order to estimate the order of magnitude of the RFI power levels contributed by the Earth-facing hemisphere. Incoming radio flux from the Sun due to various time-varying solar activities are not taken into account but will surely be detected by the experiment, even in an ideal case where the satellite would be eclipsed from the Sun, since the Sun’s apparent size at L2 is slightly larger than the Earth (by around 0.022). As seen in Figure 12, the RFI brightness temperature lies in the order of a few hundred mK, increasing up to 1 K, thus adding significant spectral features to the sky spectrum. This would result in poor sensitivity for all CD/EoR models.

5.2. Sources of Uncertainty

1. All simulations demonstrated in this paper are presented with the caveat that the database procured from FMLIST is a reasonably accurate representation of FM transmission. There are a few discrepancies including missing information. A total of 18.75% of transmitters obtained from the FMLIST database have missing EIRP information, which were replaced with power values drawn from a global transmitter power distribution. However, the interpretation of the STARFIRE-2 algorithm using the FOM metric is insensitive to the exact value of EIRP introduced since flagging is based on the presence or absence of RFI in a channel and not the RFI power. A total of 0.35% transmitters have missing location information that are filled using the statistical methods described herein. While nonzero, this is not a very large number, and we acknowledge this to be a source of uncertainty.
2. The exact location, frequency, and power of the missing database entries populated using statistical methods are representative and reflect the underlying statistical distribution of the population.
3. Shapefiles used for generating synthetic transmitter information can have missing spatial information (archipelagos, small territories with lesser angular extent), thus affecting the performance of the probabilistic model of transmitter distribution.
4. The radiation pattern of all FM transmitters are assumed to be isotropic to reduce the complexities involved in obtaining the beam pattern of each FM transmitter. This is a conservative assumption compared to omnidirectional antennas that are typically employed by FM transmitters. In practice the transmitting antennas may be more directive, and because of this there may be a larger number of cleaner patches in the output heatmaps of STARFIRE-2 and thus more locations in orbit that are classified as low-RFI zones.
5. Other sources of RFI in the band, including but not limited to dynamic sources, such as other satellites in orbit, maritime radars, harmonics, and wideband noise from power systems from other sources, are not included in the current database. We also do not include unintended and hence undocumented sources of emission, such as from power systems on board satellites that have been observed by low frequency telescopes on Earth (F. Di Vruno et al. 2023; C. Bassa et al. 2024; X. Zhang et al. 2025). These are best determined empirically due to the unregulated nature of emission,

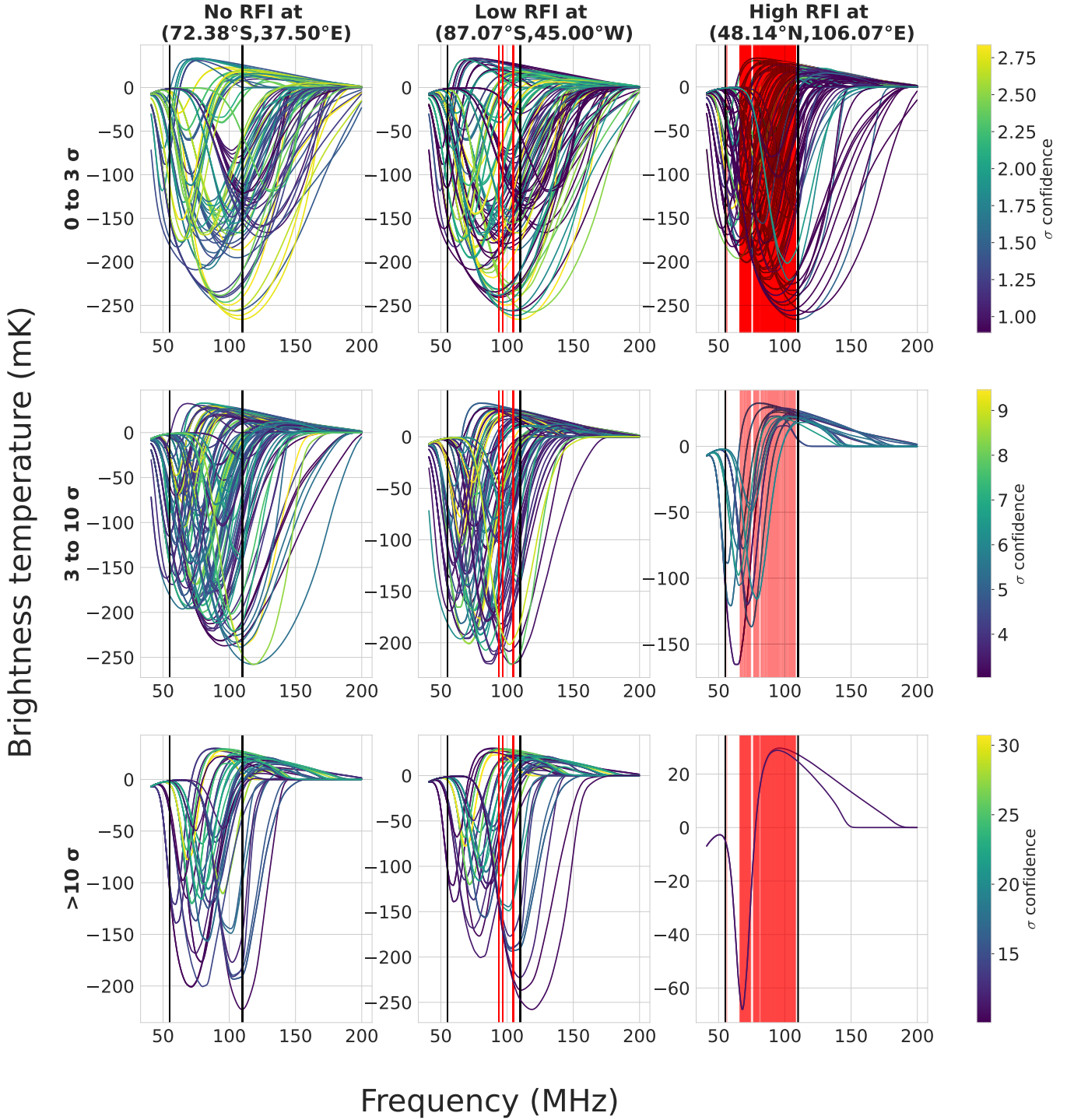


Figure 11. Confidence in detecting different CD/EoR signal models at different locations. Each column of the panel represents an RFI scenario encountered in the optimized orbit, while each row represents a confidence interval. The solid black lines indicate the PRATUSH-I frequency window, and the solid red lines indicate frequency channels lost to RFI.

Table 3

Percentage of Detecting Different CD/EoR Models at Different Confidence Levels

Confidence Level	No RFI	Low RFI	High RFI
0–3 σ	34.6%	34.6%	90.9%
3–10 σ	45.6%	45.6%	8.3%
>10 σ	19.7%	19.7%	0.76%

which motivates a dedicated measurement of the radio environment in orbit around Earth.

- The results presented in the paper will be dramatically different for an antenna that has zero/high suppression beam response in the lower half of the azimuthal plane. For instance, it is possible to achieve a narrow beam that is pointed away from Earth, say via beamforming, causing significant suppression in RFI levels. Such an antenna would

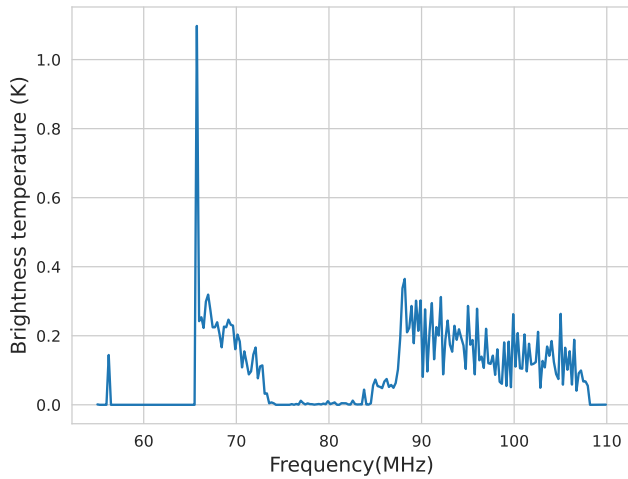


Figure 12. RFI power levels at L2 of the Sun–Earth system.

need to be very large to achieve the directive beam at the frequencies in question, and achieving a stable achromatic beam over the full frequency range is extremely difficult. Thus, for PRATUSH-I, antenna design prioritizes frequency-independent performance across the band. Hence a narrow-beam response is not plausible.

7. The statistical methods employed in this paper inevitably introduce a positive bias. Consequently, the model might fail to populate locations with missing transmitters if none were present in the prior distributions, even though they may exist in reality.

6. Conclusions and Future Work

It is worth exploring an Earth orbiting radiometer as a precursor to a lunar farside experiment for detecting the global redshifted 21 cm signal from the CD. The relatively low cost, low complexity, technology-readiness level of electronics, and launch opportunities all make an Earth-orbiting satellite an attractive and feasible option. Such a radiometer would perform better than a ground-based counterpart on two fronts, namely ionosphere-induced chromaticity and systematics in antenna behavior from coupling to dielectrics in the near field. However, it will certainly receive RFI from Earth, such as from FM transmitters in its FOV. Focusing only on FM-based RFI, in this work we have investigated the presence of any low-RFI zones in Earth orbit, the number of frequency channels that remain usable for signal detection after RFI flagging, and hence CD signal detection prospects in three Earth orbits—namely, polar LEO, equatorial LEO, and GEO. We consider RFI only from FM transmitters on Earth, building a global database of the same, using free data from <https://www.fmlist.org/> and populating gaps in the data using a scalable statistical model. We find that there are indeed regions, such as around the Antarctic region and over the Pacific Ocean, where within simulation resolution there is no FM-seeded RFI, assuming a wide antenna beam. While this is a lower limit based on FM-seeded RFI alone, it is worth exploring this region for detecting CD signals in space around Earth. A polar LEO orbit maximizes time spent over this FM radio quiet zone and is the preferred orbit for maximizing signal detection sensitivity. Mock sky observations show that recovering a CD/EoR signal over these FM radio quiet zones with high confidence is feasible for a majority of the standard theoretical models. The work carried out in this paper only considers stationary FM-based RFI

sources for simulations. However, time-varying RFI sources, including other satellites in Earth orbit, pose a more complex problem to the experiment and must be considered for future work carried out for CD/EoR science in Earth orbit. Temperature-dependent effects on the properties of the antenna and any residual ionospheric effects (based on the altitude) are not considered in these simulations.

The FOM metric used in the simulations assume that there is no amplitude leakage from channels having very high RFI temperature. Accounting for this fact would decrease the percentage of CD/EoR models that can be detected with high confidence at high RFI regions. In all, we motivate the utility of a pilot experiment to map the RFI around Earth in space, particularly in polar orbit, to also capture the effect of other transient sources of RFI. This information of total RFI spectrum in space around Earth can pave the way for lower-cost and lower-complexity radiometer (than lunar farside) for CD detection radiometer orbiting Earth.

Acknowledgments

The authors would like to thank Sonia Ghosh for providing assistance with the STARFIRE algorithm. The authors would also like to thank Vishakha Pandharpure for providing electromagnetic simulation data for the PRATUSH antenna. Y.P. extends special appreciation to Adarsh Kumar Dash and Tejas Oak for their helpful inputs during the project.

Appendix

Tests for Potential Correlation of RFI with Human Activity

Since the STARFIRE algorithm operates using the HEALPix pixelization scheme, STARFIRE-2 aimed to conduct statistical studies using the same scheme. Population density was suspected to be correlated to the number of FM transmitters installed and hence the total RFI power in the FM band (R. Umar et al. 2014). Hence studies were done to measure the extent of correlation between the two factors by comparing HEALPix heatmaps of population density with pixel-wise FM transmitter population and power. Obtaining a scalable statistical model mandates the usage of high-resolution data, such as worldwide-population-density datasets available in the form of GEOTiff files. The pixels in a GEOTiff image are georeferenced (S. Mahammad & R. Ramakrishnan 2003), indicating that they represent a scale affine transformation of the World Geodetic System-84 (WGS-84). One can confirm the accurate alignment of the right data points from the original image with the corresponding HEALPix pixel by ensuring that the defined spherical coordinates match the shape of the image. A 15' resolution population-density raster dataset was obtained from NASA-SEDAC⁷ and processed into a HEALPix array using a Python algorithm. $N_{\text{Side}} 128$ was chosen as the resolution of the HEALPix map throughout the correlation studies (see Figure 13).

The STARFIRE-2 FM transmitter database contains information about various parameters of transmitters, including their geodetic coordinates. This information was utilized to create an $N_{\text{Side}} 128$ HEALPix map of the transmitters on the globe with various parameters, such as total transmitter power per pixel and number of transmitters per pixel.

⁷ <https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density-adjusted-to-2015-unwpp-country-totals-rev11/data-download>

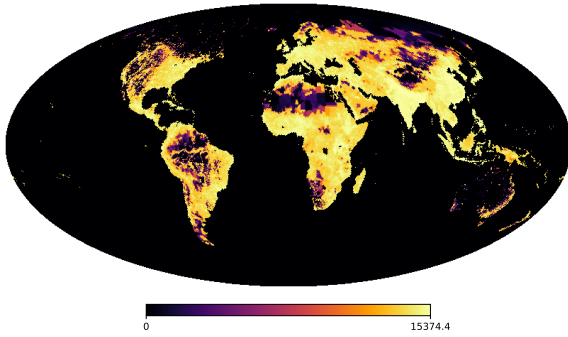


Figure 13. Global human population density HEALPix map of N_{Side} 128.

As a preliminary test, the correlation studies were conducted for the limited countries' database present in the original STARFIRE database.

A.1. Using Linear Combinations of Parameters for Correlation Studies

Since direct comparison between population density and transmitter parameters seemed to be generally uncorrelated (Pearson's $r=0.29$), different factors such as the Human Development Index (HDI) and country-wise land area seemed to contribute to these parameters as well. Hence the independent variable of the correlation studies was modified into a linear combination of the previously mentioned factors and analyzed for potential trends.

A raster GEOTiff image of country-wise HDI map was created in QGIS using a world-map shapefile. The shapefiles were connected to the HDI delimited text by utilizing the country name as the shared identifier between the two datasets. The resulting shapefile was rasterized with an HDI feature as pixel weight and run through the same Python algorithm, thus creating an N_{Side} 128 HEALPix map of country-wise HDI. Hence the map was readily usable as a direct multiplicative factor with the population-density HEALPix map for potentially better correlations since it contained exactly the same number of country-wise HDI weighted pixels as any other HEALPix map in use. However, there was no significant increase in Pearson's correlation coefficient ($r=0.292$). Hence it was concluded that the spatial distribution had no strong correlation with factors like population density but were rather randomly distributed with country-specific transmitter power and frequency values.

A.2. Correlation Studies with Complete Database

Further correlation studies were conducted using the newly complete database. Analyzing the trends between pixel-wise transmitter power and pixel-wise population density further solidified the notion that there was no strong correlation between these two parameters (Pearson's $r=0.0099$), but were rather randomly distributed with country-dependent location and scale parameters.

ORCID iDs

Mayuri Sathyanarayana Rao  <https://orcid.org/0000-0002-9761-3676>

References

Afifi Nasuddin, K., Abdullah, M., & Hamid, N. S. A. 2019, *NPGeo*, **26**, 25
Agrawal, Y., Kavitha, K., & Singh, S. 2024, *ApJ*, **974**, 137

- Asikainen, T., & Mursula, K. 2008, *JASTP*, **70**, 532
Bassa, C., Di Vruno, F., Winkel, B., et al. 2024, *A&A*, **689**, L10
Bernardi, G., McQuinn, M., & Greenhill, L. J. 2015, *ApJ*, **799**, 90
Bowman, J., Rogers, A., Monsalve, R., Mozdzen, T., & Mahesh, N. 2018, *Natur*, **555**, 67
Bradley, R. F., Tauscher, K., Rapetti, D., & Burns, J. O. 2019, *ApJ*, **874**, 153
Burns, J. O., Bale, S., & Bradley, R. F. 2019, AAS Meeting, **234**, 212.02
Burns, J. O., Lazio, J., Bale, S., et al. 2012, *AdSpR*, **49**, 433
Chen, X., Yan, J., Deng, L., et al. 2020, *RSPTA*, **379**, 20190566
Clampin, M. 2008, *AdSpR*, **41**, 1983
Cohen, A., Fialkov, A., Barkana, R., & Lotem, M. 2017, *MNRAS*, **472**, 1915
Condon, J. J., & Ransom, S. M. 2016, *Essential Radio Astronomy* (Princeton Univ. Press)
Datta, A., Bradley, R., Burns, J. O., et al. 2016, *ApJ*, **831**, 6
de Lera Acedo, E., de Villiers, D. I. L., Razavi-Ghods, N., et al. 2022, *NatAs*, **6**, 984
Di Vruno, F., Winkel, B., & Bassa, C. 2023, *A&A*, **676**, A75
Freedman, D., & Diaconis, P. 1981, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **57**, 453
Gagnon-Hartman, S., Cui, Y., Liu, A., & Ravanbakhsh, S. 2021, *MNRAS*, **504**, 4716
Ghosh, S., Rao, M. S., & Singh, S. 2023, *A&C*, **44**, 100727
Gorski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, **622**, 759
Goward, S. N., Masek, J. G., Williams, D. L., Irons, J. R., & Thompson, R. 2001, *RSEnv*, **78**, 3
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Natur*, **585**, 357
Hechler, M., & Cobos, J. 2003, *Libration Point Orbits and Applications* (World Scientific), 115
Heitzler, J. 2002, *JASTP*, **64**, 1701
Knuth, K. H. 2006, arXiv:physics/0605197
Mahammad, S., & Ramakrishnan, R. 2003, Map India Conf. 2003 <https://geospatialworld.net/wp-content/uploads/images/pdf/117.pdf>
Moran, P. A. P. 1950, *Biometrika*, **37**, 17
Njoku, E. G., & Smith, E. K. 1985, *RaSc*, **20**, 591
Offringa, A. R., Wayth, R. B., Hurley-Walker, N., et al. 2015, *PASA*, **32**, e008
Philip, L., Abdurashidova, Z., Chiang, H. C., Ghazi, N., et al. 2018, *JAI*, **8**, 1950004
Price, D. C., Greenhill, L. J., Fialkov, A., et al. 2018, *ApJ*, **478**, 4193
Pritchard, J. R., & Loeb, A. 2012, *RPPH*, **75**, 086901
Rao, M. S., Subrahmanyam, R., Shankar, N. U., & Chluba, J. 2017, *ApJ*, **840**, 33
Rhodes, B. 2019, Skyfield: High precision research-grade positions for planets and Earth satellites generator, Astrophysics Source Code Library, **1907.024**
Sathyanarayana Rao, M., Singh, S., K. S., S., et al. 2023, *ExA*, **56**, 741
Sathyanarayana Rao, M., Subrahmanyam, R., Udaya Shankar, N., & Chluba, J. 2017, *AJ*, **153**, 26
Seeber, G. 2003, *Satellite Geodesy: Foundations, Methods, and Applications* (2nd ed.; Walter de Gruyter)
Shen, E., Anstey, D., de Lera Acedo, E., Fialkov, A., & Handley, W. 2021, *MNRAS*, **503**, 344
Shimazaki, H., & Shinomoto, S. 2007, *NeCom*, **19**, 1503
Singh, K. P., Tandon, S., Agrawal, P., et al. 2014, *SPIE*, **9144**, 517
Singh, S., Nambissan, T., Subrahmanyam, R., et al. 2022, *NatAs*, **6**, 607
Singh, S., Subrahmanyam, R., Shankar, N. U., et al. 2018a, *ExA*, **45**, 269
Singh, S., Subrahmanyam, R., Shankar, N. U., et al. 2017, *ApJ*, **845**, L12
Singh, S., Subrahmanyam, R., Shankar, N. U., et al. 2018b, *ApJ*, **858**, 7
Spinelli, M., Kyriakou, G., Bernardi, G., et al. 2022, *MNRAS*, **515**, 1580
Umar, R., Zainal Abidin, Z., Abidin Ibrahim, Z., Rosli, Z., & Noorazlan, N. 2014, *RAA*, **14**, 241
Vallado, D., & Cefola, P. 2012, Proc. 63rd Int. Astronautical Congress (IAC-12) (International Astronautical Federation) <https://iafastro.directory/iac/archive/browse/IAC-12/C1/6/13640/>
Vedantham, H. K., Koopmans, L. V. E., De Bruyn, A. G., et al. 2014, *MNRAS*, **437**, 1056
Wales, D. J., & Doye, J. P. K. 1997, *JPCA*, **101**, 5111
Wang, F., & Agrawal, V. D. 2008, in 21st Int. Conf. on VLSI Design (VLSID 2008) (IEEE), 429
Wang, R., Liu, J., & Zhang, Q. M. 2009, *AdSpR*, **43**, 1065
Wilensky, M. J., Morales, M. F., Hazelton, B. J., et al. 2019, *PASP*, **131**, 114507
Xia, K., & Wei, C. 2008, in 2008 Int. Conf. on Embedded Software and Systems Symp. (IEEE), 174
Xiao-li, X., & Yong-qing, X. 2019, *ChA&A*, **43**, 563
Zhang, X., Zarka, P., Viou, C., et al. 2025, *A&A*, **698**, A244
Zonca, A., Singer, L. P., Lenz, D., et al. 2019, *JOSS*, **4**, 1298