Draft version July 8, 2024
Typeset using LaTeX **twocolumn** style in AASTeX631

Bioverse: Origins of Life

Martin Schlecker[1] and et al.

[1]*Steward Observatory, The University of Arizona, Tucson, AZ 85721, USA; schlecker@arizona.edu*

ABSTRACT

tbd.

1. INTRODUCTION

Introduce OOL, the importance of planetary contexts

1.1. *Origins of Life Scenarios and their Predictions*

Present widely discussed OOL scenarios and their predictions on exoplanet observables; derive testable hypotheses.

In this section, we present some of the most prominent origins of life scenarios and their observational predictions. We focus on the necessary environmental conditions for the processes and reactions inherent to each scenario, and aim to identify distinct observables that are accessible via present and near-future remote sensing techniques.

A widely regarded origins-of-life scenario is that abiogenesis happens in hydrothermal vents (e.g., **?**). ... The hydrothermal vents scenario requires a direct contact of an ocean and the planetary mantle/crust. This requirement is not met on an ocean world with large amounts of water, where the water pressure on the ocean floor is high enough to form high-pressure ices (Noack+2016).

see discussion in Kite & Ford 2018 Sect. 6.4

SR: The sealing away of the planetary interior from the ocean due to high-pressure ice layers is a common assumption for water world exoplanets (in addition to the references above, see e.g. Hu et al. 2021). I'm not convinced it is correct, because of relatively recent evidence showing the possibility of molecular assimitation into such ices and subsequent transport, e.g., https://iopscience.iop.org/article/10.1088/0004-637X/769/1/29/meta, https://iopscience.iop.org/article/10.3847/1538-4357/acb49a/meta, https://iopscience.iop.org/article/10.3847/1538-4357/aa5cfe/meta (I'm sure there are other workers in this area, this is just the group with which I am familiar). Exoplaneteers mostly uniformly accept this proposition, so it's not an unreasonable assumption if you want to run with it so long as its acknowledged and caveated reasonably; I'm just highlighting this for your attention so that you can make an informed decision.

Prediction: Planets with high-pressure ices do not show biosignatures.

A different scenario is the emergence of life in hot springs or ponds that are exposed to the planet's atmosphere (e.g., **?**). ... By its nature, the subaerial ponds scenario relies on rock surfaces exposed to the planetary atmosphere. Water worlds that have their entire planetary surface covered with water contradict this requirement and do not allow for the wet-dry cycling inherent to this origin of life scenario. The competition of tectonic stress with gravitational crustal spreading (Melosh 2011) sets the maximum possible height of mountains, which in the solar system does not exceed $\sim 20\,\mathrm{km}$. Such mountains will be permanently under water on water worlds. Another impediment to wet-dry cycles is tidal locking of the planet as it stalls stellar tide-induced water movement and diurnal irradiation variability.

Prediction: Biosignatures occur outside the tidal locking zone and at bulk densities consistent with exposed rock.

SR: I'm not a priori sold that tidal locking means that no wet-dry cycles occur. you can still have cycling driven by transient changes in instellation due to flares, for example (e.g., https://iopscience.iop.org/article/10.3847/1538-4357/aadfd1/meta). Similarly, I wonder if 3D effects might not give rise to variability (https://iopscience.iop.org/article/10.3847/PSJ/acc9c4/meta). I argue that it is more robust to establish a correlation between biosignatures and planets which show evidence of continents/land. I think that Ty Robinson in our department has done some work in this area, his papers might be a good starting point. Other papers which look relevant (but with which I am not familiar, as this is not my area): https://academic.oup.com/mnras/article/511/1/440/6501216, https://academic.oup.com/mnras/article/495/1/1/5733176, https://iopscience.iop.org/article/10.3847/1538-3881/aad775/meta, https://iopscience.iop.org/article/10.3847/1538-3881/aaed3a/meta, https://iopscience.iop.org/article/10.3847/1538-3881/ab2df3/meta

A major hypothesis in the origin of life is that UV light played a constructive role in getting life started on Earth (see Ranjan et al. 2016, 2017c; Rimmer et al. 2018; Rapf & Vaida 2016; Pascal et al. 2012; Green et al. 2021; and sources therein).

If UV light is required to get life started, then there is a minimum planetary UV flux requirement to have an inhabited world. This requirement is set by competitor thermal processes; if the photo-reaction does not move forward at a rate faster than the competitor thermal process(es), then the abiogenesis scenario cannot function. On the other hand, abundant UV light vastly in excess of this threshold does not increase the probability of abiogenesis, since once the UV photochemistry is no longer limiting, some other thermal process in the reaction network will be rate-limiting process instead. Therefore, a putative dependence of life on UV light is best encoded as a step function (see, e.g., Ranjan et al. 2017c; Rimmer et al. 2018; Rimmer, Ranjan & Rugheimer 2021).

One origin-of-life scenario has been refined to the point where the threshold flux has been measured. The cyanosulfidic scenario has been shown to require a mean flux of at least $F_{\mathrm{NUV,min}} = (6.8 \pm 3.6) \times 10^{10}\,\mathrm{photons\,cm^{-2}\,s^{-1}\,nm^{-1}}$ integrated from 200–280 nm at the surface in order to function

(Rimmer et al. 2018; Rimmer et al. 2021 Astrobiology; Rimmer et al. 2023).

SR: This is an interesting number, because it is below what was available on early Earth (so this scenario could have worked on early Earth) but until recently it was below what was thought to be available on habitable zone M-dwarf exoplanets. So it was thought that identification of biosignatures on M-dwarf planets could therefore falsify the cyanosulfidic scenario, with a potential caveat for transient UV from flares. Two recent developments have complicated the picture. First, Rimmer et al. 2018 had an error in their radiative transfer routines. Correcting for this error, early M-dwarfs and highly active M-dwarfs emit enough UV to meet the Rimmer et al. 2018 criterion (Ranjan et al. 2023). Second, a recent publication argues that /all/ estimates of M-dwarf UV are underestimates, and that late M-dwarf stars have similar emission to Sunlike stars (Rekhi et al. 2023). I suspect this is incorrect, because it contradicts a lot of work from e.g. the MUSCLES collaboration and the HAZMAT project, but it's worth keeping an eye on in case it is correct after all.

We use this threshold value as our baseline case.

Prediction: Past UV flux and the occurrence of biosignatures are correlated.

... Figure 1 shows the hypothesis and null hypothesis derived from the predictions of the cyanosulfidic scenario. ...

2. METHODS

2.1. *Fraction of inhabited planets with detectable biosignatures*

Presumably, not all habitable worlds are inhabited and not all inhabited worlds develop detectable biosignatures. The fraction of exo-Earth candidates (EEC) that are both inhabited and harbor detectable biosignatures at the time when we observe them remains speculative; we aggregate them in the unitless parameter $f_{\rm life}$.

Günther et al. (2020) relate U-band energy to bolometric flux.

OPTIONAL: "We further test the scenario of a linear correlation of past UV flux and biosignature occurrence rate. This test requires the detection of multiple biosignatures."
Test for negative correlations as well?

Here, we conduct a theoretical experiment on the UV irradiance requirement (Sect. 1.1) by relating the occurrence of life on an exo-earth candidate with a minimum past quiescent stellar UV flux, focusing on the prebiotically interesting near-Ultraviolet (NUV) range from 200–280 nm.

SUKRIT: got a good reference for this?

Our concrete hypothesis shall be that life only occurs on planets that at some point in their history have received such radiation exceeded a minimum flux $F_{\rm NUV,min}$.

2.2. *Semi-analytical approach*

First, we apply a Bayes Factor Design Analysis (**?**) to assess the expected probabilities of obtaining true negative or true positive evidence for the hypothesis above, as well as the probability for misleading or inconclusive evidence, under idealized conditions. This serves as a first-order estimate of the information content of a survey, before we take into account impacts from exoplanet demographics, sample selection, and survey strategy.

Let our observable be the inferred past NUV flux of the planet $F_{\rm NUV}$. Under Hypothesis H_1 (Equation 9), there exists a special unknown value of $F_{\rm NUV}$, noted $F_{\rm NUV,min}$ such that

$$P(L|\theta, H_1) = f_{\rm life} \quad \text{if } \theta > F_{\rm NUV,min} \tag{1}$$

$$P(L|\theta, H_1) = 0 \quad \text{otherwise} \tag{2}$$

where $f_{\rm life}$ is the unknown probability of abiogenesis. The corresponding null hypothesis is that there exists no such special value of $F_{\rm NUV}$ and that

$$P(L|\theta, H_{\rm null}) = f_{\rm life}. \tag{3}$$

If we now define a sample of size n as $X = \{F_{\rm NUV,i}, L_i\}_{i\in[1,n]}$ where L_i is equal to 1 if life is detected and 0 otherwise, we can calculate the evidence for hypothesis H_i against H_j through the Bayes factor

$$BF_{H_i,H_j} = \frac{P(X|H_i)}{P(X|H_j)}, \tag{4}$$

with $P(X|H_i)$ and $P(X|H_j)$ likelihoods of obtaining the sample X under either hypothesis.

If we define $k = \sum L_i$ and denote Y the random variable that describes it, $H_{\rm null}$ represents the likelihood that the number of planets with life in the sample follows the binomial distribution

$$P(Y = k|H_{\rm null}) = \binom{n}{k} f_{\rm life}^k (1 - f_{\rm life})^{n-k}. \tag{5}$$

Under H_1, Y also follows a binomial distribution, however it is conditioned by $n_\lambda = Card(\{F_{\rm NUV,i} \text{ if } F_{\rm NUV,i} > F_{\rm NUV,min}\}_{i\in[1,n]})$ the number of values of $F_{\rm NUV}$ in the experiment that exceed $F_{\rm NUV,min}$

$$P(Y = k|H_1) = \binom{n_\lambda}{k} f_{\rm life}^k (1 - f_{\rm life})^{n_\lambda - k}. \tag{6}$$

Hence,

$$BF_{H_1,H_{\rm null}} = \frac{P(Y = k|H_1)}{P(Y = k|H_{\rm null})} = \frac{\binom{n_\lambda}{k}}{\binom{n}{k}} (1 - f_{\rm life})^{n_\lambda - n}. \tag{7}$$

Given a sample of planets, where for some of them we have convincing biosignature detections but remaining agnostic on $f_{\rm life}$: What evidence for H_1 and $H_{\rm null}$ can we expect to get? Our Bayes factor (Equation 7) is determined by the unknown variables $f_{\rm life}$ and $F_{\rm NUV,min}$, as well as the number of planets with biosignature detections in the sample k. To compute the distribution of evidences, we repeatedly generated samples under H_1 and $H_{\rm null}$ and computed the Bayes factors $BF_{H_1,H_{\rm null}}$ and $BF_{H_{\rm null},H_1}$. We then evaluated the fraction of Monte Carlo runs in which certain evidence thresholds (Jeffreys 1939) were exceeded.

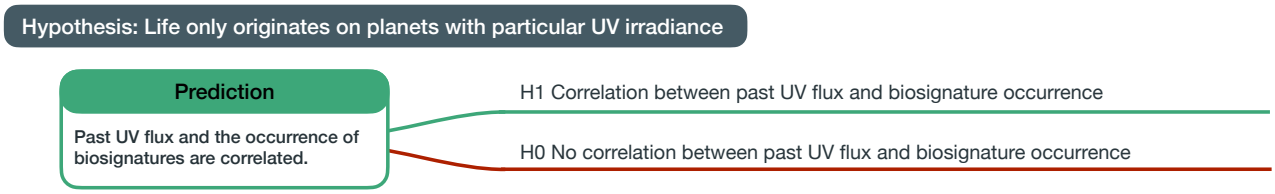

Figure 1. Population-level hypothesis and null hypothesis on UV irradiance derived from the cyanosulfidic scenario.

2.3. *Exoplanet survey simulations with Bioverse*

A real exoplanet survey will be subject to observational biases, sample selection effects, and the underlying demographics of the planet sample. To assess the information gain of a realistic exoplanet survey, we employed Bioverse (Bixel & Apai 2021), a framework that integrates multiple components including statistically realistic simulations of exoplanet populations, a survey simulation module, and a Bayesian hypothesis testing module to evaluate the statistical power of different observational strategies. The general approach is as follows:

1. **Exoplanet population synthesis:** We populate the Gaia Catalogue of Nearby Stars (Smart et al. 2021) with synthetic exoplanets whose orbital parameters and planetary properties reflect our current understanding of exoplanet demographics (Bergsten et al. 2022).

2. **Survey simulation:** We simulate the detection and characterization of these exoplanets with a hypothetical survey, taking into account the survey's sensitivity, target selection, and observational biases. To model the sensitivity of the information gain of a proposed mission to sample selection and survey strategy, we conduct survey simulations with Bioverse using different sample sizes and survey strategies.

3. **Hypothesis testing:** We employ Bayesian statistical methods to evaluate the likelihood that a given survey would detect a specified statistical trend in the exoplanet population and estimate the precision with which the survey could constrain the parameters of that trend. A common definition of the null hypothesis H_0, which is also applied here, is that there is no relationship between the independent variable (here: maximum NUV flux) and the dependent variable (here: biosignature occurrence). The alternative hypothesis H_1 proposes a specific relationship parameterized by θ. If the dependent variable is of binary nature such as in the case of biosignature detections, we model the likelihood function as

$$\mathcal{L}(y \mid \boldsymbol{\theta}) = \prod_i^N \left[y_i h\left(\boldsymbol{\theta}, x_i\right) + \left(1 - y_i\right)\left(1 - h\left(\boldsymbol{\theta}, x_i\right)\right)\right], \tag{8}$$

where here $y_i \in \{0, 1\}$ is the biosignature detection variable, x_i is the maximum NUV flux, and $h(\boldsymbol{\theta}, x_i)$ is the probability of detecting a biosignature given the maximum NUV flux and the model parameters $\boldsymbol{\theta}$. We then use a nested sampling method (Speagle 2020) to compute the Bayesian evidence for the null and alternative hypotheses and estimate the strength of evidence for the alternative hypothesis.

To determine the diagnostic capability of a given survey, Bioverse runs multiple iterations of the simulated survey and calculates the fraction of realizations that successfully reject the null hypothesis. We use this metric, known as the statistical power, to quantify the potential information content of the survey, identify critical design trades, and find strategies that maximize the survey's scientific return. The posterior samples obtained from the nested sampling runs further allows us to estimate the precision with which the survey could constrain the parameters of the hypothesized trend.

2.3.1. *Simulated star and planet sample*

We generated two sets of synthetic exoplanet populations, one for FGK-type stars and one for M-type stars. The stellar samples are drawn from the Gaia Catalogue of Nearby Stars (Smart et al. 2021) with a maximum Gaia magnitude of 16 and a maximum stellar mass of 1.5 $M_\odot$. We included stars out to a maximum distance $d_{\max}$ that depends on the required planet sample size. Planets were generated and assigned to the synthetic stars following the occurrence rates and size/orbit distributions of Bergsten et al. (2022). Following Bixel & Apai (2021), we considered only transiting EECs with radii $0.8\,S^{0.25} < R < 1.4$ that are within the habitable zone (see Section 2.3.2). The lower limit was suggested as a

minimum planet size to retain an atmosphere (Zahnle & Catling 2017). To generate planet samples larger than what the stellar catalog in combination with these occurrence rates yields, we scaled up the occurrence rates by a constant factor that yields the desired number of planets. This was in particular necessary for the FGK sample, where the rate of transiting planets that occupy the habitable zone is low.

For all survey simulations and hypothesis tests, we repeated the above in a Monte Carlo fashion to generate randomized ensembles of synthetic star and planet populations (compare Bixel & Apai 2021).

2.3.2. *Habitable zone occupancy and UV flux*

To test the UV hypothesis, we require that life occurs only on planets with sufficient past UV irradiation exceeding $F_{\mathrm{NUV,min}}$. Further, we require this flux to have lasted for a minimum duration ΔT_{min} to allow for a sufficient "origins timescale" (Rimmer 2023). All common Origin of Life scenarios require water as a solvent; we thus consider only rocky planets that may sustain liquid water on their surface, i.e., that occupy their momentary habitable zone (HZ) during the above period.

Different formulations of habitable zones as regions around a star where a planet with Earth's atmospheric composition can maintain liquid water on its surface exist (e.g., Mol Lous et al. 2022; Spinelli et al. 2023; Tuchow & Wright 2023). CITE! Ramirez & Kaltenegger 2017, 2018 Here, we adopt the popular estimates of Kasting et al. (1993) and Kopparapu et al. (2013, 2014) that define a temperate zone between the runaway greenhouse transition CITE! and the maximum greenhouse limit CITE!. We use the parametrization in Kopparapu et al. (2014) to derive luminosity and planetary mass-dependent edges of the HZ a_{inner} and a_{outer}.

A commonly discussed biosignature is molecular Oxygen (O_2), which on Earth emerged as a byproduct of photosynthesis during the Proterozoic era. No individual component of an atmosphere has been identified as a reliable biosignature in isolation (**?**), and the detection of Oxygen alone will certainly not be sufficient to confirm the presence of life. Nevertheless, we focus here on detecting this key absorber as it represents the general inherent observational challenges and trends.

To determine HZ occupancy, we interpolated the stellar luminosity evolution grid of Baraffe et al. (1998) using a Clough Tocher interpolant (Nielson 1983; Alfeld 1984, see left panel of Figure 3) to compute the evolution of the inner (runaway greenhouse) and outer (maximum greenhouse) edges as a function of planet mass and stellar spectral type (Kopparapu et al. 2014). This provides each planet's epochs within and outside the HZ. For the NUV flux, we use the age- and stellar mass-dependent NUV fluxes in the HZ obtained by Richey-Yowell et al. (2023). We linearly interpolate in their measured grid, where we convert spectral type to stellar mass using the midpoints of their mass ranges ($0.75\,\mathrm{M}_\odot$ for K stars, $0.475\,\mathrm{M}_\odot$ for early-type M stars, and $0.215\,\mathrm{M}_\odot$ for late-type M stars). Outside the age and stellar mass range covered in Richey-Yowell et al. (2023), we extrapolate using nearest simplex (see right panel of Figure 3).

We then determined which planets were both in the HZ and had NUV fluxes above $F_{\mathrm{NUV,min}}$ for $\Delta T_{\mathrm{min}} \geq 10\,\mathrm{Myr}$. We assigned the development of life to a random fraction f_{life} of all temperate planets fulfilling these requirements.

For the hypothesis tests, we then define our alternative hypothesis as

$$H_1 = f_{\mathrm{life}}(\theta, F_{\mathrm{NUV}}) = \begin{cases} 0, & F_{\mathrm{NUV}} < F_{\mathrm{NUV,min}} \\ f_{\mathrm{life}}, & F_{\mathrm{NUV}} \geq F_{\mathrm{NUV,min}} \text{ and in HZ for } \Delta t \geq \end{cases} \tag{9}$$

and the corresponding null hypothesis $H_{null} = f_{\mathrm{life}}(\theta)$, i.e., no correlation with UV flux. We imposed log-uniform priors on the parameters $\boldsymbol{\theta}$, sampling f_{life} from a log-uniform distribution between 10^{-3} and 1 and $F_{\mathrm{NUV,min}}$ from a log-uniform distribution between 10^1 and 10^5.

As shown in Figure 2, the majority of EECs orbit lower-mass stars. The fraction of inhabited planets is highest in the M dwarf sample due to the higher NUV fluxes in the HZ of these stars.

2.3.3. *Transit survey (e.g., Nautilus)*

... measure the maximum past NUV flux with a precision of 5% and the instellation that a planet receives with a precision of 5%.

3. RESULTS

3.1. *Semi-analytical assessment*

In Section 2.2 we computed the probability for true positive evidence for H_1 and H_{null} respectively. Figure 4 shows how these evidences are distributed for sample sizes 10 and 100, and how likely we are to obtain strong evidence ($BF_{H_i,H_j} > 10$). For $n = 10$, strong true evidence for H_1 (H_{null}) can be expected in $\sim 30\,\%$ ($\sim 40\,\%$) of all random experiments. In the majority of cases, the outcome of the survey will be inconclusive. The situation improves with larger samples: for $n = 100$, 80 % of random experiments yield true strong evidence under either H_1 or H_{null}.

The expected resulting evidence further depends on the a priori unknown abiogenesis rate f_{life} and on the NUV flux threshold. Figure 5 illustrates this depen-

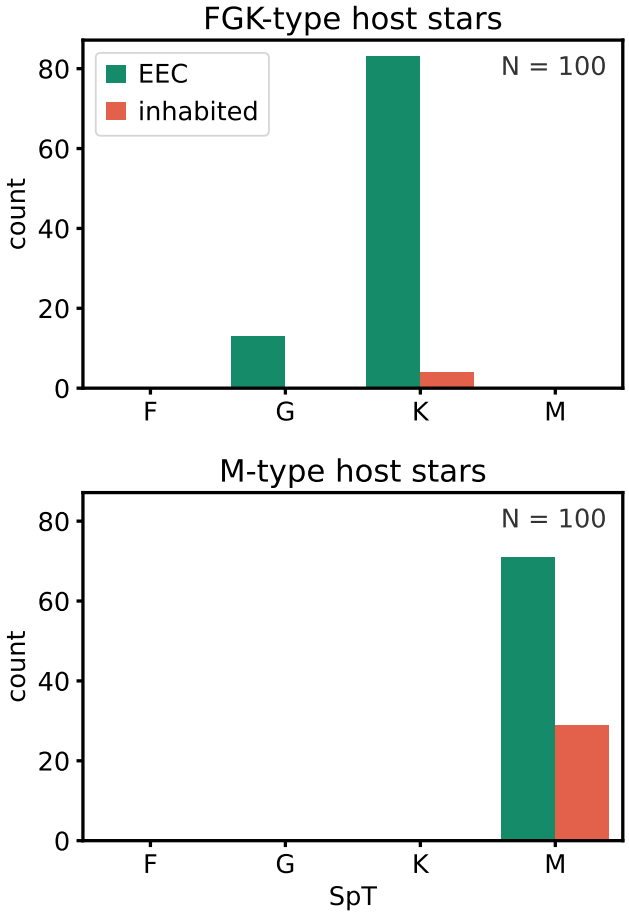

Figure 2. Host stars of all transiting EECs and inhabited planets in a simulated transit survey. In the FGK sample, most EECs and all inhabited planets orbit K dwarfs. In an M dwarf sample of the same size, the fraction of inhabited planets is larger.

dency: For very low values of either parameter, samples drawn under the null or alternative hypotheses are indistinguishable and the Bayesian evidence is always low. Both higher $f_{\rm life}$ and higher NUV flux thresholds increase the probability of obtaining strong evidence. Larger sample sizes enable this at lower values of these parameters.

So far, we have assumed random, uniform distributions of $f_{\rm life}$, $F_{\rm NUV,min}$, and $F_{\rm NUV}$. A high biosignature detection rate $f_{\rm life}$ increases the evidence (cmp. Equation 7) but we cannot influence it. The same is true for $F_{\rm NUV,min}$, where again higher values increase the evidence as the binomial distribution for H_1 gets increasingly skewed and shifted away from the one for $H_{\rm null}$. The distribution of $F_{\rm NUV}$ in the planet sample, on the other hand, can be influenced by the survey strategy. A targeted sampling approach could be to favor extreme values of $F_{\rm NUV}$ in the sample selection. Figure 6 shows how the probability of obtaining true strong evidence for H_1 scales with selectivity s, where $s \in]-1, 1[$ such that $F_{\rm NUV} \sim Beta(1/10^s, 1/10^s)$. Here, $s = 0$ corresponds to a random uniform distribution. Compared to this case, a high selectivity can increase the probability of obtaining true strong evidence to $\gtrsim 85\,\%$ for large samples.

3.2. *Survey simulations with Bioverse*

In a magnitude- and volume-limited sample of a transit survey, the host star distribution will be skewed toward later spectral types and dominated by M dwarfs (see Figure 2). Due to how the HZ scales with spectral type, by far most transiting EECs occur around M dwarfs. Their NUV fluxes are generally highest at early times $\lesssim 100\,$Myr. These host stars, in particular late subtypes, also provide extended periods of increased NUV emission that overlap with times when some of these planets occupy the HZ (see Figure 3), our requirement for abiogenesis (compare Equation 9). Because of that, all inhabited planets in our magnitude- and volume-limited sample orbit M dwarfs.

Here, we are interested in the statistical power of a transit survey with a realistic sample selection and size. In the following, we fix the sample size to 100 and consider two different survey strategies targeting FGK and M dwarfs, respectively. We further investigate the sensitivity of the survey to the a priori unknown threshold NUV flux $F_{\rm NUV,min}$ and the abiogenesis rate $f_{\rm life}$.

3.2.1. *Selectivity of simulated transit surveys*

In Section 3.1, we demonstrated that the probability of obtaining true strong evidence for the hypothesis that life only originates on planets with a minimum past NUV flux is sensitive to the distribution of sampled past NUV fluxes, i.e., the selectivity of the survey (compare Figure 6). For both surveys targeting M dwarfs and those targeting FGK dwarfs, the maximum NUV distribution is rather unimodal. Applying the approach from Sect. 3.1 of fitting a Beta function to the distribution, we find rather low selectivities (see Figure 7).

3.2.2. *Expected biosignature pattern in a transit survey*

A representative recovery of the injected biosignature pattern is shown in Figure 8. There, we assumed an abiogenesis rate of $f_{\rm life} = 0.99$ and a minimum NUV flux of $F_{\rm NUV,min} = 250.0\,{\rm erg\,s^{-1}\,cm^{-2}}$. All injected biosignatures are assumed to be detected, and the maximum NUV flux is estimated from the host star's spectral type and age with an uncertainty corresponding to the intrinsic scatter in the NUV fluxes in Richey-Yowell et al. (2023). This leads to a distribution of biosignature de-

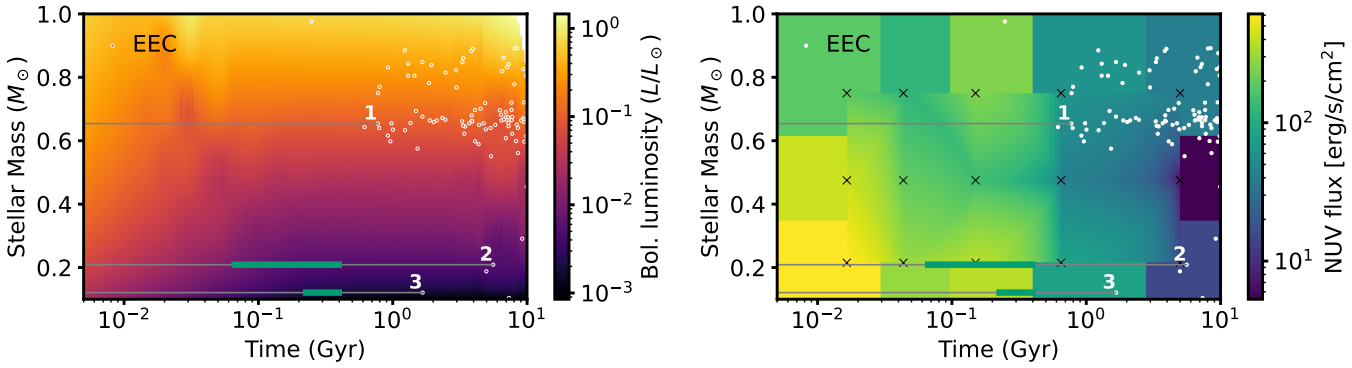

Figure 3. Interpolated stellar luminosity evolution (left) and evolution of the NUV flux in the HZ (right) as a function of host star mass. The scatter plots show age and host star mass of the transiting planets in the synthetic FGK sample; crosses denote the estimated NUV values in Richey-Yowell et al. (2019). A few example tracks for an example threshold flux of $F_{\mathrm{NUV,min}} = 250.0\,\mathrm{erg\,s^{-1}\,cm^{-2}}$ are shown; extended overlap of HZ occupancy and high NUV flux (green sections) fulfills our requirement for abiogenesis. Planet 1 is an EEC that never receives sufficient NUV flux for abiogenesis. Planet 2 and Planet 3 enter the HZ at different times and receive sufficient NUV flux for different durations until their respective host star evolves below the threshold.

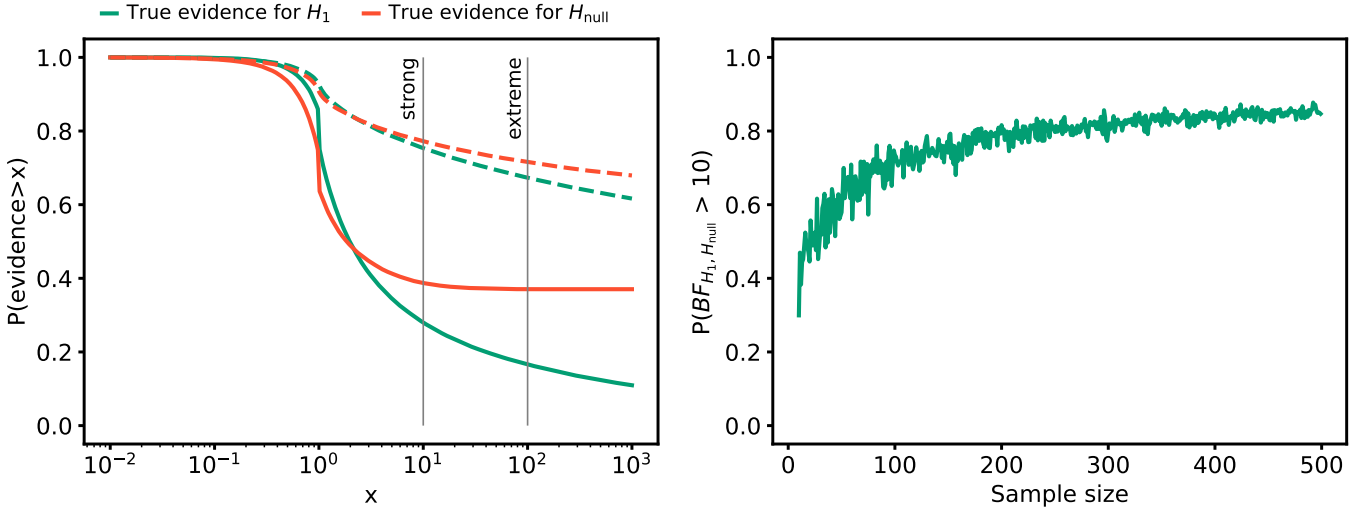

Figure 4. Probability to obtain true strong evidence. Left: evidence levels for H_1 and H_{null} under sample sizes $n = 10$ (solid) and $n = 100$ (dashed). The vertical lines denote the thresholds for "strong" evidence, $BF_{H_i,H_j} > 10$, and "extreme" evidence, $BF_{H_i,H_j} > 100$. Right: Probability of true strong evidence for H_1 as a function of sample size n.

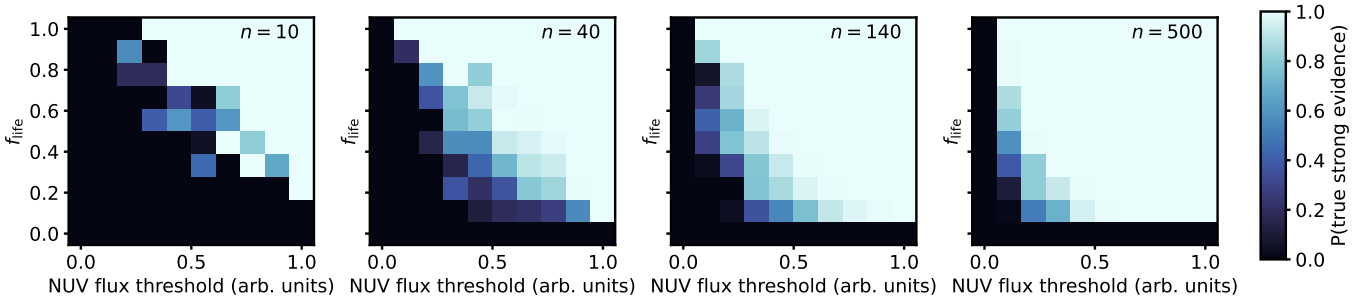

Figure 5. Probability of obtaining true strong evidence for different abiogenesis rates, NUV flux thresholds, and sample sizes. For each of these parameters, higher values increase the probability of yielding strong evidence.

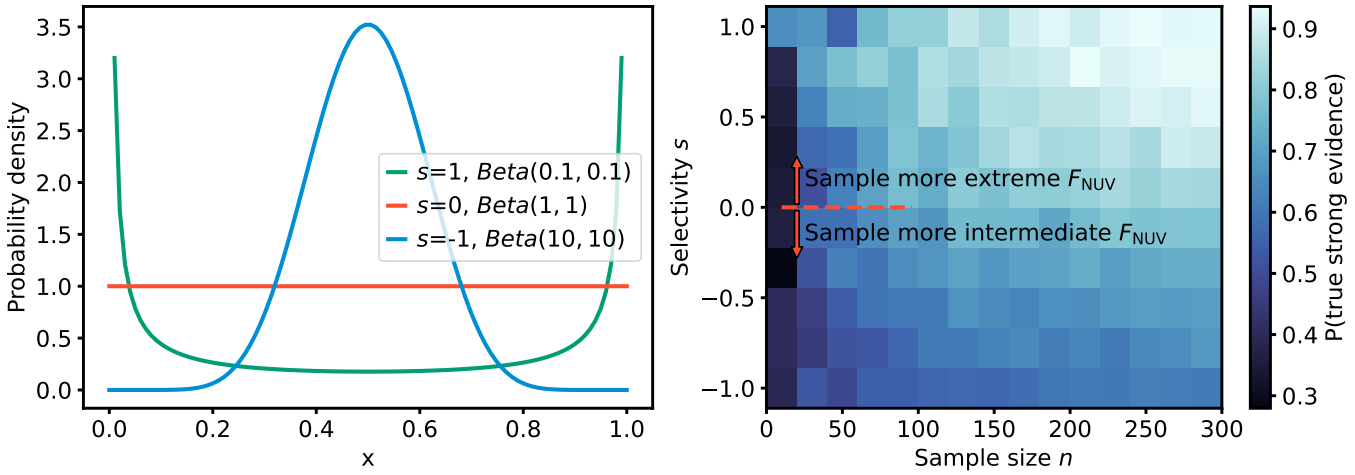

Figure 6. Scaling of the probability of obtaining true strong evidence with sample selectivity. Left: Sampling distribution for different selectivity parameters s. Right: Resulting P(true strong evidence) (random, uniform $f_{\rm life}$, $F_{\rm NUV,min}$). Sampling more extreme values of $F_{\rm NUV}$ is more likely to yield strong evidence.

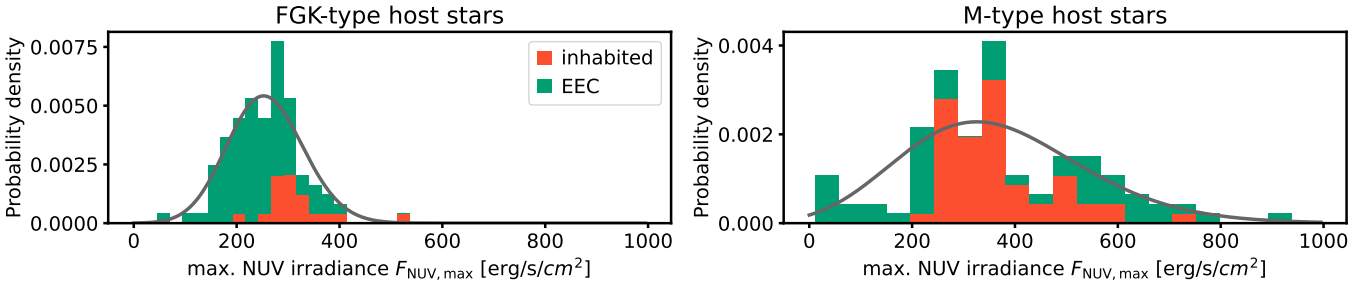

Figure 7. Distribution of maximum past NUV flux in transit surveys targeting FGK and M stars, respectively. The best-fit beta distributions (gray) correspond to selectivities of $s_{\rm FGK}$ =−3.0 and $s_{\rm M}$ =−8.3.

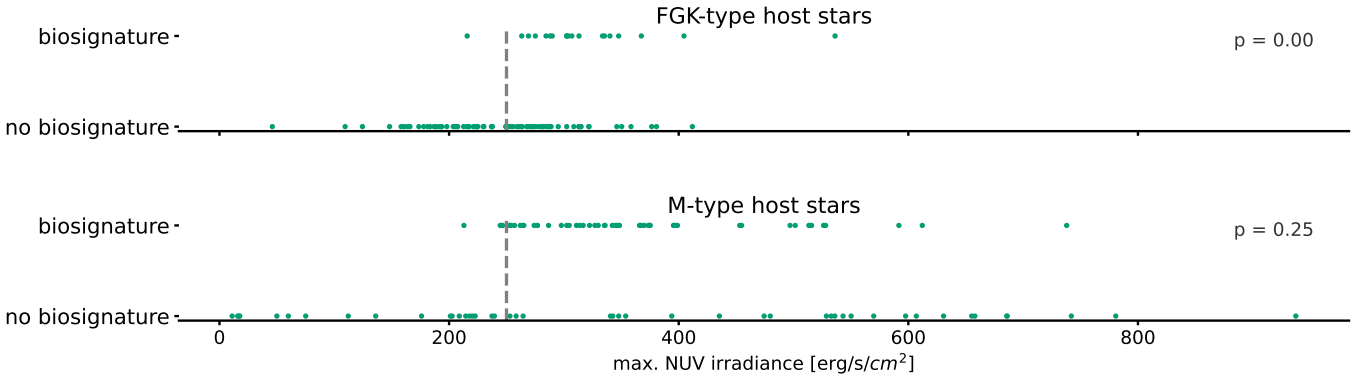

Figure 8. Recovered biosignature detections in the NUV flux-biosignature occurrence space. The dashed line denotes a generic threshold NUV flux $F_{\rm NUV,min} = 250.0\,{\rm erg\,s^{-1}\,cm^{-2}}$.

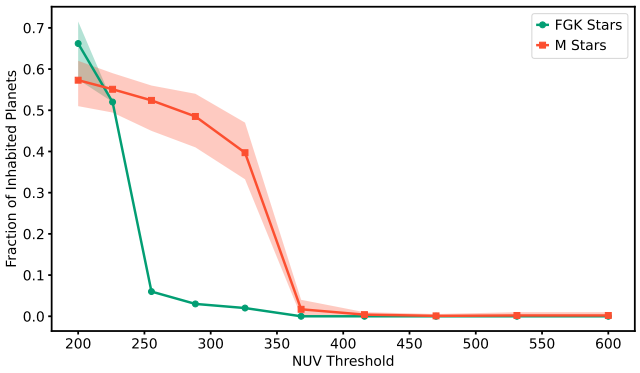

Figure 9. Fraction of inhabited planets for different threshold NUV fluxes under the UV hypothesis if the abiogenesis rate is 0.99.

tections with detections increasingly occurring above a threshold inferred NUV flux. In this example case, the few biosignature detections in the FGK sample lead to a higher evidence ($dlnZ_{\rm FGK} = -0.09$) than in the M dwarf sample ($dlnZ_{\rm M} = -0.42$), where the majority of planets are above the threshold NUV flux.

Figure 9 shows the fraction of inhabited planets under the UV hypothesis for different threshold NUV fluxes and for a high abiogenesis rate of 0.99. This fraction decreases sharply with increasing threshold flux, as fewer planets receive sufficient NUV flux for abiogenesis. For the FGK sample, the fraction of inhabited planets drops at lower threshold fluxes than for the M dwarf sample.

3.2.3. *Statistical power for a transit survey and sensitivity on astrophysical parameters*

We now investigate the sensitivity of the achieved statistical power of our default transit survey to the a priori unconstrained threshold NUV flux $F_{\rm NUV,min}$ and the abiogenesis rate $f_{\rm life}$. Figure 10 shows the statistical power as a function of these parameters for a sample size of $N = 100$. Values of $F_{\rm NUV,min}$ that lie between the extrema of the inferred maximum NUV flux increase the achieved statistical power of the survey, as in this case the dataset under the alternative hypothesis H_1 differs more from the null hypothesis. The same is true for the abiogenesis rate $f_{\rm life}$, where higher values increase the evidence for H_1.

4. DISCUSSION

4.1. *Constraining power for the origins of life as a function of biosignature location*

How does the location of biosignature detections impact the credibility of OOL scenarios?

biosignature on M dwarf planet vs. FGK: Impact on UV flux requirement?

The cyanosulfidic scenario, in particular its predicted existence of a minimum NUV flux required for prebiotic chemistry, offers an opportunity to test an origins of life hypothesis with a statistical transit survey sampling planets with varying NUV flux histories. It comes to no surprise that the success rate of such a test is sensitive to the sample size of the survey and to the occurrence of life on temperate exoplanets. As we have shown, the statistical power of this test also depends on the distribution of past NUV fluxes in the sample and on the required threshold flux. Optimizing the survey to sample a wide range of NUV flux values, particularly at the extremes, can enhance the likelihood of obtaining strong evidence for or against the hypothesis. Intermediate values of the threshold NUV flux are more likely to yield strong evidence than extreme values, as the dataset under the alternative hypothesis H_1 differs more from the null hypothesis in this case while still being sufficiently populated. The required threshold flux is, of course, a priori unknown and we cannot influence it. If, however, better theoretical predictions for the required NUV flux for abiogenesis become available, the survey strategy can be further optimized, for instance by targeting planets that are estimated to have received a NUV flux slightly below and above this threshold.

An interesting aspect lies in the distribution of host star spectral types: Under the NUV hypothesis, the occurrence of life is expected to correlate with the host star's spectral type, with late-type M dwarfs being favored due to their higher NUV fluxes in the HZ (Richey-Yowell et al. 2023). However, the distribution of maximum past NUV fluxes in this sample is narrow in this sample, which may limit the constraining power of the survey depending on the (unknown) threshold NUV flux. An M dwarf sample may help to test the high NUV flux end of the hypothesis; a higher occurrence of biosignatures here would support the hypothesis that a higher NUV flux is beneficial or necessary for life.

FGK stars, on the other hand, show a wider distribution of maximum past NUV fluxes in the HZ, which may increase the likelihood of obtaining strong evidence for or against the hypothesis. With this sample, the survey will be more sensitive to the low NUV flux end of the hypothesis. A lack of biosignatures on these planets would support the NUV hypothesis, whereas their presence might suggest that lower NUV fluxes are also sufficient for abiogenesis or indicate different abiogenic pathways. CITE

4.2. *What do we learn from a single biosignature detection?*

discuss constraining power on OOL of a convincing biosignature detection on a single planet, depending on the position of the planet in the parameter space we explored.

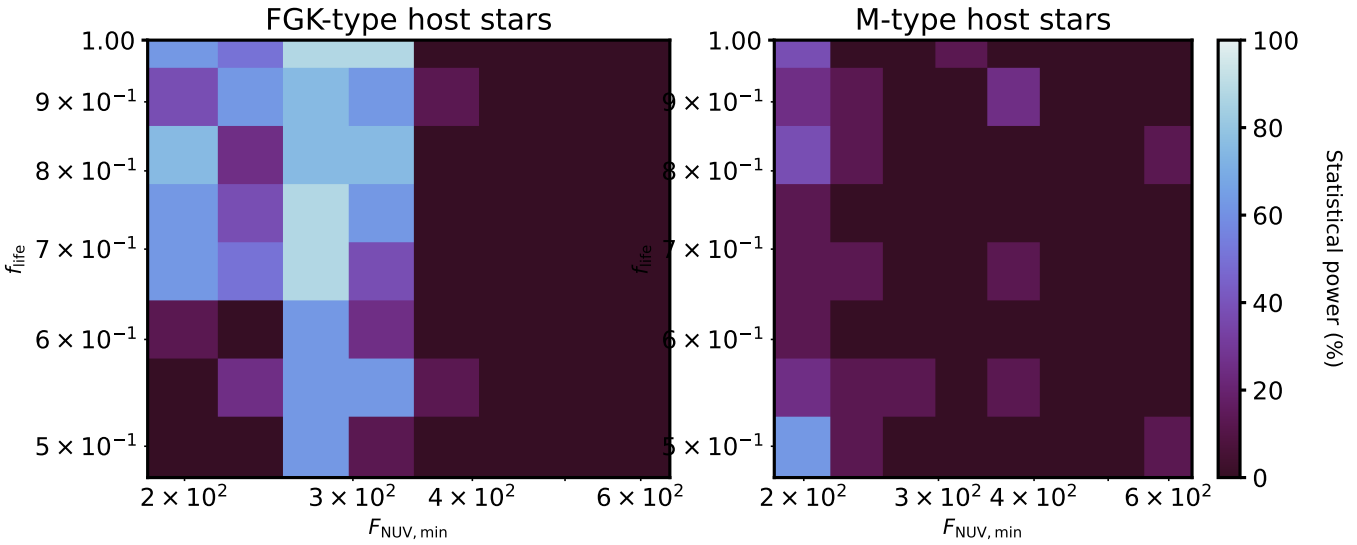

Figure 10. Statistical power as a function of threshold NUV flux and abiogenesis rate. For a given sample size (here: $N = 100$), the achieved statistical power of the survey is enhanced for higher values of $f_{\rm life}$ and intermediate values of $F_{\rm NUV,min}$.

4.3. *Sampling strategy for testing a predicted minimum NUV flux*

We show in Sect. 3 that the constraining power for testing the hypothesis of a minimum past NUV flux required for abiogenesis is sensitive on the occurrence of life, the value of this threshold flux, the sample size, and the distribution of sampled past NUV fluxes. In particular the last parameter offers an opportunity to optimize the survey strategy: Although constraints on a planet's UV history have generally large uncertainties (e.g., Richey-Yowell et al. 2023), the likely maximum flux of a planet can be inferred and used as a proxy. Sampling more extreme (low and high) values for the maximum flux increases the probability of obtaining true strong evidence for the hypothesis. For large sample sizes $\gtrsim 200$, this strategy can push this probability into the 90 % range.

and what about the probability of getting true strong evidence for H0?

4.4. *Contextual support for potential biosignature detections*

discuss how planetary context may impact credibility of tentative biosignature detections (e.g., if there is a good fit with a predicted OOL pattern; or the opposite: the planetary context does not fit well to any OOL scenarios)

As we have shown in Section 3, the interplay of NUV evolution and HZ occupancy strongly favors late spectral types for abiogenesis via the cyanosulfidic scenario. This strong predicted correlation between stellar spectral type and the occurrence of life can be used to falsify Origins of Life (OoL) scenarios. The constraint of this context is particularly strong if a candidate biosignature is detected on a planet orbiting an earlier type star, i.e., where it is unexpected in the context of this OoL scenario.

4.5. *Caveats*

4.5.1. *Atmosphere transmission*

Theoretical work suggests that the atmosphere of prebiotic Earth was largely transparent at near-UV wavelengths with the only known source of attenuation being Rayleigh scattering (Ranjan & Sasselov 2017; Ranjan et al. 2017). We thus approximated surface UV flux using top-of-atmosphere fluxes. This represents a conservative approach, since any planet that fails to meet the irradiance criterion receives even lower near-UV radiation at its surface.

SUKRIT please specify as neccessary

4.5.2. *Stellar flares*

Our assumptions on past UV flux neglect the contribution of stellar flares, which may be hypothesized as an alternative source of UV light (Ranjan & Sasselov 2017). This concerns mainly ultracool dwarfs, due to their low quiescent emission and high pre-main sequence stellar activity (**??**). Recent work indicates that the majority of stars show inadequate activity levels for a sufficient contribution through flares (Glazier et al. 2020; Ducrot et al. 2020; Günther et al. 2020). The biosignature surveys we simulated here may test the hypothesis of sufficient UV radiation via stellar flares.

5. CONCLUSIONS AND FUTURE WORK

We have investigated the potential of upcoming exoplanet surveys to test the hypothesis that a minimum past NUV flux is required for abiogenesis. To this end,

we first employed a semi-analytical Bayesian analysis to estimate probabilities of obtaining strong evidence for or against this hypothesis. We then used the Bioverse framework to assess the diagnostic power of realistic transit surveys, taking into account exoplanet demographics, time-dependency of habitability and NUV fluxes, observational biases, and target selection.

Our main findings are:

1. The required NUV radiation in the cyanosulfidic scenario should lead to a correlation between past NUV flux and current occurrence of biosignatures that will be observationally testable.

2. The required sample size for detecting this correlation depends on the occurrence of life on temperate exoplanets and the distribution of host star spectral types in the sample; in particular their past maximum NUV fluxes.

3. If the predicted NUV correlation exists, yielding strong evidence for it is likely ($> 90\,\%$) for sample sizes ≥ 100, and a survey strategy that targets extreme values of inferred past NUV irradiation.

4. Abiogenesis in the cyanosulfidic scenario prefers late host star spectral types due to the higher NUV fluxes in their habitable zones. However, due to a wider distribution of maximum past NUV fluxes in the HZ of FGK stars, a transit survey targeting FGK systems is more likely to obtain strong evidence for or against the NUV hypothesis.

Overall, our work demonstrates that upcoming exoplanet surveys have the potential to test the hypothesis that a minimum past NUV flux is required for abiogenesis. More generally, we found that models of the origins of life provide hypotheses that may be testable with near-future exoplanet surveys. Our work highlights the importance of understanding the context in which a biosignature detection is made, which can not only help to assess the credibility of the detection but also to test competing theories of the origins of life on Earth and beyond.

REPRODUCIBILITY

REFERENCES

Alfeld, P. 1984, Computer Aided Geometric Design, 1, 169, doi: 10.1016/0167-8396(84)90029-3

Baraffe, I., Chabrier, G., Allard, F., & Hauschildt, P. H. 1998, Astronomy and Astrophysics, v.337, p.403-412 (1998), 337, 403

Bergsten, G. J., Pascucci, I., Mulders, G. D., Fernandes, R. B., & Koskinen, T. T. 2022, The Astronomical Journal, 164, 190, doi: 10.3847/1538-3881/ac8fea

Bixel, A., & Apai, D. 2021, The Astronomical Journal, 161, 228, doi: 10.3847/1538-3881/abe042

Ducrot, E., Gillon, M., Delrez, L., et al. 2020, Astronomy & Astrophysics, 640, A112, doi: 10.1051/0004-6361/201937392

Glazier, A. L., Howard, W. S., Corbett, H., et al. 2020, The Astrophysical Journal, 900, 27, doi: 10.3847/1538-4357/aba4a6

Günther, M. N., Zhan, Z., Seager, S., et al. 2020, The Astronomical Journal, 159, 60, doi: 10.3847/1538-3881/ab5d3a

Jeffreys, H. 1939, Theory of Probability

Kasting, J. F., Whitmire, D. P., & Reynolds, R. T. 1993, Icarus, 101, 108, doi: 10.1006/icar.1993.1010

Kopparapu, R. K., Ramirez, R. M., SchottelKotte, J., et al. 2014, The Astrophysical Journal Letters, 787, L29, doi: 10.1088/2041-8205/787/2/L29

Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, The Astrophysical Journal, 765, 131, doi: 10.1088/0004-637X/765/2/131

Mol Lous, M., Helled, R., & Mordasini, C. 2022, Nature Astronomy, 6, 819, doi: 10.1038/s41550-022-01699-8

Nielson, G. M. 1983, Mathematics of Computation, 40, 253, doi: 10.1090/S0025-5718-1983-0679444-7

Ranjan, S., & Sasselov, D. D. 2017, Astrobiology, 17, 169, doi: 10.1089/ast.2016.1519

Ranjan, S., Wordsworth, R., & Sasselov, D. D. 2017, Astrobiology, 17, 687, doi: 10.1089/ast.2016.1596

Richey-Yowell, T., Shkolnik, E. L., Schneider, A. C., et al. 2019, The Astrophysical Journal, 872, 17, doi: 10.3847/1538-4357/aafa74

—. 2023, The Astrophysical Journal, 951, 44, doi: 10.3847/1538-4357/acd2dc

Rimmer, P. B. 2023, in Conflicting Models for the Origin of Life (John Wiley & Sons, Ltd), 407–424, doi: 10.1002/9781119555568.ch16

Smart, R. L., Sarro, L. M., Rybizki, J., et al. 2021, Astronomy & Astrophysics, 649, A6, doi: 10.1051/0004-6361/202039498

Speagle, J. S. 2020, Monthly Notices of the Royal Astronomical Society, doi: 10.1093/mnras/staa278

Spinelli, R., Borsa, F., Ghirlanda, G., Ghisellini, G., & Haardt, F. 2023, The Ultraviolet Habitable Zone of Exoplanets, doi: 10.48550/arXiv.2303.16229

Tuchow, N. W., & Wright, J. T. 2023, The Astrophysical Journal, 944, 71, doi: 10.3847/1538-4357/acb054

Zahnle, K. J., & Catling, D. C. 2017, The Astrophysical Journal, 843, 122, doi: 10.3847/1538-4357/aa7846