University of Groningen

Kapteyn Astronomical Institute

THE STELLAR ENCOUNTER RESEARCH POTENTIAL OF THE EUCLID MISSION

Towards Integrative Understanding of Encounter-Comet-Climate Events
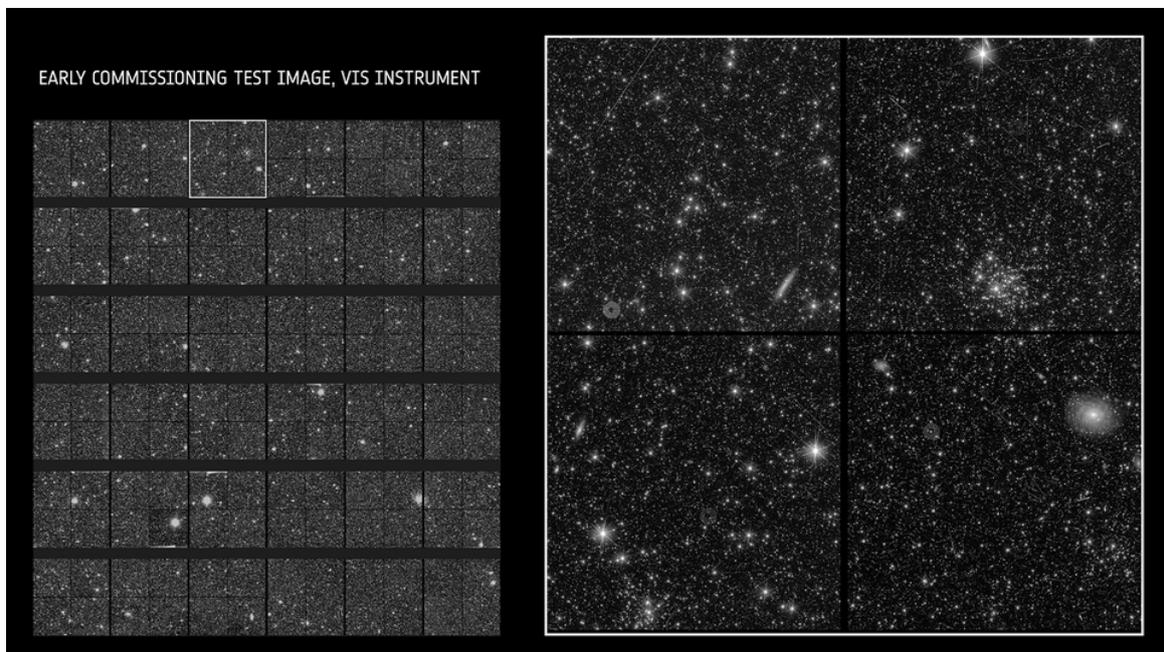
Oscar Stolk, s2768089[2,3]

Supervisors:
Dr. Gijs Verdoes Kleijn[2]
PD Dr. Georg Feulner[4]

Second Examiner:
Dr. Tim Lichtenberg[2]

MSc Coordinator:
Prof. Dr. Eline Tolstoy[2]

February 15, 2024

Abstract

Context
This thesis examines the Earth's climate as an embedded system within the Solar System and the Milky Way. It highlights the influence of comet impacts on Earth's climate, emphasizing the interconnectedness of these systems. The study focuses on the creation of comet showers by Oort Cloud perturbations from close stellar encounters and their subsequent climatic effects on Earth. The thesis presents the Sequence Equation, a framework in which these complex interactions can be interpreted and quantified.

Aims
The potential of the Euclid Mission for contributing to stellar flyby research is analyzed, including a quantification of the astrometric, photometric and spectroscopic data that is necessary to improve significantly on the current state-of-the-art of cataloging 15% of the encounters with perihelia smaller than 5 pc, occurring within ±5 Myr.

Methods
A mock stellar catalog is built with the *Galaxia* galaxy simulation. The completeness is assessed of close stellar encounter candidate samples that can be compiled using the Euclid Wide Survey data in combination with spectroscopic follow-up.

Results
Euclid-only radial distance and proper motion (combining ground- and space-based surveys) measurements suffice to create a subset of selected observations that has a true positive rate, for flyby's within 1 (5) pc of 71% (77%), and a positive predictive value of 7.7×10^{-6} (1.4×10^{-4}). Combining with spectroscopy and precise proper motion measurements potentially yields a true positive rate of 49% (62%) and a positive predictive value of 1.1×10^{-3} (6.2×10^{-3}).

Conclusions
Deep wide-field imaging surveys such as Euclid can contribute significantly to the research on close stellar encounters. This would be useful for refining the outcome of the Sequence Equation, which in this thesis is estimated to indicate between none and a few dozen encounter-comet-climate change events in the Earth's Main Sequence life time.

---

1   Introduction

1.1   Context and Background: Understanding the Earth as an Embedded System

Our home planet is embedded in larger systems. Many significant elements of the Earth System, such as the climate and ecology, are influenced by factors of the Solar System. Besides the effects of the Sun and the Moon, impacting asteroids and comets also contribute to the Earth System. These objects brought water and atmospheric gases to the Earth (Marty 2012; Morbidelli et al. 2012), led to severe extinctions and the evolution of new species[5] (Alvarez et al. 1980; Schulte et al. 2010); and extreme changes in the Earth's climate (Toon et al. 1997). These are only a few examples that indicate that viewing the Earth System as embedded in the Solar System is useful for understanding it.

Likewise, the Solar System—the home of our home planet—is embedded in a larger system. Interactions with other stars and stellar systems contribute to shaping the Solar System. For example, the Sun's birth environment influenced its formation and development (Adams 2010). The focus of this thesis is on the terrestrial effects, through mediation of comets. Within this topic, the Solar System is influenced by the gravitational force of the Milky Way and stars that pass close by or through its outer regions. These influences can disturb the orbits of outer Solar System bodies, such that they are headed towards the inner Solar System. In this way they can impact our planet, affecting the Earth System (Feng & Bailer-Jones 2015; Bobylev & Bajkova 2020). It is also possible that periphery objects of encountering stellar systems fly through the Solar System, possibly hitting and influencing Earth (The 'Oumuamua ISSI Team et al. 2019; Portegies Zwart 2021). So, the embeddedness of the Earth System is two-leveled.

It is valuable for understanding the history, future and current processes of the Earth system and Solar System by looking at their connection with the system that they are embedded in. The investigated systems contain a great number of elements (to name a few examples, the gravitational interactions between myriad stars and between comets and other Solar System bodies, and the oceanic and atmospheric currents) that are involved in even more interactive processes, and researching these elements and processes will increase our understanding of the systems. But looking beyond their confines adds a whole new body of knowledge that yields different insights in the systems. By understanding the place of the Earth in the Solar System and the place of the Solar System in the Milky Way, our understanding of each embedded system will be enhanced.

As a case study in these embedded systems, this thesis delves into the chain of events starting from close stellar encounters, leading to comet influxes, and possibly resulting in global climatic changes on Earth.

Stellar encounters, although relatively rare on a human timescale, have perhaps played a significant role in shaping the Earth's environment throughout its history. These encounters can lead to disturbances in the Oort Cloud, a vast sphere of icy objects surrounding our solar system, potentially sending comets into the inner solar system (Rickman et al. 2008).

The potential impact of these comets on Earth can have profound climatic consequences. The immediate effects of a comet impact, such as a "fire ball" and a tsunami (Toon et al. 1997), are immediately apparent. But also the secondary effects, particularly on the Earth's climate, are significant. These effects can range from short-term climatic disturbances—in the order of years—to long-term shifts in global climate patterns—in the order of millennia, potentially setting a new climatic status quo (Vellekoop et al. 2018). The K-Pg mass extinction (famously including the dinosaurs) triggered

---

[5]There are also theories that posit the origin of life on Earth to be (proto-)lifeforms being transported to here by impactors; the "panspermia" hypothesis. Currently, the scientific consensus is against this hypothesis, however (Wesson 2010; Gobat et al. 2021).

by the Chicxulub asteroid impact clearly shows how large the effect on the Earth System can be (Alvarez et al. 1980; Schulte et al. 2010).

Given the potential significance of these astronomical events on Earth's climate, there is a need for comprehensive research that not only examines each of these processes individually but also explores their interconnectedness. This thesis aims to provide an exploration of some of the factors that connect the individual processes. It distills the processes involved in the flyby-comet-climate chain, outlining the relevant scientific domains and frameworks, and their interfaces. This sets the stage for fine tuned investigations into the details of those processes.

1.2   The Sequence Equation Framework

Researchers in the fields of astronomy and earth sciences recognize the potential interconnections between celestial events and terrestrial climate changes. The concept of linking stellar flyby's, cometary impacts, and their terrestrial consequences, while acknowledged in several studies (for example, Bailer-Jones et al. 2018; Vokrouhlickỳ et al. 2019; Toon et al. 1997), has not been thoroughly explored in an integrated manner. This thesis provides a framework that spans stars, comets, and Earth, investigating these events from multiple disciplinary perspectives.[6]

In this thesis, an analytical framework is developed that offers a quantitative method to assess the likelihood and significance of impact-induced climate change. The framework culminates in an equation that encapsulates each factor of the complex chain of events that begins with close stellar encounters, leads to comet impacts, and potentially leads to global climate effects.

The equation, which I call the Sequence Equation, integrates three critical probabilities: the probability of a flyby occurring ($P(\text{encounter})$), the conditional probability of a comet impact given a flyby ($P(\text{comet} \mid \text{encounter})$), and the probability of global climate change given a comet impact ($P(\text{climate change} \mid \text{comet})$). By multiplying these probabilities, the Sequence Equation provides an estimate of the overall probability of a flyby-comet-climate change event:

$$
\begin{aligned}
P(\text{encounter} \cap \text{comet} \cap \text{climate change}) = {} & P(\text{climate change} \mid \text{comet}) \\
& \times P(\text{comet} \mid \text{encounter}) \qquad (1) \\
& \times P(\text{encounter})
\end{aligned}
$$

The Sequence Equation is inspired by, and functions similarly to, the famous Drake Equation (Drake 1965). It aims to streamline discussions about the influence of larger astronomical systems on the Earth System. It provides handles for relevant factors within this research topic, it serves as a basis to compare results, and it can be used as a tool for identifying research gaps.

The Sequence Equation is a useful tool in understanding Earth's embeddedness in larger astronomical systems. It provides a structured and consistent approach to analyzing how astronomical events influence our planet's climate, thereby contributing to our broader understanding of the Earth system in the context of the universe.

---

[6]The need for a strong and structured interdisciplinary framework is further exposed by uncareful attempts to link terrestrial and astronomical phenomena. Since the development of hypotheses involving impact-induced extinction mechanisms, various publications have attempted to correlate historical mass extinctions in the Earth's geological record with extraterrestrial impacts (for example Davis et al. 1984; Randall & Reece 2014; Rampino & Caldeira 2015). However, these studies have consistently been critiqued for methodological and statistical shortcomings (Bailer-Jones 2009, 2011). A recent example of a publication suffering from questionable multi- and interdisciplinary scientific rigor is Firestone (2021). Georg Feulner (priv. comm.) pointed out that it contains biases in its statistics, an underestimation of volcanic activities' role in mass extinctions, and factual inaccuracies, such as misattributing findings related to the K-Pg extinction to the Permian-Triassic (P-Tr) extinction. The paper's acceptance in an astronomical journal, despite these flaws, underscores the need for more comprehensive, cross-disciplinary research in the context of this topic.

1.3 Objectives

1.3.1 Assessing the Contributions of deep wide-field imaging surveys to Close Stellar Encounter Research

The larger part of the quantitative research of this thesis project is devoted to an analysis of the research possibilities for the *P*(encounter) component of the Sequence Equation. With a thorough investigation of the progress that can be made in close stellar encounter research, the thesis contributes to the line of research outlined by the Sequence Equation framework.

Against this backdrop, the first concrete aim of this thesis is to investigate the extent to which deep wide-field imaging surveys, such as Euclid (Laureijs et al. 2011), can contribute to researching close stellar encounters.

Specifically, the Euclid mission, with its advanced astronomical instrumentation and deep-field observational capabilities, is posited as a case study. The thesis research aims to determine to what extent the photometric and astrometric data collected by Euclid can enhance our knowledge of close stellar encounters. The analysis entails pointing out how many close encounters can be detected with Euclid, how well their perihelion distance and time can be measured, and the required accuracy of supportive radial velocity and proper motion measurements to complement the Euclid data to significantly improve on the current knowledge of the frequency and dynamics of close stellar encounters.

Currently, the most accurate and complete catalog of stellar flyby's consists of Gaia's observations. There are, however, still many close encounters that are not contained in this catalog. In Gaia DR2, the completeness of stars with a perihelion distance smaller than 5 pc within the past and future 5 Myr is about 15% (Bailer-Jones et al. 2018).[7] The potential of Euclid to contribute to a significant improvement of this completeness, possibly at least observing all stars that have a close stellar encounter in the past and future 15 million years, is a key focus of this thesis research.

1.3.2 Estimating significance of stellar flyby's for the Earth System

Secondly, this research aims to crudely quantify the values of the Sequence Equation factors, and estimating the uncertainty range of these values. This includes analyzing the frequency of stellar flyby's, their ability to increase the cometary flux, and the potential of these comets to induce climatic shifts. The number of stellar flyby's is analyzed with the use of a galaxy simulation, building forth on the work of Rybizki et al. (2018). An estimation of the encounter-induced comet flux is distilled from published long period comet population estimates in combination with numerical simulations of Oort Cloud disturbances. The climate consequences caused by an impact are synthesized from various sources discussing the products and effects of asteroids and comets.

Concretely, this objective entails estimating the probability of global climate effects as a result of an encounter-induced comet impact and highlighting the major sources of research uncertainty in applying the Sequence Equation. In this thesis, global climate effects are defined to occur if there is a planet-wide change in near-surface air temperature. The thesis disregards other climate variables such as precipitation and wind speed in order to keep the definition of climate change simple and because the tracers of those other variables are hard to find for past impact-induced climate changes. The definition is consistent with the interpretation of climate change in most geological research. I take the duration of the temperature change to last at least a few years. This is in line with general definitions of the climate and the body of research on the climate consequences of the Chicxulub impact, which makes claims about the post-Chicxulub impact winter of a few years (Brugger et al. 2017) and the subsequent warming of millennia (Vellekoop et al. 2018).

---

[7]No quantitative analyses of the completeness of later Gaia data releases are published, but the recent DR3 (Gaia Collaboration et al. 2023) up to DR5 (expected in 10 years) will add a significant number of close encounters to the catalog, mainly because of enhanced radial velocity measurements (Bailer-Jones 2022).

A crucial part of this research goal involves the estimation of the produced amount and effect of each type of impact ejecta; dust, sulfates, and soot. How much of these climate-active materials is injected into the atmosphere and how strong their climatic influence is, depends mainly on the impact energy of the projectile (Toon et al. 1997, 2016; Artemieva et al. 2017). Therefore, the thesis aims to estimate the impact energy "threshold" above which global climate effects occur.

The uncertainty associated with this threshold is significant. There is considerable spread in the scientific results concerning the production of ejecta and their climatic influence for a given impact energy. Moreover, the climatic consequences of an impact significantly depend on the impact location (Pankhurst et al. 2021; Morgan et al. 2022), adding a layer of irreducible uncertainty. The objective is to indicate the major sources of uncertainty and to find the range of impact energies that could possibly be high enough to cause global climate effects.

Note that most of the published research focuses on the climate effects of an asteroid impact and specifically the Chixculub impact. However, since impact energy is the most important variable for the resulting climate effects, and there are no clear reasons that comets would have different consequences than asteroids (Toon et al. 1997; Pierazzo et al. 2003; Quintana & Schultz 2019), the thesis assumes that the climatic consequences of asteroid impacts of a given energy are indicative for cometary impacts of the same impact energy.

Since there is also a significant uncertainty in the number of close stellar encounters and the distribution of comets that they trigger, the uncertainties for each factor in the Sequence Equation are important in this thesis research. Therefore, the final outcome of this research objective is a range of values (in orders of magnitude) for the flyby-comet-climate chain, the result of the Sequence Equation. In this way, the thesis provides both an estimation of the value of the Sequence Equation, but also an indication of the key sources of uncertainties, which provides guidance for future research within the Sequence Equation framework.

To conclude the introduction, this thesis aims to provide a set-up and start towards a greater understanding of the climate part of the Earth System as embedded in larger systems. It introduces an analytical framework in which to quantify the flyby-comet-climate chain of events and interpret their probabilities, and it gives an estimation of the Sequence Equation value. Moreover, the thesis explores the stellar observation potential of deep-field missions and Euclid in particular, and hence contributes to identifying the opportunities and challenges in close stellar encounter research.

The thesis is structured as follows. Section 2 serves as a primer for the wider thesis research; it describes the processes that constitute and connect the phenomena in the encounter-comet-climate chain. Section 3 gives an account of the methods used for the analysis of the potential of wide-field imaging surveys for researching close stellar encounters; it describes the numerical simulation of the Milky Way and the statistical analysis that was performed. Section 4 contains the results of this analysis. Section 5 discusses the limitations of the analysis. Section 6 points out the key sources of uncertainty in each factor of the Sequence Equation and gives an estimate of its value. Section 7 concludes the thesis by revisiting its highlights and their context.

2 Chain of Events: Close Stellar Encounters, Comets, and Climate Change

2.1 Oort Cloud Disturbance by Close Encounter

The first step in the sequence consists of the potential of close stellar encounters to disturb the Oort Cloud, a distant cloud of icy objects at the edge of our solar system (Oort 1950). These disturbances are crucial in understanding the influx of long period comets into the inner solar system (Rickman et al. 2008; Feng & Bailer-Jones 2014).
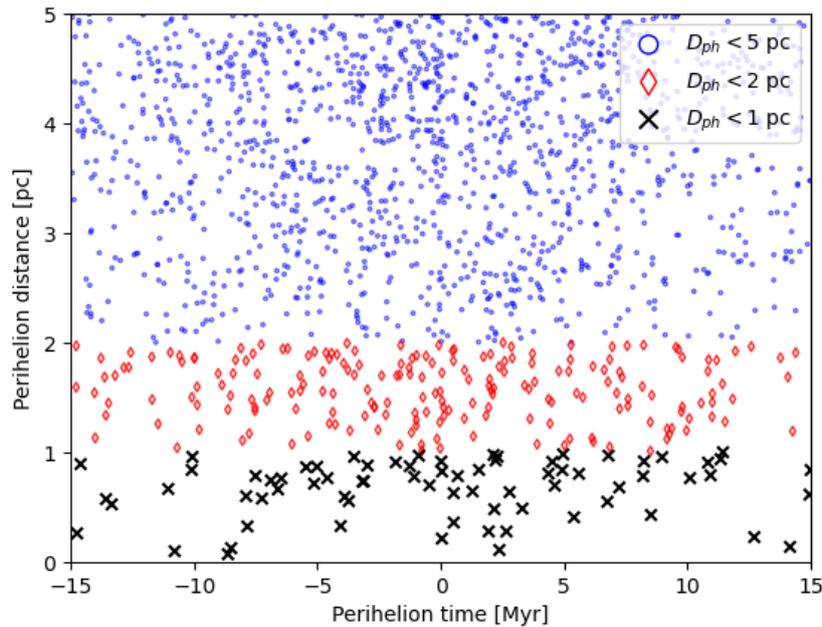
Figure 1: Encounter diagram showing the time and distance of perihelion of (simulated) encountering stars. The encounters are categorized according to their distance of closest approach: 5 pc > d_{ph} > 2 pc (blue circles), 2 pc > d_{ph} > 1 pc (red diamonds), and 1 pc > d_{ph} (black crosses). The data are generated with the *Galaxia* numerical simulation of the Milky Way.

When a star passes in close proximity to our solar system (Bailer-Jones et al. (2018) estimate that at the current rate, each million years about 20 stars pass within 1 pc of the Sun; also see figure 1), its gravitational influence can perturb the orbits of objects in the Oort Cloud. This perturbation can cause a large decrease in perihelion distance of a significant number of these objects, their orbits thereby potentially crossing the inner solar system, including Earth (Dybczyński 2002).

The disturbance of the Oort Cloud by a close encounter can result in a temporarily increased flux of comets, elevating the likelihood of comet impacts on Earth. The frequency and intensity of such comet showers are directly related to the mass of the passing star, and its distance and relative velocity to the Sun (Feng & Bailer-Jones 2015).

The exact significance of close stellar encounters for the production of long period comets in the Oort Cloud has been debated. Some authors suggest that the gravitational influence of the galactic plane, as the Sun moves up and down through it, is the major source of Oort Cloud perturbation and hence comets (Gardner et al. 2011; Wickramasinghe & Napier 2008). Others postulate a massive body inside the Oort Cloud as the factor that tips the objects out of their orbit (Matese & Whitmire 2011).

However, setting aside the other potential origins of long period comets, recent results show that close encounters of other stars are definitely able to produce significant comet showers. Feng & Bailer-Jones (2015) show with detailed numerical experiments in which the Oort Cloud is perturbed by passing stars that close stellar encounters can spawn streams of comets over relatively short periods of time (see figure 2). As a concrete example, Berski & Dybczyński (2016) estimate with a detailed numerical simulation that Gliese 710, a well-studied future close encounter, will raise the comet flux in the inner Solar System with about an order of magnitude for several million years. Synergy effects from combinations with sequences of flyby's or with the galactic tide potentially further increase the comet production of close stellar encounters (Feng & Bailer-Jones 2014).
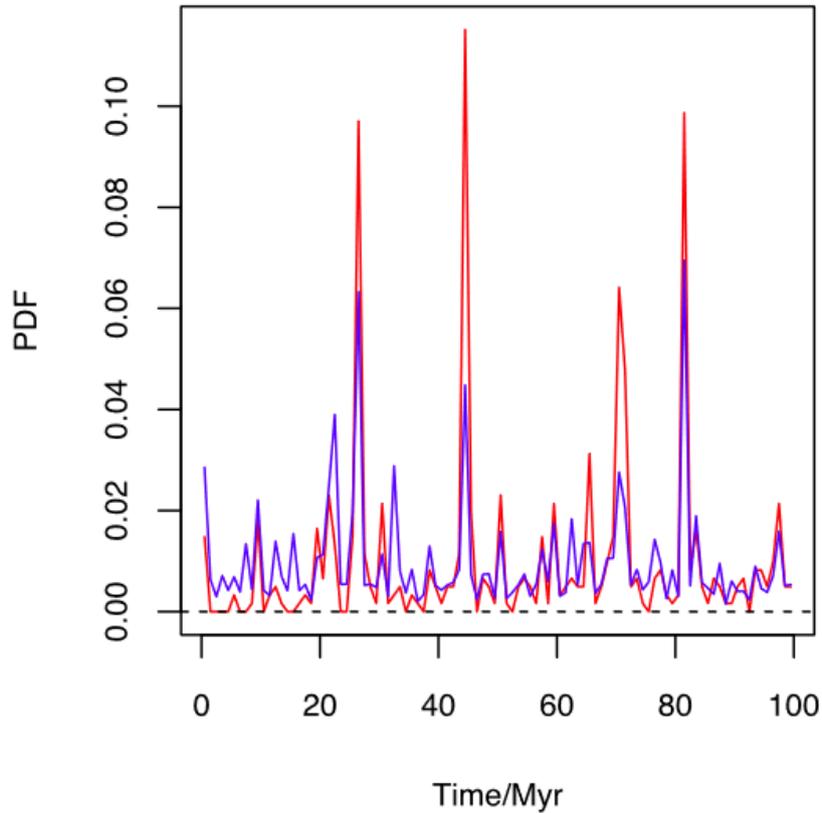
Figure 2: Simulated encounter-induced comet flux probability density over time (red line). The blue line shows the predicted comet flux probability density on basis of each simulated flyby's value of $\gamma = M/(v \cdot r)$, where M is the mass of the encountering star, v its velocity with respect to the Sun and r its minimum distance. The authors note that γ is a reasonable proxy for the number of comets produced by a flyby, and further refine the quantity by replacing the r with r^2 in a later publication (Bailer-Jones et al. 2018). The vertical axis represents the relative injected comet flux in a time bin with a width of 1 Myr, or $N_{inj}/(N_{tot}\Delta t)$ with N_{inj} is the number of injected comets in this bin and N_{tot} is the total number of comets. Figure from Feng & Bailer-Jones (2014).

2.2   Comet Impact Flux

The second step in understanding the processes and significance of the flyby-comet-climate chain of events is the long period comet flux. A detailed account of the number, size and impact frequency of comets that cross Earth's orbit is needed to investigate their influence on the Earth System. The climatic consequences of an impacting body are mainly governed by the composition of the impact location and the projectile's impact energy (Toon et al. 1997, 2016; Artemieva et al. 2017). It is hard to quantify statements about impact location characteristics, but the impact energy distribution can be derived from the impactors' mass and velocity distribution.

Since the terrestrial effects of asteroid impacts are similar to those of comet impacts (Pierazzo et al. 2003; Toon et al. 1997), and hence the research of both topics overlaps, it is useful to compare the long period comet impact flux with the asteroid impact flux. However, the different typical impact probabilities and impact energies of asteroids and comets, due to their different typical orbital characteristics, have to be taken into account.

As indicated in figure 21, the size distribution of asteroids roughly follows a power law. Large asteroids break up in multiple smaller ones, increasing their number exponentially. A similar process occurs with comets, but an elimination mechanism causes their size distribution to consists of three parts (Sosa & Fernández 2011; Vokrouhlický et al. 2019).

Comet nuclei, being loosely bound aggregates of ice and dust, experience strong gravitational forces from the giant planets and the Sun during their journey through the Solar System. Smaller comets are especially vulnerably to disintegration and elimination by these effects and hence are relatively scarce.

Larger comets, with their stronger internal gravity and more stable orbits, are less prone to disintegration, explaining the shallow slope of the distribution for larger diameters. These comets are more likely to remain intact and, thus, stay within the observable population. Conversely, the steep part of the distribution is comprised of many fragments of larger comets, accounting for the sudden rise in numbers (Vokrouhlický et al. 2019; Fernandez & Sosa 2012).

The shape of the cometary size distribution indicates that large ($\gtrsim$ 10 km) comets are relatively abundant compared to asteroids.[8] Those large objects are the ones of interest for this thesis, since they are the ones that are able to cause global climate effects (Toon et al. 1997, 2016).

Furthermore, the impact velocity and the spread in velocity of comets is larger than that of asteroids. The latter have impact velocities between 13 and 41 km s^{-1}, with a mean around 17–22 km s^{-1}; and the former have impact velocities between 16 and 73 km s^{-1}, with a mean around 52–55 km s^{-1} (Steel 1998; Jeffers et al. 2001). Therefore, even though comets have a lower mass density (about 0.5–0.6 g cm^{-3} (Carry 2012; Lowry & Weissman 2003) versus 1–3 g cm^{-3} (Carry 2012)) than asteroids, they have a significant abundance among the energetic terrestrial impactors.

The last factor that increases the relevance of comet impact research is the existence of interstellar comets. The first of these objects, 'Oumuamua, was discovered in 2017 (The 'Oumuamua ISSI Team et al. 2019). Two years later, Borisov, another such interstellar traveler was observed (Guzik et al. 2020). These are probably not rare instances, it is expected that many objects leap from other stellar systems to ours. Estimates based on the above two instances of interstellar interlopers suggest that each year 2–12 such objects come closer than 1 au from the Sun (Portegies Zwart et al. 2018) and they hit Earth every $\sim$ 30 Myr (Do et al. 2018). Especially objects that reside in the outer regions of their stellar system and are weakly bound to their star, like our Oort Cloud bodies, can be easily forced out of their system and encounter the Sun in their trajectory through space. It is also possible that the Sun captures comets from other stellar systems during close stellar encounters (Lingam &

---

[8]Salotti (2022) even estimates that giant comets (larger than 100 km) are around 1.5 times more probable to hit the Earth than asteroids of the same size.It might be good to note, however, that such objects probably do not have significant impact probability in the lifetime of the Earth.

Loeb 2018; Portegies Zwart 2021).

2.3   Impact-Induced Climate Change

The consequences for the Earth System can be enormous when an energetic projectile impacts, potentially including severe changes in the climate. There is one abrupt climate change in the geological record that is confidently linked to our planet being hit by an extraterrestrial body: the crater near Chicxulub, Mexico, associated with the boundary between the Cretaceous (K) and Paleogene (Pg) geological periods (Alvarez et al. 1980; Schulte et al. 2010).

The asteroid strike at Chicxulub caused a global temperature drop of more than twenty degrees Celsius, lasting for several years (Brugger et al. 2017; Vellekoop et al. 2014). The subsequent mass extinction and changes in the planet's carbon cycle left traces in the rock layers of the Earth crust and caused lasting climatological changes, shifting the equilibrium of climate variables (D'Hondt 2005; Kaiho & Oshima 2017; Schulte et al. 2010; Vellekoop et al. 2018).

One of the immediate outcomes of a comet or asteroid impact is the ejection of substantial amounts of dust and debris into the atmosphere. This material first forms an ejecta curtain—mainly solid and molten fragments of the projectile and target rock, expanding outwards from the impact site in a conical shape—and an impact plume—projectile and target material that is vaporized due to the high energy of the impact, rising upwards from the impact site. Consequently, the ejected materials spread through the atmosphere, potentially over the whole globe. This material can have a significant short-term cooling effect by blocking sunlight (Artemieva et al. 2017; Svetsov & Shuvalov 2019). The Chicxulub impact, for instance, led to the formation of a dense, opaque shroud that significantly reduced solar radiation reaching Earth's surface (Morgan et al. 2022; Kaiho et al. 2016).

In addition to dust, sulfate aerosols formed from the sulfur compounds released during an impact can lead to atmospheric cooling. These aerosols are capable of scattering sunlight, contributing to a decrease in Earth's surface temperature.[9] The magnitude of this cooling effect is contingent on the amount of sulfur released and the resulting concentration of sulfate aerosols in the atmosphere. The Chicxulub impact, in particular, injected a massive amount of sulfate aerosols due to the sulfur-rich target rock at the site of impact, which significantly contributed to global cooling (Pierazzo et al. 2003; Ohno et al. 2014).

Soot, resulting from wildfires ignited by the impact or from the burning of organic material in the target rock, also plays a crucial role in climate change. In the case of the Chicxulub impact, the thermal radiation emitted by the ejecta plume was intense enough to set off wildfires, adding enormous quantities of soot to the atmosphere (Svetsov & Shuvalov 2019). Soot can further contribute to short-term cooling when in the upper atmosphere by blocking solar radiation (Kring & Durda 2002; Morgan et al. 2013).

The reduction in sunlight, caused by the combination of dust, sulfate aerosols, and soot, has a significant impact on photosynthesis. Since photosynthesizing organisms play a crucial role in shaping the composition of the atmosphere, including its concentration of oxygen and greenhouse gases such as carbon dioxide and water, they influence the Earth's climate. Therefore, a disruption of photosynthesis due to post-impact darkening can change the global temperature in the long run. A decrease in photosynthesis can also lead to broader ecological consequences, affecting terrestrial and aquatic food chains. The Chicxulub impact, for instance, resulted in a catastrophic drop in global net primary productivity (the rate at which all organisms on Earth produce chemical energy from sunlight), severely impacting both land and marine ecosystems (Schulte et al. 2010; Henehan et al. 2019; Hull et al. 2015).

Furthermore, the rapid global cooling induced by these aerosols and particulates can have direct

---

[9]The absorption of heat by aerosols might contribute to the cooling as well.

physiological effects on many organisms. Species may struggle to adapt or migrate quickly enough to survive the sudden temperature changes, leading to widespread mortality. This was evident in the case of the Chicxulub impact, where large animals, in particular, faced severe challenges due to the rapid environmental shifts (Archibald 1996; Longrich et al. 2012). Additionally, the atmospheric alterations following an impact can lead to increased acidification, primarily due to the high content of sulfates in the ejecta, further deteriorating the conditions for life on Earth (Ohno et al. 2014).

The impacts of close stellar encounters and subsequent comet flux on Earth's environment underscore the importance of understanding these astronomical events. An important step in increasing this understanding and working out the Sequence Equation is to research the frequency of stellar flyby's and their potential to perturb the Oort Cloud. The major part of the quantitative research of this thesis is dedicated to such research, exploring the potential of Euclid for identifying candidates for close stellar encounters. The next section details the methodology employed for this analysis.

3   Methods: Creating a Mock Euclid Stellar Catalog[10]

Transitioning from the theoretical background about the implications of close stellar encounters and their impact on Earth, this section provides the start of delving into the practical aspects of identifying such flyby's. The Euclid mission's Wide Survey forms the central case study of this analysis.

3.1   Properties of the Euclid mission

The Euclid mission is aimed at probing galaxies, but its photometric properties are also suited for stellar observations. It combines ground-based near-ultraviolet (near-UV) and visual observations with space-based photometers operating in the visual and near-infrared (near-IR) spectrum. This integration results in a comprehensive coverage across eight or nine bands, spanning wavelengths from approximately 3,000 Å or 4,000 Å to over 16,000 Å. The bands potentially include u, g, r, i, z from ground-based observations, and VIS, Y, J, H from the spacecraft. All surveys conducted in these bands are characterized by limiting AB magnitudes exceeding ~23 (Laureijs et al. 2011).

The survey area covered by Euclid's wide field is expansive, covering almost 15,000 square degrees. Such a vast survey area implies a significant capture of stellar data. However, the mission strategically avoids the Milky Way's stellar disk in its line of sight. This approach avoids galaxy shape measurements being complicated by overlapping PSFs due to stellar crowding. Furthermore, the survey field deliberately excludes the ecliptic plane of the Solar System (as illustrated in figure 3) to mitigate interference from nearby dust reflecting sunlight, which could impede the imaging of galaxies because the near-infrared radiation contained in zodiacal light is significant compared to the faint near-infrared signal of galaxies. Despite the measures that reduce stellar observations, I will show that Euclid detects over a hundred million stars and that it is a useful mission for identifying stellar flyby's.

3.2   The *Galaxia* model, simulation set-up and sample selection

The *Galaxia* model, as described by Sharma et al. (2011), represents a useful software tool for conducting synthetic surveys of the Milky Way. This code is adept at generating a comprehensive catalog of stars within our galaxy, configured by various parameters that define the Milky Way's structure and composition. The model's flexibility allows for input in the form of star formation rates, age-metallicity relations, and age-velocity dispersion relations, among others. Significantly, the model

---

[10]The code that was used for the analysis is available at `https://gitlab.astro.rug.nl/stolk/stellar-encounters-and-climate-change`.
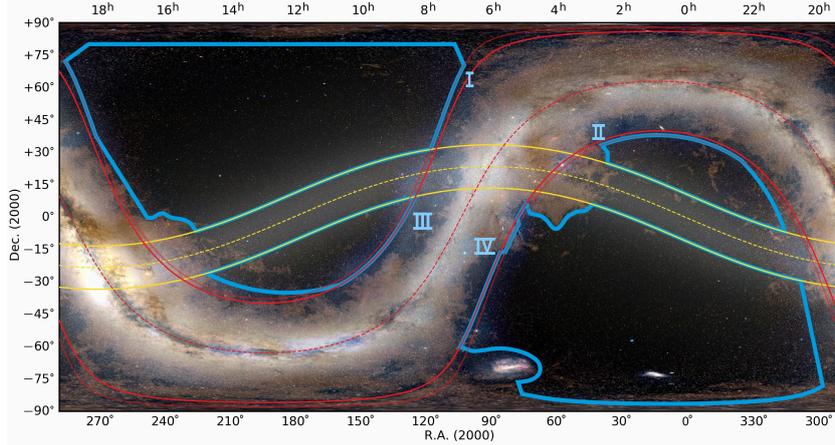
Figure 3: The planned survey area of Euclid. The boundaries of the four regions of interest draw the contours of the Milky Way disk and the ecliptic plane in order to avoid as much interfering light from stars and dust as possible. From Euclid Collaboration et al. (2022).

can specify density distributions or N-body realizations, providing a versatile approach to simulating the galactic environment. Utilizing theoretical isochrones, *Galaxia* can generate a detailed catalog of stars, inclusive of a wide range of photometric bands. This feature is particularly relevant for our research, as it aids in the precise simulation of an observational catalog.

The implementation of the Besançon Milky Way model (Robin et al. 2003) for the disk, along with the simulated stellar halo N-body models of Bullock & Johnston (2005), further demonstrates the model's comprehensiveness. Notably, *Galaxia* includes a scheme to disperse the stars that are spawned by an N-body particle in such a way that the resulting phase space density of the spawned stars is consistent with that of the N-body particles. Because of this, the simulation can sample stars from N-body simulations while maintaining accurate stellar positions and velocities.

The galaxy is modeled as comprising distinct components, including the thin disk, the thick disk, the stellar halo, and the bulge. Each of these components is characterized by a distribution function, defining the number density of stars based on various parameters such as position (r), velocity (v), Milky Way age (τ), metallicity (Z), and mass (m) of stars. This distribution function is denoted as $f_j = f_j(r, v, \tau, Z, m)$, where j represents the specific galactic component. With the distribution functions written out, the analytical model takes the form:

$$f = \frac{\Psi(\tau)}{\langle m \rangle} \xi(m, \tau) f_{\text{pos}}(r, \tau) f_{\text{vel}}(v, r, \tau) f_Z(Z, r, \tau) \tag{2}$$

The model accounts for the star formation rate ($\Psi(\tau)$) and the initial mass function (IMF) as a function of galactic age ($\xi(m, \tau)$). The spatial distribution of stars ($f_{\text{pos}}(r, \tau)$) is also incorporated as a function of age. Furthermore, the velocity distribution ($f_{\text{vel}}(v, r, \tau)$) and the metallicity distribution ($f_Z(Z, r, \tau)$) are considered as functions of both age and position. The IMF in the model is normalized, with the mean stellar mass represented as $\langle m \rangle$.

In the *Galaxia* model, the Sun is assumed to be located 8 kpc away from the Galactic Center. This assumption slightly differs from the Besançon model, which places the Sun at a distance of 8.5 kpc from the Galactic Center. Furthermore, the Sun is assumed to be situated 15 pc above the Galactic Plane. This is in the center of the most commonly noted values, ranging from 8 to 35 pc (Perryman 2009).

The exact geometry, age and metallicity of the components are noted in tables 1 and 2. The thin disk, spanning an age range of 0–10 Gyr, exhibits a density distribution that varies with age. This distribution is defined by exponential functions based on specific age groups, with adjustments for

flare effects and parameters like scale heights. The Initial Mass Function (IMF) within the thin disk changes according to mass, adhering to a power-law form. The thick disk, assumed to be of a constant age of 11 Gyr, has a density distribution that modifies with the vertical height z, integrating a radial scale length and a vertical scale height. Its IMF follows a specific power-law form.

In modeling the galactic bulge, with a constant age of 10 Gyr, the *Galaxia* model employs a bar-shaped (G2 density) distribution. This distribution has a core at the center and elongates along one axis. The bulge's kinematics are characterized by solid body rotation, with velocity dispersions detailed in cylindrical coordinates.

The stellar halo component is dual-faceted, consisting of numerical and analytic models. The simulated halo uses the Bullock & Johnston (2005) models within the ΛCDM cosmological framework, encompassing the accretion of individual satellites and their star formation histories. Additionally, an analytical smooth halo is modeled following an oblate power-law distribution with defined ellipticity and power-law index values.

In the *Galaxia* model, theoretical stellar isochrones are a fundamental component. These isochrones are crucial for assigning specific stellar parameters such as luminosity, effective temperature, magnitude, and color. These parameters are assigned based on a star's age, metallicity, and mass. *Galaxia* employs the Padova isochrones, developed by Marigo et al. (2008), Marigo & Girardi (2007), Girardi et al. (2000), and Bertelli et al. (1994). These isochrones stand out for their comprehensive coverage, spanning a wide range of photometric systems. They are notably inclusive of phases in stellar evolution such as the asymptotic giant branch and the red clump.

For my research, I utilized the CMD web interface[11] for retrieving theoretical isochrone tables. In order to align my simulations as closely as possible with the combined space and ground-based Euclid data, I selected the `Euclid VIS + NISP (ABmags)` photometric system, supplemented by the `SDSS ugriz` photometric system. The rationale behind this selection lies in the similarity of the SDSS *ugriz* photometry to the ground-based Euclid *ugriz* photometry. This alignment ensures that my simulations represent the actual observational data that are being obtained from the Euclid mission.

In my simulation, I focused on all stars within a heliocentric distance of a maximum of 3 kpc. This decision was driven by the need to reduce the computational load. For the same reason, and in order to confine the project's complexity within a master thesis research, I concentrated on investigating close stellar encounters within the past and future 15 Myr. This time window is sufficiently small to ensure that the linear motion approximation (LMA) method yields accurate enough results (see section 3.4.1), thereby avoiding the need for more complex and computationally intensive investigation methods. Furthermore, a window of 15 Myr in the past and future aligns with other research in this field, such as the work by Bailer-Jones (2018).

The range of 3 kpc is strategically chosen as it captures all close encounters within this time window, because a star must have a radial velocity exceeding 200 km s^{-1} to approach the Sun within this timeframe if it has a larger radial distance. Such a high radial velocity is not realistic, making the 3 kpc range a practical and efficient choice for our simulation.

In my simulation, I selected the stars that represent the Wide Survey field of the Euclid mission. I simulated all stars with a galactic latitude (*b*) outside +23 and -23 degrees, as the Euclid survey field avoids the Milky Way (MW) disk and does not capture stars within these latitudes. Additionally, the survey field avoids the solar system ecliptic between +10 and -10 degrees declination in equatorial coordinates. To facilitate this selection, I converted the galactic coordinates from the *Galaxia* mock catalog to ecliptic coordinates using the `astropy`[12] package, subsequently excluding stars in the aforementioned regions.

Furthermore, I applied a selection based on the brightness of the stars. Observing depth values for

---

[11]`http://stev.oapd.inaf.it/cgi-bin/cmd`, retrieved October 2024

[12]More information: `https://docs.astropy.org/en/stable/`

Component	Age (Gyr)	Density Law $\rho(r, \tau)$
Thin disk	≤ 0.15	$\frac{\rho_c \Psi(\tau)}{k_{\text{flare}} \epsilon(\tau)} \left\{ \exp\left(-\left(\frac{a}{h_{R+}}\right)^2\right) - \exp\left(-\left(\frac{a}{h_{R-}}\right)^2\right) \right\}$ where $h_{R+} = 5000$ pc, $h_{R-} = 3000$ pc IMF- $\xi(m) \propto m^{-1.6}$ for $m < 1 M_\odot$ and $\xi(m) \propto m^{-3.0}$ for $m > 1 M_\odot$
Thin disk	$0.15 - 10$	$\frac{\rho_c \Psi(\tau)}{k_{\text{flare}} \epsilon(\tau)} \left\{ \exp\left(-\left(0.5^2 + \frac{a^2}{h_{R+}^2}\right)\right) - \exp\left(-\left(0.5^2 + \frac{a^2}{h_{R-}^2}\right)\right) \right\}$ where $h_{R+} = 2530$ pc, $h_{R-} = 1320$ pc IMF- $\xi(m) \propto m^{-1.6}$ for $m < 1 M_\odot$ and $\xi(m) \propto m^{-3.0}$ for $m > 1 M_\odot$
Thick disk	11	if $|z| \leq x_l$, $\rho_c \delta(\tau - 11) \exp\left(-\frac{R - R_\odot}{h_R}\right) \times \left(1 - \frac{1/h_z}{x_l \times (2.+x_l/h_z)} \times z^2\right)$ if $|z| > x_l$, $\rho_c \delta(\tau - 11) \exp\left(-\frac{R - R_\odot}{h_R}\right) \times \frac{\exp(x_l/h_z)}{1 + x_l/2h_z} \exp\left(-\frac{|z|}{h_z}\right)$ where $h_R = 2500$ pc, $h_z = 800$ pc, $x_l = 400$ pc IMF- $\xi(m) \propto m^{-0.5}$
Spheroid	14	$\rho_c \delta(\tau - 14) \left(\frac{\text{Max}(a_c, a)}{R_\odot}\right)^{n_H}$ where $a^2 = R^2 + \frac{z^2}{\epsilon^2}$, $a_c = 500$ pc, $\epsilon = 0.64$, $n_H = -2.77$ IMF- $\xi(m) \propto m^{-0.5}$
Bulge	10	if $\sqrt{x^2 + y^2} < R_c$, $\rho_c \delta(\tau - 10) \exp(-0.5 r_s^2)$ if $\sqrt{x^2 + y^2} > R_c$, $\rho_c \delta(\tau - 10) \exp(-0.5 r_s^2) \times \exp\left(-0.5 \left(\frac{\sqrt{x^2 + y^2} - R_c}{0.5}\right)^2\right)$ where $r_s^2 = \sqrt{\left[\left(\frac{x}{x_0}\right)^2 + \left(\frac{y}{y_0}\right)^2\right]^2 + \left(\frac{z}{z_0}\right)^4}$, $R_c = 2.54$, $x_0 = 1.59$, $y_0 = z_0 = 0.424$, $\alpha = 78.9°$, $\beta = 3.5°$, $\gamma = 91.3°$ IMF- $\xi(m) \propto m^{-2.35}$
ISM		$\rho_c \exp\left(-\frac{R - R_\odot}{h_R}\right) \times \exp\left(-\frac{|z|}{h_z}\right)$ where $h_R = 4500$ pc, $h_z = 140$ pc
Dark halo		$\frac{\rho_c}{(1.+(a/R_c)^2)}$ where $R_c = 2697$ pc and $\rho_c = 0.1079$

Table 1: Geometry of Stellar Components. The formulas used are from Robin et al. (2003). Note: (R, θ, z) are the coordinates in the galactocentric cylindrical coordinate system, and $a^2 = R^2 + \frac{(z - z_{\text{warp}})}{k_{\text{flare}} \epsilon(\tau)}$ (for the thin disk).

Component	Age (Gyr)	[Fe/H]	$\sigma_{[Fe/H]}$	$\frac{d[Fe/H]}{dR}$
Thin disk	0-0.15	-0.01	0.12	
	0.15-1	-0.03	0.12	
	1-2	-0.03	0.10	
	2-3	-0.01	0.11	-0.07
	3-5	-0.07	0.18	
	5-7	-0.14	0.17	
	7-10	-0.37	0.20	
Thick disk	11	-0.78	0.30	0
Stellar halo	14	-1.78	0.50	0
Bulge	10	0.00	0.40	0

Table 2: Age and Metallicity Distribution (Mean and Dispersion) of Galactic Components. The values shown are from Robin et al. (2003).

each passband were taken from Euclid Collaboration et al. (2020), specifying the AB mag depth at $10\,\sigma$ for point sources in a 2 arcsecond diameter aperture: 24.50 for u, 24.20 for g, 23.85 for r, 22.96 for i, 22.45 for z, 23.81 for Y, 23.59 for J, 23.13 for H and 24.50 for VIS. Extinction was applied based on the Schlegel et al. (1998) model, with the extinction values for each passband taken from table 6 of Schlafly & Finkbeiner (2011). Then the distance modulus was calculated to yield apparent magnitudes, and the stars were selected that fell within the sensitivity range of at least one of the Euclid filters.

3.3   Inspection of the Euclid mock stellar catalog

To ensure the reliability of my simulated data and the integrity of my computational code, it is essential to inspect the mock stellar catalog generated using the *Galaxia* model. This inspection process involves comparing the distributions of various stellar properties in the catalog both before and after applying the Euclid mission's magnitude cut-off and (RA, Dec) selection criteria. My approach encompasses both qualitative and quantitative analyses to ascertain the consistency of the mock catalog with the expected characteristics of the Milky Way as per the input continuous functions.

3.3.1   Stellar positions

Figure 4 illustrates the distribution of mock stars in heliocentric (x, y), (x, z) and (y, z) planes, with a distinct visualization of the subset detectable by the Euclid Wide Survey. This visualization allows for a qualitative assessment of the spatial distribution of stars in the Milky Way. Clearly visible is the absence of the Milky Way disk line-of-sight direction stars in the Euclid selection (orange). The absence of the ecliptic plane in the survey field is not discernible in this figure, but was verified to be visible in a (overcrowded) version of the plot that contained many more stars. The galactic center is in the $+x$-direction, consistent with the stellar density increase. The z-dimension is perpendicular to the Milky Way plane, as can be seen from the steep drop-off in density at a few hundred parsecs above and below $z = 0$.

Figure 5 extends this analysis by plotting the positions of the Euclid sample of stars in right ascension (RA) and declination (Dec) coordinates. This plot is crucial for assessing the effectiveness of the Euclid (RA, Dec) selection in identifying a representative sample of stars. It shows that the positional selection in the mock catalog matches the survey field depicted in figure 3.
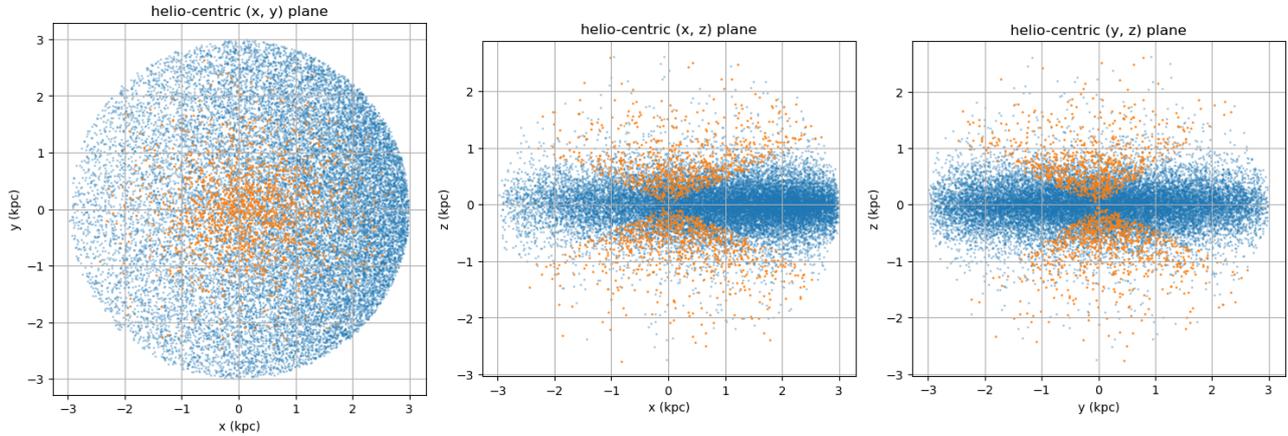
Figure 4: The simulated spatial distribution of stars in heliocentric (x, y), (x, z), and (y, z) planes. The positive *x*-axis points towards the galactic center and the *z*-dimension is perpendicular to the Milky Way plane. Stars within the Euclid Wide Survey's detection limits are indicated in orange, contrasting with the entire population of mock stars in blue. These distributions provide a visual validation of the Euclid Survey's coverage and sampling consistency relative to the modeled Galactic structure. For clarity a random 10^{-5} fraction of the mock stars are shown in this figure and in figure 5.
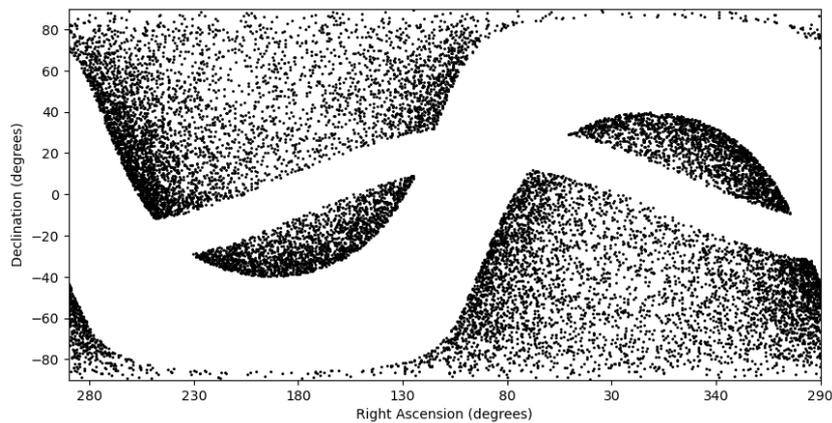
Figure 5: Simulated sky positions (RA, Dec) of stars detected in the Euclid Wide Survey. It shows the location of the mock stars produced with *Galaxia* after applying the Euclid mission magnitude cut-off and (RA,Dec) selection of the Euclid Wide Survey. Comparison with figure 3 shows that the Euclid mock sample closely resembles the planned Euclid survey area.
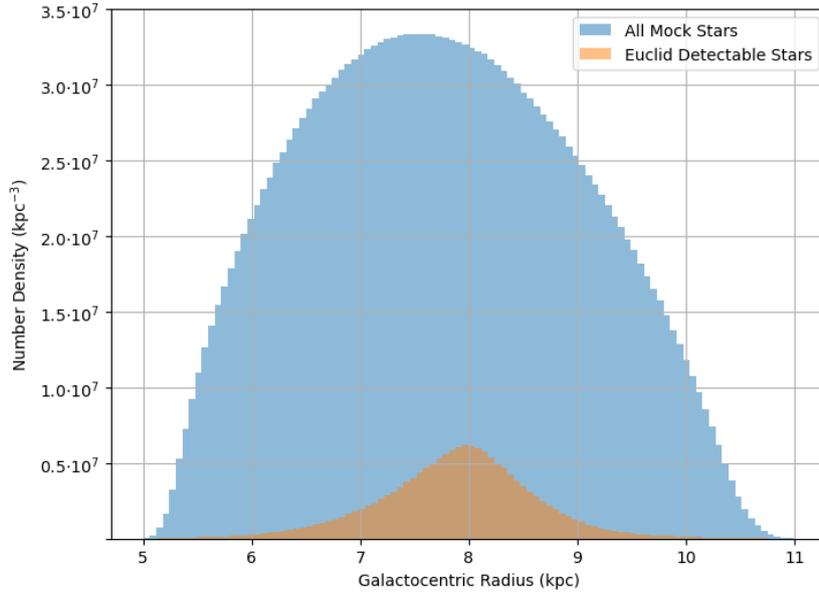
Figure 6: Number density of all simulated stars in the *Galaxia* model (in blue) as a function of Galactocentric radius, overlaid with the subset detectable by the Euclid Wide Survey (in orange). The asymmetrical distribution indicates the increasing stellar density towards the galactic center. The overall shape of the density distribution of all stars reflects the simulated spherical volume centered at the Sun with a radius of 3 kpc. The shape of the Euclid subset distribution has a steeper drop-off due to the apparent brightness of distant stars falling outside of its detection limits.

3.3.2   Number density and stellar properties

Figure 6 shows the number density of all mock stars as a function of Galactocentric radius, including a comparison with the subset detectable by the Euclid Wide Survey. This data is crucial for understanding the Euclid survey's selection function and for assessing whether our mock catalog is a representative sample of the Milky Way's stellar population.

In figure 7, I examine the metallicity distribution of all mock stars against those observable by Euclid. This comparison is vital for understanding the selection bias introduced by the Euclid magnitude cut-off and assessing whether our mock catalog accurately reflects the known metallicity gradient in the Milky Way. Comparison of the histogram shape with the observed metallicity distribution (see, for example, figure 3 of Madau 2023) suggests that the metallicity distribution shape of the simulated data set accurately reflects stellar observations.

Similarly, figure 8 focuses on the stellar mass distribution, comparing the overall mock catalog with the Euclid detectable subset. This plot helps in identifying any biases towards certain stellar masses due to the observational limits of the Euclid mission and possible simulation deficiencies.

The ensemble of these plots provides a comprehensive characterization of my mock catalog. The consistency of the mock catalog with established galactic models and empirical data implies a high degree of accuracy in the simulation process and the effectiveness of the Euclid selection criteria. This characterization strengthens the reliability of my subsequent analyses and conclusions drawn from this catalog.
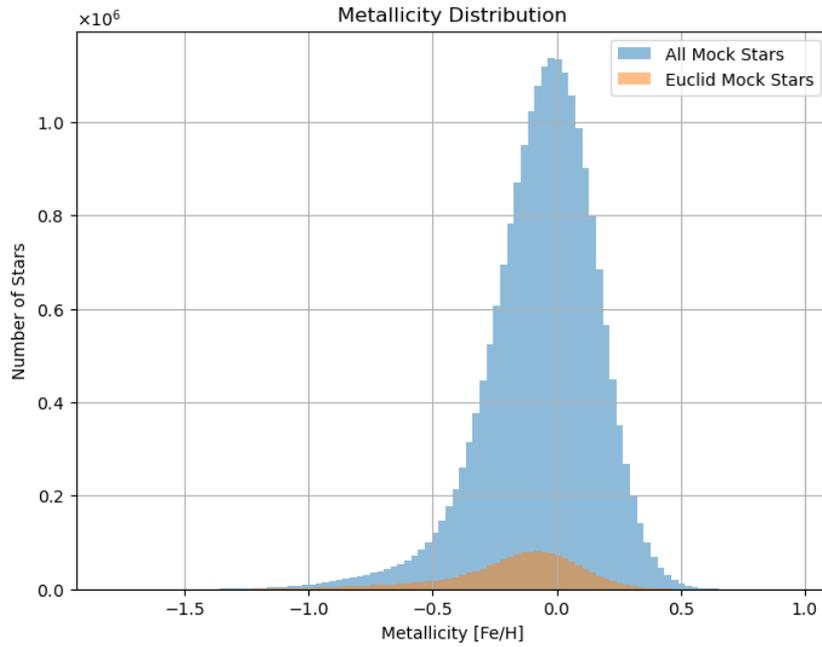
Figure 7: Histogram comparing the metallicity distribution of the entire mock star catalog (in blue) with the metallicity distribution of stars that fall within the Euclid Wide Survey's detection capabilities (in orange). The overlay illustrates the extent to which the Euclid Survey's magnitude cut-off influences the observed metallicity spread of the Galactic population.
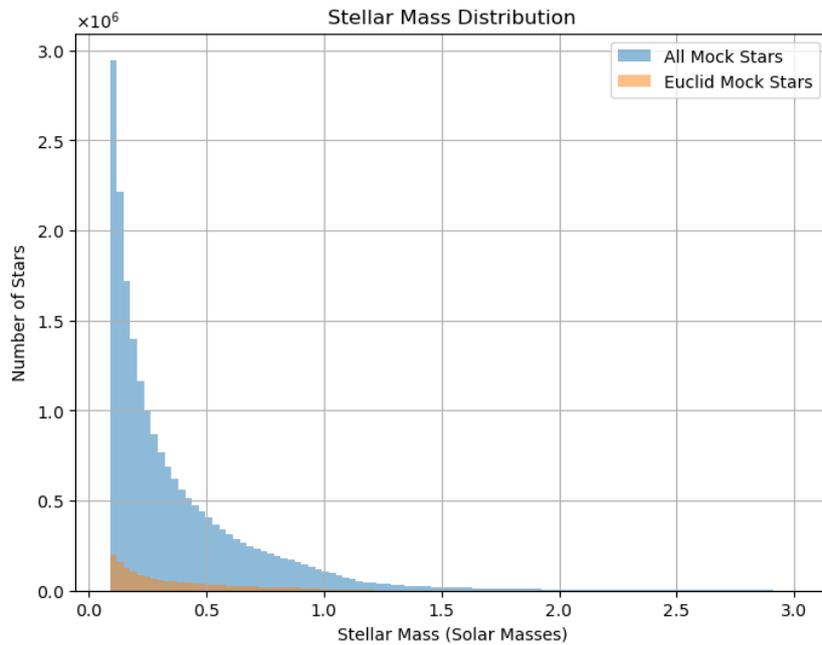
Figure 8: Distribution of stellar mass for all simulated stars (in blue) compared to the Euclid detectable subset (in orange).

3.4   Linear motion approximation and uncertainty handling

3.4.1   LMA method

To estimate the distance and time of closest approach for a passing star, the linear motion approximation (LMA) method is employed. For close encounters with a perihelion time not exceeding a few million years (Myr), the LMA method provides sufficiently accurate results, aligning closely (within about 5% for the perihelion distance) with outcomes derived from methods that integrate stellar orbits (Bobylev & Bajkova 2020; Bailer-Jones 2015).

The LMA method assumes that a passing star maintains a linear trajectory and constant velocity relative to the Sun. Under this assumption, the time and distance of the star's closest approach to the Sun can be computed using its radial and transverse velocity components and its initial distance from the Sun. As elucidated by Bailer-Jones (2015), the position $\boldsymbol{r}$ of a star, in relation to the Sun at time t, is expressed by:

$$\boldsymbol{r} = \boldsymbol{r}_0 + \boldsymbol{v}t \tag{3}$$

where $\boldsymbol{r}_0$ is the star's position at the observation time ($t = 0$) and $\boldsymbol{v}$ is its velocity with respect to the Sun. The perihelion distance, or the minimum value of $|\boldsymbol{r}|$, is achieved when the time of closest approach is determined by minimizing $|\boldsymbol{r} \cdot \boldsymbol{r}|$. This leads to:

$$t_{ph} = -\frac{\boldsymbol{r}_0 \cdot \boldsymbol{v}}{\boldsymbol{v} \cdot \boldsymbol{v}} \tag{4}$$

Transforming this equation into observational astrometric terms, we consider the radial distance r in kiloparsecs (kpc), the radial velocity v_r in kilometers per second (km s^{-1}, negative for approaching stars), and the proper motions $\mu_{\alpha*}$ and μ_δ in milliarcseconds per year (mas yr^{-1}). The transverse velocity v_T in km s^{-1} is then given by:

$$v_T = 4.74047 \text{ [AU km}^{-1} \text{ yr}^{-1}] \, r \, \sqrt{\mu_{\alpha*}^2 + \mu_\delta^2} \tag{5}$$

And the total velocity v, also in km s^{-1}, is:

$$v = \sqrt{v_T^2 + v_r^2} \tag{6}$$

Combining equations 3 and 4 with 5 and 6 allows us to derive relations for the perihelion time t_{ph} in Myr and the perihelion distance d_{ph} in parsecs (pc):

$$t_{ph} = -0.97779 \times 10^9 \text{ [10}^{-3} \text{ pc}^{-1} \text{ km yr]} \, \frac{r \, v_r}{v^2} \tag{7}$$

$$d_{ph} = 10^3 \, \frac{r \, v_T}{v} \tag{8}$$

3.4.2   Astrometric uncertainties

In assessing the potential of stellar surveys for close encounter research, it is imperative to consider the uncertainties in astrometric measurements. The variances in proper motion, radial distance, and radial velocity utilized in my analysis represent typical uncertainties for these parameters. As the study relies on simulated data, the actual values were precisely known. However, the chosen values aim to reflect realistic uncertainties in the context of Euclid data for investigating close stellar encounters.

For accurate distance measurements of stars, stellar parallax is typically employed. This method involves comparing two observations of the same star from different locations, with the background star shift indicating the parallax, and consequently, the radial distance. Unfortunately, the Euclid

mission is not designed for precise parallax measurements; it observes each patch of sky only from one location.

Nonetheless, the mission's extensive array of photometric filters offers an alternative approach for estimating radial distances. In this study, I developed a rudimentary machine learning method to demonstrate how astrometric quantities can be deduced from photometry, detailed in appendix 7. Several research groups (such as Shull & Danforth 2019; Queiroz et al. 2020; Thomas et al. 2019; Zucker et al. 2018) have developed similar machine learning algorithms for inferring astrometric data from photometric observations. While fundamentally akin to my approach, these algorithms incorporate additional observational data, multiple neural network layers, and observational errors, enhancing the accuracy and realism of their photometry-based astrometric estimates. My analysis of Euclid's stellar research possibilities considers such advanced machine learning methods. With eight photometric filters, these methods achieve a radial distance variance of approximately 7-8%. Assuming the inclusion of a ninth filter for Euclid and further advancements in machine learning, I adopt a variance δr of 5%.

For the radial velocity uncertainty δv_r, I assume a realistic value of 1 km s^{-1}, based on typical uncertainties in stellar spectroscopy, such as those from the Gaia mission (Gaia Collaboration et al. 2023). Consequently, the part of my analysis in which the radial velocity is involved should be interpreted as a synergy of Euclid's photometric data and spectroscopic data from an auxiliary telescope.

Regarding proper motion uncertainty, it is assumed that a realistic variance for combined ground- and space-based observations of Euclid is around 10 mas yr^{-1}. This value, inferred by Gijs Verdoes Kleijn (priv. comm.), is based on the smaller than 10 mas yr^{-1} precision obtained by combining ground-based and space-based Gaia measurements (Tian et al. 2017), and the achievements from one of the Euclid ground-based surveys by itself (Bernstein et al. 2017). However, to make meaningful estimations of close encounter properties, various $\delta\mu$ values are considered, envisaging a second pass of Euclid separated by ~5 years and/or a combination of Euclid's observations with precise parallax measurements from other surveys like Gaia, WEAVE and RAVE. According to Lindegren et al. (2021), Gaia's proper motion uncertainty ranges from 0.02 mas yr^{-1} for bright sources to 0.5 mas yr^{-1} for fainter ones (magnitude 20), expected to reduce by an order of magnitude over the remainder of its mission (Gaia Collaboration et al. 2023). Therefore, I examined scenarios with proper motion variances of 0.01, 0.1, 1, and 10 mas yr^{-1}.

In order to integrate the uncertainties into the mock Euclid catalog, I adopted a statistical approach that involves generating multiple instances of each star's features. For every star in the catalog, each feature (such as proper motion, radial distance, and radial velocity) is assigned 30 random values. These values are derived from a normal distribution centered around the star's true value, with the variance for each feature corresponding to the estimated uncertainties discussed previously.

This methodology creates a dataset that embodies the inherent measurement uncertainties one would expect in actual observational data. For instance, with radial distance (r), I utilized a normal distribution with a variance (δr) of 5%, reflecting the expected accuracy in distance measurements derived from Euclid's photometric data combined with advanced machine learning techniques. Similarly, for radial velocity (v_r), the normal distribution is centered around the true velocity value with a standard deviation of 1 km s^{-1}, mirroring the anticipated spectroscopic precision.

The process is analogous for proper motion, where multiple scenarios were analyzed, each characterized by a different level of uncertainty ($\delta\mu$). These ranged from an optimistic 0.01 mas yr^{-1}, as might be achieved with future advancements in spectroscopic measurements, to a more conservative 10 mas yr^{-1}, which represents what could be expected from a combination of Euclid's ground- and space-based data. By simulating 30 instances for each feature of every star, the mock catalog captures a realistic spread of possible observations, mirroring the statistical nature of observational uncertainties.

Observations	Selection within 3 kpc	
	Full sky	Euclid survey
All Stars	2.3×10^9	1.5×10^8 (7%)
CE, < 1 pc	570	73 (13%)
CE, < 2 pc	2340	272 (12%)
CE, < 5 pc	14730	1644 (11%)

Table 3: Number of stars in the mock catalog. The left column represents the number of stars within a distance of 3 kpc from the Sun. The right column contains the number of stars that are contained in the Euclid survey area. The top row represents the total number of stars, and the bottom three rows the number of close encounters with the indicated maximum perihelion distance and with a perihelion time not further than 15 Myr in the past or future. The percentages indicate the fraction of stars (residing within 3 kpc) and close encounters that Euclid observes.

4   Results: Close Stellar Encounters Observable by Euclid

4.1   Completeness of Euclid survey area

A completeness test was conducted to determine the extent of stars and close encounters missed by Euclid due to its brightness detection limit and avoidance of the disk and ecliptic. This test involved a full-sky *Galaxia* simulation, where I assessed how many stars bright enough for Euclid passbands were excluded in these regions.

The sensitivity of the photometry allows for the detection of all (100%) stars that currently reside inside the Euclid survey and that have a perihelion distance smaller than 5 pc and a perihelion time smaller than ±5 Myr (by implication, also all <1 pc encounters within this time frame are detected). For the a perihelion distance smaller than 1 pc (5 pc), and a perihelion time smaller than ±15 Myr, the detection limit based on brightness is 99% (96%).

The completeness values of the Euclid Wide Survey, considering the survey area in addition to the photometric sensitivity, are noted in table 3. The table shows that only about 7% of the stars within a 3 kpc radius sphere are contained in the survey area and that Euclid observes about 13% of the close encounters. The Euclid survey avoids a direct line-of-sight through the Milky Way disk, but it does observe in the direction 'above' and 'below' the Sun's position. Since the Sun is vertically close to the middle of the galactic disk, Euclid still captures a significant portion of a dense part of the Milky Way. Almost all of these stars are relatively close by, since the dense thin disk ends at about 1 kpc above and below the Sun. In this way, the survey includes relatively many stars that are nearby and hence more likely to have close encounters with the Sun. As a result, while its overall stellar completeness might be reduced to about 7%, its completeness regarding potential close encounters with the Sun is still quite high at about 13%.

4.2   Visualization and count of stellar encounter observations

To make a reasoned estimate of the perihelion of an observed star, its proper motion, radial velocity, and radial distance need to be suitably known. This ensures that for each observed star, the astrometric quantities are accurate enough to confidently identify whether it is a close encounter. Two critical statistics for assessing this certainty are the true positive rate (TPR) and the positive predictive value (PPV). The TPR, or sensitivity, is the fraction of true close stellar encounters correctly identified, addressing the question: "how many of the true close encounters are found?" The PPV, or precision, is the fraction of observations identified as close encounters that are indeed close encounters, answering:

"how certain are we that an estimated close encounter is true?"

The TPR and PPV are computed as follows:

$$\text{TPR} = \frac{n(\overline{\text{CE}}|\text{CE})}{n(\text{CE})} \tag{9}$$

$$\text{PPV} = \frac{n(\overline{\text{CE}}|\text{CE})}{n(\overline{\text{CE}})} \tag{10}$$

Here, CE denotes the set of close encounters, and $\overline{\text{CE}}$ indicates the observations classified as close encounters. Thus, $n(\text{CE})$ is the total number of true close encounters, $n(\overline{\text{CE}})$ is the number of observations classified as close encounters, and $n(\overline{\text{CE}}|\text{CE})$ is the count of observations correctly identified as close encounters (true positives).

Secondary to the TPR and PPV are their complements: the false negative rate (FNR) and the false discovery rate (FDR). The FNR, calculated as FNR $= 1 - $ TPR, indicates the proportion of missed stellar flyby's, while the FDR, the complement of PPV (FDR $= 1 - $ PPV), represents the extent to which the set of identified close encounters is polluted by non-flyby's.

In its wide survey field, the Euclid survey will observe more than a hundred million stars. A crucial question is whether a significant portion of stars not involved in close encounters can be filtered out from this vast dataset. Given that stellar radial distance and proper motion data provide some insight into their perihelion time and distance, it is essential to explore whether these data are sufficient to select a set of stars with a high probability of encountering the Sun closely (e.g., within 1 parsec in a time frame of ± 15 Myr). In other words, are the proper motion and radial distance data adequate to achieve a TPR near 1 (ensuring almost all close encounters are included) and a considerable PPV (indicating a significant portion of the set actually represents close encounters)?

4.2.1   Inference from radial distance and proper motion

When plotting stellar proper motion against radial distance and highlighting close encounters, a pattern emerges. Each star with a perihelion distance less than 1, 2, or 5 pc and a perihelion time within 15 Myr exhibits either little proper motion or a small radial distance. To define a boundary in the proper motion-radial distance space that encompasses all close encounters while excluding more non-flyby observations than the theoretical limit, a demarcation line is drawn (see figure 9). All stars positioned left and below this line are categorized as potential close encounters, aiming to include most close encounters as stellar flyby candidates and minimize the inclusion of non-close encounters.

I manually constructed a boundary function, and with a process of trial and error found the following to be the most effective: $\mu(r) = (r - A)^{-0.8} - B$, where A and B are parameters varying with different perihelion distance limits. In principle, machine learning techniques such as support vector machines could be used to draw this boundary, possibly one that is more effective at separating the close encounters from the non-close encounters in the proper motion-radial distance space. I did not manage to produce a more effective boundary function than the above, however. Moreover, when constructing a boundary function, weighing false positives against false negatives is inevitable. I chose to draw the boundary quite "far out" from the close encounter points in the bottom left part of the vector space in order to capture almost all close encounters (this is reflected in the high true positive rates listed in table 4).

Testing the full dataset against this boundary showed that almost all of the true close encounters fall on the correct side (left and under) of the boundary, resulting in a true positive rate close to unity (see the first column of table 4). Nevertheless, this crude selection method also classifies a large number of non-close encounters as true encounters, as indicated by the positive predictive value of
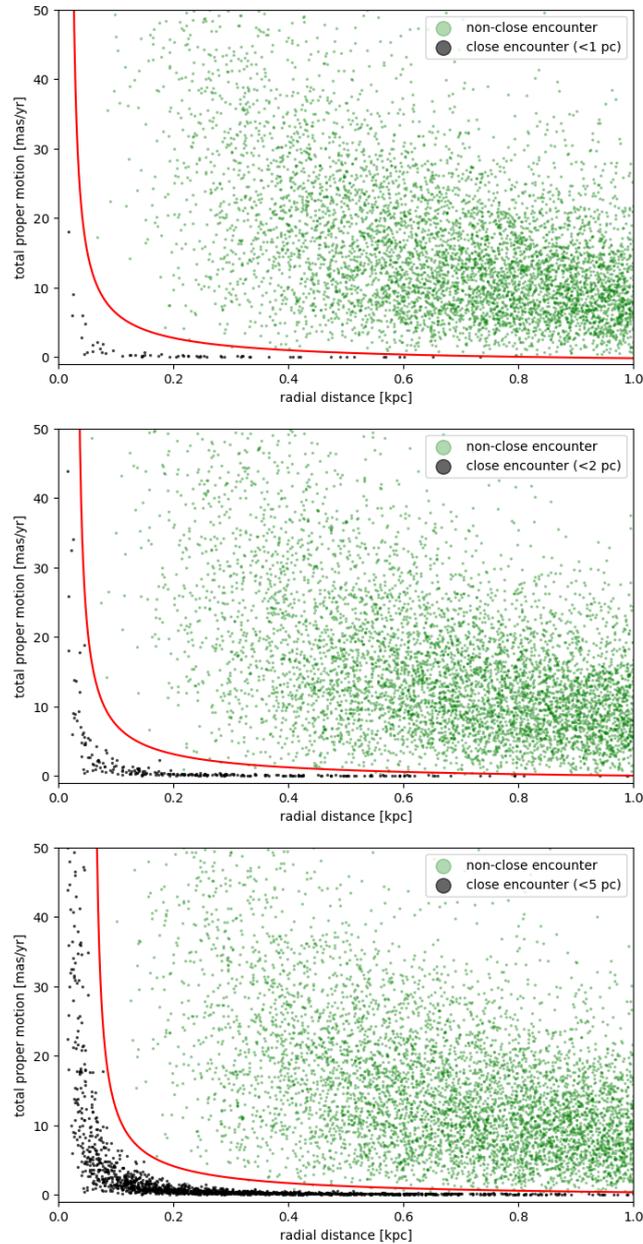
Figure 9: Proper motion versus radial distance for close encounters (black) and non-close encounter observations (green), showing the nominal values (no measurement uncertainty) for each variable. The three panels show different definitions of a close encounter, each with perihelion time smaller than 15 Myr, but with perihelion distance smaller than 1, 2, and 5 pc. All close encounters in the data set are plotted, but only 0.01% of the non-close encounters to maintain readability (this leads to the illusion that the boundary function classifies almost no non-close encounters incorrectly as close encounters, but see figure 10). The red lines indicate the categorization boundary, following the relation: $\mu(r) = (r - A)^{-0.8} - B$, with $A = 0.02, 0.03$, and 0.06, and $B = 1.2, 1.0$, and 0.7 for close encounter limits of 1, 2, and 5 pc, respectively.
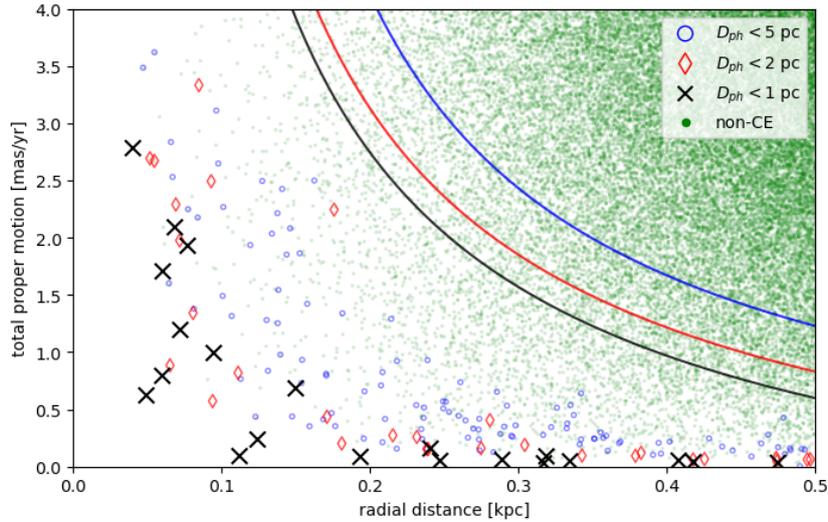
Figure 10: Proper motion versus radial distance of close encounters (blue circle, red diamond and black cross) and non-close encounters (green dot). In this plot (contrasting with figure 9) all non-close encounters are plotted. A snippet of the range is shown. All close encounters appear to be correctly classified within this range, in line with the high true positive rates listed in table 4. It is evident here that the positive predictive value is much lower (see table 5), with numerous non-close encounters lying in the designated close encounter region.

around 10^{-3} (see the first column of table 5). This implies that for each correctly identified close encounter, about a thousand non-close encounters are erroneously classified as stellar flyby's. Figure 10 visualizes this, showing many non-close encounters on the 'close encounter' side of the boundaries.

4.2.2 Incorporating measurement uncertainties

To account for uncertainties in the astrometric quantities, the mock dataset was expanded as described in section 3.4.2. The expanded mock star set was then analyzed using the crude selection method, with each random assignment classified as a close encounter or non-close encounter according to the established boundary. The classifications were then compared to the LMA classification based on the true astrometric values of the stars.

The inclusion of uncertainties in proper motion and radial distance led to slightly fewer true close encounters being correctly identified. The true positive rates decrease with increasing measurement uncertainty, but the drop in true positive rate is limited, as shown in table 4. This is also illustrated in figure 11, where some close encounters with imprecise proper motion and radial distance values fall outside the correctly classified region. Despite this, the majority were still accurately categorized. The positive predictive value decreases a bit stronger with increasing measurement uncertainty, as detailed in table 5. For the closest encounters and the most realistic proper motion variance, the positive predictive value was approximately 10^{-5}, meaning that for each correctly classified close encounter, there are about a hundred thousand false positives.

Despite these challenges, the crude selection method's success is noteworthy. The results suggest that Euclid, even with substantial uncertainty in proper motion measurements, could effectively filter many stars from its survey data without discarding a significant number of close encounter candidates. More precisely, Euclid data with 10 mas yr^{-1} uncertainty in proper motion could be used to capture 77% of all close encounters with $d_{ph} < 5$ pc and $|t_{ph}| < 15$ Myr (see table 4), which amounts to about 1266 true close encounters (77% of 1644, see table 3). The selected sample of flyby candidates contains about 9 million stars (1266 true encounters times a PPV of 0.014%, see table 5), thereby

CE threshold d_{ph}	var$_{pm}$ [mas/yr]				
	0	**0.01**	**0.1**	**1**	**10**
< 5 pc	0.98	0.95	0.94	0.88	0.77
< 2 pc	0.93	0.92	0.91	0.86	0.75
< 1 pc	0.90	0.89	0.88	0.85	0.73

Table 4: True positive rates for the crude selection method based on proper motion and radial distance. The left column indicates precisely known values for proper motion and radial distance. The other columns indicate increasing variance in proper motion. The variance in radial distance in those analyses was 5% of the nominal value.

CE threshold d_{ph}	var$_{pm}$ [mas/yr]				
	0	**0.01**	**0.1**	**1**	**10**
< 5 pc	7.0	1.5	1.4	0.73	0.14
< 2 pc	3.7	0.45	0.42	0.19	0.027
< 1 pc	1.8	0.15	0.14	0.068	0.0079

Table 5: Positive predictive value (multiplied by 1,000) for the crude selection method based on proper motion and radial distance. The left column indicates precisely known values for proper motion and radial distance. The other columns indicate increasing variance in proper motion. The variance in radial distance in those analyses was 5% of the nominal value.
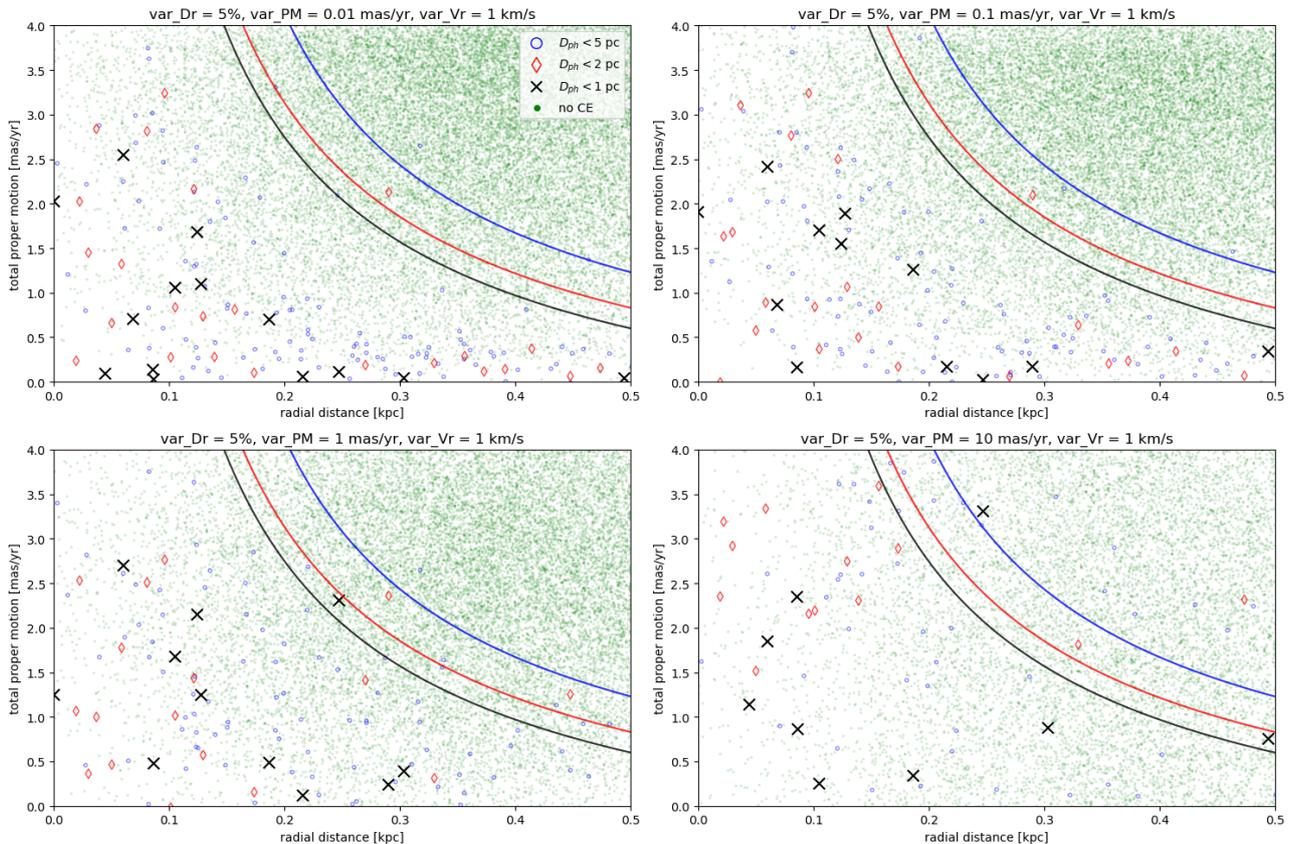
Figure 11: Proper motion versus radial distance for close encounters and non-close encounters, including uncertainties. The plot shows a snippet of the range, revealing some true close encounters on the incorrect side of the boundary, albeit the majority being correctly classified. The low positive predictive value is evident, with many non-close encounters falling into the close encounter region.

filtering out ~94% of the total number of stars from its survey (the 9 million candidates make up about 6% of the total 150 million observations, see table 3).

A detection of 77% of the close encounters would be a significant improvement over the current 15% achieved with Gaia (which is even for a smaller time frame of ±5 Myr). But there is still an enormous effort to be made to distill and determine the detailed astrometric properties of the true close encounters, which account for 0.014% of the candidate sample.

4.3   Including radial velocity

Incorporating radial velocity measurements into my analysis of Euclid's potential for stellar flyby research could significantly enhance the selection of close encounter candidates. Suppose we utilize spectroscopic data along with photometry for all stars observed by Euclid. This integrated approach would likely yield a stricter selection than the method presented in the previous section, which relied solely on proper motion and radial distance. For this enhanced scenario, I calculated the true positive rate (TPR) and positive predictive value (PPV) for distinguishing close encounters from non-close encounters, utilizing sets of proper motion, radial distance, and radial velocity measurements with the previously mentioned uncertainties.

The essence of this method is to compute hypothetical perihelion distances and times using imprecise measurements and compare these estimates to the actual close encounter data from the galaxy simulation. To understand the impact of astrometric uncertainties on perihelion time and distance, I assumed that the uncertainties in proper motion, radial distance, and radial velocity are independent. The variances propagate to perihelion time and distance as follows:

$$\delta v_T = v_T \ \sqrt{(\delta\mu/\mu)^2 + (\delta r/r)^2} \tag{11}$$

$$\delta v = v^{-1} \ \sqrt{v_T^2 \ \delta v_T^2 + v_r^2 \ \delta v_r^2} \tag{12}$$

Applying these to the perihelion time and distance equations yields:

$$\delta t_{ph} = t_{ph} \ \sqrt{(\delta r/r)^2 + (\delta v_r/v_r)^2 + (2 \ \delta v/v)^2} \tag{13}$$

$$\delta d_{ph} = d_{ph} \ \sqrt{(\delta r/r)^2 + (\delta v_T/v_T)^2 + (\delta v/v)^2} \tag{14}$$

Subsequently, I generated 30 random perihelion time and distance values for each star, drawn from a normal distribution with the actual values as means and variances inferred from the variances in proper motion, radial distance, and radial velocity.

Comparing these imprecise perihelion estimates with actual values yielded the TPR and PPV, as visualized in figures 12, 13, 14, and 15. While some actual close encounters were not identified correctly, many non-close encounters were erroneously classified as close encounters, especially with increasing uncertainty in proper motion, as shown in tables 6 and 7.

Interestingly, the inclusion of radial velocity measurements did not improve the TPR when considering astrometric uncertainties. In fact, the selection method based solely on proper motion and radial distance achieved a higher TPR. However, there was an improvement in the PPV, particularly in scenarios with small perihelion distances and large uncertainties. This is reflected in the size of the flyby candidate sample. For example, without incorporating radial velocity, the candidate sample size for encounters with $d_{ph} < 1$ pc and a proper motion uncertainty of 10 mas yr^{-1} is about 6.7 million. But for the method including radial velocity, the number of stars in the flyby candidate sample is about 58 thousand.

Adjusting the selection boundaries could potentially increase either the TPR or PPV at the expense of the other. Thus, overall, the inclusion of radial velocity seems to refine the method for selecting close encounter candidates.
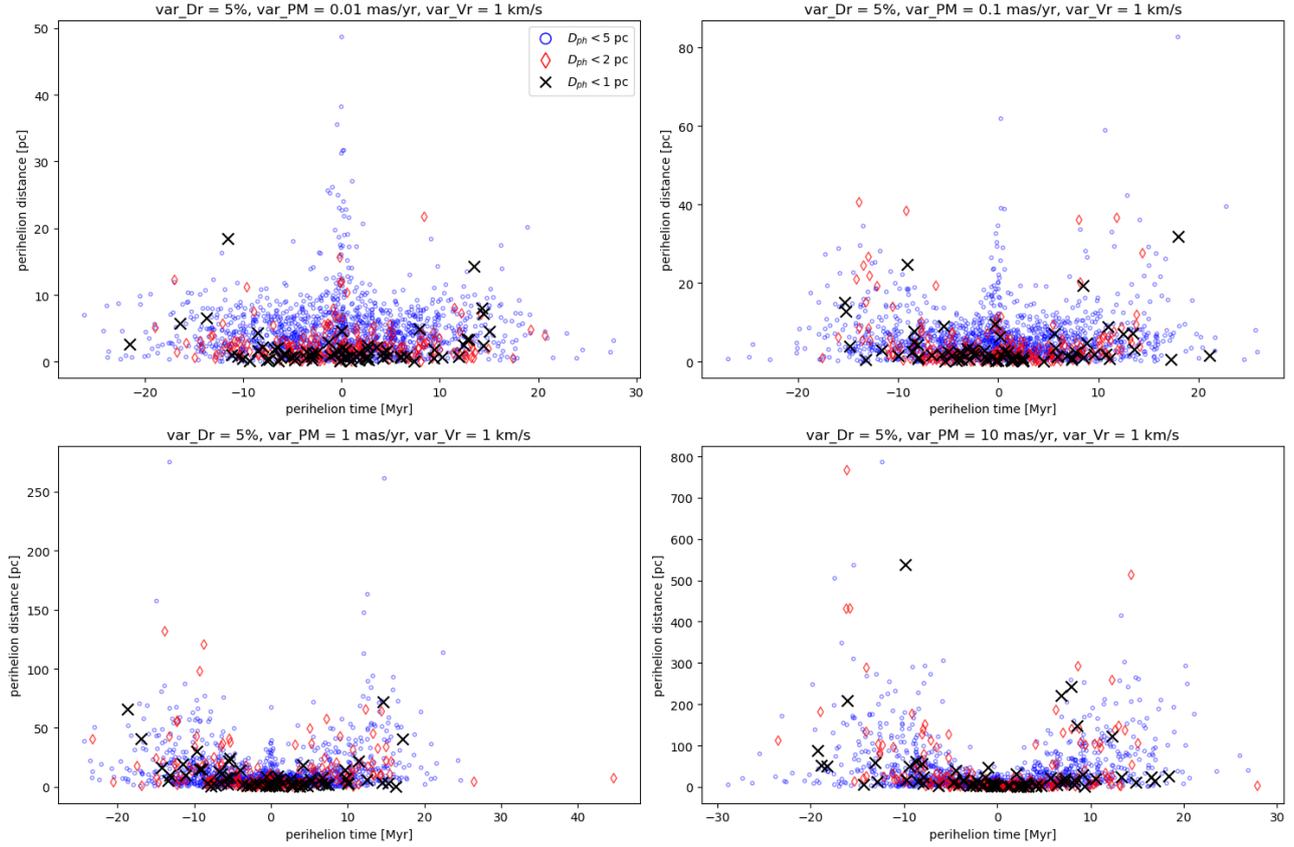
Figure 12: Estimated perihelion distance versus perihelion time for the 1644 close encounters noted in table 3. Blue circles, red diamonds and black crosses are stars with true (LMA derived) 5 pc > d_{ph} > 2 pc, 2 pc > d_{ph} > 1 pc and 1 pc > d_{ph}, respectively. The plotted values are one (random) instance of the 30 random samplings of the probability distribution functions for d_{ph} and t_{ph} specified in section 3.4.2. The measurement uncertainty is specified above each panel. Clearly visible is the large spread over perihelion distance and time in case of large measurement uncertainty (compare to the true perihelion distance and time values in figure 1). Note the difference in axes range between the four panels.

CE threshold d_{ph}	var$_{pm}$ [mas/yr]			
	0.01	**0.1**	**1**	**10**
< 5 pc	0.63	0.55	0.40	0.26
< 2 pc	0.57	0.43	0.29	0.18
< 1 pc	0.48	0.33	0.21	0.12

Table 6: True positive rates for the selection method considering proper motion, radial distance, and radial velocity.

CE threshold d_{ph}	var$_{pm}$ [mas/yr]			
	0.01	**0.1**	**1**	**10**
< 5 pc	6.3	5.4	3.5	1.5
< 2 pc	2.4	1.7	1.0	0.41
< 1 pc	1.1	0.73	0.41	0.15

Table 7: Positive predictive values (multiplied by 1,000) for the selection method considering proper motion, radial distance, and radial velocity.
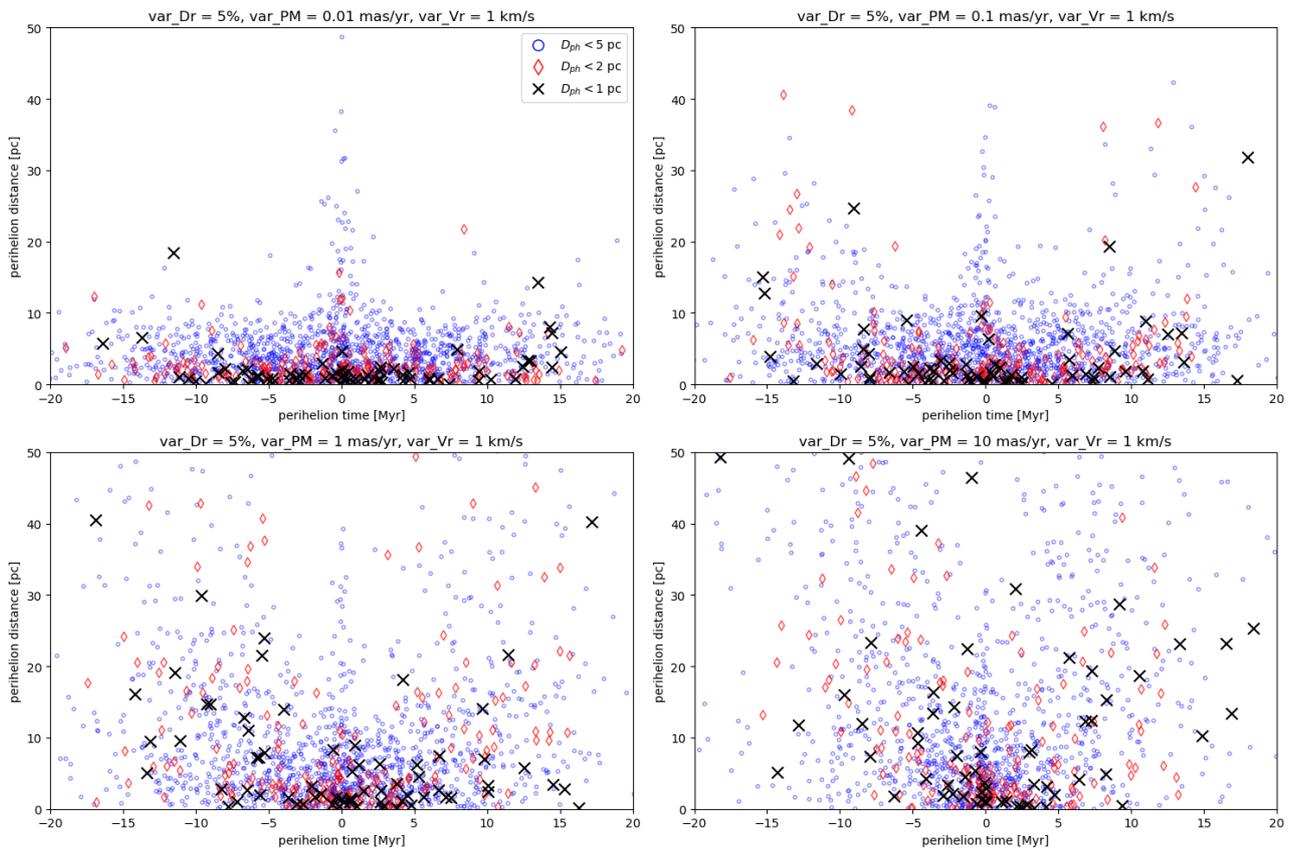
Figure 13: Zoomed view of perihelion distance and time estimates. The spread in perihelion distance highlights the large uncertainty stemming from imprecise astrometric data, while the perihelion time uncertainty is relatively smaller.
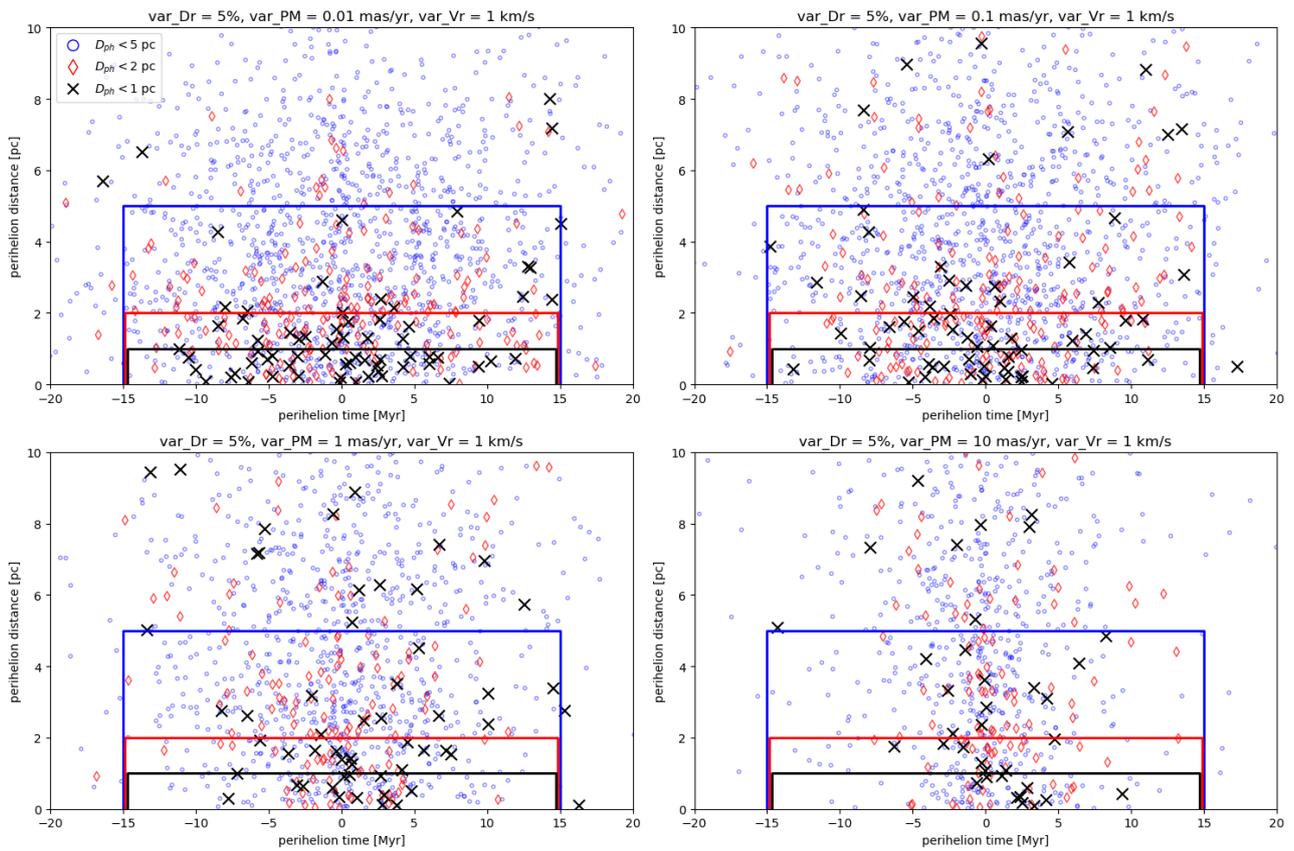
Figure 14: Further zoomed view with lines indicating the region for observed close encounters. Points within each region are classified as close encounters based on astrometric data, with colors corresponding to different close encounter thresholds.
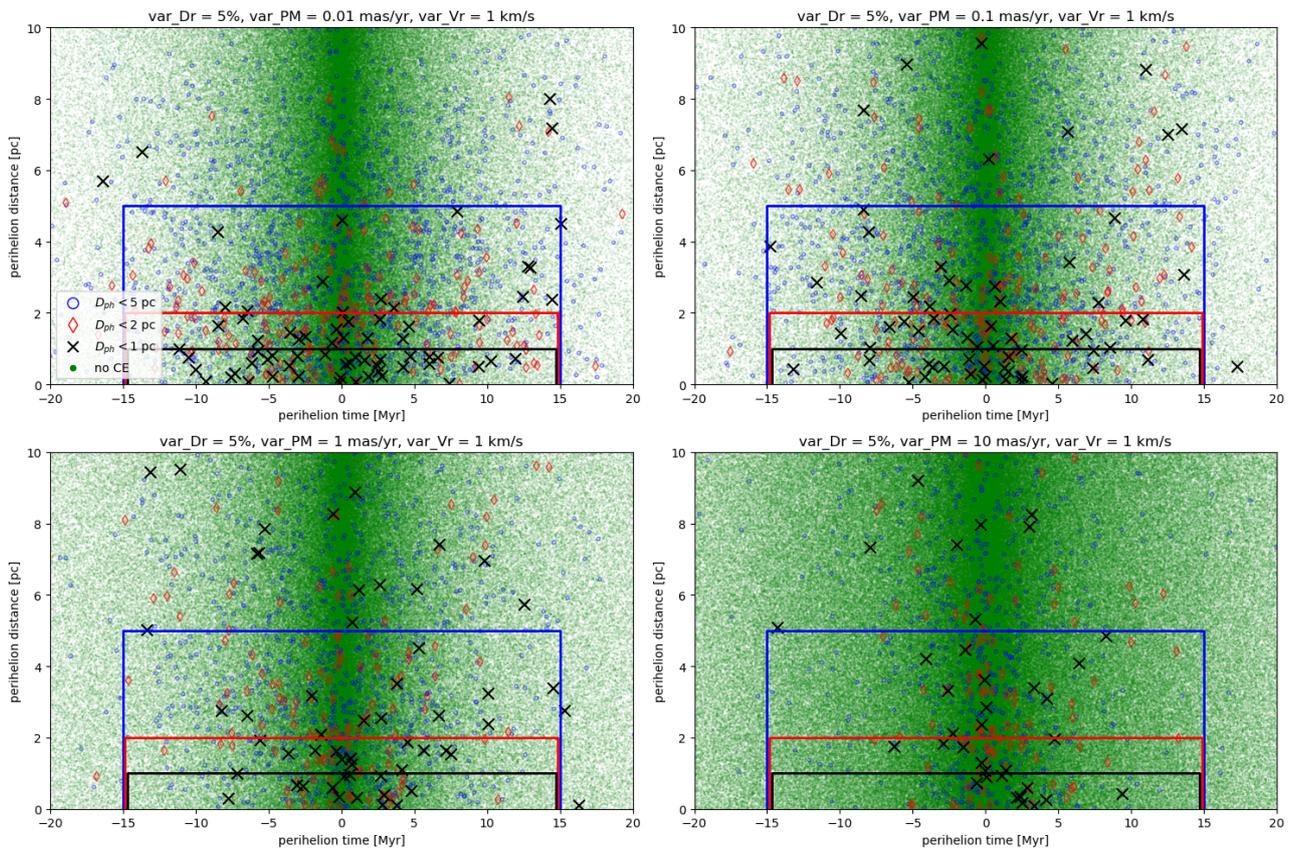
Figure 15: Similar to figure 14, but including non-close encounters. Many non-flyby's are incorrectly classified as close encounters, as indicated by the green points within the boundary lines.

5 Discussion of Euclid Analysis

5.1 Application difficulties of Euclid data

The Euclid mission, while providing valuable data for celestial research, presents several challenges that impact its application in studying close stellar encounters and their potential climatic effects.

A primary concern with the Euclid mission data is the limited survey field. This constraint means that the data covers only a portion of the sky, potentially omitting significant stellar encounters that occur outside the surveyed region. In principle it would be possible to dedicate a part of Euclid's operation time to studying parts of the sky that are densely populated with stars. The extra research time spent on the galactic plane would help the science that focuses on the galactic plane, including stellar encounter research. However, observing the galactic plane is not part of the Euclid consortium's primary research goals.

Another challenge is the large uncertainty in proper motion inherent in the Euclid data. The accuracy of proper motion measurements is crucial for determining the trajectories and potential encounters of stars with the Solar System. The potential necessity to incorporate precise proper motion measurements, to improve the accuracy of perihelion inference, adds another layer of complexity and uncertainty to the analysis.

Additionally, the Euclid mission data lacks measurements of radial velocity, which are essential for a complete understanding of stellar dynamics. The need for additional spectroscopic measurements to obtain this data presents both logistical and financial challenges. Conducting spectroscopy on a large number of stars to complement Euclid data is not only technically demanding but also potentially expensive.

So, while Euclid could contribute significantly to observing and measuring the radial distances of almost all close encounters, much additional scientific work has to be made to find the remaining 85% of stellar flyby's that is currently missing.

5.2 Limitations of data analysis

A key assumption in the data analysis is the independence and Gaussian nature of measurement uncertainties. While this assumption simplifies the statistical treatment of data, it may not accurately reflect the true nature of uncertainties in astronomical observations. For the uncertainties in proper motion and radial distance, for example, I used a folded Gaussian distribution (taking the absolute value of the distribution to avoid negative values), which might be too simple to reflect actual observational uncertainties. The complex, often non-Gaussian and correlated nature of astrometric measurement uncertainties could potentially skew the results, even though this issue seems to be small in this case (Bailer-Jones 2015).

The methodology employed in this research includes a binary definition of stellar fly-by events, categorizing encounters strictly as either flyby's or non-flyby's. This dichotomous approach may omit nuanced information about near-miss or weak encounters that could still have relevance to the Earth System, especially in a cumulative or long-term context.

The analysis is constrained by a limited time window, which relies on the Linear Motion Approximation method. While this approach is effective for short-term predictions and analysis, it may not capture the full dynamics of stellar movements and interactions over longer periods. A broader time window, accompanied by more advanced analytical methods, could provide a deeper understanding of the long-term relevance of stellar encounters for the Earth System.

6   Towards Researching the Flyby-Comet-Climate Chain

In the preceding sections, the potential of the Euclid mission in detecting close stellar encounters was explored. In this section, the focus shifts towards further application of the Sequence Equation framework. It delves into the capacity of stellar flyby's of disturbing the Oort Cloud, the resulting comet flux and associated impact energies, and the climatic effects that can be induced by an impact. Additionally, the applicability and limitations of the Sequence Equation Framework in this context are critically evaluated.

6.1   Oort Cloud perturbation strength

Close stellar encounters significantly influence the dynamics of the Oort Cloud, with their impact primarily determined by three critical parameters: the mass of the encountering star (M_{enc}), the perihelion distance of the encounter (d_{ph}), and the velocity of the star at perihelion relative to the Solar System (v_{ph}). Following the classical impulse approximation, these parameters collectively determine the gravitational influence exerted on the Oort Cloud objects (Oort 1950; Rickman 1976).

The mass of the encountering star, denoted as M_{enc}, has a direct linear relationship with the gravitational force it exerts. The perihelion distance, d_{ph}, plays a crucial role in the strength of this gravitational interaction. Generally, a smaller d_{ph} implies a closer approach to the majority of Oort Cloud objects, intensifying the gravitational effect, which inversely scales with the square of this distance. The velocity at perihelion, v_{ph}, determines the duration for which the star exerts its influence on the Oort Cloud. A higher velocity results in a shorter interaction period, thereby reducing the extent of orbital disturbances of the Oort Cloud objects.

To quantify the perturbation strength on the Oort Cloud, Feng & Bailer-Jones (2015) introduced the term g, a dimensionless factor representing the perturbation magnitude (an improvement on γ, which is shown in figure 2 and used in various Oort Cloud perturbation studies, such as Fouchard et al. 2011 and Kaib & Quinn 2009). It is formulated as:

$$g = \frac{M_{enc}}{v_{ph} \cdot d_{ph}^2} \tag{15}$$

Utilizing the *Galaxia* simulation and the mock Euclid catalog, I conducted an analysis to gauge the distribution of perturbation effects caused by close stellar encounters. This analysis aimed to yield an indication of the proportion of encounters that are strong perturbers and to evaluate the influence of the g-factor on the observation and classification of these encounters. This was achieved by calculating the g-factor for all mock stars with perihelion distances under 15 pc. The study also considered the varying astrometric uncertainties and visualized how the perturbation strength is distributed within the observational vector space, enhancing our understanding of these stellar encounters and their impacts on the Oort Cloud.

The encounter diagram (Figure 16) illustrates the relationship between perihelion time and perihelion distance, with the size of each point proportionate to its g-factor value. This factor is indicative of the Oort Cloud disturbance strength, as detailed by Feng & Bailer-Jones (2015). The diagram offers a visual representation of flyby's and their distribution in g. It reveals that there is a steep drop-off in impactful encounters; most of them have little Oort Cloud perturbation strength.

Figures 17, 18 and 19 depict the uncertainty in perihelion time ($\sigma(t_{ph})$) and perihelion distance ($\sigma(d_{ph})$), based on the astrometric uncertainty discussed in the thesis. Unsurprisingly (because g is strongly dependent on the perihelion distance), the uncertainty in perihelion distance generally increases with an increase in proper motion variance, as displayed in figures 17 and 19 by the points spreading over larger values of $\sigma(d_{ph})$ in panels with higher proper motion variance.
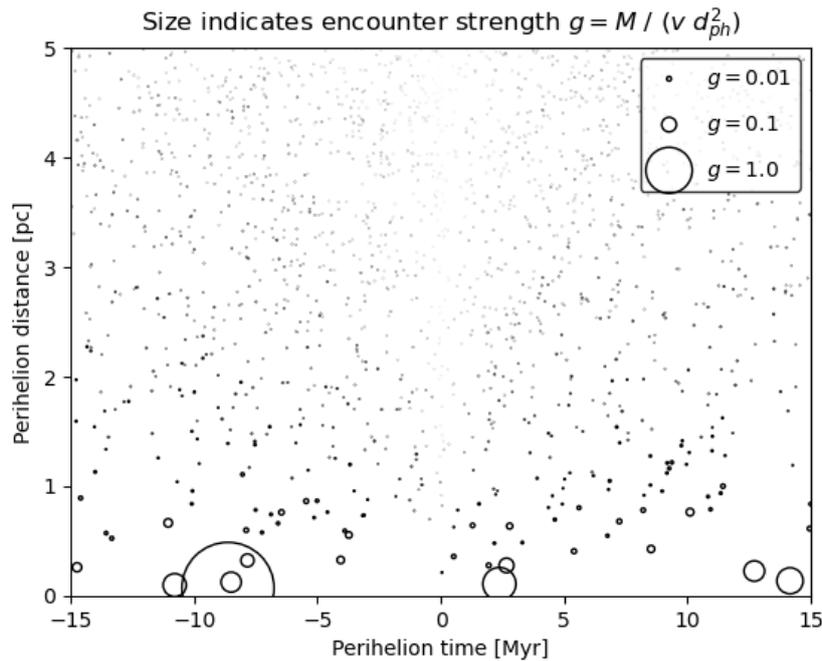
Figure 16: Encounter diagram showing perihelion time versus perihelion distance for the 1644 stellar encounters presented in table 3. Each point represents an encounter, with the size indicating the g-factor.

Notably, encounters with a large g-factor do not exhibit particular uncertainties in perihelion time and distance. In figure 17, the large circles are relatively evenly spread over the panels, hinting at a lack of correlation between g on the one hand, and $\sigma(d_{ph})$ and $\sigma(t_{ph})$ on the other hand—large g-factor values can be found relatively distant from the origin (0,0). Also in figures 18 and 19, no pattern between g and the uncertainty in perihelion time and distance is discernible. Therefore, there is no reason to assume that flyby's with much Oort Cloud perturbation strength can be more accurately measured than weak encounters.

Figure 20 shows that the uncertainty in g (computed with analytical error propagation of the measurement uncertainties in stellar mass, encounter velocity and perihelion distance) is generally large for observations with a high g value. This suggests that for the impactful flyby's, it is most difficult to precisely determine how impactful they are.

The above insights further emphasizes the need for more comprehensive detection and astrometric measurement of the Solar neighborhood. Our current knowledge of 15% of close stellar encounters within ± 5 Myr does not just exclude many flyby's, but possibly also many highly impactful ones.

6.2   Comet flux and impact energy

In this section I present an analysis of the cumulative size and impact energy distribution of near-Earth comets, and their impact probability. This analysis is portrayed alongside the same data for asteroids, to place it into a better-known context.

The estimation of the impact flux and energy of long-period comets (LPCs) and asteroids within the inner solar system is adapted from data found in the literature.

The size-frequency distribution, energy distribution and impact probability distribution of asteroids is provided by Harris & D'Abramo (2015). Their data is based on the re-detection ratio of near-Earth asteroids, pivotal in determining the proportion of all detections that are re-discoveries of previously known objects rather than new ones. The data were plotted (figures 21, 22 and 23) together
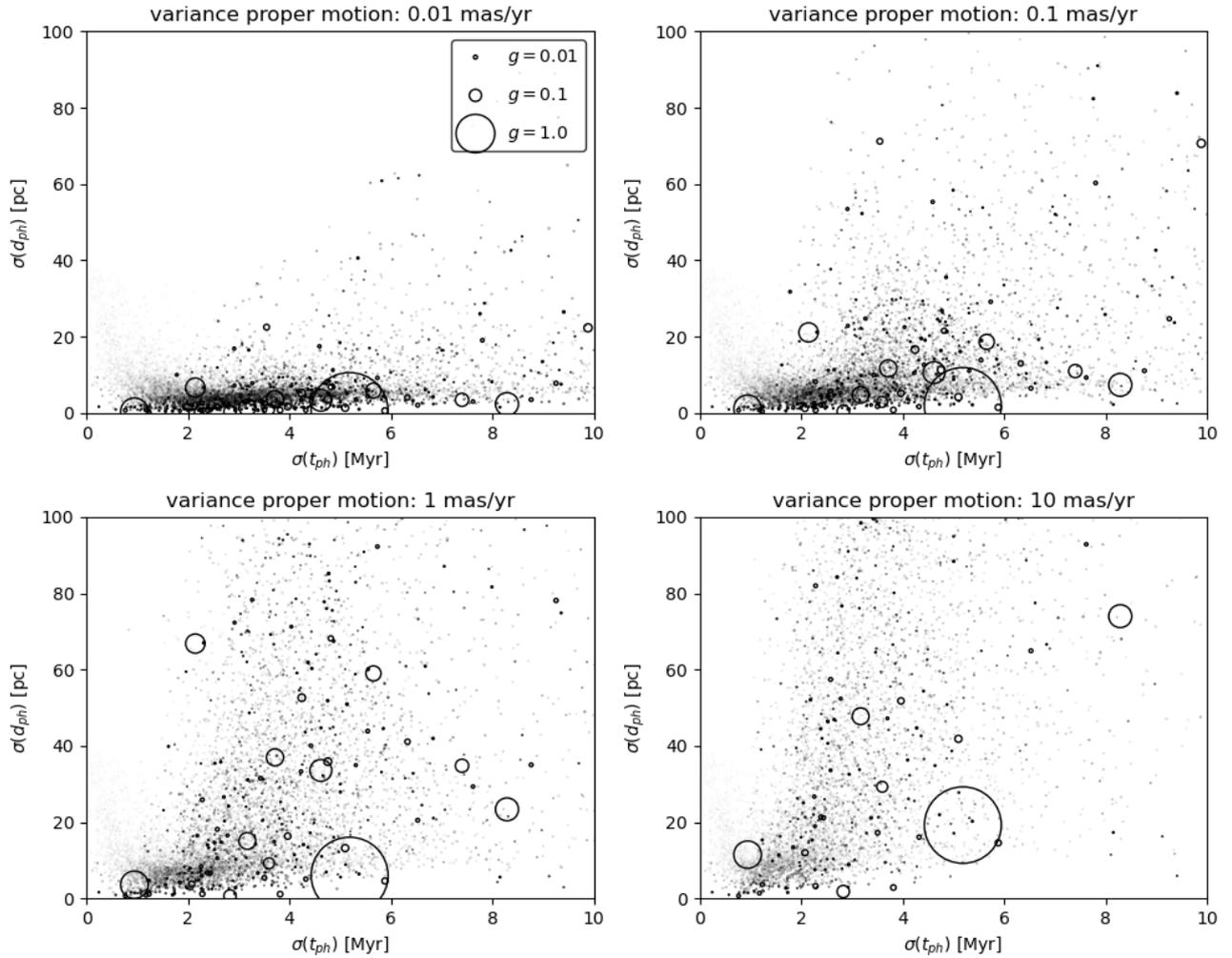
Figure 17: Uncertainty in perihelion time ($\sigma(t_{ph})$) versus the uncertainty in perihelion distance ($\sigma(d_{ph})$) for the 1644 close encounters of table 3, when using the astrometric uncertainties described in section 3.4.2. Each panel corresponds to a different value of the uncertainty in proper motion, with point size reflecting the g-factor. Note that some data points in the bottom panels have exceeded the frame.
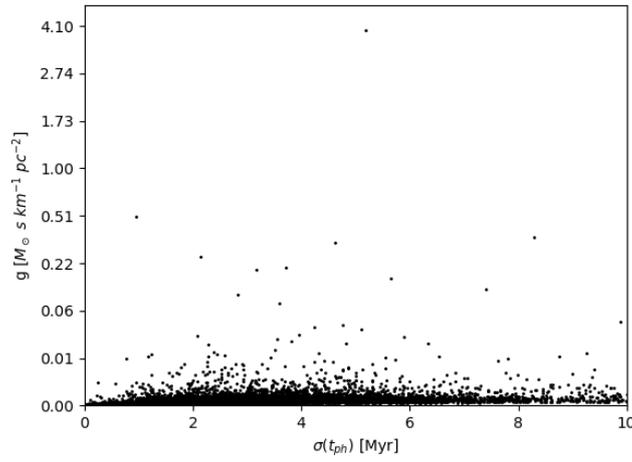
Figure 18: Follow-up on figure 17, showing $\sigma(t_{ph})$ versus g. Only values corresponding to an uncertainty of 10 mas yr^{-1} are plotted, since $\sigma(t_{ph})$ is almost not dependent on the uncertainty in proper motion. The plot shows no indication of a correlation. Note the cube root scale of the vertical axis, implemented to increase the visibility of the differences in g value.
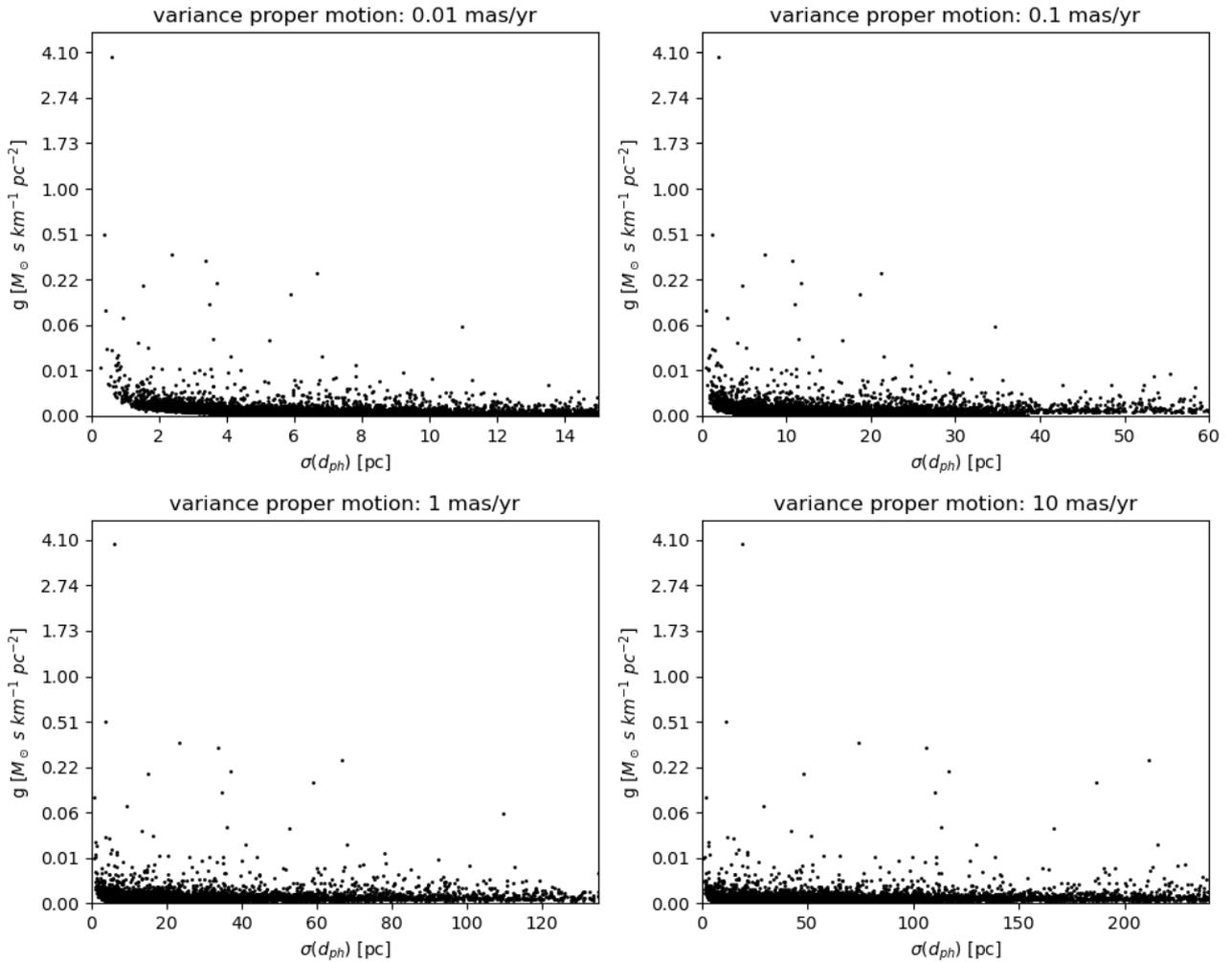
Figure 19: Uncertainty in perihelion distance ($\sigma(d_{ph})$) versus the Oort Cloud perturbation strength g. Each panel corresponds to a different value of the uncertainty in proper motion. The plot shows no indication of a correlation. Note the cube root scale of the vertical axis, implemented to increase the visibility of the differences in g value.
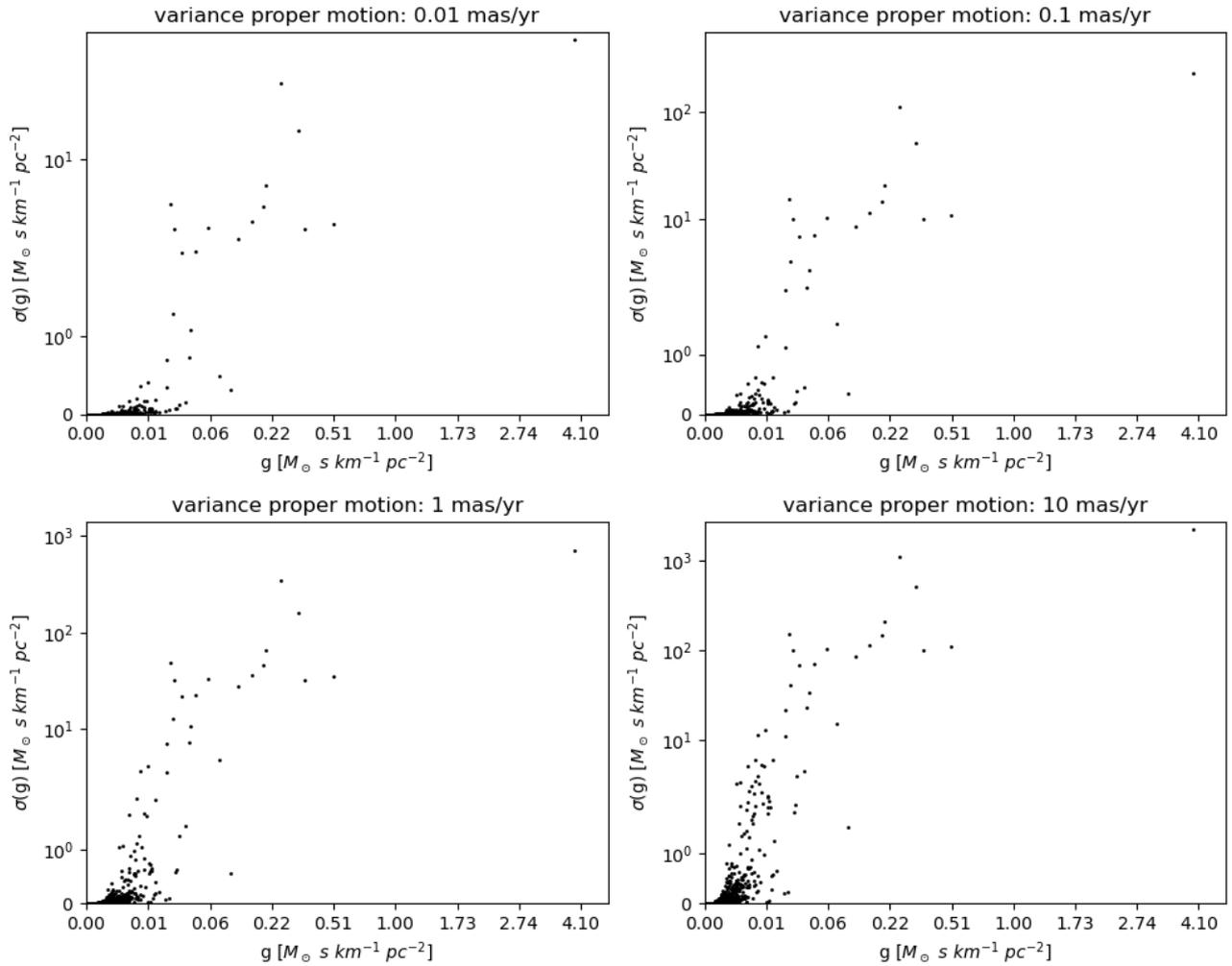
Figure 20: Oort Cloud perturbation strength *g* versus the uncertainty in *g*. Each panel corresponds to a different value of the uncertainty in proper motion. Note the logarithmic and cube root scales on the vertical and horizontal axes, respectively, to increase readability. Each panel shows a positive correlation between *g* and its uncertainty.

with the results of Stuart & Binzel (2004) to show the consistency in the distribution.

To retrieve relevant information about the distribution of LPCs, magnitude distribution data provided by Fernandez & Sosa (2012) was used. Their work encompasses LPCs discovered throughout the history of human recordings and they estimated absolute total visual magnitudes H.

I convert the LPC distribution of H-magnitude to diameter using the relation:

$$\log D[\text{km}] = 0.9 - 0.13H \qquad (16)$$

as derived by Sosa & Fernández (2011). They found this relation by estimating the mass of long period comets under consideration of their visual light curves and non-gravitational parameters, and consequently relating the masses with the observed brightness of comets. This conversion assumes a uniform bulk density (ρ) of 0.6 g/cm^3, which is the most common value for cometary bodies.

To estimate the distribution of larger asteroids and comets than found in the literature, simple power laws are applied to the established data. They have the form $N(D) \propto D^{-q}$, where D is the diameter of the object, with $q \approx 3.5$ for asteroids (Dohnanyi 1969; Pena et al. 2020) and $q \approx 1.4 - 2.6$ for long period comets (Meech et al. 2004; Fernandez & Sosa 2012; Boe et al. 2019). In this way I extrapolate the frequency and potential impact energy to the objects related to the extreme impact events that might be relevant in my research.

Figure 21 illustrates the cumulative size distribution for asteroids and LPCs. For larger object sizes the distributions are extrapolated conforming to the power law $N(D) \propto D^{-q}$, using $q = 3.5$ for asteroids and $q = 2$ for comets. Notably, the cometary size distribution demonstrates a 'knee' around a diameter of 12 km and flattens for sizes below approximately 8 km, a pattern that can be attributed to the natural break-up and elimination processes of LPCs, as described in section 2.2.
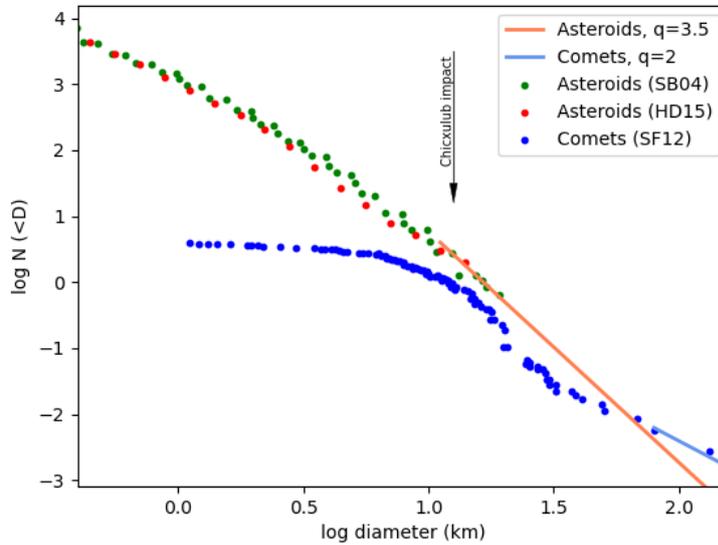
Figure 21: Cumulative size distribution of asteroids (in red and green) and comets (in blue). The red dots are from Harris & D'Abramo (2015), the green from Stuart & Binzel (2004). The blue dots are adapted from the H-magnitude data of Fernandez & Sosa (2012), representing an estimate of the full population of near-Earth long period comets. The red and blue lines are extrapolations of the estimated data based on a simple power law. The size of the K-Pg impactor is indicated for reference.

In order to obtain the impact probability distribution of LPCs, I multiply the total comet flux within 1.3 AU of the Sun, of 0.9 yr^{-1} identified by Fouchard et al. (2017) with the probability of impact per perihelion passage of 3×10^{-9} as computed by Steel (1993). I further refine this probability

value by adjusting it for earth-crossing orbits. For this, I incorporate the assumption of a uniform perihelion distribution within the inner solar system, which entails a linear relationship between perihelion distance and flux (Vokrouhlický et al. 2019). In this case, it pertains to dividing the total flux of 0.9 yr^{-1} by 1.3.

The comparison of impact frequencies reveals that approximately 2400 asteroids cross Earth's orbit annually, with a much smaller number of LPCs doing so, leading to respective impact probabilities of $\sim 10^{-6}$ and $\sim 10^{-9}$ per year. However, LPCs are notably more prevalent among larger potential impactors, particularly those with diameters exceeding 2 km. The variability in estimates for the number flux of these larger objects is primarily due to the scarcity of observations from which to derive size-frequency distributions. For instance, the slope of the size-frequency distribution of Earth-crossing objects is approximately 10^{-4} for asteroids greater than 10 km in diameter and $10^{-3} - 10^{-1}$ for comets of similar size. During comet showers induced by stellar flyby's the flux of LPCs can increase significantly, raising the background flux with around an order of magnitude (Berski & Dybczyński 2016; Feng & Bailer-Jones 2014).

The frequency of Earth's encounters with asteroids and LPCs is quantified in Figure 22. While asteroids, especially smaller ones, encounter Earth with greater frequency, they largely disintegrate in the atmosphere. The impact frequencies of LPCs, depicted in blue, were computed based on the total comet flux and the impact probability per perihelion passage.
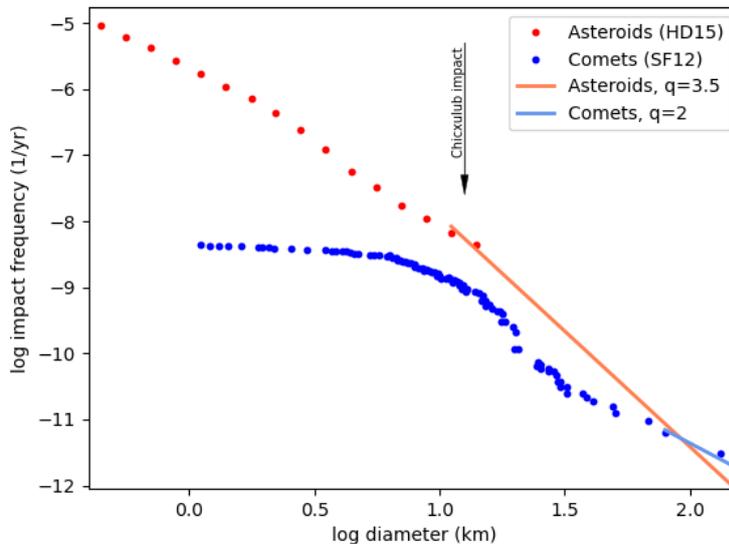
Figure 22: Impact frequency versus diameter for comets and asteroids. The estimates for asteroid impact frequencies (in red) are taken from Harris & D'Abramo (2015). The impact frequencies of comets (in blue) were computed based on the total comet flux and the impact probability per perihelion passage. The size of the K-Pg impactor is indicated for reference.

Impact energy depends mainly on the mass and velocity of the impacting object. It is linearly proportional to the mass and scales quadratically with the velocity. Consequently, LPCs, due to their significantly higher velocities when compared to asteroids, can possess substantially greater impact energies.

To yield the distribution of impact energy, I employ the simple kinetic energy relation:

$$E_{\text{kin}} = 0.5 \cdot m \cdot v^2 \tag{17}$$

where the mass (m) is inferred from the bulk density and the diameter distribution, and the velocity (v) is based on the mean velocity of 55 km s^{-1} as reported in the literature, particularly by Steel (1998).

As depicted in Figure 23, the computed impact energies are presented in megatons of TNT equivalence (Mt), a common unit in the field of impact science. It is important to note, however, that the actual impact energy can vary from the plotted estimates, as both the densities and velocities of impactors are not uniform and can differ from the assumed averages.
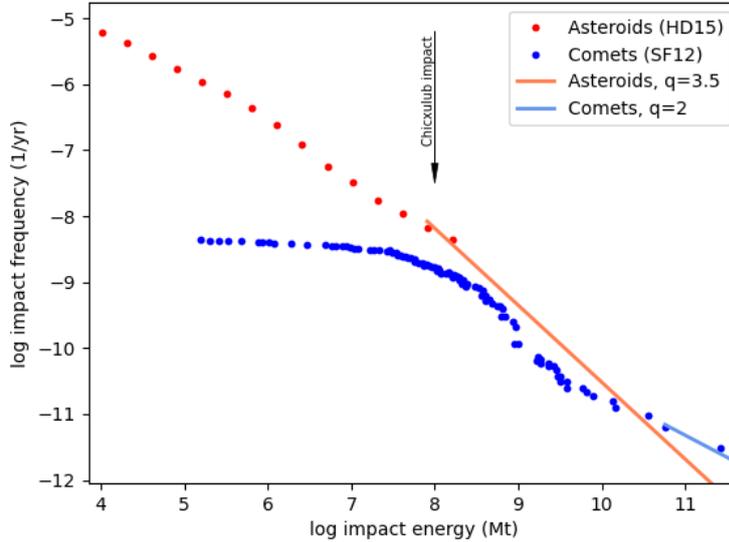
Figure 23: Impact frequency versus impact energy in units equivalent to megatons of TNT (Mt). The impact energy of the asteroids (in red) is taken from Harris & D'Abramo (2015). The impact energy of the comets (in blue) is computed as a function of their mass and velocity, assuming a bulk density of 0.6 g cm^{-3} and an average impact velocity of 55 km s^{-1}. The impact energy of the K-Pg impactor is indicated for reference.

6.2.1   Estimating P(encounter) and P(comet | encounter)

The data in table 3 suggest that there are approximately 20 close stellar encounters with a perihelion distance less than 1 parsec, per million years—570 per 30 Myr is roughly 20 per Myr. This is in line with the analysis of Bailer-Jones et al. (2018). So I will posit a value of 20 per Myr as the value of P(encounter).

Determining whether the Earth is currently within a comet shower, possibly a weak one, remains challenging due to incomplete data regarding recent close stellar encounters (Bailer-Jones et al. 2018). However, Feng & Bailer-Jones (2015) estimate on the basis of known recent past flyby's that only about 5 percent of the current long period comet flux is induced by stellar encounters. In general, comet showers last for several million years and can raise the impact frequency by an order of magnitude (Berski & Dybczyński 2016; Feng & Bailer-Jones 2014).

Therefore, I estimate the lower limit of the encounter-induced cometary impact probability based on the aforementioned 5 percent of the current observed flux of LPCs. The current impact frequency (as can be inferred from the leftmost cometary data point in figure 23) is a bit less than 10^{-2} impacts per Myr. A conservative inference (5% of a bit less than 10^{-2}) gives a lower limit of the order of magnitude 10^{-4} Myr^{-1}. Conversely, the upper limit presupposes that the Earth is not currently experiencing a comet shower and that during several million years after a stellar flyby, the Earth will witness an elevated comet flux, potentially a magnitude larger than present observations. Hence, I will posit P(comet | encounter) to be in the range of 10^{-4} to 10^{-3} per Myr.

6.3   Impact-Induced Climate Change

In the context of terrestrial impacts and their climatic consequences, the Chicxulub impact event is unparalleled in its contribution to scientific understanding. It remains the only impact event in Earth's history that is confidently linked to global climate effects. The major part of the most comprehensive research investigating the cratering process including the production of ejecta, and their subsequent influence on the global climate, focuses on Chicxulub.

In order to make general claims about the consequences of cometary impacts and an estimation of P(climate change — encounter & comet), I untangle the processes that contributed to the climate effects due to Chixculub and estimate their significance for varying comet impact energy. Furthermore, the few publications that make more general (that is, beyond Chixculub alone) statements about the consequences of impacts—Artemieva et al. (2017); Toon et al. (1997, 2016)—also proved useful for my analysis.

The three types of ejecta that determine the climate effect of an impact are dust, sulfates and soot. For each of these ejecta, I figure out the amount that is needed to cause a global climate effect and the impact energy required to produce that threshold amount of ejecta. I also consider the influence of photosynthesis on the terrestrial climate. Subsequently, I synthesize the thresholds for each factor into a threshold impact energy that is enough to cause planet-wide climate consequences.
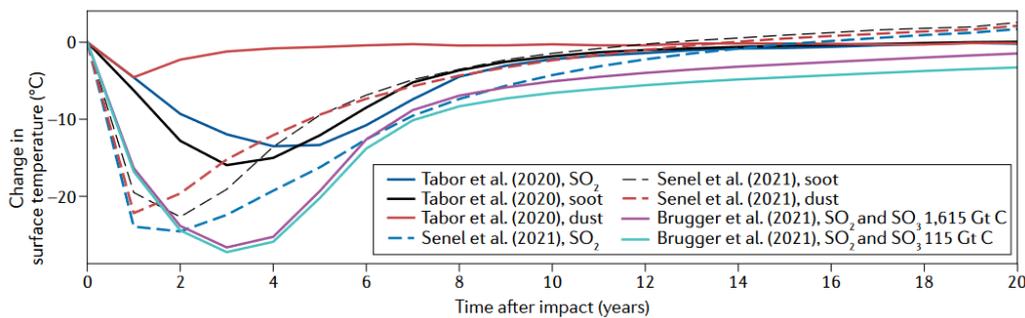
Figure 24: Surface air temperature change over time under the influence of different variables (soot, dust, and two different sulfate aerosols; SO_2 and SO_3). Results from three separate simulations (Tabor et al. 2020; Senel et al. 2021; Brugger et al. 2021) are shown, each simulation was set up to model the global surface air temperature consequences of only one (or a few) climate active impact ejecta to show their individual contribution to the temperature change. Figure adapted from (Morgan et al. 2022).

6.3.1   Dust

The dust produced during the Chicxulub impact event is a key factor in understanding the climatic aftermath of such events. The total mass of dust generated by the Chicxulub impact is estimated to range between 10^{16} and 10^{19} g, predominantly consisting of pulverized sedimentary material (Toon et al. 2016).

There is considerable uncertainty in the composition of this dust, particularly concerning iron-rich nanoparticles. Different studies propose a wide range of estimates for the amount of iron-rich dust, from none to as high as 10^{18} g. This variation significantly influences the darkening effect of the dust, leading to disparate outcomes in different models that simulate the climatic consequences of the impact.

Two notable simulations, those by Tabor et al. (2020) and Senel et al. (2021), despite using similar total amounts of dust, arrive at vastly different conclusions regarding the cooling timeline post-impact (see figure 24). In the model by Tabor et al. (2020), iron-rich particles are hypothesized to coagulate

rapidly and fall out of the atmosphere, resulting in minimal impact cooling. Conversely, the Senel et al. (2021) model assumes the absence of iron and coagulation, leading to a more pronounced and prolonged dust-induced cooling effect.[13]

Another key aspect of the influence of dust on the climate is its spreading process over the globe. Impact spherules, which are solidified melt droplets ejected in the ejecta curtain, are found up to 6,000 km from the Chicxulub crater. Beyond this distance an iridium-rich clay layer is observed, which is distributed globally at the K-Pg boundary (Belza et al. 2017). The global nature of this iridium enrichment indicates widespread distribution of certain ejecta components.

The initial leading theory proposed that the global K-Pg layer was formed by impact ejecta spreading over the Earth after being ballistically injected into the stratosphere, settling over a few years to create a uniformly thick layer (Morgan et al. 2022). This theory is supported by findings such as the presence of microkrystites in the K-Pg layer (Smit 1999), which are condensed spherical particles formed in the impact plume.

However, this theory faces two main challenges. Firstly, the presence of shocked minerals globally in the K-Pg layer suggests very high ejection velocities, which are inconsistent with the maximum ejection velocities for minerals sourced from deep layers at the impact site (Alvarez et al. 1995). Secondly, the extraordinary consistency in the thickness of the K-Pg layer implies a majority of the ejecta mass in the impact plume, conflicting with numerical simulations that suggest most ejecta mass resides in the ejecta curtain (Artemieva & Morgan 2020).

A plausible resolution to these challenges is the consideration of a fast-moving cloud, as shown in Figure 25, containing soot, shocked rocks, dust, and sulfate aerosols from both the target material and the impactor (Artemieva & Morgan 2009, 2020). This theory aligns with the post-ballistic skidding process described by Toon et al. (1997). The cloud, formed by the heating of the atmosphere due to interactions with the ejecta curtain, spreads rapidly across the globe, depositing materials as it travels. Within hours of the impact, this cloud envelops the Earth, explaining the global distribution of ejecta-related materials and the consistency in the K-Pg layer thickness.

For large impacts in general, Toon et al. (2016) estimate that about 40% of the projectile mass, in addition to the same amount of target material, is ejected as dust in the atmosphere. They estimate that an impact of the order 10^5 Mt would produce dust with a considerable optical depth of 1.5. With a ballistic spreading mechanism, it would spread over half the globe. However, if we consider the cloud spreading effect, a similar amount of energy, around 10^5 Mt, might suffice for global coverage and potential wildfire ignition.

Despite the potential for widespread dust distribution, the optical depth of the order unity appears insufficient for causing significant climate effects, especially in the time span of multiple years. This is primarily because dust particles are rapidly removed from the atmosphere, diminishing their climatic impact. I reckon that a global optical depth in the order of 10 is potentially enough to trigger planet-wide cooling that lasts several years.

Given these factors, and assuming a linear relation between optical depth and ejected dust mass, I estimate that the impact energy threshold for significant dust influence on global climate is of the order of 10^6 Mt.

6.3.2   Sulfates

Sunlight-blocking droplets containing sulfur were probably crucial for the strong consequences of the Chicxulub impact. Compared to sulfates, an impact produces more dust and the darkening properties

---

[13]Georg Feulner (priv. comm.) pointed out that the Senel et al. (2021) model might actually compute surface temperature instead of near-surface *air* temperature, regarding the surprisingly high impact winter Arctic Ocean temperatures depicted in Senel et al. (2023). This concern does not undermine the general pattern of temperature change that is shown in figure 24, however.

a Numerical simulation at 300 s

300 s

Top of atmosphere

Interactions between the ejecta curtain and atmosphere form a fast-moving cloud

Ejecta curtain

450 km

Top of stratosphere

1,000 km

b Numerical simulation at 3,600 s

Two theoretical travel paths for high-velocity plume ejecta, which arrive at the top of the atmosphere <1 hour after impact

6.4 km s⁻¹ 45°

8.0 km s⁻¹ 35°

3,600 s

Ejecta being deposited from cloud at proximal sites

Front of cloud travels above the stratosphere at 2–4 km s⁻¹

Cloud

350 km

Chicxulub   2,000   4,000   6,000   Europe   Top of stratosphere

Distance (km)

Cloud density (kg m⁻³): 1 to 10⁻⁷

Atmosphere density (kg m⁻³): 1 to 10⁻⁹

Figure 25: a) 300 seconds after the impact, illustrating the division of the ejecta curtain and the formation of a fast-moving cloud. b) 3,600 seconds after the impact, showing the extensive travel and ongoing velocity of the cloud. Adapted from (Artemieva & Morgan 2020).

of soot are stronger. But dust rains out of the atmosphere quickly and it is questionable whether soot could have been produced globally. Sulfate aerosols easily spread over the planet and can reside for years in the atmosphere.

Advanced numerical simulations tell that they caused a global mean temperature drop of as much as 15 to 26 degrees Celsius in two to four years after the impact. It took about 15 to 30 years for the global mean temperature to return to the pre-impact level (Brugger et al. 2017, 2021; Tabor et al. 2020). The target rock contributed almost all of the sulfur ejected into the atmosphere (Kaiho & Oshima 2017), so the large sulfate production was mainly due to the impact location. The amount of sulfur produced during the Chicxulub impact is in the range of 10^{17}–10^{18} g (Toon et al. 2016; Artemieva et al. 2017).

The ratio of sulfur dioxide (SO_2) to sulfur trioxide (SO_3) could be significant for the climate effect of the aerosols that they form due to their different residence time and particle sizes. Most prominent simulations of the Chicxulub impact's climate effects have assumed that sulfur dioxide production exceeds that of sulfur trioxide (D'Hondt et al. 1994). However, more recent theoretical and experimental research suggests the contrary, indicating that SO_3 production may be higher during impacts (Ohno et al. 2014). Particularly in large impact plumes, the SO_3/SO_2 ratio can be considerably high due to SO_3's stability at lower temperatures and the slower cooling process of larger plumes. Chemical reactions with ocean water vapor further influence the ratio, adding to the sensitivity to impact location (Toon et al. 1997).

The amount of sulfate aerosols produced in impacts smaller than Chicxulub scales roughly linearly with the impact energy (Artemieva et al. 2017). So for a 10^5 Mt impactor (the smallest long period comets, as can be seen in figure 23), the estimated amount of sulfur produced, given the same sulfur-rich target rock as for Chicxulub, is 10^{14} g (Toon et al. 2016).

The temperature change resulting from sulfate aerosols does not depend heavily on the precise amount of sulfur released, as radiative forcing plateaus beyond a sulfur mass of 30 Gt (Pierazzo et al. 2003). Even relatively small amounts can cause a global temperature dip, as demonstrated by the 1991 Pinatubo volcanic eruption (Hoff 1992). However, for the cooling effect to persist beyond a year, approximately 10^{15} g of sulfur is necessary.

Assuming a linear relation between sulfur production and impact energy, an impact of around 10^6 Mt would suffice for lasting global effects, especially if it hits a sulfur-rich target like the one at the Chicxulub event (Pierazzo et al. 2003). In comets, the maximum sulfur content is about one percent, equating to a sulfur mass of 10^{15} g for a comet mass of 10^{17} g. This corresponds to an impact energy on the order of 10^8 Mt, so a very energetic impact is needed for sulfates to cause a global climate effect if the target rock is sulfur-poor.

Given that most comets have a lower sulfur abundance and most terrestrial rocks contain less sulfur than the Chicxulub target, sulfate aerosols are unlikely to be the primary determinant of the impact energy threshold for global climate change. Consequently, the probability of global sulfate effects from an impact is estimated to be akin to that of a Chicxulub caliber event. Impacts on the order of 10^6 to 10^8 Mt, particularly on sulfur-rich targets, are likely necessary for sulfate aerosols to induce global climate effects.

6.3.3   Soot

The thermal radiation of the impact plume could ignite wildfires in a large local range, but to produce enough soot to significantly reduce the global incoming sunlight, other mechanisms have to be considered. Re-entry of ejecta into the atmosphere after ballistic ejection could ignite global wildfires (Morgan et al. 2013; Robertson et al. 2013), but the Chicxulub impact was probably not energetic enough for such a process to account for the amount of organic particles present in the corresponding geological layer. The rapid hot dust cloud scenario described in section 6.3.1 might explain ignition

of global wildfires (Artemieva & Morgan 2020). The found abundance of organic particles, however, could also originate from the target rock, which is rich in polycyclic aromatic hydrocarbons (PAHs) (Shelby et al. 2020).

The amount of soot produced during the Chicxulub event is estimated to be in the range between 10^{14} and 10^{16} g (Kaiho et al. 2016; Bardeen et al. 2017). With soot injections of the lower end of this range, moderate global climate effects occur (Bardeen et al. 2017).

The production of organic particles, or soot, during the Chicxulub impact did not depend on the impact location in orders of magnitude. I assume that that soot production is primarily dependent on the size and temperature of the ejecta plume or cloud, which in turn scales linearly with the impact energy. Under this assumption, the amount of soot produced is correlated with the impact energy.

To estimate the threshold for soot-induced climate change, I presume, in line with Bardeen et al. (2017), that approximately 10^{14} g of soot is needed. Extrapolating from the Chixculub event, this translates into an impact energy of about 10^7 Mt.

6.3.4   Photosynthesis and the carbon cycle

Phytoplankton is responsible for approximately half of the Earth's net primary productivity (NPP), a critical ecological metric representing the rate at which all plants in an ecosystem synthesize new biomass through photosynthesis, minus the rate of carbon they respire. NPP is a key indicator of the amount of carbon sequestered from the atmosphere and plays a vital role in the carbon cycle. The impact winter in the first year after the Chicxulub event almost completely halted photosynthesis on Earth (Brugger et al. 2021). However, once the darkness receded and nutrients such as iron, phosphorus, and nitrogen settled in the oceans, a surge in phytoplankton growth likely occurred (Behrenfeld et al. 2006). These nutrients could have come from the target rock, the asteroid, and post-impact wildfires (Ardyna et al. 2022; Vellekoop et al. 2018; Morgan et al. 2022). The bloom was further enhanced by ocean upwelling due to the sudden cooling of the ocean surface (Brugger et al. 2021).

Over longer timescales, the impact produced a global warming of 1 to 2 degrees Celsius, lasting for about 100,000 years, primarily driven by high carbon dioxide (CO_2) levels (Vellekoop et al. 2018). This CO_2 originated in the vaporization of carbonate rock, burning of organic material in wildfires, and soil decay during the productivity hiatus caused by the impact winter (Tyrrell et al. 2015; Giardina & Ryan 2000). The increased photosynthesis post-impact acted to lower atmospheric CO_2 levels, thus moderating the global warming effect.

Terrestrial productivity and biomass returned to pre-impact levels within a decade and a century, respectively. However, the recovery in the oceans was potentially much slower, with some models suggesting it could have taken millions of years for the organic carbon flux to the deep ocean to be restored (Lomax et al. 2001; Henehan et al. 2019; Sepúlveda et al. 2019), although there is some contention over this, with other models proposing a much shorter recovery period (Sepúlveda et al. 2009).

Any decrease in sunlight directly affects photosynthesis, thereby influencing the NPP. The degree of sunlight reduction, whether caused by dust, soot, or sulfate aerosols, varies with the scale of the impact event. The same holds for the deposition of nutrients like iron and phosphorus, either from the impact ejecta or secondary sources such as wildfires.

Deep sea upwelling caused by sea surface cooling, which can increase productivity, might be triggered by a certain amount of temperature drop (Brugger et al. 2021). However, the specific threshold for this phenomenon is not well-defined and is likely too complex to establish accurately for impact events in this thesis. Furthermore, since the increase in photosynthesis and the greenhouse effect from elevated CO_2 levels counterbalance each other, this consequence of an impact does not result in a pronounced global temperature shift.

Considering the lack of a clear threshold effect and the relatively limited influence on the global

near-surface air temperature, I do not include photosynthesis in my estimation of the P(climate change | comet) range of values.

6.3.5   Estimating P(climate change | comet) and P(encounter $\cap$ comet $\cap$ climate change)

From the above analysis I conclude that the range of impact energies that potentially cause global climate effects is from 10^6 to 10^8 Mt. On the lower end of this range, only dust has a global influence and on the upper end, all three climate-active ejecta have planet-wide effects. This indicates that an impact with an energy of the order 10^8 Mt or more, such as the Chixculub impact, almost certainly produces global climate change. And impacts with an energy less than the order 10^6 Mt probably do not cause global temperature change lasting for a few years.

The cometary impact frequency distribution (see figure 23) suggests that virtually all of the long period comets possess sufficient energy to exceed the lower climate effect threshold of 10^6 Mt. The upper end of the impact energy threshold, 10^8 Mt, is reached by approximately half of the encounter-induced comets. Therefore, I posit the probability value of P(climate change | comet) to range from 10^{-1} to 1.

Note that the contribution of long period comets to the total number of impacts with an energy higher than the lower end of the threshold—in other words, all impacts more energetic than 10^6 Mt—is small; many more asteroids than comets are expected to impact with this energy (see figure 23). However, for the upper end of the impact energy threshold (impacts of more than 10^8 Mt), the cometary contribution is significant (accounting for about one third of the total). Therefore it is highly relevant to research long period comets in the light of impact-induced climate change.

Multiplying 20 Myr^{-1} for P(encounter) with 10^{-4}–10^{-3} for P(comet | encounter) and with 10^{-1}–1 for P(climate change | comet), yields a range of 10^{-4} to 10^{-2} encounter-comet-climate change events per million years. This corresponds to an average time of 0.1 to 10 Gyr in between such events, or maximally several dozen such events in the Earth's history. To place this number into context, the expected number of asteroid impacts in the global climate change impact energy range (10^6–10^8 Mt) is of the order 10^{-2}–10^{-1} per million years. So, even though it is expected that more climate changing impacts result from asteroids than from long period comets, the contribution of the latter is not insignificant.

6.4   Critical Evaluation of the Sequence Equation Framework

There are several aspects of the Sequence Equation framework that warrant a critical examination. One of the primary challenges in applying the Sequence Equation is the substantial uncertainty inherent in the scarcely available data. Additionally, numerical models often focus on specific cases (such as the Chicxulub event), which limits their generalizability and further amplifies the uncertainties involved in such calculations.

Even though I provided a range of values for each of the factors in the equation, and a crude "ball park" estimate can already be a valuable result of the application of this framework, it can be questioned if the accuracy of the results is high enough to have significant scientific value. On the other hand, the framework is a useful tool for highlighting the key sources of uncertainty that are present in the research on this topic. The components of the flyby-comet-climate research that I outlined and their associated probability ranges that I presented can be valuable for future investigations in the context of this topic.

Furthermore, the framework may be too simplistic to effectively capture the intricacies of the involved processes.[14] Trying to distill a complex chain of events, which includes various astronomical

---

[14]Similar commentary is delivered at the Drake equation. For example, Madau (2023) suggests adjustments of the

and climatic phenomena, into a single equation can lead to the omission of relevant information. This simplification might be akin to "using a cannon to kill a mosquito", where the broad scope of the equation overshadows the nuanced details of each component.

Moreover, the outcome of the Sequence Equation, a singular probability value, might not provide a comprehensive or deep understanding of the Earth's embeddedness in larger astronomical systems. A single number, especially if it is very small, may not effectively convey the scale or significance of the impact of these events on a scale relatable to planetary time scales.

The above reflection is meant to indicate the limitations of the Sequence Equation framework as it is presented in this thesis. I do have the hope, however, that those rough edges can be further refined. I still think that it is a valuable framework to assess and discuss the flyby-comet-climate sequence of events and, in a wider sense, the embeddedness of the Earth system. It gives the opportunity to quantify the probability of the occurrence of such a sequence and the current research uncertainties associated with it in a segmented way, and it can serve as a scientific "conversation starter" to enhance research on this topic.

7 Conclusion

The capacity of Euclid to provide observational data that can be used to select a subset of stellar flyby candidates is promising. I will summarize the main results of my analysis of the Euclid mock stellar catalog.

Based on only radial distance and proper motion measurements, even with a variance of 10 mas yr^{-1} (which should be attainable with Euclid) in proper motion, the selected subset has a true positive rate of 71% for encounters closer than 1 parsec and 77% for encounters within 5 parsec, occurring in ± 15 Myr. That means that only 29% and 23% of the close encounters are excluded from this subset.

With the same proper motion measurement accuracy, the selected subset has a positive predictive value of 7.7×10^{-6} (encounter within 1 pc and ± 15 Myr) and 1.4×10^{-4} (within 5 pc). This indicates that, if one would precisely measure the perihelion distances and times of this subset, they would have to measure, respectively, 1.3×10^{5} and 7.1×10^{3} stars for each close encounter that they want to obtain.

If accurate proper motion measurements are included, with a variance of 0.01 mas yr^{-1}, the true positive rate of the selected subset grows to 86% (95%) for encounters within 1 (5) pc, such that it misses only 14% (5%) of flyby's. The associated positive predictive values, of 1.5×10^{-4} (1.5×10^{-3}) indicate that one would need to scrutinize 6.7×10^{3} (6.7×10^{2}) to find a close encounter.

In the context of the current state-of-the-art of close encounter research, Euclid provides the potential for a significant improvement. The most complete catalog of close stellar encounters contains approximately 15% of encounters closer than 5 pc within ± 5 Myr (Bailer-Jones et al. 2018). According to my analysis, Euclid's instruments are sensitive enough to observe all encounters closer than 5 pc within ± 5 Myr, and 99% (96%) of the <1 pc (<5 pc) encounters within ± 15 Myr. This completeness reduces to 13% (11%) due to the limited survey area of Euclid.

Combining the high photometric sensitivity with the true positive rate of the close encounter candidate selection on basis of radial distance and proper motion yields a great rise in completeness with respect to the Gaia-based catalog. In the scenario in which Euclid (with a proper motion variance of 10 mas yr^{-1}) would be employed to survey the full sky, it could thus create a subset of observations that has a completeness of approximately 70% (74%) of <1 pc (<5 pc) encounters within ± 15 Myr.

This analysis points at a large potential for deep wide-field imaging surveys to strengthen stellar research. Combining the qualities of multiple apparatuses can yield data that adds to the current

---

Drake equation such that it considers the time dependency of several of its factors.

knowledge. Probing outside of the scope of the initial aims of a mission, such as the observation of distant galaxies in the case of Euclid, can unleash even more potential of the project.

It is good to note, however, that for detailed astrometric measurements the Euclid observations must be combined with spectroscopic and proper motion measurements. And even if that would be accomplished, gaining accurate astrometric information about the potential close encounters in ±15 Myr would still be challenging. In a straightforward application of typical measurement uncertainties as employed in this thesis, the true positive rate and positive predictive value of the best case scenario are 49% for encounters within 1 pc (62% for encounters within 5 pc) and 1.1×10^{-3} (6.2×10^{-3}), respectively.

The analysis in this thesis provides a basis for further exploring the possibilities of the application of Euclid to stellar investigations in future research. It could be useful to use an early Euclid data release as an indication of its potential for encounter research.

The general conclusion of the application of the Sequence Equation, is that the number of expected encounter-comet-climate change events per million years ranges between 10^{-4} and 10^{-2}, or between none and a few dozen in the Earth's history.

Even though this means that only a few encounter-induced comets have impacted Earth in its history—and will in its future, with the Sun being roughly halfway its Main Sequence lifetime—these events are extremely influential for the whole Earth System. This is testified by the severe consequences of the Chicxulub impact, which ended the Cretaceous and initiated the Paleogene geological period.

That famous asteroid impact further highlights the relevance of encounter-induced post-impact climate change. One instance does not give strong statistical backing—which, as I pointed out above, is a general difficulty for the research on climate change as a consequence of impacts—but the Chicxulub event does show that there is a positive probability of impact-induced climate change. The analysis above indicates that climate changing asteroid impacts are expected to occur only ten times as often, or even have a frequency of the same order of magnitude as flyby-comet-climate events.

The Sequence Equation, which is developed to quantify the probability of a chain of events leading from a stellar encounter to significant climate change, offers a framework for situating and interpreting the findings from the Euclid mission. This framework facilitates a systematic approach to quantifying, integrating and analyzing the data collected, and it provides an outline for future research and discussion about this topic. As such, it allows for a more comprehensive understanding of the significance of close encounters for the Earth's climate.

There are, however, some limitations to the Sequence Equation framework as presented in this thesis. The climatic consequences of comet impacts are extremely complex, there are still uncertainties in some of the relevant physical and chemical processes, and the research relies on sparse data, which leads to large uncertainties regarding the whole encounter-comet-climate chain. Another limitation is posed by the simple format of the probabilistic equation: by reducing the complex sequence of events to a probability value, arguably a lot of valuable insight is lost.

For future research that further applies the Sequence Equation, it would be useful to consider the various research gaps and uncertainties that were pointed out in this thesis. In summary, each component of the Sequence Equation has a major source of uncertainty. The current knowledge of recent and near-future close stellar encounters is limited; observations of fainter stars can significantly improve on this. To gain more insight in the production of comets by stellar flyby's, more numerical simulations of stellar Oort Cloud perturbations are needed, employing the increasing observational knowledge of close stellar encounters. But more information about the composition of the Oort Cloud, both the solar one and those of encountering stars, would also contribute to the understanding of the second component of the Sequence Equation. In order to improve on the understanding of impact-induced climate change, more generalizable research would be valuable. There is much to gain in research on the production of ejecta by impacts of varying energy and impact location, and on

the climatic consequences of varying ejecta inputs.

The encounter-comet-climate change sequence of events is a distinctive and significant manifestation of the embeddedness of the Earth in larger systems, particularly the Solar System and the Milky Way. The Sequence Equation provides a quantitative framework to interpret the chain of events, and the thesis' analysis of the research potential of deep wide-field imaging surveys an initial application of that framework. As such, this thesis is a step towards an integrative understanding of encounter-comet-climate events and our home planet in general.

Acknowledgments

References

Adams, F. C. 2010, Annual Review of Astronomy and Astrophysics, 48, 47

Alvarez, L. W., Alvarez, W., Asaro, F., & Michel, H. V. 1980, Science, 208, 1095

Alvarez, W., Claeys, P., & Kieffer, S. W. 1995, Science, 269, 930

Archibald, J. D. 1996, Dinosaur extinction and the end of an era: what the fossils say (Columbia University Press)

Ardyna, M., Hamilton, D. S., Harmel, T., et al. 2022, Communications Earth & Environment, 3, 201

Artemieva, N. & Morgan, J. 2009, Icarus, 201, 768

Artemieva, N. & Morgan, J. 2020, Geophysical Research Letters, 47

Artemieva, N., Morgan, J., & Expedition 364 Science Party. 2017, Geophysical Research Letters, 44, 10

Bailer-Jones, C. A. L. 2009, International Journal of Astrobiology, 8, 213

Bailer-Jones, C. A. L. 2011, Monthly Notices of the Royal Astronomical Society, 416, 1163

Bailer-Jones, C. A. L. 2015, Astronomy & Astrophysics, 575, A35

Bailer-Jones, C. A. L. 2022, The Astrophysical Journal Letters, 935, L9

Bailer-Jones, C. A. L., Rybizki, J., Andrae, R., & Fouesneau, M. 2018, Astronomy & Astrophysics, 616, A37

Bardeen, C. G., Garcia, R. R., Toon, O. B., & Conley, A. J. 2017, Proceedings of the National Academy of Sciences, 114, E7415

Behrenfeld, M. J., O'Malley, R. T., Siegel, D. A., et al. 2006, Nature, 444, 752

Belza, J., Goderis, S., Montanari, A., Vanhaecke, F., & Claeys, P. 2017, Geochimica et Cosmochimica Acta, 202, 231

Bernstein, G., Armstrong, R., Plazas, A., et al. 2017, Publications of the Astronomical Society of the Pacific, 129, 074503

Berski, F. & Dybczyński, P. A. 2016, Astronomy & Astrophysics, 595, L10

Bertelli, G., Bressan, A., Chiosi, C., Fagotto, F., & Nasi, E. 1994, Astronomy and Astrophysics, 106, 275

Bobylev, V. V. & Bajkova, A. T. 2020, Astronomy Letters, 46, 245

Boe, B., Jedicke, R., Meech, K. J., et al. 2019, Icarus, 333, 252

Brugger, J., Feulner, G., Hofmann, M., & Petri, S. 2021, Geophysical Research Letters, 48

Brugger, J., Feulner, G., & Petri, S. 2017, Geophysical Research Letters, 44, 419

Bullock, J. S. & Johnston, K. V. 2005, The Astrophysical Journal, 635, 931

Carry, B. 2012, Planetary and Space Science, 73, 98

Davis, M., Hut, P., & Muller, R. A. 1984, Nature, 308, 715

D'Hondt, S. 2005, Annual Review of Ecology, Evolution, and Systematics, 36, 295

D'Hondt, S., Pilson, M. E. Q., Sigurdsson, H., Hanson, Alfred K., J., & Carey, S. 1994, Geology, 22, 983

Do, A., Tucker, M. A., & Tonry, J. 2018, The Astrophysical Journal Letters, 855, L10

Dohnanyi, J. S. 1969, Journal of Geophysical Research, 74, 2531

Drake, F. 1965, Current aspects of exobiology, 32, 323

Dybczyński, P. 2002, Astronomy & Astrophysics, 396, 283

Euclid Collaboration, Desprez, G., Paltani, S., et al. 2020, Astronomy & Astrophysics, 644, A31

Euclid Collaboration, Scaramella, R., Amiaux, J., et al. 2022, Astronomy & Astrophysics, 662, A112

Feng, F. & Bailer-Jones, C. A. L. 2014, Monthly Notices of the Royal Astronomical Society, 442, 3653

Feng, F. & Bailer-Jones, C. A. L. 2015, Monthly Notices of the Royal Astronomical Society, 454, 3267

Fernandez, J. A. & Sosa, A. 2012, Monthly Notices of the Royal Astronomical Society, 423, 1674

Firestone, R. B. 2021, Monthly Notices of the Royal Astronomical Society, 501, 3350

Fouchard, M., Froeschlé, C., Rickman, H., & Valsecchi, G. 2011, Icarus, 214, 334

Fouchard, M., Rickman, H., Froeschlé, C., & Valsecchi, G. B. 2017, Astronomy & Astrophysics, 604, A24

Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2023, Astronomy & Astrophysics, 674, A1

Gardner, E., Nurmi, P., Flynn, C., & Mikkola, S. 2011, Monthly Notices of the Royal Astronomical Society, 411, 947

Giardina, C. P. & Ryan, M. G. 2000, Nature, 404, 858

Girardi, L., Bressan, A., Bertelli, G., & Chiosi, C. 2000, Astronomy & Astrophysics, 141, 371

Gobat, R., Hong, S. E., Snaith, O., & Hong, S. 2021, The Astrophysical Journal, 921, 157

Guzik, P., Drahus, M., Rusek, K., et al. 2020, Nature Astronomy, 4, 53

Harris, A. W. & D'Abramo, G. 2015, Icarus, 257, 302

Henehan, M. J., Ridgwell, A., Thomas, E., et al. 2019, Proceedings of the National Academy of Sciences, 116, 22500

Hoff, R. M. 1992, Geophysical Research Letters, 19, 175

Hull, P. M., Darroch, S. A. F., & Erwin, D. H. 2015, Nature, 528, 345

Jeffers, S. V., Manley, S. P., Bailey, M. E., & Asher, D. J. 2001, Monthly Notices of the Royal Astronomical Society, 327, 126

Kaib, N. A. & Quinn, T. 2009, Science, 325, 1234

Kaiho, K. & Oshima, N. 2017, Scientific Reports, 7, 14855

Kaiho, K., Oshima, N., Adachi, K., et al. 2016, Scientific reports, 6, 28427

Kring, D. A. & Durda, D. D. 2002, Journal of Geophysical Research: Planets, 107, 6

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv preprint arXiv:1110.3193

Lindegren, L., Klioner, S. A., Hernández, J., et al. 2021, Astronomy and Astrophysics, 649, A2

Lingam, M. & Loeb, A. 2018, The Astronomical Journal, 156, 193

Lomax, B., Beerling, D., Upchurch, G., & Otto-Bliesner, B. 2001, Earth and Planetary Science Letters, 192, 137

Longrich, N. R., Bhullar, B.-A. S., & Gauthier, J. A. 2012, Proceedings of the National Academy of Sciences, 109, 21396

Lowry, S. C. & Weissman, P. R. 2003, Icarus, 164, 492

Madau, P. 2023, The Astrophysical Journal, 957, 66

Marigo, P. & Girardi, L. 2007, Astronomy & Astrophysics, 469, 239

Marigo, P., Girardi, L., Bressan, A., et al. 2008, Astronomy & Astrophysics, 482, 883

Marty, B. 2012, Earth and Planetary Science Letters, 313, 56

Matese, J. J. & Whitmire, D. P. 2011, Icarus, 211, 926

Meech, K. J., Hainaut, O. R., & Marsden, B. G. 2004, Icarus, 170, 463

Morbidelli, A., Lunine, J. I., O'Brien, D. P., Raymond, S. N., & Walsh, K. J. 2012, Annual Review of Earth and Planetary Sciences, 40, 251

Morgan, J., Artemieva, N., & Goldin, T. 2013, Journal of Geophysical Research: Biogeosciences, 118, 1508

Morgan, J. V., Bralower, T. J., Brugger, J., & Wünnemann, K. 2022, Nature Reviews Earth & Environment, 3, 338

Ohno, S., Kadono, T., Kurosawa, K., et al. 2014, Nature Geoscience, 7, 279

Oort, J. H. 1950, Bulletin of the Astronomical Institutes of the Netherlands, 11, 91

Pankhurst, M. J., Stevenson, C. J., & Coldwell, B. C. 2021, Journal of the Geological Society

Pena, J., Fuentes, C., Frster, F., et al. 2020, The Astronomical Journal, 159, 148

Perryman, M. 2009, Astronomical applications of astrometry: ten years of exploitation of the Hipparcos satellite data (Cambridge University Press)

Pierazzo, E., Hahmann, A. N., & Sloan, L. C. 2003, Astrobiology, 3, 99

Portegies Zwart, S. 2021, Astronomy & Astrophysics, 647, A136

Portegies Zwart, S., Torres, S., Pelupessy, I., Bédorf, J., & Cai, M. X. 2018, Monthly Notices of the Royal Astronomical Society: Letters, 479, L17

Queiroz, A. B. d. A., Anders, F., Chiappini, C., et al. 2020, Astronomy & Astrophysics, 638, A76

Quintana, S. N. & Schultz, P. H. 2019, Icarus, 326, 176

Rampino, M. R. & Caldeira, K. 2015, Monthly Notices of the Royal Astronomical Society, 454, 3480

Randall, L. & Reece, M. 2014, Physical review letters, 112, 161301

Rickman, H. 1976, Bulletin of the Astronomical Institutes of Czechoslovakia, 27, 92

Rickman, H., Fouchard, M., Froeschlé, C., & Valsecchi, G. B. 2008, Celestial Mechanics and Dynamical Astronomy, 102, 111

Robertson, D. S., Lewis, W. M., Sheehan, P. M., & Toon, O. B. 2013, Journal of Geophysical Research: Biogeosciences, 118, 329

Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, Astronomy & Astrophysics, 409, 523

Rybizki, J., Demleitner, M., Fouesneau, M., et al. 2018, Publications of the Astronomical Society of the Pacific, 130, 074101

Salotti, J.-M. 2022, Futures, 138, 102933, elsevier

Schlafly, E. F. & Finkbeiner, D. P. 2011, The Astrophysical Journal, 737, 103

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, The Astrophysical Journal, 500, 525

Schulte, P., Alegret, L., Arenillas, I., et al. 2010, Science, 327, 1214

Senel, C., Temel, O., Kaskes, P., et al. 2021, GSA Connects, 53, 6

Senel, C. B., Kaskes, P., Temel, O., et al. 2023, Nature Geoscience, 16, 1033

Sepúlveda, J., Alegret, L., Thomas, E., et al. 2019, Paleoceanography and Paleoclimatology, 34, 1195

Sepúlveda, J., Wendler, J. E., Summons, R. E., & Hinrichs, K.-U. 2009, Science, 326, 129

Sharma, S., Bland-Hawthorn, J., Johnston, K. V., & Binney, J. 2011, The Astrophysical Journal, 730, 3

Shelby, L. L., Allison, T. K., Timothy, J. B., et al. 2020, Proceedings of the National Academy of Sciences, 117, 25327

Shull, J. M. & Danforth, C. W. 2019, The Astrophysical Journal, 882, 180

Smit, J. 1999, Annual Review of Earth and Planetary Sciences, 27, 75

Sosa, A. & Fernández, J. A. 2011, Monthly Notices of the Royal Astronomical Society, 416, 767

Steel, D. 1993, Monthly Notices of the Royal Astronomical Society, 264, 813

Steel, D. 1998, Planetary and Space Science, 46, 473

Stuart, J. S. & Binzel, R. P. 2004, Icarus, 170, 295

Svetsov, V. & Shuvalov, V. 2019, Meteoritics & Planetary Science, 54, 126

Tabor, C. R., Bardeen, C. G., Otto-Bliesner, B. L., Garcia, R. R., & Toon, O. B. 2020, Geophysical Research Letters, 47, e60121

The 'Oumuamua ISSI Team, Bannister, M. T., Bhandare, A., et al. 2019, Nature Astronomy, 3, 594

Thomas, G. F., Annau, N., McConnachie, A., et al. 2019, The Astrophysical Journal, 886, 10, iOP Publishing

Tian, H.-J., Gupta, P., Sesar, B., et al. 2017, The Astrophysical Journal Supplement Series, 232, 4

Toon, O. B., Bardeen, C., & Garcia, R. 2016, Atmospheric Chemistry and Physics, 16, 13185

Toon, O. B., Zahnle, K., Morrison, D., Turco, R. P., & Covey, C. 1997, Reviews of Geophysics, 35, 41

Tyrrell, T., Merico, A., & Armstrong McKay, D. I. 2015, Proceedings of the National Academy of Sciences, 112, 6556

Vellekoop, J., Sluijs, A., Smit, J., et al. 2014, Proceedings of the National Academy of Sciences, 111, 7537

Vellekoop, J., Woelders, L., van Helmond, N. A., et al. 2018, Geology, 46, 683

Vokrouhlický, D., Nesvorný, D., & Dones, L. 2019, The Astronomical Journal, 157, 181

Wesson, P. S. 2010, Space Science Reviews, 156, 239

Wickramasinghe, J. & Napier, W. 2008, Monthly Notices of the Royal Astronomical Society, 387, 153

Zucker, C., Schlafly, E. F., Speagle, J. S., et al. 2018, The Astrophysical Journal, 869, 83

Appendix A: Simple Machine Learning Method for Estimating Stellar Properties from Photometry

The Euclid mission's array of photometric filters provides a unique opportunity to estimate the radial distance of stars. By analyzing the relative luminosity across these filters, a crude emission spectrum for each star can be constructed. This spectrum allows for the estimation of the star's spectral type, which in turn is linked to a typical absolute luminosity. Consequently, the observed luminosity can be used to infer the distance from the star to the observer.

In this study, I developed a simple machine learning algorithm to demonstrate how astrometric quantities can be inferred from photometry without explicitly projecting emission spectra. Utilizing the *Galaxia* simulation, I created a 9-dimensional space with each dimension representing one of Euclid's photometric filters. A test data set of 500 random stars was extracted from the simulation, while the remaining stars formed the training set within this space.

The estimation of the desired quantities for the test stars was conducted using the k-nearest neighbor algorithm. For each test star, the algorithm identified the k nearest training stars in the photometric space. The test star was then assigned the average value of the required quantity from these k nearest neighbors, weighted by their proximity to the test star. Essentially, this method assigns a quantity value to each test star based on the properties of training stars with similar photometric profiles.

This approach proved effective for estimating stellar properties closely related to photometry, such as effective temperature and stellar mass. Notably, it also provided a reliable prediction for radial distance, where the translation from observed photometry to emission spectra and distance occurs implicitly (as illustrated in figure 26). However, attempts to estimate perihelion distance based solely on photometric data were less successful, which is logical given the sensitivity of perihelion distance to proper motion, a factor that does not significantly influence a star's photometry. Interestingly, when using a large k value, the algorithm did not tend to underestimate perihelion distance. This observation may be attributed to the quadratic increase in the number of stars with perihelion distance.
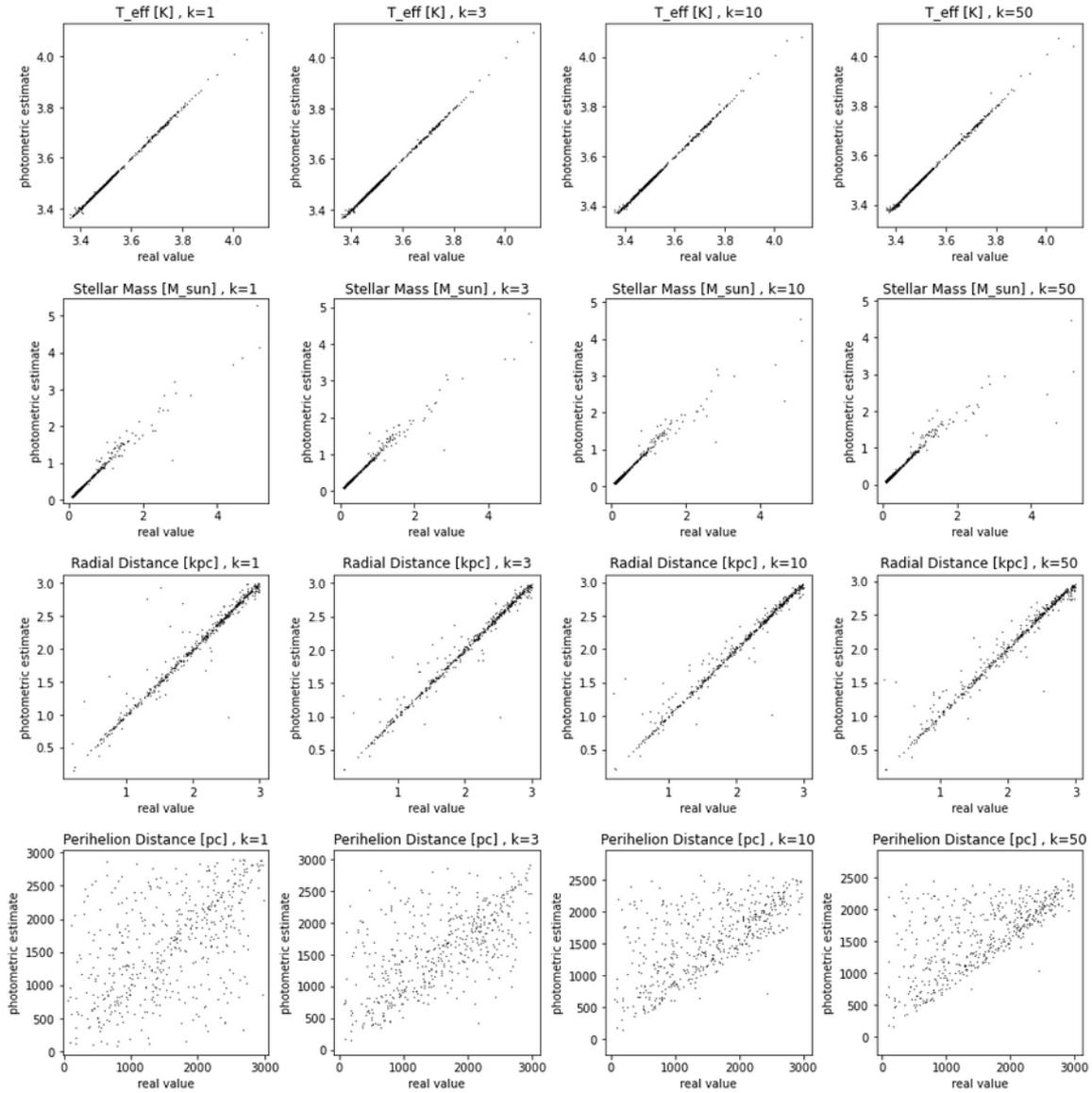
Figure 26: Results of the k-nearest neighbors machine learning algorithm applied to the test set of 500 stars. The top three rows demonstrate the alignment of estimated values (vertical axis) with their true values (horizontal axis) from the *Galaxia* simulation for effective temperature, stellar mass, and radial distance. The bottom row highlights the less accurate estimation for perihelion distance, with an interesting absence of underestimation in cases with large *k* values.