The systematic detection of AGN flares in the era of time domain astronomy

Summer A. J. McLaughlin

School of Mathematical and Physical Sciences
The University of Sheffield

A dissertation submitted in candidature for the degree of Doctor of Philosophy at the University of Sheffield

31st March 2025

“We are all in the gutter, but some of us are looking at the stars.”

— Oscar Wilde

Declaration

I declare that, unless otherwise stated, the work presented in this thesis is my own.

No part of this thesis has been accepted or is currently being submitted for any other qualification at the University of Sheffield or elsewhere.

The following chapters are based wholly or in part on previous publications:

- Chapter 3 (Using Gaussian Processes to detect AGN flares): McLaughlin et al. (2024)

Acknowledgements

It has genuinely been the privilege of my life to be able to do this PhD at Sheffield.

First and foremost, I give my utmost thanks to my supervisor Dr James Mullaney. Your unwavering support and faith in me has kept me going strong. You have dealt with a student with a serious case of impostor syndrome which can't have been easy and I'm sure it was more than you thought you had signed up for when I landed on your doorstep. You have been a mentor, a teacher, a friend, a listening ear, and had my back since day one. You have taught me how to be a scientist, a statistician, a scientific writer, and not to take things too seriously (though this last bit I'm still working on). It has genuinely been an honour to work with you.

Secondly, a big thank you goes out to my two examiners, Professor Clive Tadhunter and Professor Brooke Simmons, for an enjoyable *viva voce*. Your comments and encouragement have improved the final version of this thesis.

Thank you to Professor Stuart Littlefair for your phenomenal expertise in statistics and Gaussian Processes. Professor Clive Tadhunter, thank you for your mentorship, unparalleled wisdom and for letting me crash your observing run at the INT in La Palma. It was one of the best experiences of my life so far.

I extend my thank you to all of the staff who make up the Sheffield Astronomy Group. What an amazing group of people I have had the privilege of working with. Special thanks go out to Paul Kerry for his encyclopaedic knowledge of computers and for being an all round ray of sunshine, and Professor Paul Crowther for being my cheerleader. I

must also thank Professor Jenny Clark who has been my mentor throughout — I think part-time therapist should be in your job description!

I would not be the person and scientist I am today without my fellow students and postdocs. I am so lucky to have worked with you all and to be able to call you great friends. The Nerf guns, the oversharing, the pub trips where we put the world to rights, I have never met a group of people so welcoming and accepting in my life. Thank you for letting me be unapologetically myself. Rebecca and Alex, my astronomy Mum and Dad, what on Earth would I have done without you? Martin and Mark, two of the cleverest people on the planet, you have taught me so much. Emma, Umar, George, Luke, Adam, Thaer, Dustin, James, Eleanor, Megan, Dan, Freya, Amalie, Heloise and Joe, you are stars. I have no doubt that each of you will succeed with flying colours in your future endeavours, whatever they may be.

Mum and Dad, you nurtured my curiosity from a young age. Growing up, whichever career I had suddenly declared that I wanted to do that particular week: artist, doctor, lawyer, rocket scientist, I was always met with unconditional support. You raised me to know that the world was my oyster and that I could achieve whatever I put my mind to. That gift is priceless, and probably the reason I am here today. You have believed in me always, even when I haven't believed in myself. Thank you.

Thank you to my wonderful boyfriend Tyler who has been by my side throughout all of this. You have dealt with the brunt of my stress: the late night tears, the cancelled plans, the panic stations. You too, are someone who has never doubted me for a second. I can't thank you enough. My Universe changed when I met you.

At the risk of sounding narcissistic, there are two important people still left to thank. To 10 year old Summer, thank you for your curiosity. Your spark has carried me all this way. To 20 year old Summer, thank you for your grit. You never gave up, even when you told yourself you weren't good enough. I can only hope you are both proud of me.

Summary

A characteristic feature of AGN is their variable luminosity, which is observed at all wavelengths. Their lightcurves typically vary by a few tenths of a magnitude or more over periods lasting from weeks to years. This variability is caused by the turbulent processes occurring within the accretion disk, meaning that the study of AGN variability is an effective means of studying the accretion process. Recently, a growing number of AGN have been observed to show extreme variability, whereby their lightcurves exhibit luminosity changes that represent a significant departure from their baseline variability. These rare events are known as AGN flares. Since their discovery, AGN flares have quickly become an active area of research with the intention of better understanding accretion physics. Looking ahead to high-cadence surveys such as the Legacy Survey of Space and Time (LSST), which promises millions of transient detections per night in the coming decade, there is a need for the fast and efficient detection and classification of AGN flares. The problem with the systematic detection of AGN flares, however, is the requirement to detect them against a stochastically variable baseline; the ability to define a signal as a significant departure from the ever-present variability is a considerable statistical challenge. In this thesis, I investigate the use of a statistical tool called Gaussian Processes (GPs) to systematically detect AGN flares in optical lightcurves, and I demonstrate that GPs are a viable means of transient detection in the coming era of time domain astronomy.

Contents

Figures

Tables

Acronyms and Abbreviations

AGN	Active Galactic Nucleus
ALPAKA	AGN Line Profile And Kinematics Archive
ANTs	Ambiguous Nuclear Transients
CRTS	Catalina Real-Time Transient Survey
GP	Gaussian Process
LSST	Legacy Survey of Space and Time
MCMC	Markov Chain Monte Carlo
MHD	Magnetohydrodynamical
MILLIQUAS	Million Quasars Catalog
NGC	New General Catalogue
PSF	Point Spread Function
SDSS	Sloan Digital Sky Survey
SED	Spectral Energy Distribution
SMBH	Supermassive Black Hole
TDE	Tidal Disruption Event
ZTF	Zwicky Transient Facility

Chapter 1

Introduction

1.1 Active Galactic Nuclei

1.1.1 Background

The term Active Galactic Nuclei (AGN) refers to the accreting supermassive black holes at the centres of galaxies. AGN are some of the most powerful sources of energy in the Universe, often outshining their entire host galaxy (Peterson & Burbidge, 1998). It was Fath in 1909 who made the first discovery of AGN when he characterised the spectra of what was then termed "spiral nebulae" to determine if these objects were collections of unresolved stars or if they showed similarities to nearby gaseous nebulae (Fath, 1909). In that work, he had unknowingly made the first detection of an AGN in the galaxy NGC 1068, observing that its spectrum was composite, showing both bright emission and absorption lines (Fath, 1909). Despite this discovery, and despite the advent of systematic studies of galaxies such as that published by Seyfert in 1943, the study of AGN as we know now was not pioneered until some of the first radio surveys of the sky in the 1950s (*e.g.,* Ryle et al., 1950; Hanbury Brown et al., 1952; Baade & Minkowski, 1954). By that time, the angular resolution of radio observations was sufficient to identify the strongest

radio sources in the sky (and associate them with optical sources: Peterson & Burbidge, 1998). This led to the discovery of quasars by Schmidt in 1963 (extremely luminous AGN: Schmidt, 1963), which were initially believed to be a separate phenomenon from the Seyfert-type objects discovered previously and were then known as "quasi-stellar objects" (Greenstein & Schmidt, 1964). Later, however, it was realised that both quasars and Seyfert galaxies are sub-types of AGN that differ in the amount of radiation emitted by the central source (with quasars being typically 100 times more luminous: Kristian, 1973; Shields, 1999). The overarching question during these early AGN detections was the nature of the central object emitting this massive amount of energy; explaining this extreme energy release from such a compact region was not straightforward (*e.g.,* Salpeter, 1964). As a result, it was suggested that the central object must be a massive black hole (since stellar fusion would not produce enough energy: see Salpeter, 1964). In this case, the spiral infall of material forms an accretion disk and the energy release comes from angular momentum transfer outwards from the centre of the disk (Zel'dovich & Novikov, 1964; Salpeter, 1964; Lynden-Bell, 1978). This is now the widely accepted AGN paradigm (Rees, 1984; Antonucci, 1993; Urry & Padovani, 1995; Peterson & Burbidge, 1998). Today, AGN are a focus of observational effort in every frequency band across the entire electromagnetic spectrum from radio to gamma rays (Shields, 1999; Padovani et al., 2017). The research field is more active than ever before, spanning a wide range of areas including accretion disk physics (*e.g.,* King, 2008; Fausnaugh et al., 2017; Guo et al., 2022), the relationship between the black hole and its host galaxy (*e.g.,* Peterson, 2008; Bessiere et al., 2024), jets and outflows (*e.g.,* Di Matteo et al., 2005; Blandford et al., 2019) and their impact on galaxy evolution (*e.g.,* Fabian, 2012; Kormendy & Ho, 2013; Harrison, 2017), and the evolution and growth of supermassive black holes (*e.g.,* Alexander & Hickox, 2012; Ricci et al., 2025).

AGN have been observed to exhibit the following properties (Padovani et al., 2017; Peterson & Burbidge, 1998):

1. **High luminosities**, with bolometric (*i.e.,* integrated over all wavelengths) luminosities up to $\approx 10^{48}\,\mathrm{erg\,s^{-1}}$ (Baade & Minkowski, 1954; Padovani et al., 2017).

2. **Broad emission lines**, where broad widths of emission lines (*e.g.,* H_α, H_β, Ly_α) are seen in their optical and UV spectra, ranging from 2000 to 10 000 $\mathrm{kms^{-1}}$ (Seyfert, 1943; Ho & Kormendy, 2000; Peterson, 2006).

3. **Time-variable continuum flux**, where AGN lightcurves exhibit stochastic variability in all wavelength bands (Sandage, 1964; Ulrich et al., 1997; MacLeod et al., 2010; Kovacevic et al., 2025).

For the purpose of this introduction, I will be focussing on the third item in the above list: the time-variable continuum flux of AGN, which is the cornerstone of the work undertaken in the following thesis chapters. The study of AGN variability provides insight into the processes that power supermassive black holes (*e.g.,* Peterson, 2001) and this is the main motivation behind my research. Variability was one of the first properties of AGN identified via optical photometry (e.g Matthews & Sandage, 1963; Smith & Hoffleit, 1963). It is thought that this variability is a result of the turbulent and unstable nature of accretion (Ulrich et al., 1997). This is the reason why the study of AGN variability is now regarded as an indirect means to investigate the accretion process when the accretion disk cannot be directly spatially resolved (Peterson, 2001). I continue this discussion of AGN variability in more detail in §1.1.3.

In the following sections, I will describe the internal structure of AGN (§1.1.2) and provide deeper insight into AGN variability (§1.1.3). Then in §1.2 I will describe and explain the phenomena termed AGN flares (extreme variability of AGN). In §1.3 I introduce and describe a key statistical technique used throughout the thesis to systematically detect AGN flares (Gaussian Processes). Then finally I describe how the work undertaken in this thesis is aptly timed for the Legacy Survey of Space and Time (LSST) in §1.4 and I outline the key open questions and intentions for the thesis in §1.5.

1.1.2 Structure of an AGN

In this section, I outline the structure of an AGN. Understanding AGN structure is key to understanding their observational characteristics, including their variability. The various structural components each contribute to the observed properties of AGN including their continuum variability (Osterbrock, 1993; Netzer, 2015; Padovani et al., 2017). AGN can now be classified into many different types with varying characteristics, but the Unified Model (see *e.g.,* Antonucci, 1993; Netzer, 2015) can constrain the general properties of the wide range of AGN types. In this model (see Fig. 1.1), AGN are characterised by a central supermassive black hole surrounded by a gaseous accretion disk, an optically-thick dusty torus, broad and narrow line regions, and (in approximately ten percent of them) the presence of radio jets. The Unified Model gives rise to the two main AGN classifications known as Type 1 and Type 2, where Type 1 exhibit both narrow (forbidden) and broad (permitted) emission lines, and Type 2 only exhibit the former due to obscuration from the dusty torus (Netzer, 2015; Liu et al., 2018; MacLeod et al., 2019). It is widely accepted that these two subtypes differ because of the angle at which we observe them: in the case of Type 1 AGN the central engine is observed at low inclination angles (*i.e.,* almost face on) whereas Type 2 AGN are observed at high inclination angles – though this paradigm has come into question in recent years in the advent of what have become known as changing-look AGN, where an AGN is observed to seemingly "change type" from Type 1 to Type 2 or vice versa as their spectra exhibit the rapid disappearance or appearance of broad lines (*e.g.,* LaMassa et al., 2015; Gezari et al., 2017; Yang et al., 2023, or see review by Ricci & Trakhtenbrot 2023).

Below is a breakdown of the properties of the various components that make up an AGN, based on the Unified Model (Antonucci, 1993), listed in order of moving outwards from the central black hole:

- *The central supermassive black hole.* The "central engine" upon which accretion of

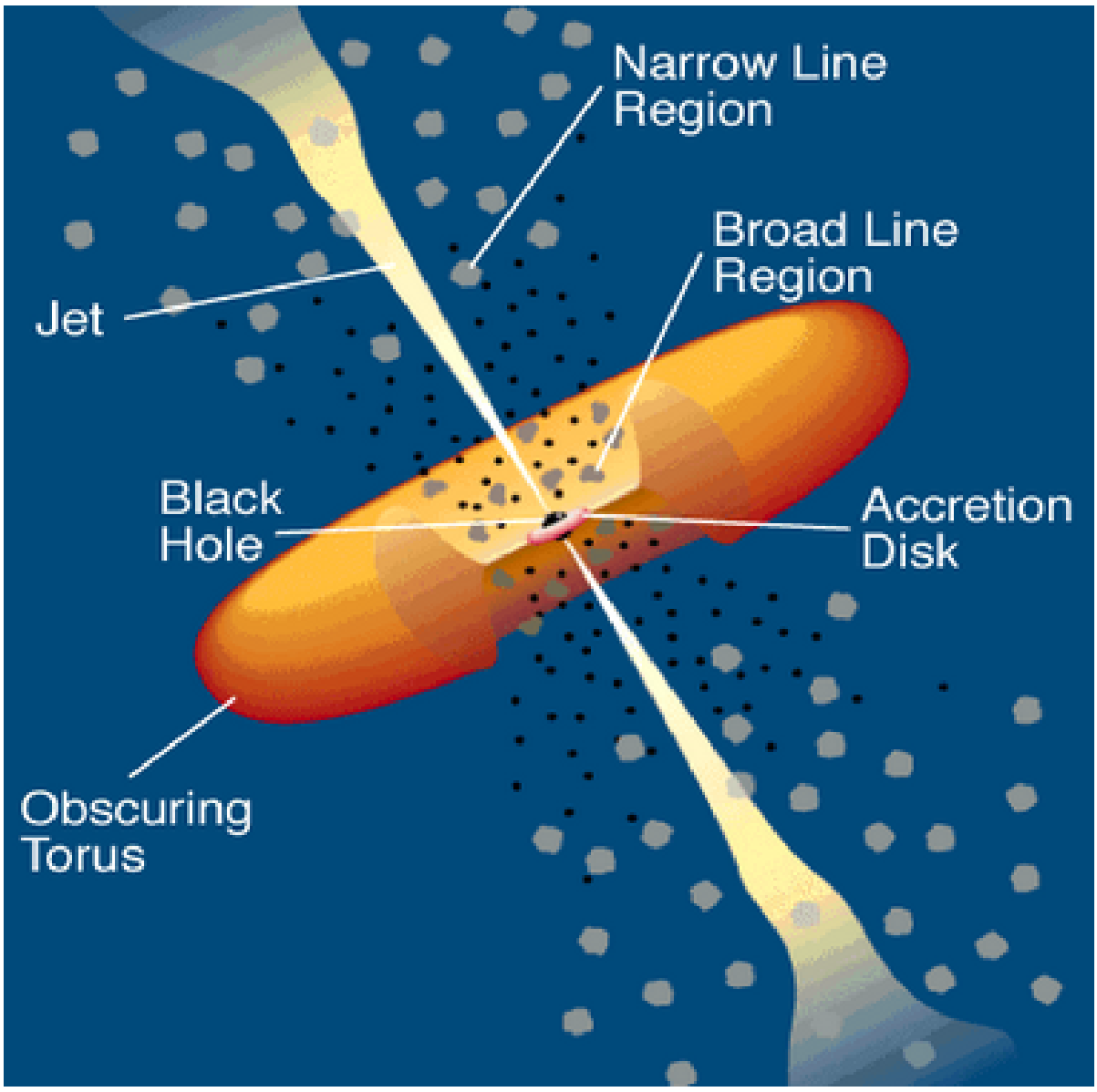

Figure 1.1: Schematic diagram of the current paradigm for AGN from Urry & Padovani (1995). The central black hole, dusty torus, accretion disk, broad and narrow line regions and jets are labelled. Type 1 AGN are observed at an approximate angle of 45° from the axis of the dusty torus, and Type 2 AGN are observed along the line of sight of the torus.

material takes place, with mass typically $> 10^5\,M_\odot$ (Netzer, 2015).

- *The accretion disk.* The gravitational infall of material into the black hole results in the generation of significant energy in the form of a hot accretion disk (Peterson & Burbidge, 1998). The accretion disk is a turbulent structure, the physics of which is not well understood and is likely governed by magnetohydrodynamics (MHD: *e.g.,* Kawaguchi et al., 2000; Begelman & Silk, 2017; Dexter & Begelman, 2019; Hossein Nouri, Fatemeh & Janiuk, Agnieszka, 2024). The general principles of the accretion disk have been understood since the 1970s, whereby the infall of material into the black hole results in the formation of an accretion disk, and there is an outward transfer of angular momentum, which facilitates this accretion (Shakura & Sunyaev, 1973; Peterson & Burbidge, 1998). This happens because the disk is differentially rotating and so different layers move over each other causing friction, which allows for this loss of angular momentum and infall into the black hole (Shakura & Sunyaev, 1973; Lawrence, 2018). This popular accretion model is commonly known as an "optically thick, geometrically thin disk model", "viscous thin disk model" or the Shakura-Sunyaev model named after their seminal paper in 1973. This notion of "viscosity" is needed for outwards angular momentum transfer, and though it is not well understood, it is agreed that this viscosity is likely a result of turbulence or magnetic fields (Lawrence, 2018). Later, in the 1990s, the idea of X-ray reprocessing was introduced to justify short-timescale variability in the UV and optical – since the viscous model would not allow for such short-term variability (McHardy et al., 2016). Now, it is generally accepted that there is an X-ray corona as part of the accretion disk, which shines on it and heats it up, resulting in short-timescale variability in UV and optical wavelengths. The accretion disk is estimated to be only a small fraction of a parsec across (Hawkins, 2007), of the order of light-days.

- *The dusty torus.* Aforementioned, a small (0.1 - 10 parsecs across: Netzer 2015), clumpy toroidal mass of dust and gas that obscures the nucleus from some lines of sight while collimating the radiation that escapes (Mason, 2015). The torus is responsible for the strong infrared emission observed from many AGN (Urry & Padovani, 1995; Antonucci, 1993).

- *The broad line region (BLR).* An emitting region of high-density gas clouds close to the accretion disk (between 0.01 and 1 parsec from the black hole), thought to be optically thick and photoionised (Peterson, 2006; Netzer, 2015).

- *The narrow line region.* The furthest emitting region of gas from the central source (Peterson & Burbidge, 1998). The gas is of a lower density than the BLR, but still ionised, extending from the torus to hundreds or thousands of parsecs away (Netzer, 2015).

Each of the above components is responsible for key observational properties of AGN, not least the form of observed spectral energy distribution (SED) itself (Padovani et al., 2017). A useful schematic of the SED of a typical AGN is shown in Fig. 1.2, which demonstrates the effect of the different contributions of some of the components described above. At mid-infrared wavelengths the AGN spectrum is dominated by the emission from the dusty torus. As wavelength is decreased to optical and UV, the accretion disk emission dominates as a result of reprocessing of X rays from the corona. The X-ray corona itself contributes from wavelengths ranging from optical to X-ray.

1.1.3 AGN variability

A trademark feature of AGN is the variability in their lightcurves across all wavelengths (Peterson & Burbidge, 1998). Despite the fact that the exact physical mechanisms for this variability are unknown, their varying luminosities can be well-modelled (Kozłowski,

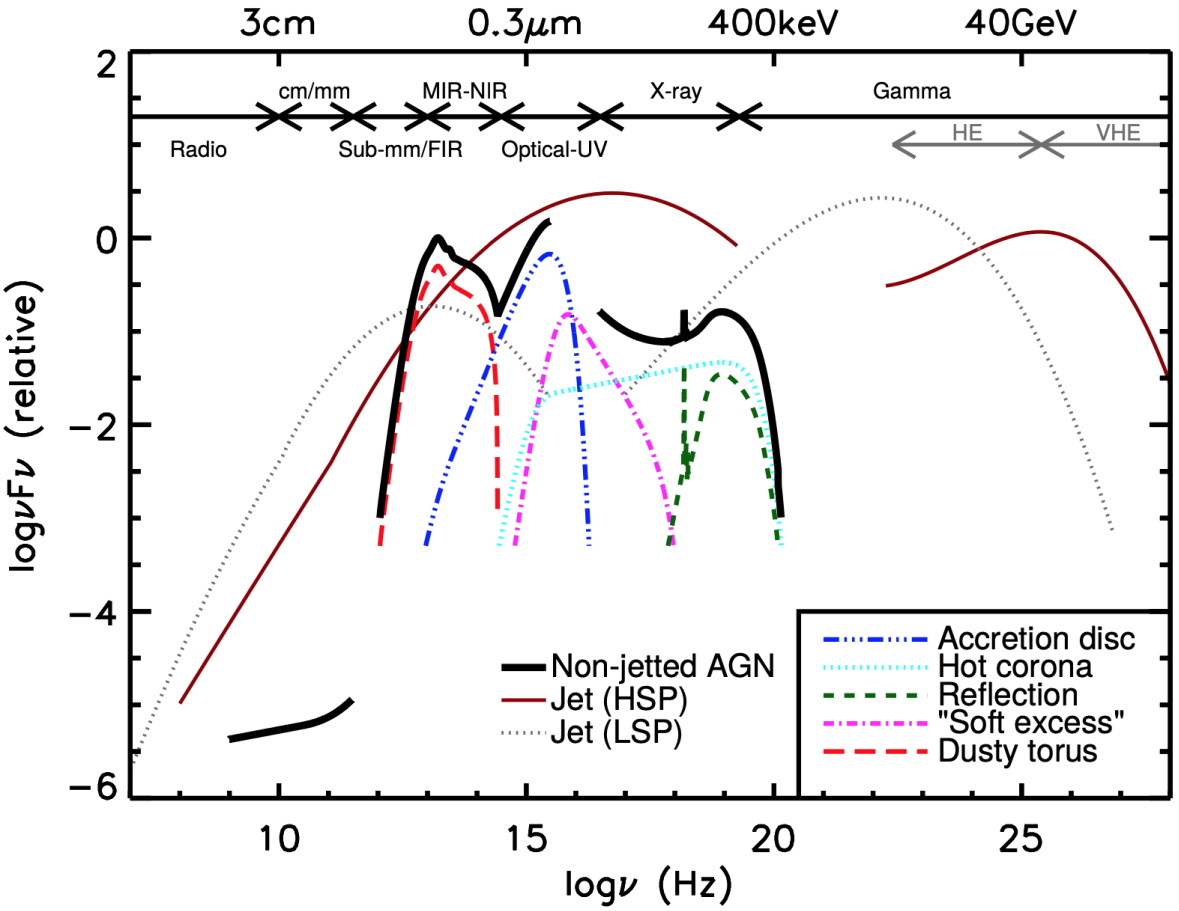

Figure 1.2: Figure 1 from Padovani et al. (2017): A schematic representation of an AGN spectral energy distribution (SED). The black curve is the total emission and the coloured curves show the contribution in overall emission from the individual components of the AGN. The jet SED is also shown for a high synchrotron peaked blazar (HSP) and a low synchrotron peaked blazar (LSP). Image credit: C. M. Harrison (Harrison, 2016) adapted by Padovani et al. (2017).

2017). AGN exhibit a stochastic variability in their luminosities, which is well-described statistically by a damped random walk (*e.g.,* MacLeod et al., 2010). This observed optical variability is typically a few tenths of a magnitude in amplitude with a characteristic timescale (the average time period over which variability occurs) of several months, though this is correlated with black hole mass and anti-correlated with luminosity and Eddington ratio (*i.e.,* how close the AGN's accretion rate is to its maximum value whilst radiation pressure is still balanced; Graham et al., 2017; Sesar et al., 2007). A key agreement in the literature is that more luminous AGN vary with smaller amplitudes (Hawkins, 2007), and the variations in brightness have amplitude of the order of 20 percent of their baseline luminosity on timescales of months to years (MacLeod et al., 2010). The random variability of a typical AGN lightcurve in comparison to the constant host galaxy flux is emphasised in Fig. 1.3.

The most common functional form for the analysis of AGN variability is known as the structure function. The structure function, $S(\Delta t)$ is usually defined as:

$$S(\Delta t) = \sqrt{\frac{1}{N(\Delta t)} \sum_{i<j} [m(t_j) - m(t_i)]^2}. \tag{1.1}$$

where Δt is equal to $t_i - t_j$, $m(t_i)$ is the magnitude measure at a given time epoch t_i (where t_j is a subsequent time epoch), and the sum runs over the total number of N epochs/observations (Hawkins, 2007). Kozlowski (2016) approximate the distribution of characteristic timescales at a median of $\tau = 0.98$ years and the amplitudes of variations $\sigma_\tau = 0.41$ mag (Kozlowski, 2016). As mentioned above, this timescale is correlated with the black hole mass (a higher black hole mass results in an increased variability timescale) and it is anti-correlated with luminosity (Kozlowski, 2016).

Despite their prevalence in AGN physics, the relationship between AGN type and variability timescale, or the physical causes of their stochastic variability, are not well understood (Creque-Sarbinowski et al., 2021). Although the exact mechanism is still

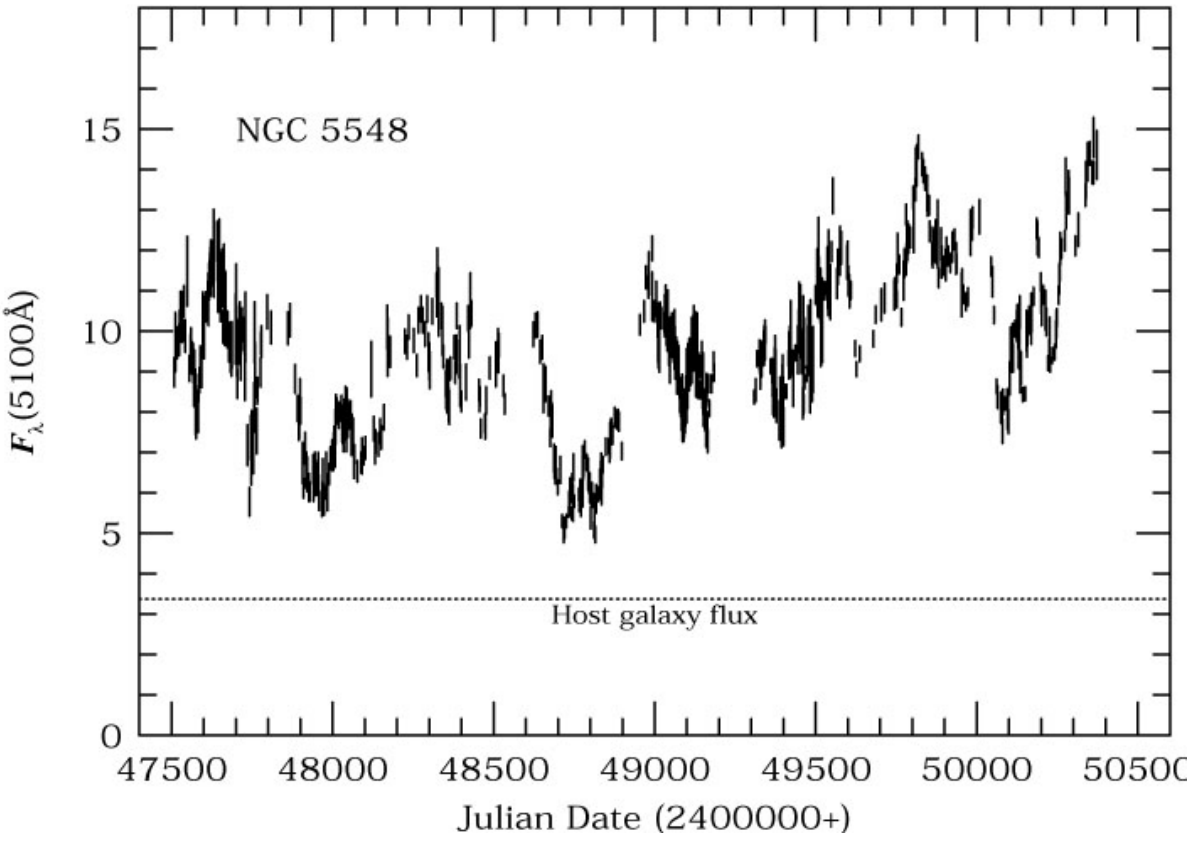

Figure 1.3: Lightcurve of NGC 5548 showing the stochastic variability of the AGN compared to the constant host galaxy flux. Figure from Peterson et al. (1999).

unknown, it is thought that continuum variations are caused by magnetohydrodynamic instabilities in the accretion disk (Peterson et al., 1999; Kawaguchi et al., 2000). Furthermore, a number of other models have been suggested to describe AGN variability, including supernova bursts, microlensing, and accretion disk instabilities (*i.e.,* an extreme example of typical AGN variability, MacLeod et al. 2010; Hawkins 2007). It has been proposed that the variability timescale of AGN is related to the thermal timescale of accretion disks (Kelly et al., 2009; Collier & Peterson, 2001), which would explain why AGN become bluer as they brighten (MacLeod et al., 2010).

Understanding and characterizing the origins of AGN variability is paramount to further constraining the physics of black hole accretion, especially at spatial scales that are beyond the resolving power of most current and future AGN observing facilities (Sartori et al., 2019). Furthermore, detection of AGN using their variability has proved highly effective – especially in lower mass, lower luminosity cases, which has helped provide a more representative AGN catalogue (MacLeod et al., 2010). This is particularly important to note when considering the fact that upcoming time-domain surveys (such as the Legacy Survey of Space and Time: Ivezić et al., 2019) will provide unprecedented high-cadence observations of AGN and related transients (see §1.4), which will facilitate the detection of highly variable AGN in real-time, across shorter timescales than ever before.

1.2 AGN flares

In addition to showing stochastic variability, there is growing evidence that AGN can exhibit extreme variability that differs significantly from the variable baseline (*e.g.,* Meusinger et al., 2010; Lawrence et al., 2016; Graham et al., 2017; Wiseman et al., 2025). These events are known as AGN flares. Current flare detections indicate that they occur over timescales of several hundreds of days (*e.g.,* Graham et al., 2017), but their rarity

brings up questions about how representative our existing samples are. It is thought that AGN flares are a separate phenomenon from AGN variability, though the exact cause is unknown (Lawrence et al., 2016; Zabludoff et al., 2021), with different studies suggesting that they could be caused by extreme instabilities in the accretion disk (*e.g.,* Hawley & Krolik, 2001), microlensing (*e.g.,* Lawrence et al., 2016; Bruce et al., 2017), tidal disruption events (TDEs: Chan et al. 2019; Frederick et al. 2021), superluminous supernovae (*e.g.,* Drake et al., 2011), mergers of stellar mass black holes (*e.g.,* Graham et al., 2017), or changes in accretion state (*e.g.,* Lawrence et al., 2016; MacLeod et al., 2019). Furthermore, (MHD) simulations suggest that AGN flares may be caused by energy dissipation following magnetic reconnection in the accretion disk (Nathanail et al., 2020). In the past decade, there have been attempts to organise the various AGN-related transient events into separate classes although it is acknowledged that, due to a lack of understanding of the underlying physics, there is overlap between these classes. Various types of AGN-related transients recently observed include Ambiguous Nuclear Transients (ANTs: *e.g.,* Graham et al., 2017; Wiseman et al., 2025; Kankare et al., 2017), Bowen fluorescence flares (*e.g.,* Trakhtenbrot et al., 2019; Frederick et al., 2021) and Changing-look AGN (*e.g.,* Gezari et al., 2017; Yang et al., 2023; Ricci & Trakhtenbrot, 2023).

An example of what AGN flares look like in optical lightcurves is shown in Fig. 1.4. This is a figure taken from Frederick et al. (2021), showing the optical r and g band lightcurves of five transients associated with Type 1 AGN that were originally classified as supernovae by the Zwicky Transient Facility (ZTF: Bellm et al., 2019). The underlying stochastic variability characteristic of a "normal" AGN is present alongside a smooth, large-amplitude, and rapidly rising flare.

The detection of AGN flaring events provides a window into studying the accretion physics of the disk. The timescales and magnitude changes of AGN flares put constraints on MHD simulations of the accretion disk and act as important probes for AGN accretion

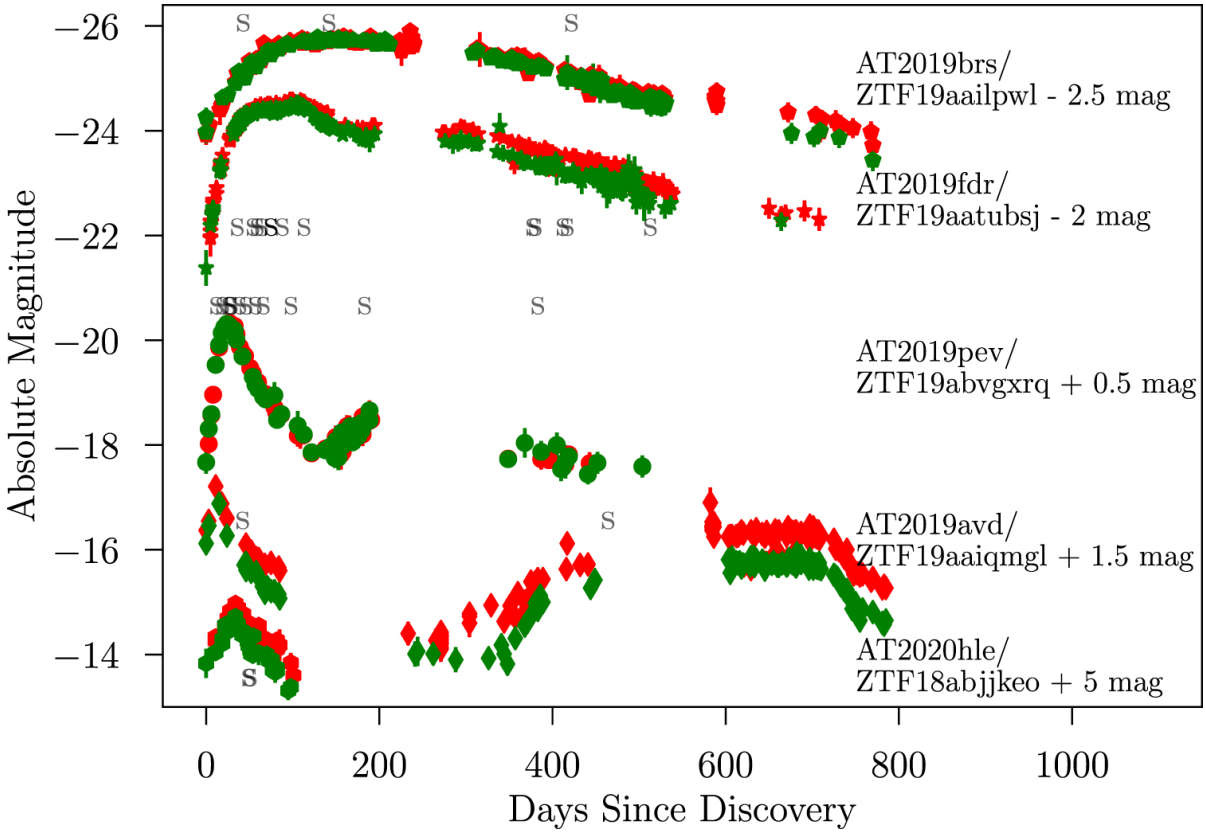

Figure 1.4: Figure 1 from Frederick et al. (2021): Comparison of the g and r band difference imaging lightcurve shapes and absolute magnitudes of flare candidates identified from ZTF. Spectroscopic epochs are labelled for each lightcurve with an "S".

rates over short timescales (Lodato & Rossi, 2010; Blagorodnova et al., 2016; Graham et al., 2017). In addition, if follow-up spectra are acquired, reverberation mapping can enable a calculation of the size of the flaring region within the disk (Payne et al., 2022; Zhang et al., 2013; Somalwar et al., 2022). Moreover, it is important to note that one of the reasons why the field of AGN flares has exploded in recent years is due to the fact that the extreme brightness changes in the optical across such short timescales (observations that have been enabled by high-cadence surveys such as ZTF; see Bellm et al. 2019) pose problems for current accretion disk models (§1.1.2). The viscous accretion disk model in which the accretion disk reprocesses X-rays is not able to explain such extreme short-timescale changes above the underlying variability, which has resulted in significant interest in the underlying physics behind these phenomena (Wiseman et al., 2025).

Given their importance in probing the accretion physics of AGN, it is somewhat frustrating that detecting and identifying AGN flares has proven to be such a challenge. In the past, they have been difficult to detect and characterise against an intrinsically variable AGN lightcurve (Zabludoff et al., 2021), but improvements in imaging and machine-learning techniques are now starting to enable their detection in significant numbers (Mattila et al., 2019). Distinguishing a valid flare detection from the background variability presents a statistical and observational challenge (Zabludoff et al., 2021; Gezari, 2021). Even once a detection is identified, then without extensive follow-up observations, it is difficult to categorise these events with certainty (see review by Zabludoff et al., 2021 regarding how to distinguish between types of transient). The ability to identify and distinguish between these events in real-time will be crucial to further understand not only the origins of these phenomena, but also better constrain the physics of supermassive black holes and their accretion disks, and the relationship between them.

In the coming years, it is hoped that future high-cadence surveys will facilitate the detection and monitoring of flaring AGN in real-time (Creque-Sarbinowski et al., 2021).

The Zwicky Transient Facility (ZTF; Bellm et al. 2019) and the The Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST: see §1.4 and Ivezić et al., 2019) are examples of facilities that will undertake surveys to help mobilise this area of research in the coming decade with regular, high-cadence time-domain observations (Graham et al., 2019). These facilities will not only expand the current catalogue of AGN but will also provide insights into key unanswered questions in this field, facilitating a better understanding of AGN variability (Creque-Sarbinowski et al., 2021), the rates of different nuclear transient events, and what distinguishes them from each other. Real-time processing of nuclear transients detected by such surveys is critical to identifying shorter-lived, rare events and allocating follow-up resources efficiently (Soraisam et al., 2020).

1.3 Detecting AGN flares in lightcurves

Since AGN variability is stochastic, it is not straightforward to detect a transient signal amongst it. For this reason, there is a requirement to develop a robust means of detecting AGN flares in already variable lightcurves, one that perhaps accounts for the stochastic variability in a statistical manner. As the reader will see in Chapter 2, I attempt to detect AGN flares using a brute-force approach that replicated the technique used by Graham et al. (2017), however it quickly became apparent that this method is unsuitable in this context. This was due to an inability to distinguish between typical variability and extreme variability, which resulted in a significant number of spurious detections. A technique that showed promise in solving this issue (and its use is explored in detail in Chapter 3) is called a Gaussian Process (GP), which is sometimes considered a form of machine learning and often used in the context of regression (Seeger, 2004). In the following section, I describe the theory behind GPs and their uses within astronomy.

1.3.1 Gaussian Processes

A GP is often defined as a prior over functions, which generates a probability distribution over all possible functions that fit a data-set (Rasmussen & Williams, 2006). In simpler terms, a GP takes a user-defined covariance function (known as a kernel, a function that quantifies the changing relationship between two variables) and, *given* the data, it generates the full set of functions that are both: (a) possible fits to the data, and (b) feasibly produced by the chosen covariance function. From this set of functions that describe the data, the GP uses the likelihood of each function to then calculate the mean function (*i.e.,* the most likely or best-fit function) along with the uncertainty about this mean (Aigrain et al., 2012). With the user's choice of covariance function, the GP asks the question: "bound by the behaviour of this chosen covariance function, what is the most likely function generated, given the available data?" Formally, a GP is a collection of random variables, any finite number of which have joint Gaussian distributions. The GP is fully specified by its mean function and covariance function, which is a generalization of the Gaussian distribution. Rasmussen & Williams (2006), which is the superlative text for a comprehensive description of GPs, use the following mathematical formality for defining a GP:

$$f(x) \sim GP(m(x), k(x, x')) \tag{1.2}$$

where $m(x)$ is the mean function, x and x' are two different observations. $k(x, x')$ is the covariance function of a real process $f(x)$, and hence the GP is fully described by its mean and covariance functions.

Gaussian Processes are an effective non-parametric, non-linear form of regression that is powerful at handling heteroskedastic (non-uniform) uncertainties. They are regularly used in the context of astronomy for a number of tasks such as regression, modelling and classification in a variety of contexts including quasi-periodic oscillations, transi-

ent classification, AGN variability and exoplanet transits (Aigrain & Foreman-Mackey, 2022). For example, GPs have been used to de-trend variable exoplanet lightcurves from transit surveys (Crossfield et al., 2016) and also to model quasi-periodic stellar activity (Aigrain et al., 2012; Angus et al., 2017; Nicholson & Aigrain, 2022). With regards to AGN variability, GPs are commonly used to model lightcurves with a damped random walk kernel (Kozłowski et al., 2010; MacLeod et al., 2010), and indeed one of the first uses of GPs in astronomy was by Press & Rybicki (1998) to model the variability of gravitationally-lensed quasar 0957+561.

GPs are a means of parameterising the covariance of a dataset, hence quantifying the similarity between data points. The covariance function in the context of GPs is called a kernel, and it encodes the assumptions (priors) about the underlying predictive function: for example, whether it is periodic or highly variable (Rasmussen & Williams, 2006). While the Gaussian Process optimises the coefficients of the kernel, it is important to choose a kernel with a functional form that is appropriate for the data in hand. In the context of GPs, the kernel function is referred to as the *prior* as in Bayesian statistics, as the kernel encodes the assumptions about the function that we wish to learn. Commonly used kernels or covariance functions, especially in astronomy, include (Rasmussen & Williams, 2006):

- *The squared exponential covariance function.* A stochastic variability kernel. The most widely used kernel in astronomy and beyond (Aigrain & Foreman-Mackey, 2022), defined as:

$$k(r) = \exp(-\frac{r^2}{2l^2}) \tag{1.3}$$

  where r is defined as the difference between x-values (x-x'), and l is the (temporal) length scale. In other words, increasing the parameter l increases the duration of variations in the function.

- *The Matérn Class of Covariance Functions* (Matérn, 1966). A more generalised

case of the squared exponential kernel, but is less smooth:

$$k(r) = \frac{2^{1-\nu}}{\Gamma^{\nu}}(\frac{\sqrt{2\nu}r}{l})^{\nu} K_{\nu}(\frac{\sqrt{2\nu}r}{l}) \tag{1.4}$$

where ν and l are chosen positive parameters and K_{ν} is a modified Bessel function.

- *The periodic covariance function* (Görtler et al., 2019). A periodic kernel that replicates repeating functions defined as:

$$k(r) = \sigma^2 \exp(-\frac{2}{l^2}\sin^2(\pi\frac{r}{p})) \tag{1.5}$$

where σ^2 is the amplitude and p is the period of the oscillations.

I recommend reading Görtler et al. (2019) for a visual exploration of these different kernels listed above and how they work in practise. In this work, I use the Matérn-3/2 kernel, which depends on the variability amplitude σ, the variabiity timescale ρ[1], and the difference between all pairs of points r (in this case time, *i.e.*, $r = |\mathbf{t} - \mathbf{t}'|$, where I use bold lettering to represent a vector containing all values of time, thus ensuring that r, and hence $k(r)$ is a square matrix). In this kernel, the covariance k is defined as (Rasmussen & Williams, 2006; Foreman-Mackey et al., 2017):

$$k(r) = \sigma^2 \left(1 + \frac{\sqrt{3}r}{\rho}\right) \exp\left(\frac{-\sqrt{3}r}{\rho}\right) \tag{1.6}$$

Fig. 1.5 shows three different function realisations that have been sampled from a GP with a Matérn 3/2 kernel, and Fig. 1.6 shows a GP fit to an AGN lightcurve with this

[1]As a cautionary note, it is important to realise that the variability amplitude and timescale, σ and ρ, do not necessarily directly correspond to the physical characteristics of the lightcurve. Whilst they are an indication of the scale of a lightcurve's overall variability, direct comparisons between the values of these hyperparameters and the quantitative behaviour of a lightcurve should be avoided. Throughout this thesis, the hyperparameter distributions are used solely as a summary statistic for which to compare different lightcurves, rather than to ascertain physical properties from the variability.

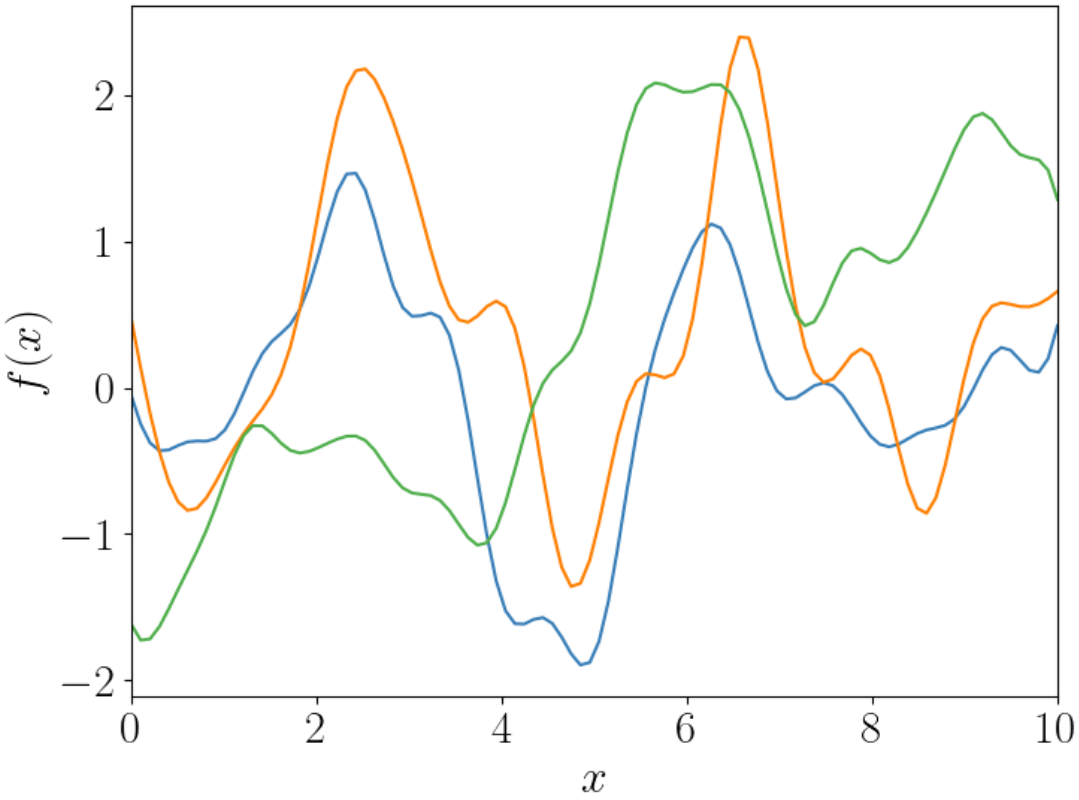

Figure 1.5: Three different function realisations that have been sampled from a Gaussian Process with a Matérn 3/2 kernel, using values of 1 for both σ and ρ in Eq. 3.1. Because Gaussian Processes are probabilistic in nature, the same kernel can produce different functions. Similarly, different functions (in this case lightcurves) can have the same, or very similar, kernel coefficients. Figure from McLaughlin et al. (2024), generated using Roelants (2019).

same kernel. Note that the reasoning behind this choice of kernel is discussed in Chapter 3.

GPs are widely used in the context of transient classification, but primarily used as an interpolation tool for priming sparse or noisy time series data for machine learning algorithms (*e.g.,* Villar et al., 2020). Within the field of astronomy research, GPs have typically been used as a pre-processing step in machine learning methods and, to my knowledge, have not been used for flare detection directly (Aigrain & Foreman-Mackey, 2022). I do, however, note that Graham et al. (2023) used GPs to *confirm* that suspected

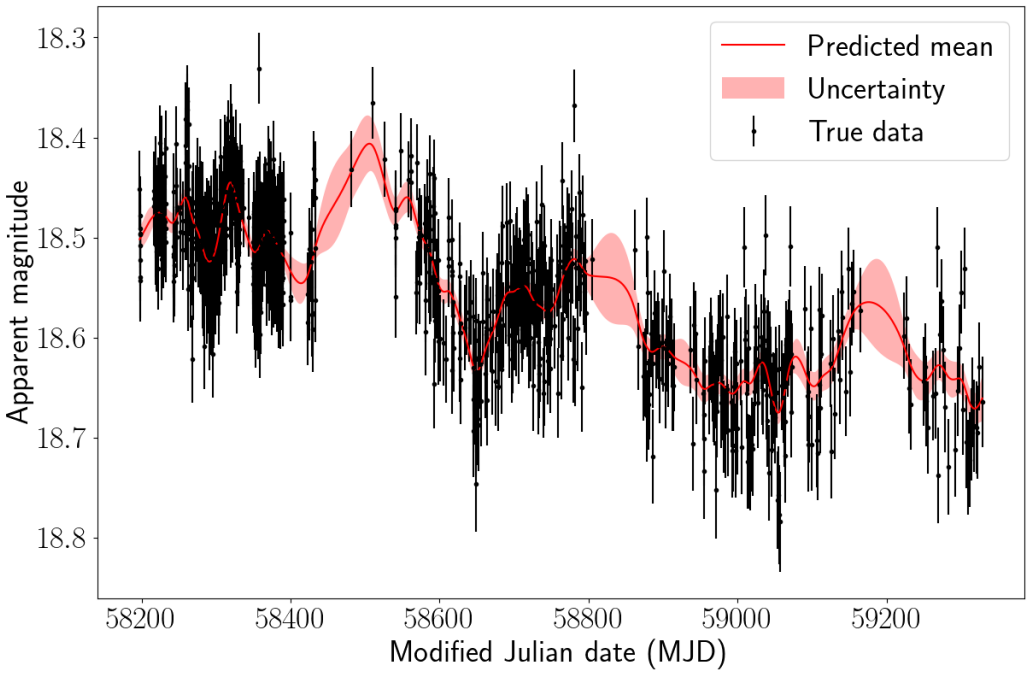

Figure 1.6: A Gaussian Process fit to a ZTF Type 1 AGN lightcurve. The red line shows the posterior mean of the Gaussian Process, given the observed data and kernel parameters. The red shaded region shows the 1-sigma uncertainty around this mean. Note how the uncertainties change depending on the density of data points in a certain region. The Gaussian Process effectively "learns" how variable the data is, which allows it to make reasonable predictions for regions of sparse data.

flares detected via other means do, indeed, represent significant departures from the underlying AGN variability. For flare classification directly, since the covariance function describes how all of the data points in a lightcurve are related to each other, it can be used as a summary statistic of the variability. This includes whether the lightcurve is periodic or a one-off outlier event. This therefore motivates an exploration into whether GPs can be used as a tool to classify transient astronomical events – and specifically AGN flares – directly. This forms the basis of Chapter 3.

1.4 Detecting AGN flares in the advent of large high-cadence surveys

In Section 1.2, I outlined that effective detection and characterisation of AGN flares is crucial in advent of upcoming high-cadence surveys where we will potentially be able to observe these fast-changing events as they happen. Throughout this thesis, I use a sample of 9035 Type 1 AGN lightcurves (ALPAKA: Mullaney et al., 2013) from the Zwicky Transient Facility (ZTF: Bellm et al., 2019), and further details of this sample are outlined in Chapter 2. In the coming years, however, astronomers will have access to groundbreaking amounts of data provided by upcoming high-cadence surveys. One such survey is the Legacy Survey of Space and Time (LSST) at the Vera C. Rubin Observatory equipped with an 8.4 meter primary mirror and the largest digital camera ever built (Ivezić et al., 2019). The telescope, with first-light scheduled for late 2025, will take repeated detailed images of the southern hemisphere sky for 10 years, covering the entire sky every few nights. This exceptional cadence is ideal for trying to detect AGN flares in real-time and also for characterising AGN variability in general. LSST will allow us to probe unprecedented regions of AGN luminosity and redshift parameter space, which is important when trying to characterise "normal" AGN variability, and

hence be able to determine when an AGN is deviating from this behaviour. AGN flares are extremely rare (§1.2) and so in order to cultivate statistically meaningful samples of these events, we need the large amount of data that LSST will provide.

The work undertaken in this thesis is aptly timed to coincide with LSST first light and the findings published in McLaughlin et al. (2024) and in Chapters 4 and 5 will be useful and informative for the wider transient detection community looking ahead to high-cadence surveys such as LSST.

Since LSST is not projected to release data for many months at the time of writing, the samples of AGN studied throughout this thesis were acquired from another high-cadence, wide-field survey, the Zwicky Transient Facility (ZTF: Bellm et al., 2019). ZTF is an ideal survey to use in the meantime for the purpose of developing transient detection techniques; it is described by Graham et al. (2019) as a "stepping stone for LSST for transient object astronomy". In the following subsection, I provide a technical overview of the survey in order to contextualise the data and findings presented in the following chapters.

1.4.1 The Zwicky Transient Facility

The Zwicky Transient Facility (ZTF) is a time-domain, optical survey located at the Palomar Observatory in California (Bellm et al., 2019; Graham et al., 2019). The Samuel Oschin Schmidt telescope at the observatory has a 48-inch aperture and a 47-square-degree, 600 megapixel science camera (Dekany et al., 2020). The survey images the entire observable northern sky every three nights (Bellm & Kulkarni, 2017; Graham et al., 2019), with the intention of characterising the transient sky, focussing on fast-changing astrophysical phenomena such as supernovae, variable stars, AGN, tidal disruption events, and multi-messenger (gravitational-wave) astronomy (Graham et al., 2019).

ZTF has a complement of three filters, g (ranging from $\sim$ 400-550 nm), r (ranging from $\sim$ 550-700 nm), and i (ranging from $\sim$ 700-850 nm) with 5σ limiting magnitudes of 20.8, 20.6, and 19.9 respectively for typical exposure times of 30 seconds (Bellm et al., 2019).

Images are created by difference imaging using the ZOGY algorithm (Zackay et al., 2016), where a science image is compared to a template image by subtracting the latter from the former, resulting in an image depicting moving or changing sources (transients). In addition to a transient alert stream[2] (Smith et al., 2019), ZTF also offers a forced photometry service (Masci et al., 2023), which was used to acquire the large samples of AGN lightcurves analysed in this thesis. Forced photometry is a method of measuring flux at a given set of coordinates, regardless of whether the source is detected above the threshold noise. This is achieved through using point spread function (PSF) models to fit the flux at a fixed location. Forced photometry is beneficial for lightcurve completeness in the context of transient follow-up as it can prevent cases of non-detection in surveys where there is low signal-to-noise. Therefore, meaningful upper limits can be obtained for objects that may not have been detected in a given survey (Makrygianni et al., 2021). In addition, forced photometry prevents cases of source confusion in crowded regions and complex backgrounds (Masci et al., 2023).

Despite such benefits, the use of forced photometry is not without its drawbacks, which the reader may wish to refer back to throughout this thesis. The main caveats of using forced photometry are:

- In crowded fields, source confusion is still possible using an aperture at a fixed position (Portillo et al., 2020; Masci et al., 2023).
- There is a reliance on assumptions about the source's PSF, which can cause flux estimates to be biased (Portillo et al., 2020; Makrygianni et al., 2021).

[2]https://lasair-ztf.lsst.ac.uk.

- When the source is faint or undetected, flux values can be dominated by noise (Masci et al., 2023).

The final point listed is the most important limitation of forced photometry to consider when interpreting the findings of the preceding thesis chapters; when observations are close to the magnitude limit of the wavelength band in question (*i.e.,* low signal-to-noise), the data can be unreliable and noise-dominated.

1.5 Open questions and intentions for the thesis

Over the past three and a half years I have been developing and refining a robust technique for the systematic detection of AGN flares in optical lightcurves, with the aim to produce a pipeline that can be used to sift transients from the large amounts of data that astronomers will be facing in the advent of high-cadence, time-domain surveys such as the LSST. The thesis seeks to present Gaussian Processes as a feasible means of achieving this whilst attempting to answer some open questions in the field currently, outlined below:

- Can systematic detection techniques distinguish between an AGN flare and typical AGN variability in optical lightcurves?
- At what point can AGN flares be detected in lightcurves outside of the underlying variability?
- Can we determine exactly when a flare actually occurs, *i.e.,* its location within a lightcurve?
- How early into a flare are we able to detect it?

To answer the above questions and demonstrate the process of developing such a statistical flare-detection technique, the outline of this thesis is as follows:

- In Chapter 2, I outline my initial approach to systematic flare detection in AGN lightcurves using first principles. After an investigation into what a flare might look like in the time-series of an optical AGN lightcurve, I decided to use a brute-force approach whereby a statistically robust median was subtracted from each lightcurve and any contiguous sets of residuals were flagged as potential flares. I discuss the methodology behind this straight-forward approach and present the results of applying such approach to the ALPAKA catalogue of Type 1 optical AGN lightcurves (Mullaney et al., 2013).

- In Chapter 3, I move from the brute-force approach for systematic flare detection in Chapter 2 to the statistically robust, non-parametric regression that is a Gaussian Process. Informed by the limitations of the first-principles approach in Chapter 2, I embark on a feasibility study to determine the efficacy of using a Gaussian Process for automated, systematic flare detection in AGN lightcurves. I present the results of this feasibility study on simulated AGN lightcurves with injected simulated flares, before applying the technique to the real AGN lightcurves of the ALPAKA catalogue and I present lightcurves that contain flare candidates. The work in this chapter is based on my publication, McLaughlin et al. (2024).

- Since Chapter 3 results in the ability to determine whether a given AGN lightcurve *contains* a flare but not its position within the lightcurve (*i.e.,* the exact point at which the flare occurs), in Chapter 4 I discuss how I alter the GP-based flare detection technique of Chapter 3 to enable for real-time flare detection, which I refer to as flare localisation. I demonstrate that this localisation technique is successful in simulated AGN lightcurves and then I present the results of systematic flare localisation analysis on the ALPAKA catalogue.

- In Chapter 5, I expand my sample from the 9035 Type 1 AGN lightcurves of the ALPAKA catalogue to the Million Quasar Catalogue containing 907 144 Type 1

quasars and AGN (MILLIQUAS: Flesch, 2023). I apply the Gaussian Process flare detection technique from Chapter 3 to this sample and I compare the results to those from the ALPAKA catalogue.

- In Chapter 6, I outline my conclusions of the thesis and the implications of my findings for the broader community. I also discuss potential future work that would be valuable to further the progress made in the thesis for the purpose of real-time transient detection in the era of high-cadence surveys.

Chapter 2

A "brute force" approach to systematic flare detection

2.1 Introduction

Extreme variability of AGN, showing large luminosity changes that are statistically significant from the baseline variability, are known as AGN flares (*e.g.,* Graham et al., 2017; Lawrence et al., 2016; Frederick et al., 2021). Since the study of AGN variability, and indeed the study of the extreme variability of AGN flares, is an indirect probe of the accretion physics of the disk when it cannot be resolved directly (Peterson, 2001), there is a need to be able to systematically detect AGN flares amongst large amounts of data. The detection of AGN flares presents two key problems: 1. they are thought to be rare (Lawrence et al., 2016; Graham et al., 2017) and 2. disentangling a flare from the stochastic baseline variability is not straightforward. This "needle in a haystack" problem has gained a lot of traction in recent years, with the intention to seek a better understanding of supermassive black hole (SMBH) accretion physics, to identify the physical phenomena at the heart of these flaring events, and to put constraints on accretion models (Wiseman et al., 2025). In the coming era of time-domain astronomy,

with surveys such as the LSST and others promising unprecedented amounts of data (Ivezić et al., 2019), it will be possible to detect AGN flares in significant numbers and collate a statistically representative sample of these events to study. To achieve this, however, it is imperative that an effective detection technique is developed that is able to successfully sift AGN flares from typical AGN variability. This challenge is at the crux of this thesis.

To commence investigation into this needle in a haystack problem, in this chapter I present a "first principles" approach to systematically detect AGN flares and extreme variability in optical lightcurves. The motivation behind this approach is straightforward and asks the question: what is it, mathematically, about the time series of a flaring lightcurve that sets it apart from a typical AGN lightcurve? I undertake this systematic search for AGN flares in 20 000 optical AGN lightcurves from the ALPAKA catalogue (Mullaney et al., 2013) by calculating a statistically robust median across pre-determined sections of lightcurves of various sizes. After subtracting this median from the lightcurve, I identify sections where there are multiple contiguous data points below this median magnitude value(s) and define these sections as potential flares.

This chapter is organised as follows: in §2.2, I describe the data used in this chapter and the methodology of this technique. In §2.3, I present the results of this systematic search for AGN flares. I discuss the implications and limitations of these results in §2.4 and finally I present concluding remarks in §2.5.

2.2 Methodology

Systematic searches for AGN flares in the literature are limited, with most of the literature focussing on individual events. Notably, Graham et al. (2017) undertook a systematic search for AGN flares in the Catalina Real-Time Transient Survey (CRTS: Drake et al., 2009). They identified 51 events from over 900 000 known quasars and high-probability

quasar candidates, typically lasting 900 days with a median peak amplitude of 1.25 magnitudes above the median variability. I expanded on this method and developed my own algorithm to systematically search for short-duration AGN flares in ZTF lightcurves. I define "short-duration" as a timescale of < 1000 days, but since Graham et al. (2017) ignore flares shorter than 300 days in their investigation, I focused on the prospect of identifying flares with shorter timescales than this (< 300 days).

Using ZTF lightcurves of over 20 000 known AGN (acquired using forced photometry) from the ALPAKA catalogue developed by Mullaney et al. (2013) – see §2.2.1 for more detail – I wrote a Python algorithm using the ASTROPY[1] Python package (Astropy Collaboration et al., 2013, 2018) to identify AGN flares from the data. My algorithm works in three steps: first, it splits the lightcurve into sections of user-determined window size (in this case, 100, 300, 500, 800, and 1000 days). The second step invokes the use of the Theil-Sen estimator, which is a robust form of regression, first proposed by Theil (1950) and later expanded by Sen (1968). It is essentially the median slope between all pairs of points; the Theil-Sen estimator of a set of two-dimensional points (x_i, y_i) is given by the median of the slopes $(y_j - y_i)/(x_j - x_i)$ determined by all pairs of points in the set. The algorithm calculates the Theil-Sen median slope across each lightcurve section, along with the intercept using SCIPY[2] (Virtanen et al., 2020). From these parameters, the Theil-Sen median magnitude can be calculated, which is then subtracted from the original lightcurve. Finally, the algorithm identifies sections of lightcurve where there are multiple (10) contiguous data points brighter than the Theil-Sen median. For each flare that is detected, the algorithm calculates the number of points in the flare, the flare duration, and the median, minimum and maximum amplitude of the flare. I chose to ignore flare candidates with fewer than 10 data points in the flaring region, and flare candidates lasting less than 10 days. This was due to a large number of false-positives

[1] https://www.astropy.org
[2] https://docs.scipy.org/doc/scipy/index.html

caused by noise in the ZTF lightcurves, which is likely the result of observational effects and the limitations of forced-photometry.

Graham et al. (2017) used the Theil-Sen estimator to calculate the median slope of each entire lightcurve, but my algorithm splits each lightcurve into sections across which the Theil-Sen median is calculated. By using window sizes of 100, 300, 500, 800 and 1000 days, I investigated how the distribution of flare duration vs flare magnitude change is related to which window size is chosen. By using the median of the entire lightcurve, it risks ignoring more local, shorter term variability and hence risks the loss of shorter-duration flare candidates. By splitting up a lightcurve into smaller sections, flare candidates with shorter timescales are identified as the algorithm is focused more on local magnitude changes rather than long-term trends of the lightcurve.

2.2.1 Data

The ALPAKA catalogue was developed by Mullaney et al. (2013) and comprises 24 264 optically selected AGN from the Sloan Digital Sky Survey (SDSS) Data Release 7 database. These AGN were selected spectroscopically using a multicomponent fitting routine and were classified into Type 1 or Type 2 depending on the presence of broad lines in their spectra; see Mullaney et al. (2013) for a detailed description of the SDSS query, selection criteria and contaminant removal. Optical, r-band lightcurves of the ALPAKA catalogue were obtained from the Zwicky Transient Facility (ZTF) Data Release 6 (see Chapter 3, §3.3 for more detail).

2.3 Results

In this section I present the results found from a systematic search for AGN flares obtained using the method described above. First, I present typical distributions of flare duration vs flare amplitude for different window sizes and then I present lightcurves of

potential flare candidates.

2.3.1 Typical flare duration and magnitude distributions

Using the "brute force" flare detection algorithm (see §2.2), I produced plots showing distributions of flare duration vs apparent magnitude change for Type 1 and Type 2 AGN; these are shown in Fig. 2.1 and Fig. 2.2. The algorithm output a list of AGN flare candidates and the location of the flare within its lightcurve, the flare duration, the number of data points within the flare, and the peak, median and minimum flare magnitude. I produced these distributions using the output of my own algorithm, which makes use of a window-based approach in order to identify shorter term flares. Flare duration vs magnitude distributions for both Type 1 and Type 2 AGN are presented in the following sections.

Type 1 flare candidates

Distributions of typical flare duration vs flare amplitude for Type 1 AGN are shown in Fig. 2.1, and median flare duration and flare amplitude for each window size are outlined in Table 2.1. The shorter the window size used, the shorter the duration and smaller the amplitude is of the flares. It can be seen from the high density in the bottom-left corner of each plot in Fig. 2.1 that the majority of AGN flares are extremely short duration and small amplitude, with magnitude changes of less than 0.2 on timescales of tens of days. These flare candidates are ruled out and are considered to be part of the intrinsic AGN variability. It is clear that the majority of "flares" detected with this technique are simply examples of standard AGN variability. Indeed, the median flare amplitudes identified in Table 2.1 are typical of AGN variability (*e.g.,* MacLeod et al., 2010).

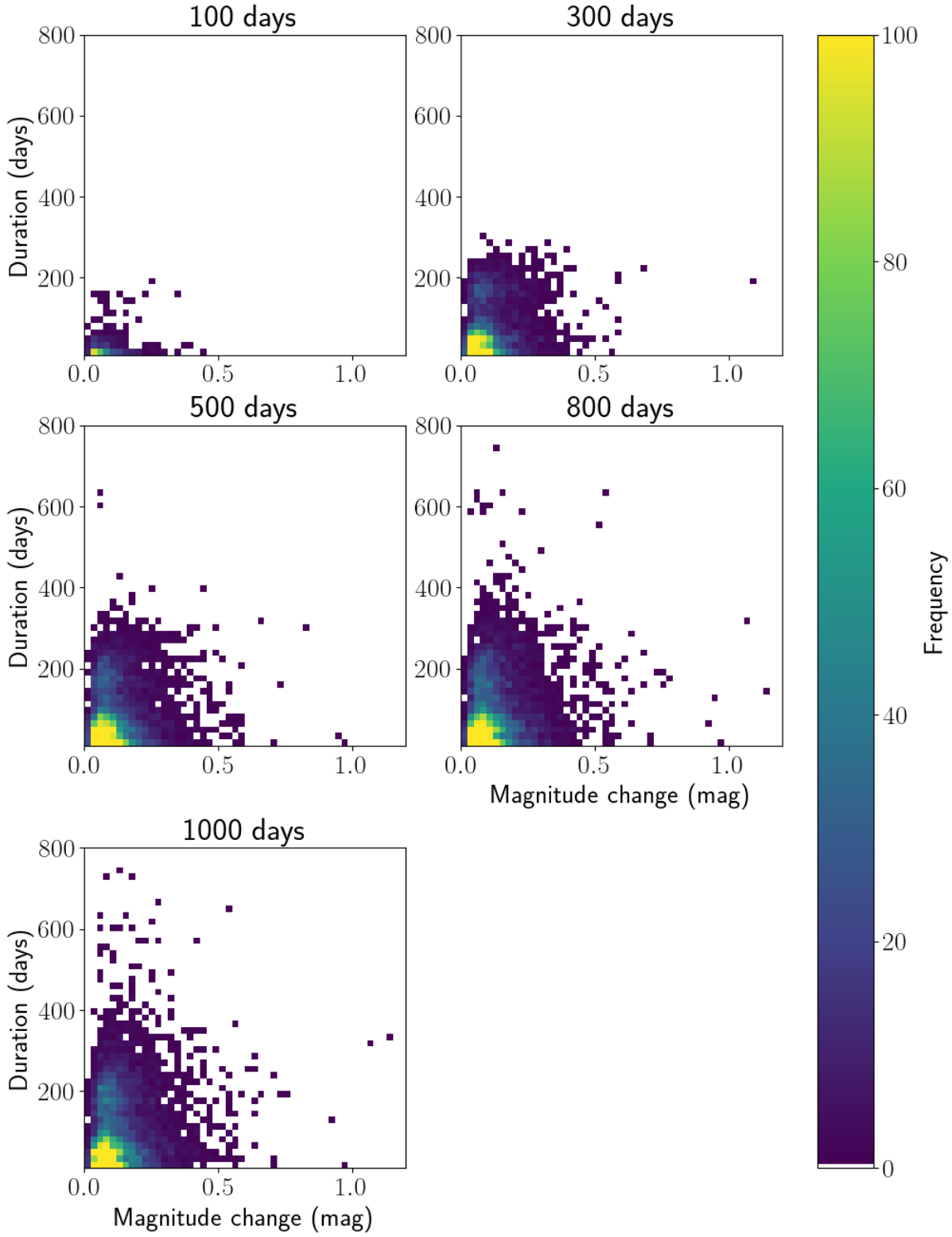

Figure 2.1: Distributions of flare duration vs flare magnitude change for Type 1 AGN across timescales of 100, 300, 500, 800 and 1000 days.

Table 2.1: Median flare duration and median flare amplitude for Type 1 AGN flare candidates.

Window size (days)	Median flare duration (days)	Median flare amplitude (mag)
100	20	0.067
300	49	0.083
500	53	0.090
800	59	0.097
1000	63	0.10

Table 2.2: Median flare duration and median flare amplitude for Type 2 AGN flare candidates.

Window size (days)	Median flare duration (days)	Median flare amplitude (mag)
100	19	0.18
300	30	0.14
500	28	0.14
800	32	0.16
1000	29	0.15

Type 2 flare candidates

Distributions of typical flare duration vs flare amplitude for Type 2 AGN are shown in Fig. 2.2, and median flare duration and flare amplitude for each window size are outlined in Table 2.2. It is important to note that the median flare amplitude values of the Type 2 AGN are greater than those found in the Type 1 AGN. At first glance, this seemingly contradicts the paradigm of the Unified Model of AGN (see Antonucci, 1993) where, because the accretion disk is obscured by the dusty torus, suppressed variability is observed in Type 2 AGN (*e.g.,* Yip et al., 2009; Kovacevic et al., 2025). In spite of that, I consider these findings to be the result of a selection effect whereby – due to the very fact that the accretion disks of the Type 2 AGN *are* obscured by the torus – they are dimmer than their Type 1 counterparts. As a result, their lightcurves are subject to increased noise levels, causing both greater flare amplitudes to be measured, and shorter flare durations.

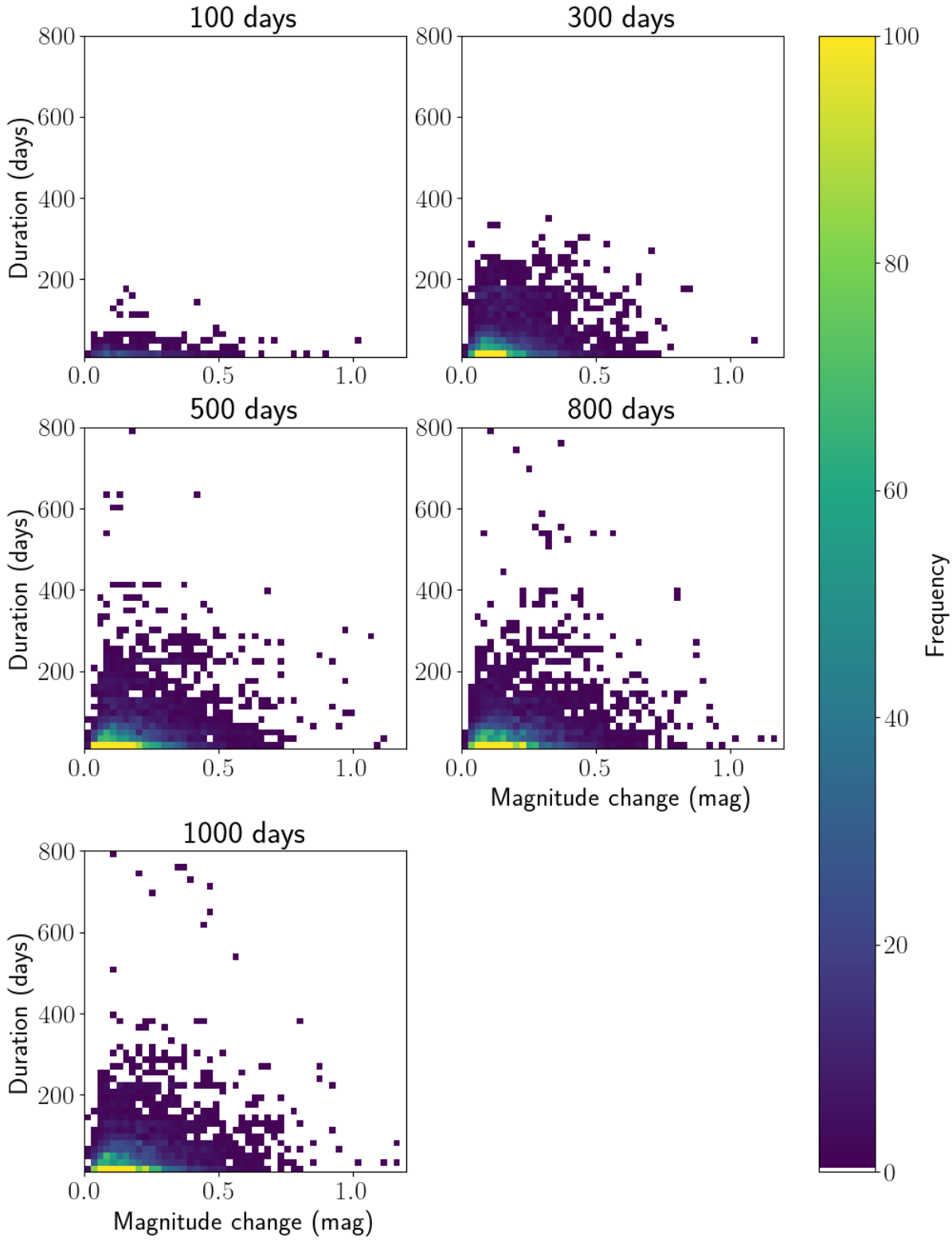

Figure 2.2: Distributions of flare duration vs flare magnitude change for Type 2 AGN.

2.3.2 Lightcurves of flare candidates

In this section I present ZTF r- and g-band lightcurves of the AGN flare candidates found by my flare-finder algorithm.

Type 1 flare candidates

The flare-finder algorithm output contained all instances of contiguous sets of data points below the windowed Theil-Sen median, although as explained in §2.3.1, the majority of these are part of the background AGN variability. Flare candidates were selected to have at least 10 data points in the flare, to be greater than 10 days in duration, and to have a flare magnitude change of at least 0.5. These thresholds were chosen to ensure that there were enough data points in the flaring region to justify a detection, and to mitigate the impact of false-positives caused by noise in the ZTF lightcurves, which is likely the result of observational effects and the limitations of forced-photometry (see §1.4.1). The minimum magnitude change of 0.5 was chosen to be a lower threshold than the requirements set in previous studies such as the 1.5 magnitude threshold used by Graham et al. (2017), to account for the fact that I am probing for the shortest duration flares (which will correspond to lower amplitudes).

Figs. 2.3 and 2.4 show the optical r-band lightcurves of the identified Type 1 AGN flare candidates. Where available, g-band data is also shown, though in many cases the g-band lightcurve is sparse in comparison to the r-band lightcurve. For the objects without g-band data available, or in the cases where the g-band does not exhibit variability in the same region as the r-band "flare", I consider that there is insufficient evidence to make a flare detection. Follow-up aperture photometry could be obtained to confirm these flare candidates due to the risk of anomalous data in ZTF lightcurves (see §1.4.1 for a warning regarding forced photometry), along with cross-matching against other transient databases to rule out supernovae or blazars. It is also important to note that

the majority of these flare candidates exhibit much lower amplitudes than recorded in the literature (*e.g.,* Graham et al., 2017; Lawrence et al., 2016) and it is possible that they are simply the tail-end of the distribution of typical AGN variability (*e.g.,* MacLeod et al., 2010). It is apparent that this "brute force" technique is mistaking the underlying AGN variability with flare activity.

Type 2 flare candidates

Fig. 2.5 shows the r-band lightcurves of the Type 2 AGN flare candidates identified. Where available, corresponding g-band data is also provided. In these cases, whilst these multi-band lightcurves show a potential flare in the r-band, they do not exhibit the same behaviour in the g-band.

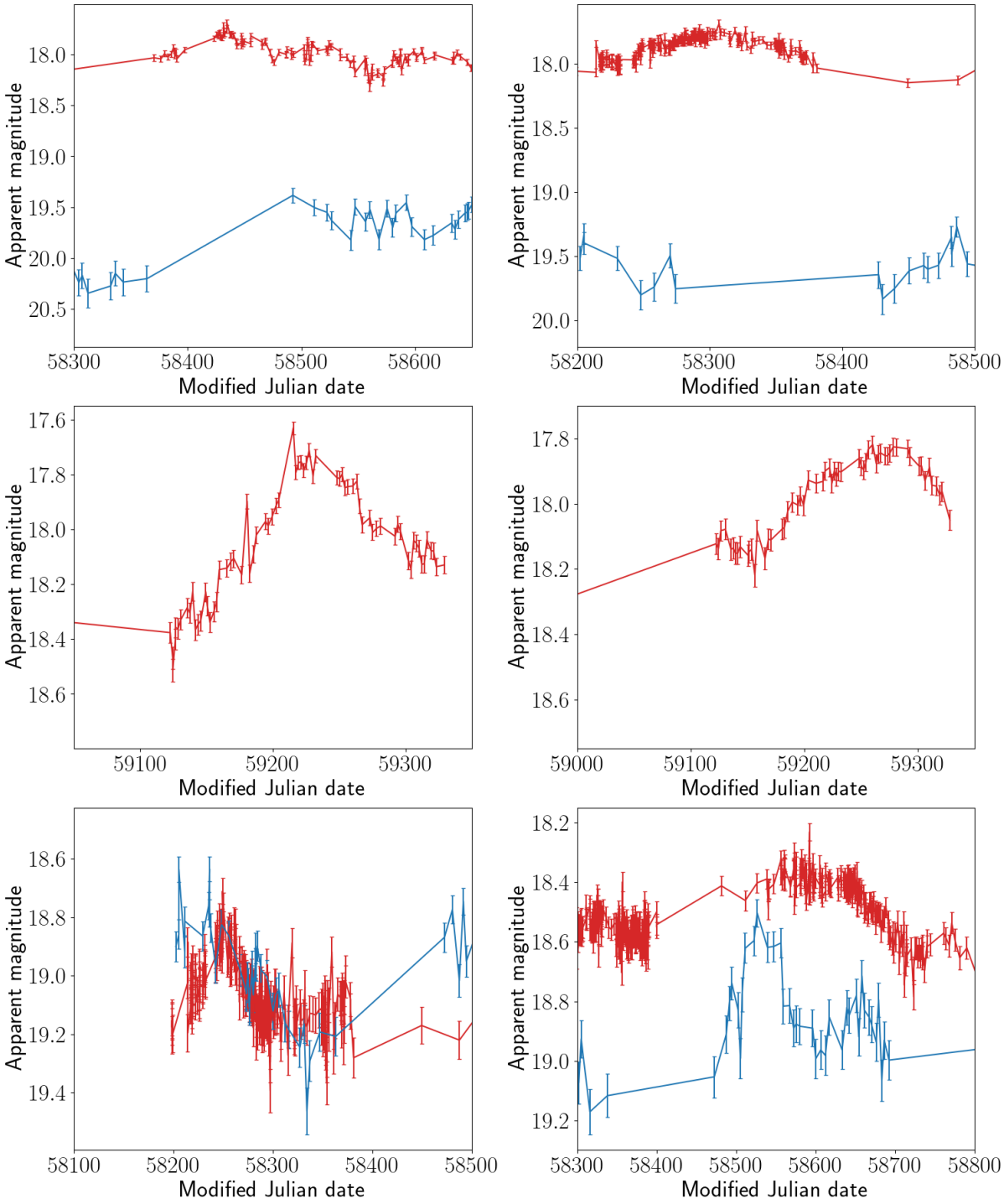

Figure 2.3: Lightcurves of Type 1 flare candidates. The r- and g-bands are shown in red and blue respectively. Lightcurves with only r-band data are plotted in red. As a reminder, the limiting magnitudes of ZTF in the r- and g- bands are 20.6 and 20.8 respectively.

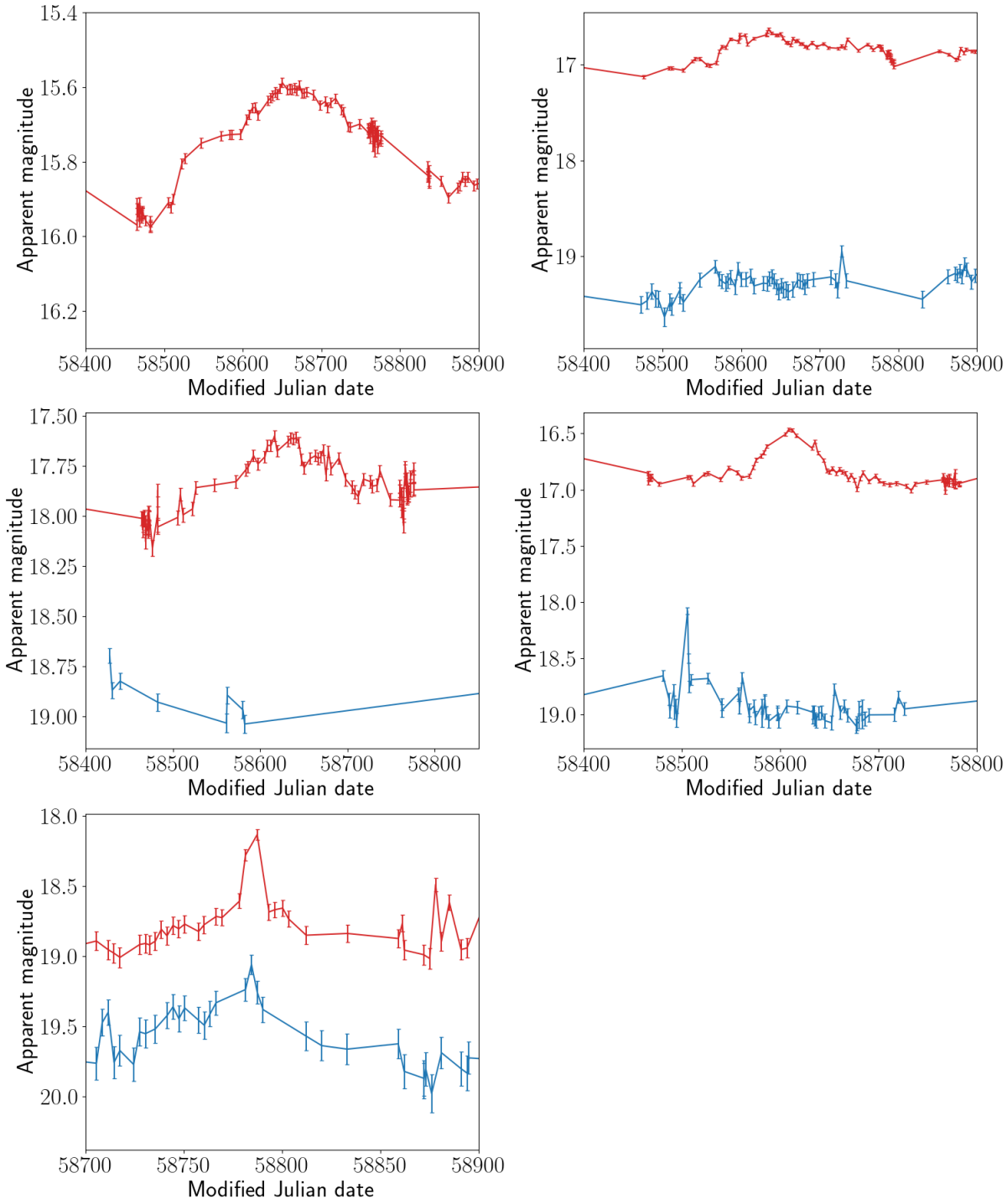

Figure 2.4: Lightcurves of Type 1 flare candidates (cont.) The r- and g-bands are shown in red and blue respectively. Lightcurves with only r-band data are plotted in red. As a reminder, the limiting magnitudes of ZTF in the r- and g-bands are 20.6 and 20.8 respectively.

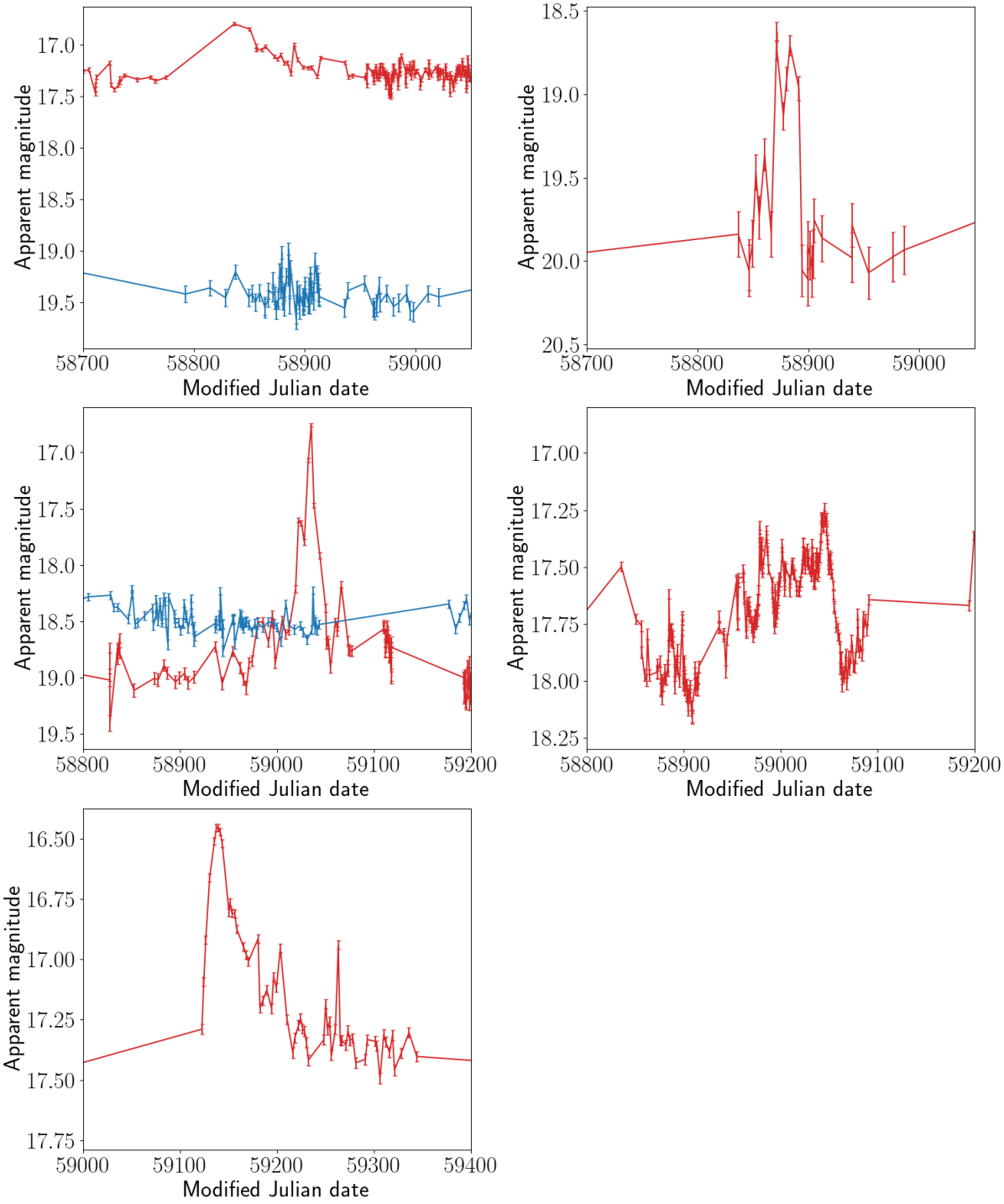

Figure 2.5: Lightcurves of Type 2 flare candidates. The r- and g-bands are shown in red and blue respectively. Lightcurves with only r-band data are plotted in red. As a reminder, the limiting magnitudes of ZTF in the r- and g-bands are 20.6 and 20.8 respectively.

2.4 Discussion

I have presented histograms of flare duration vs flare magnitude change for Type 1 and Type 2 AGN (Figs. 2.1 and 2.2), generated using the flare-finder algorithm described in §2.2 with window sizes of 100, 300, 500, 800 and 1000 days. These distributions can serve as a means to compare new observations of AGN flares against the broader population, which will enable for follow-up of sources that are outliers. It is apparent from these distributions that it is typical for AGN to vary with amplitudes of 0.5 magnitudes or less on these timescales.

The flare finder algorithm identified a total of 16 potential flare candidates shown in Figs. 2.3, 2.4 and 2.5. 11 of these are Type 1 and 5 are Type 2 AGN. From lightcurve selection, objects with g-band data that do not show a flare corresponding to the r-band flare can be ruled out. After removing these sources, there are a total of 11 flare candidates. On inspection, even the most robust flare candidates with corresponding flares in both r- and g-bands show little evidence of flaring activity on inspection of their photometric images (*i.e.,* due to poor seeing and difficulty distinguishing the source from the background). This could be a result of calibration issues with ZTF images, weather effects, or the fact that shorter-duration flares are typically in dimmer AGN.

From these results it is apparent that this brute-force technique results in a significant number of false positives and is sensitive to the underlying variability of an AGN lightcurve, erroneously identifying typical AGN variability as flaring. At this stage, there are two key limitations of this approach: (a) the significant numbers of false positives as a result of the algorithm being unable to distinguish between flares and the underlying AGN variability and (b) the fact that, to identify flares in this way, one must place arbitrary thresholds on the flare properties (*e.g.,* its amplitude must exceed 1 magnitude). Any arbitrary limits placed on the flare amplitude or timescale are not statistically motivated and by definition do not take into account the stochastic nature of the underlying

variability.

2.5 Conclusions

In this chapter I have demonstrated that this “brute force” or “first principles” approach to flare detection is not sufficient for systematic flare detection as it results in a significant number of spurious detections. Furthermore, the technique is not able to take into account the underlying variability of a lightcurve in a statistical manner. Like systematic studies that have come before (*e.g.,* Lawrence et al., 2016; Graham et al., 2017), one has to make arbitrary cuts on either the duration or amplitude (or both) of the flare one intends to detect, which ignores the stochastic nature of variability (see §2.3.2). The findings of this chapter, whilst they have been enlightening in terms of characterising typical AGN variability across a given timescale, motivate the requirement for a statistical means of systematically detecting AGN flares, where no assumptions are made about the nature of the flares and the underlying variability is taken into account. This forms the basis of the following Chapter 3.

Chapter 3

Using Gaussian Processes to detect AGN flares

3.1 Introduction

In Chapter 2, I presented a "first-principles" approach to detecting AGN flares in optical lightcurves, by subtracting the median from each lightcurve and identifying contiguous sets of points below this median. I showed that whilst this technique is valid, it results in a significant number of false positives. The technique is not statistically robust as it cannot determine in a statistical manner what constitutes a significant departure from the baseline variability. Furthermore, as with studies undertaken previously (Lawrence et al., 2016; Graham et al., 2017), the technique makes assumptions about the duration and amplitude of the flares one intends to detect, which can introduce selection effects and false positives. For the remainder of this thesis, I investigate the use of Gaussian Processes (GPs) to systematically detect AGN flares as a statistically robust alternative; by employing a GP to parameterise the covariance of a lightcurve, there is no need to *a priori* assume anything about the properties of the flare.

With the above in mind, the aim of this chapter is to assess the viability of using

GPs to identify AGN flares. To do so, I first simulate the lightcurves of a population of variable AGN, including flaring events, then apply GP analysis to assess how successfully it identifies the latter. Next, I apply this analysis to real AGN lightcurves as a systematic search for flaring events. This data was obtained from the ZTF Public Data Release 6 and the sample comprises optical, r-band lightcurves of Type 1 AGN (Masci et al., 2018; Bellm et al., 2019).

The outline of this chapter is as follows: I describe the theory behind GPs (§3.2), the data used to investigate the efficacy of GPs (§3.3), the GP kernel hyperparameter distributions (§3.4) and the methodology behind using GPs for the classification of AGN flares (§3.5) before presenting the retrieval rates of the GP analysis when dealing with different types of simulated lightcurves, and, finally, the lightcurves of real AGN (§3.6). I discuss my findings and future directions of studies in §3.7, and provide some brief concluding remarks in §3.8.

3.2 Gaussian Processes

A detailed explanation of the theory behind GPs is provided in Chapter 1.3, but I will provide a brief reminder of the technique and explain the reasoning behind the decision to employ them in the context of flare detection. GPs are an interpolation technique that is able to predict the underlying function given a set of data (Rasmussen & Williams, 2006). To reiterate the description in Chapter 1, a GP is a prior over functions that takes a user-defined covariance function and, *given* the data, it generates the full set of functions that are both: (a) possible fits to the data, and (b) feasibly produced by the chosen covariance function. In addition to its function as an interpolation tool, it is able to provide the parameters of the covariance function (kernel) that produces the most likely description of the data. In this capacity, GPs are a means of encapsulating the variability of a

lightcurve, and these optimum kernel parameters (known as hyperparameters[1]) can be used as a summary statistic of the variability of a given dataset. It is for this reason that GPs were chosen as a potential technique for AGN flare detection: since they can be used to provide a description of how a given lightcurve is varying, it is reasonable to expect that they can be used to differentiate between lightcurves that contain flares and those that do not. This therefore motivates an exploration into whether GPs can be used as a tool to classify transient astronomical events – and specifically AGN flares – directly, which is the premise of this chapter. The choice, then, is to decide the most suitable covariance function that represents typical AGN variability, and there are a number of candidates that are suitable, such as the squared exponential kernel and the Matérn class of kernel, which both represent stochastic variability, and their main difference being that the squared exponential kernel results in smoother function realisations. In this work, based on the findings of Griffiths et al. (2021) that the Matérn-3/2 kernel is more effective at reproducing AGN lightcurves, I choose the same kernel in which the covariance k is defined as follows:

$$k(r) = \sigma^2 \left(1 + \frac{\sqrt{3}r}{\rho}\right) \exp\left(\frac{-\sqrt{3}r}{\rho}\right) \tag{3.1}$$

where r is equal to the difference between all pairs of values of the independent ordinate (in this case time, *i.e.,* $r = |\mathbf{t} - \mathbf{t}'|$), σ is the variability amplitude and ρ is the variability timescale (Rasmussen & Williams, 2006; Foreman-Mackey et al., 2017). Visualisations of the types of function generated by this kernel and an example of a GP with this kernel fit to an AGN lightcurve are shown in Chapter 1 §1.3.

[1]The *parameters* of the GP, *i.e.,* the quantities that are allowed to vary given the data, are the *functions* produced by the kernel. Therefore, the coefficients of the kernel are defined as *hyperparameters* as they describe the parameters of the model.

3.3 Data

Prior to using GPs to classify real AGN lightcurves, I wanted to determine whether they are even a feasible means to detect flaring events. The problem with using real data for such a feasibility study is that, with AGN flares being so rare, I would need to use a large sample of AGN lightcurves (*i.e.,* numbering tens to hundreds of thousands) to ensure it contains even a small handful of true flaring events. For such a large sample, however, it is unfeasible for us to know which real lightcurves contain true flaring events, so I cannot evaluate success rates. To overcome this, I turned to simulating lightcurves, which allows us to inject flares. Since I know which of the simulated lightcurves contain injected flares, I can determine true and false positive and negative rates. Once I have assessed the feasibility of using GPs to detect AGN flares in simulated data, I then apply it to real AGN lightcurves to determine whether it can, indeed, detect real AGN flaring events. I note, however, that this final step is simply an exploratory exercise; I cannot easily assess success rates on large samples of real data for the reasons outlined above. It is also important to note that while I use the simulated datasets to assess the feasibility of using GPs to identify flares, I *do not* use the simulated datasets to inform my priors for analysing the real ZTF lightcurves; instead I use the GP analysis to determine the range of typical variability parameters of each sample independently. In doing so, I ensure that any deviations from that range – which potentially highlight the presence of a flare – are specific to that sample. In this section I outline how I produced my sample of simulated AGN lightcurves, describe how I employed GPs to analyse these simulated data, and then explain how I applied Bayesian hypothesis testing to the output of the GP analysis to calculate the probability of a lightcurve containing a flare.

3.3.1 Simulated lightcurves

It has been known for over a decade that non-flaring AGN lightcurves are well-described by a one-dimensional damped random walk (*e.g.,* Kelly et al., 2009; MacLeod et al., 2010). This involves adding a correctional term (*i.e.,* a damping term) to a random walk to encourage extreme deviations back to the mean value. Kelly et al. (2009) first showed that a damped random walk can statistically explain the observed lightcurves of AGN; they analysed 100 quasar lightcurves and, using a Bayesian approach, showed that this stochastic process is capable of modelling AGN lightcurves at an accuracy level of 0.01 - 0.02 magnitudes.

MacLeod et al. (2010) modelled the time variability of 9000 quasars in SDSS Stripe 82 as a damped random walk and confirmed previous results (*e.g.,* Kelly et al., 2009; Kozłowski et al., 2010) that this model describes quasar lightcurves well. Therefore, I used this damped random walk model to simulate my own AGN lightcurves. I drew values of the variability parameters (SF^{∞} and τ, which are defined as the difference in magnitude across the longest time steps and the damping timescale respectively – see Eq. 3.2) from the distributions of best-fit variability parameters presented in MacLeod et al. (2010). To achieve this, I randomly drew values of $\log(SF^{\infty})$ from a normal distribution with a mean of -0.8 mag and standard deviation of 0.2 mag. Then, I calculated the values of $\log(\tau)$ based on the best-fit power law in MacLeod et al. (2010). By selecting the full range of these parameters, I intrinsically include the variability of the entire quasar population. It should also be noted that since I am assuming that AGN flares are extremely rare, I assume that the MacLeod et al. (2010) sample represents "normal" AGN variability. In their work, MacLeod et al. (2010) make use of the structure function, $S(\tau)$, to express the long-term variability of an AGN lightcurve. In the context of a damped random walk, the structure function is defined as follows:

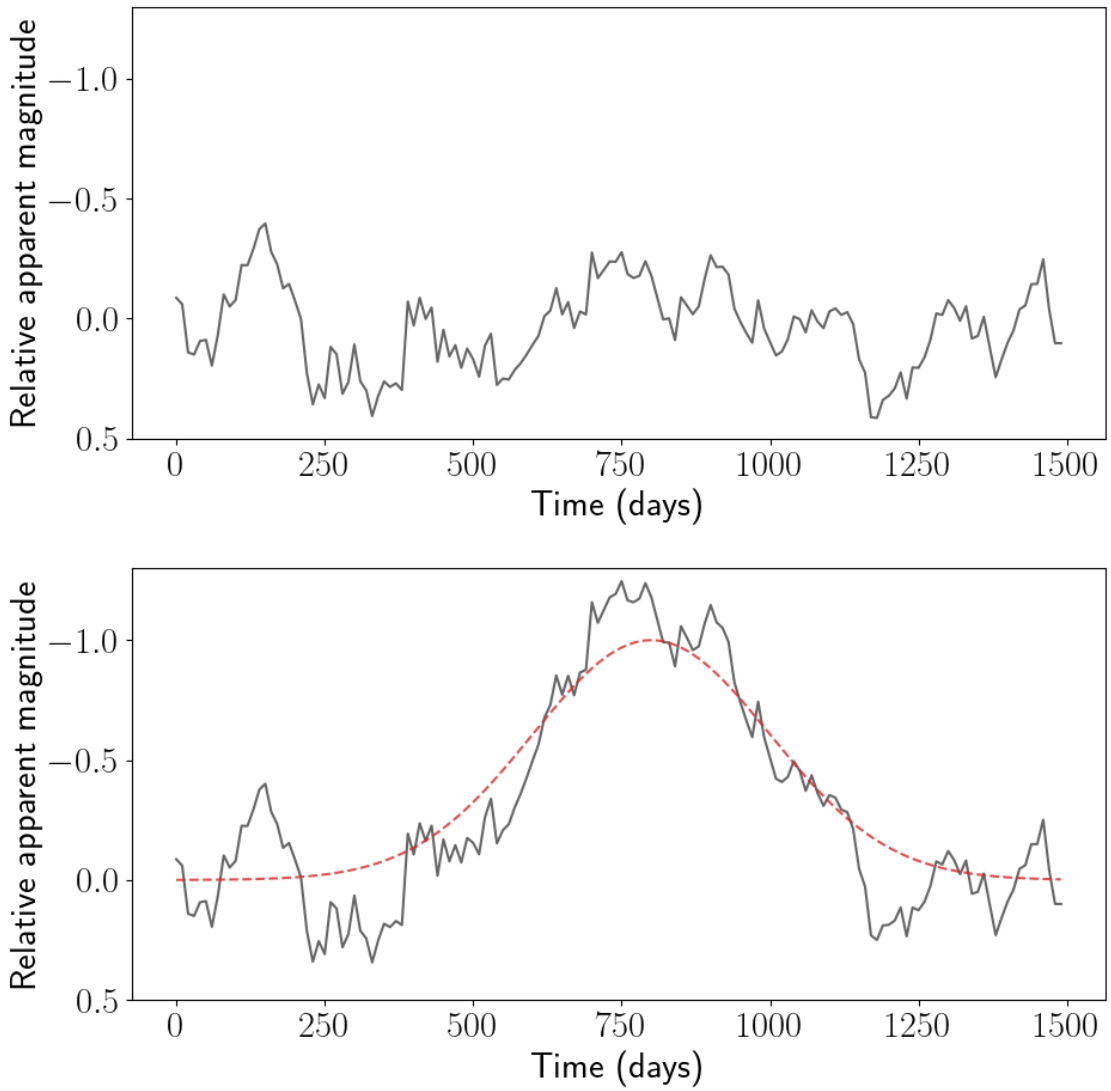

Figure 3.1: Top: a simulated AGN lightcurve using a 1D damped random walk. Bottom: the same simulated lightcurve with an injected 1 magnitude Gaussian flare at 800 days with a width of 200 days. The dotted red line shows this underlying Gaussian function. Figure from McLaughlin et al. (2024).

$$\mathrm{SF}(\Delta t) = \mathrm{SF}_\infty \left(1 - \exp\frac{-|\Delta t|}{\tau}\right)^{\frac{1}{2}}. \tag{3.2}$$

Here I introduce two variability parameters: SF^∞ and τ. The former is the difference in magnitude calculated across the longest time steps, while τ can be thought of as the damping timescale in days, upon which the value of the lightcurve returns to its mean. The structure function is computed by collecting the differences in magnitude for all points in the lightcurve separated by a given time lag, Δt. In addition, the function asymptotes at very large time steps (tending to SF^∞), which corresponds to a power-law fit (MacLeod et al., 2010).

A lightcurve is generated by selecting values for SF^∞, τ and the mean value of the lightcurve, μ (in my case, the values from the full distributions presented in MacLeod et al. 2010). The magnitude $X(t)$ at a given time step δt from a previous value $X(t-\delta t)$ is drawn from a normal distribution with a mean and variance given by (Kelly et al., 2009; MacLeod et al., 2010):

$$E(X(t)|X(t-\Delta t)) = \exp\left(-\frac{\Delta t}{\tau}\right) X(t-\Delta t) + \mu\left(1 - \exp\left(-\frac{\Delta t}{\tau}\right)\right) \tag{3.3}$$

and

$$\mathrm{Var}(X(t)|X(t-\Delta t)) = 0.5(SF^\infty)^2\left(1 - \exp\left(-\frac{2\Delta t}{\tau}\right)\right). \tag{3.4}$$

Using this approach, I simulated 10 000 AGN lightcurves with a cadence of 10 days and uniform uncertainties of 0.1 magnitudes. These represent my "perfect" simulated AGN lightcurves since they are regularly sampled and do not contain any anomalies[2] (nor flares). An example of one such lightcurve is shown in Fig. 3.1. In the following subsections I discuss the steps undertaken to modify and filter these perfect simulated

[2]Hereafter, I refrain from placing quotation marks around "perfect".

lightcurves to include flares and to make them more representative of real, irregularly-sampled AGN lightcurves. To summarise, these are:

1. injected flares and simulated with a constant 10-day cadence;
2. injected flares and sub-sampled to match the cadence of real ZTF lightcurves;
3. as 2, but with added outliers;
4. real ZTF with injected flares;
5. real ZTF lightcurves.

The objective was to investigate the ability of GPs to classify flares and non-flares in each of these cases, with each step becoming progressively more representative of observed AGN lightcurves.

3.3.2 Lightcurves with injected flares

To determine how GPs would handle perfect, uniformly-sampled data without outliers, I simulated 10 000 AGN lightcurves with a cadence of ten days, and a baseline of 1500 data points. This would act as a control sample. A copy of this sample was created, and a simulated flare was injected into each lightcurve in the copied sample. This resulted in a control sample and a flare sample of uniformly-sampled AGN lightcurves, totalling 10 000 lightcurves per sample.

The flares were simulated in two ways: (1) as Gaussian functions and (2) as gamma functions, to investigate the effect of the shape of the flare on the GP fit (resulting in a total sample of 30 000 lightcurves including the control, Gaussian and gamma samples). Gaussian flares are symmetrical and gamma flares have a short rise-time and a longer decay. These flares were simulated with amplitudes ranging from 1–2.5 magnitudes and durations of between 100 and 1000 days. To achieve this, flare amplitudes were randomly

drawn from a uniform grid ranging between 1–2.5 magnitudes with a spacing of 0.1, and the durations were randomly chosen from an array containing durations of 100, 200, 300, 500 and 800 days. The term "duration" in this context refers to the full width at half maximum of the flare function. These flares were injected at random locations (drawn from a uniform grid with spacing of 100 days) into each simulated lightcurve such that their peak lies after the first 300 days but before the last 300 days. This is to ensure that in all cases the rise and fall of the flare was included. These injected flare properties were decided in consideration of the findings from previous systematic searches for AGN flares (e.g. Lawrence et al., 2016; Graham et al., 2017).

3.3.3 Sub-sampled lightcurves

In reality, AGN lightcurves are not uniformly sampled. One way to achieve non-uniformity is by randomly sub-sampling each lightcurve, however this would not faithfully represent real, observed lightcurves due to weather effects, differences between filters and large gaps in the data. For this reason, I instead interpolated the simulated lightcurves onto the time axis of real ZTF lightcurves to sub-sample them. This enabled us to investigate how GPs would handle sparsely sampled data and ensured the simulated lightcurves have realistic cadences. These ZTF lightcurves are described in §3.3.5.

3.3.4 Lightcurves with added outliers

In real AGN lightcurves, it is not uncommon to see systematic outliers in the data due to uncorrected atmospheric effects, bad pixels, etc. The GP must be robust against these effects if they are to be used as a classifying tool. Therefore, to further construct simulated lightcurves that were as representative of real data as possible, I added systematic outliers. To achieve this, once the lightcurves had been sub-sampled, I added to each lightcurve a contiguous pair of outliers that were five standard deviations above the

variability of the individual lightcurve; visual inspection shows that this level of outlier is typical of ZTF lightcurves.

3.3.5 ZTF lightcurves

As well as using simulated data, I also used real ZTF lightcurves in the r-band; these data were downloaded in August 2021 from Public Data Release 6, which was the most current data release at the time[3] (ZTF: Masci et al., 2018; Bellm et al., 2019). These ZTF lightcurves were acquired from spectroscopically-selected AGN from SDSS DR7, forming the ALPAKA catalogue (Mullaney et al., 2013). Of this sample, 9035 AGN are Type 1, and it is these AGN whose lightcurves I utilised in this work. Note that this is the same dataset that was used and described in more detail in Chapter 2.[4] For the sake of a proof of concept demonstration, only the r-band was considered, though it would be possible to use Gaussian Processes to perform a multi-band analysis (see §3.7). First, I made a copy of each AGN's ZTF lightcurve, and a Gaussian flare was injected into each copy. This was repeated for the injection of gamma flares. These flares were simulated as in §3.3.2. This created a control sample and two "flare" samples (Gaussian and gamma) of real ZTF lightcurves. This represents as close a sample to real AGN flaring lightcurves as possible, without being true flaring events.

Finally, the original sample of 9035 ZTF lightcurves were processed using the method outlined in the following section to determine if any of these AGN lightcurves would be classified as containing flares.

[3]Later, in January 2024 I re-downloaded the same sample from the more recent data release DR 20. I present brief results using this more up-to-date sample in Appendix A.

[4]I excluded Type 2 AGN from my analysis as, under the unified AGN model (*e.g.,* Antonucci 1993), I do not have a direct view of the nuclear region, meaning they are less variable and I should not – in theory – observe flaring events in such cases.

3.4 GP kernel parameter distributions

With my various simulated and real lightcurves in-hand, I next analysed them with a GP in order to calculate the optimised kernel coefficients (hereafter, hyperparameters) of a Matérn-3/2 kernel. For this, I made use of the open-source Python library CELERITE[5] (Foreman-Mackey et al., 2017), which enables fast and scalable Gaussian Process modelling. Since CELERITE provides a pair of optimised hyperparameters (σ,ρ) given the data in each of my lightcurves, I can plot distributions of these hyperparameters. This enables us to assess whether the distributions for flaring and non-flaring lightcurves reside in different regions of parameter space. If they do, then this opens up the prospect of using GP analysis to classify a lightcurve. In what follows, I consider the hyperparameter distributions for each of my five different classes of lightcurves (*i.e.,* those described in §3.3).

3.4.1 Perfect lightcurves

The distribution of the optimised hyperparameters for my sample of perfect lightcurves are shown in Fig. 3.2 (for Gaussian flares) and Fig. 3.3 (for gamma flares). In these and all following plots in this section, the variability amplitude σ increases as the variability of the lightcurve increases, while the timescale ρ increases with the timescale across which the variability is occurring. The figures show that the hyperparameters for flares and non-flares exist in different but partially overlapping regions of parameter space. This is the case for both Gaussian and gamma flares. The covariances of these well-sampled, flaring and non-flaring lightcurves are statistically different. To further illustrate this, Fig. 3.2 shows the locations of the simulated lightcurves from Fig. 3.1 in hyperparameter space, demonstrating that simply the injection of a flare into a simulated lightcurve moves its hyperparameters from the non-flare distribution to the flare distribution.

[5]https://celerite.readthedocs.io/en/stable/

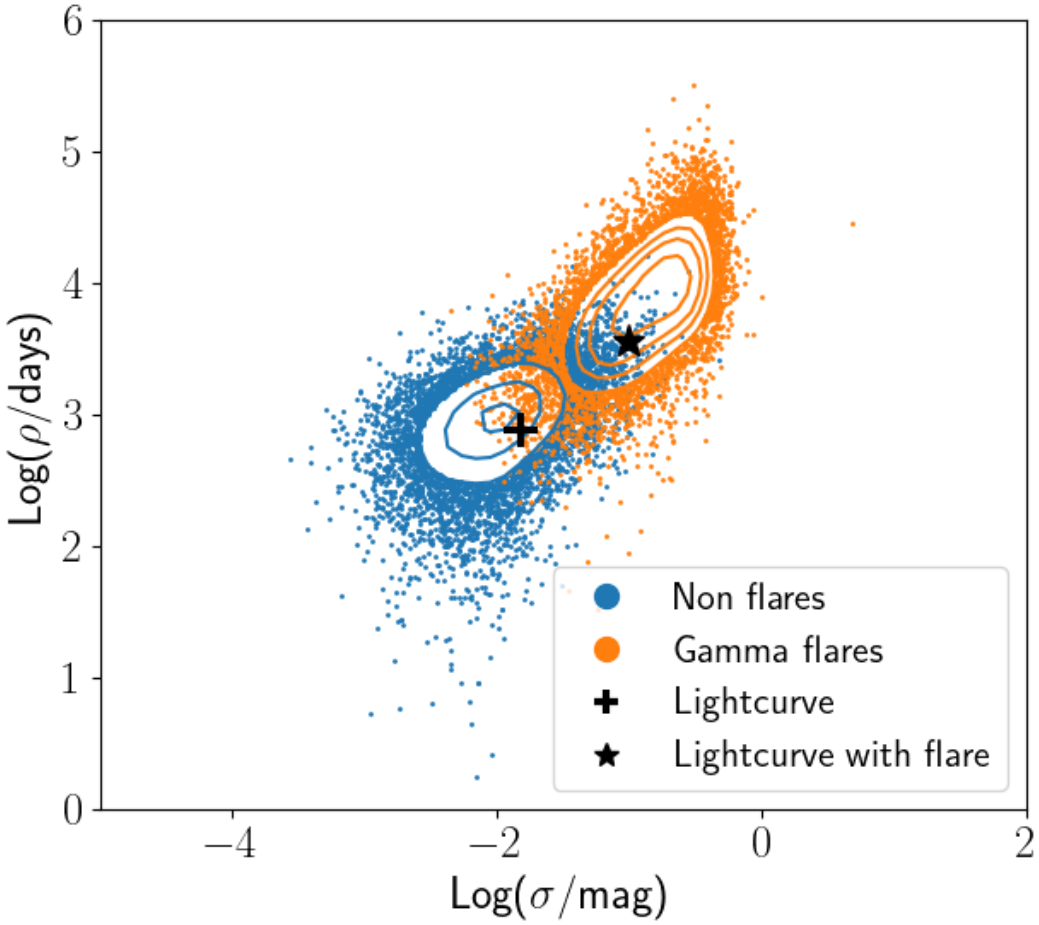

Figure 3.2: Distributions of flare and non-flare hyperparameters for perfect simulated lightcurves with injected Gaussian flares. It is clear that the kernel hyperparameters of lightcurves containing flares and lightcurves without flares exist in distinct but partially overlapping regions of parameter space. This demonstrates that the GP analysis finds that the covariances of these lightcurves are statistically different. The contours are representative of the density of the data points. The locations of the simulated lightcurves from Fig. 3.1 are shown, demonstrating that simply the injection of a flare can move a lightcurve's hyperparameters from the non-flare distribution (blue points and contours) to the flare distribution (orange points and contours). The contours are representative of the density of the data points.

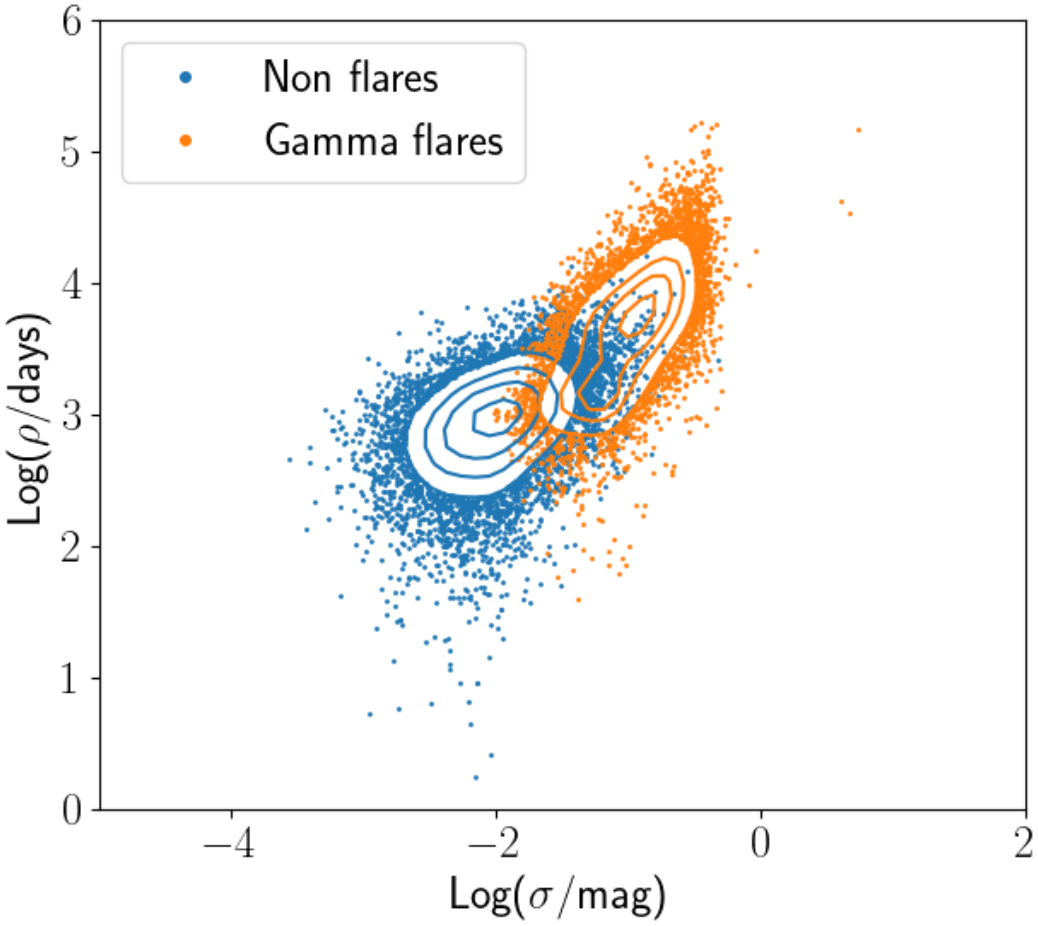

Figure 3.3: Distributions of flare and non-flare hyperparameters for perfect simulated lightcurves with injected gamma flares. Again, it is clear that the kernel hyperparameters of lightcurves containing flares and lightcurves without flares exist in distinct but partially overlapping regions of parameter space. The contours are representative of the density of the data points. Figure from McLaughlin et al. (2024).

3.4.2 Sub-sampled lightcurves

For sub-sampled lightcurves, the distributions of hyperparameters (see Figs. 3.4 and 3.5 for Gaussian and gamma flares, respectively) overlap more than those of the perfect lightcurves. It should be noted, however, that this is partly due to some flares being removed by the sub-sampling, and also due to the GP analysis finding it more difficult to fit lightcurves with irregular cadence and gaps in the data; this is a result of the sparsity of data causing the maximum likelihood estimate of the hyperparameters to show more scatter around the true values. In cases where the flare is largely removed by the sub-sampling, it is incorrect to regard them as false negatives, since almost all evidence of a flare has been removed from the lightcurve and it is important to consider it a non-flare. To determine the number of lightcurves in which this is the case, I summed the magnitude values of the flare points that remained post sub-sampling and ignored those lightcurves in which this sum was less than one. Though this choice of one is arbitrary, I investigated changing this cutoff to 0.5 and 1.5 and the results were not materially different. The use of this magnitude threshold is to ensure that lightcurves containing flares of which a significant proportion of the flare has been removed by the sub-sampling are treated as effectively non-flaring lightcurves.

3.4.3 Lightcurves with added outliers

The hyperparameter distributions for the simulated, sub-sampled lightcurves with added outliers are shown in Figs. 3.6 and 3.7 for the Gaussian and gamma flares respectively. There is still significant overlap between the flare and non-flare distributions, and it is clear that the flaring lightcurves tend to show higher values of σ than non-flaring lightcurves. In both the non-flaring and flaring cases, there is a greater variance across the y-axis tending towards smaller values of ρ compared to the previously discussed classes of lightcurve. This larger spread is a result of the addition of the outliers reducing

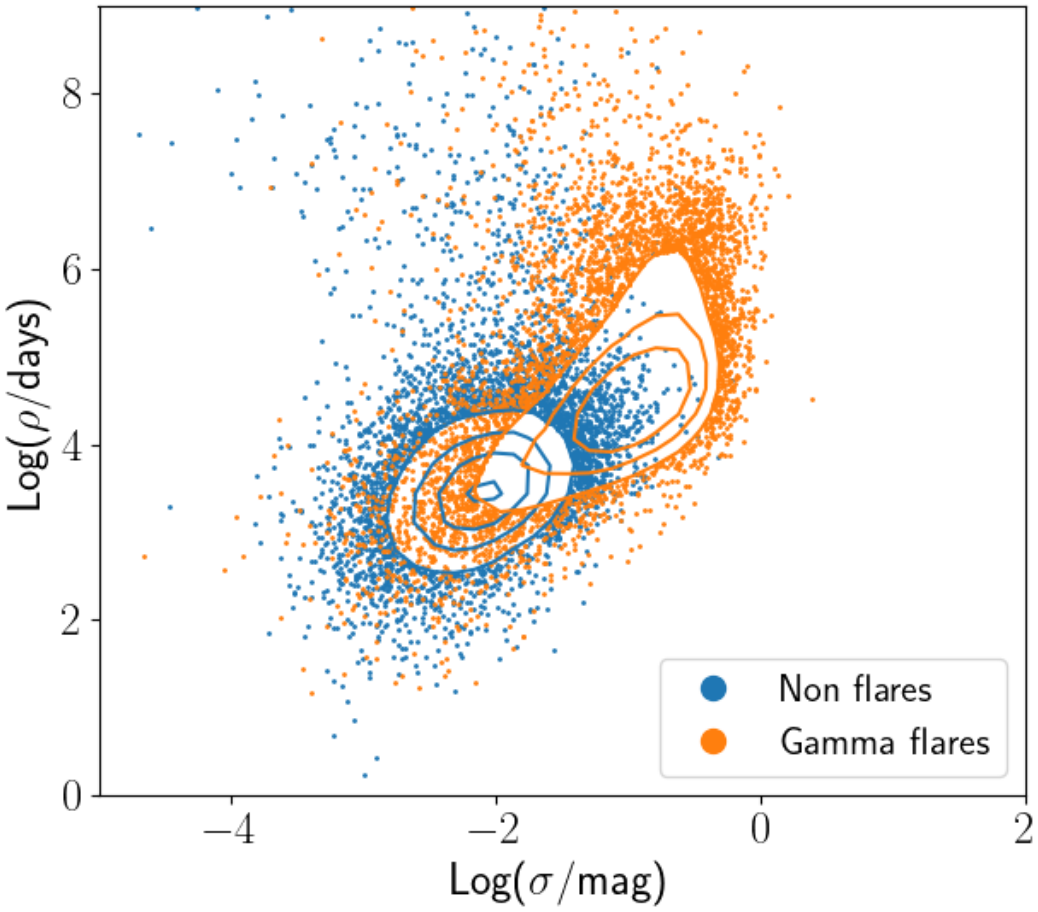

Figure 3.4: Distributions of flare and non-flare hyperparameters for simulated lightcurves with ZTF-like cadence with Gaussian flares. Compared to the hyperparameters of the well-sampled lightcurves, the distributions of lightcurves containing flares and lightcurves without flares are significantly overlapping. The contours are representative of the density of the data points. Figure adapted from McLaughlin et al. (2024).

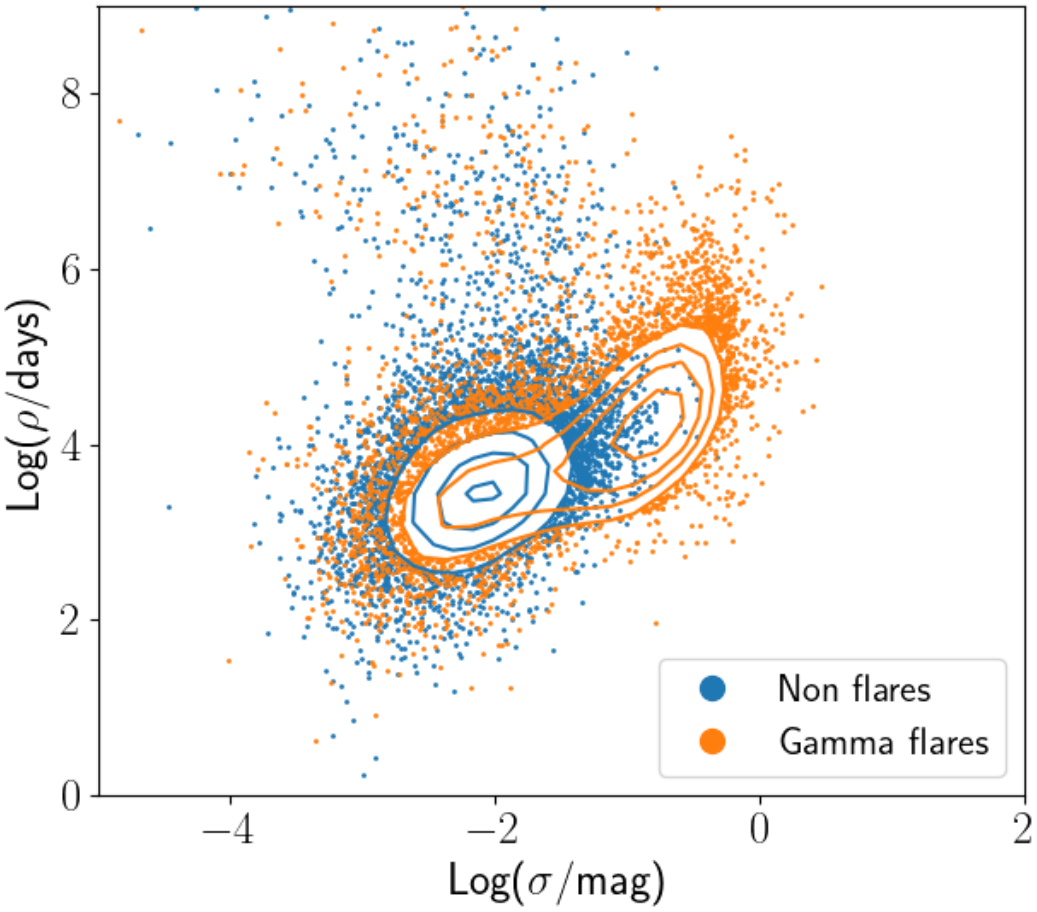

Figure 3.5: Distributions of flare and non-flare hyperparameters for simulated lightcurves with ZTF-like cadence with gamma flares. Again, compared to the hyperparameters of the well-sampled lightcurves, the distributions of lightcurves containing flares and lightcurves without flares are significantly overlapping. The contours are representative of the density of the data points. Figure adapted from McLaughlin et al. (2024).

the timescale of variability. In addition, compared to the previously-discussed classes of lightcurves, the lightcurves with added outliers tend towards greater values of σ. Again, this is a result of the injected outliers increasing the amplitude of variability. Finally, there are no notable differences between the Gaussian and gamma flare distributions.

3.4.4 ZTF lightcurves with injected flares

The hyperparameter distributions for the ZTF lightcurves, including those with injected Gaussian and gamma flares, are shown in Figs. 3.8 and 3.9, respectively. In this case, I have assumed that the prevalence of *real* flares within the ZTF sample is low enough that it is reasonable to label them all as non-flaring for this part of the study. As we shall see in §3.6.2, it is likely that some ZTF lightcurves *are* flaring, but that their numbers are so low that (*i.e.*, $\ll 1$ per cent) that they do not affect how I use these distributions to identify potential flares (see §3.5).

The distributions of (injected) flaring and (assumed) non-flaring ZTF lightcurves most closely resemble those of the lightcurves with added outliers, displaying a larger spread across the y-axis compared to the "perfect" and sub-sampled lightcurves. However, while some overlap between hyperparameters for flaring and non-flaring lightcurves is clearly present, it is somewhat less than that seen in the case of the lightcurves with added outliers. It is difficult to know for certain why this is the case; it may be due to the fact that I have created my simulated lightcurves in a way that makes them more variable than the ZTF (the median σ value of my simulated lightcurves is a factor of 3 greater than that of the ZTF lightcurves, because the structure function values – SF^{∞} and τ – used to simulate my lightcurves were taken from MacLeod et al. 2010 who modelled a sample of quasars rather than AGN). This means that – all else being equal – the injection of a flare into a simulated lightcurve has a smaller impact on its overall variability than the injection of the same flare into a ZTF lightcurve.

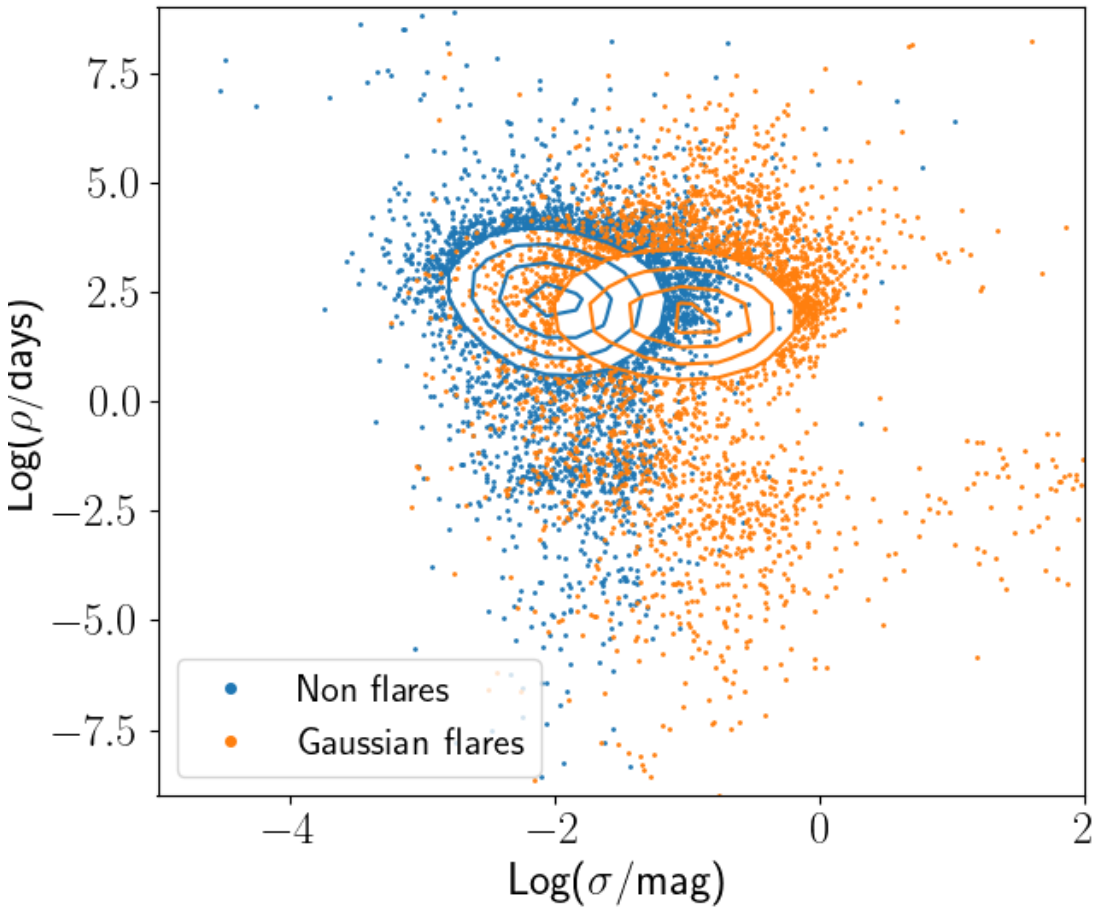

Figure 3.6: Distributions of flare and non-flare hyperparameters for simulated light-curves with ZTF-like cadence with added outliers and Gaussian flares. These distributions are significantly overlapping as in Fig. 3.4, and also the values of ρ have been reduced. This is likely due to the injection of outliers reducing the timescale of variability calculated by the GP. The contours are representative of the density of the data points. Note that the y-axis scaling is different to the previous figures to include all of the data points. Figure from McLaughlin et al. (2024).

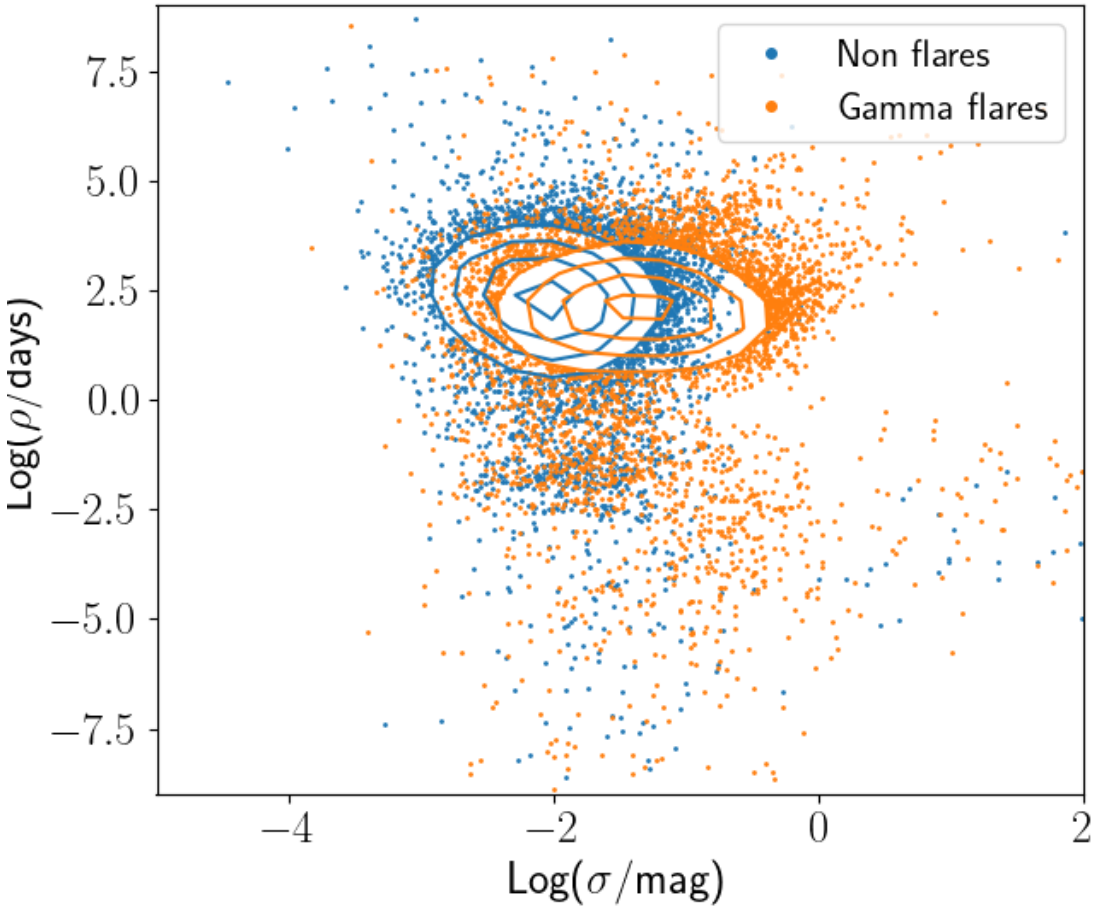

Figure 3.7: Distributions of flare and non-flare hyperparameters for simulated light-curves with ZTF-like cadence with added outliers and gamma flares. Again, these distributions are significantly overlapping as in Fig. 3.7, and also the values of ρ have been reduced. This is likely due to the injection of outliers reducing the timescale of variability calculated by the GP. The contours are representative of the density of the data points. Figure from McLaughlin et al. (2024).

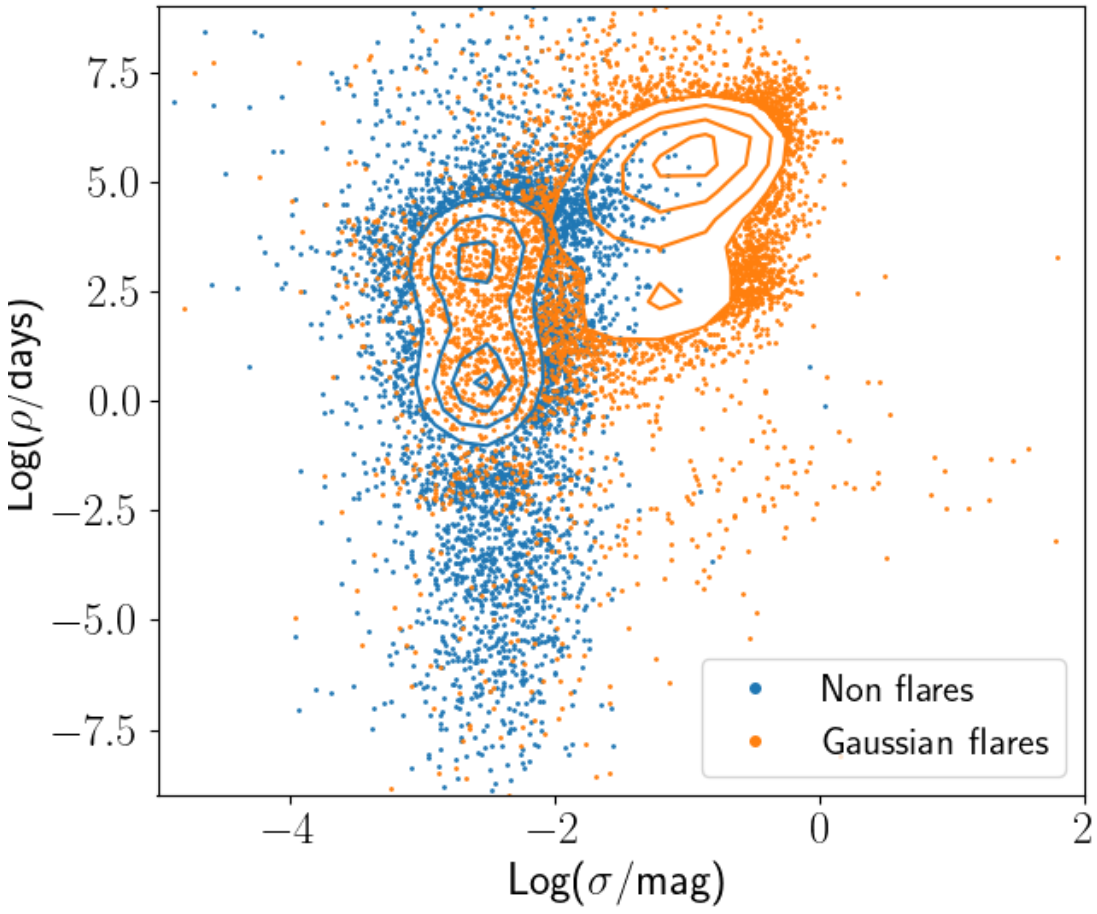

Figure 3.8: Distributions of hyperparameters for ZTF lightcurves with injected Gaussian flares. The contours are representative of the density of the data points. Compared with Figs. 3.2, 3.4 and 3.6, there is a much greater spread of ρ values although there is still significant overlap between the distributions of lightcurves containing flares and lightcurves without flares. Figure from McLaughlin et al. (2024).

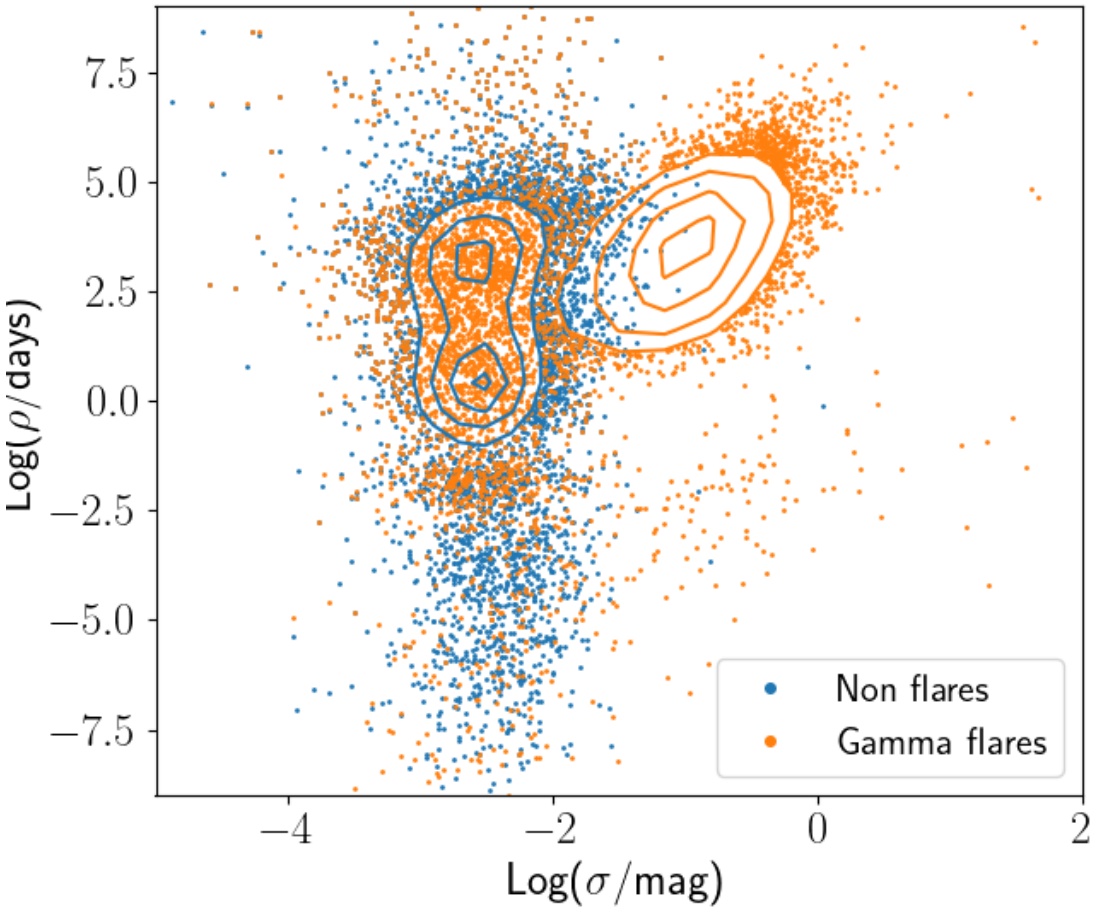

Figure 3.9: Distributions of hyperparameters for ZTF lightcurves with injected gamma flares. The contours are representative of the density of the data points. Compared with Figs. 3.3, 3.5 and 3.7, there is a much greater spread of ρ values although there is still significant overlap between the distributions of lightcurves containing flares and lightcurves without flares. Figure from McLaughlin et al. (2024).

3.5 Using GPs to identify flaring lightcurves

I have demonstrated that the kernel hyperparameters of flaring and non-flaring lightcurves reside in different regions of parameter space, which overlap to a greater or lesser extent, depending on the class of lightcurve (*i.e.,* perfect, sub-sampled, etc.). This therefore opens up the prospect of using GP analysis to identify flaring lightcurves. Given that the hyperparameter distributions overlap, however, the best one can do is to assign a probability that a lightcurve contains a flare (*i.e.,* $\theta = 1$) or not (*i.e.,* $\theta = 0$), where θ acts as a binarized representation of whether a lightcurve is flaring. To achieve this, once the kernel hyperparameters had been optimized for each lightcurve, I used Bayesian hypothesis testing to determine the probability of a new lightcurve belonging to either the flare ($\theta = 1$) or non-flare ($\theta = 0$) populations. In this method, the posterior probability of a lightcurve containing a flare or not can be described (according to Bayes' theorem) as:

$$P(\sigma, \rho, \theta|y) \propto P(y|\sigma, \rho)P(\sigma, \rho|\theta)P(\theta), \tag{3.5}$$

where σ and ρ are the kernel hyperparameters, y is the data, and $P(\theta)$ is defined as a "hyper prior". In principle, $P(\theta = 1)$ could be regarded as representing my *a priori* belief of a given lightcurve containing a flare, and we are thus free to choose a value we see fit. For example, $P(\theta = 1) = 0.5$ would imply a prior belief that a given lightcurve has a 50:50 chance of containing a flare. In practice, however, I used the results of the analysis of my simulated lightcurves to inform us of what value of $P(\theta = 1)$ gives the best compromise between numbers of false and true positives. I found, for example, that adopting $P(\theta = 1) = 0.001$ (which may be considered to be a reasonable estimate of the frequency of flares in AGN lightcurves, for example, see: MacLeod et al. 2012; Lawrence et al. 2016) led to a large number of simulated flares to be missed. I also investigated using $P(\theta = 1) = 0.5, 0.1, 0.01$, and found $P(\theta = 1) = 0.5$ resulted in a large false

positive rate, while $P(\theta = 1) = 0.01$ suffered from a low true positive rate. Based on these results, I chose a value of $P(\theta = 1) = 0.1$.

I used Markov chain Monte Carlo sampling methods to sample the posterior probability distribution. I first perform an initial GP analysis of the lightcurve I wish to classify. This gives us the optimised values for σ_c and ρ_c for this lightcurve (where the subscript c is used to denote the lightcurve I wish to classify). Next, I assign a value of 1 to θ_c if $P(\theta_c = 1|\sigma_c, \rho_c) > P(\theta_c = 0|\sigma_c, \rho_c)$ based on the distributions of σ and ρ obtained from my GP analysis, and 0 otherwise. Next, I calculate the posterior probability using:

- the appropriate value for $P(\theta_c)$ (*i.e.,* 0.1 or 0.9, depending on the value of θ_c);
- the value of $P(\sigma_c, \rho_c|\theta_c)$ based on a 2D-Gaussian approximation of the hyperparameter distributions obtained from the Gaussian Process analysis of the corresponding *non-flaring* lightcurve class.[6] It is important to note that I do not use the hyperparameter distributions obtained for simulated *flaring* AGN as a prior for that class. Instead, I use a 2D Gaussian that encompasses a much larger region of parameter space than both the flaring and non-flaring lightcurves (*i.e.,* it is a non-informative prior). This is done to ensure that I am making minimal assumptions regarding the properties of the flares since I do not know whether my simulated flares do, indeed, fully represent the true diversity of real flares[7]; and
- the likelihood $P(y|\sigma_c, \rho_c)$, which I obtain from the Gaussian Process fit of the lightcurve I am classifying.

For the next step in the MCMC I randomly propose (with equal chance of choosing 0 or 1) a new value of θ_c ($= \theta'_c$), and recalculate the posterior using θ'_c. I accept this

[6]By "corresponding", I mean that if I am attempting to classify a ZTF lightcurve then I used the hyperparameter distribution I obtained by analysing my sample of ZTF lightcurves (which I assume to be dominated by non-flaring lightcurves), approximated using multiple 2D Gaussians.

[7]In this regard, my analysis is agnostic to how I simulate the flares (see §3.3.2) since the analysis is only ascertaining whether the GP parameters of a given lightcurve deviate significantly from those of the non-flaring population and therefore likely to contain a flare, irrespective of its properties.

value of θ_c' with probability:

$$\min\left(1, \frac{P(\sigma_c, \rho_c, \theta_c'|y)}{P(\sigma_c, \rho_c, \theta_c|y)}\right) \tag{3.6}$$

(*i.e.,* I always accept if the proposed posterior probability is greater than the current posterior probability, but accept with a probability equal to the ratio of the two posterior probabilities if $P(\sigma_c, \rho_c, \theta'|y) < P(\sigma_c, \rho_c, \theta|y)$). Next, I simultaneously propose new values of ρ_c and σ_c (*i.e.,* ρ_c', σ_c') and recalculate the posterior probability using these new parameters, which includes calculating $P(y|\sigma_c', \rho_c')$ using a GP. Again, I accept these values of σ_c' and ρ_c' with probability:

$$\min\left(1, \frac{P(\sigma_c', \rho_c', \theta_c|y)}{P(\sigma_c, \rho_c, \theta_c|y)}\right). \tag{3.7}$$

Using Markov chain Monte Carlo (MCMC), I repeat the process of proposing (and, when appropriate, accepting) new θ and (σ_c, ρ_c) values in order to sample the posterior parameter space. I chose 12 000 steps with a burn-in of 2000 as this was sufficient for the trace to converge.

Each time I propose a new value of θ_c, I add the accepted value (whether the newly-proposed value, or the old one) into a 1D array; this results in a vector of length 10 000 (excluding the burn-in) of zeroes and ones corresponding to the accepted θ_c value. The relative numbers of zeros and ones give the relative probabilities of the lightcurve being labelled as a flare or non-flare. As such, the final probability of the lightcurve containing a flare, $P_{\rm Flare}$, is thus given by the sum of this vector, divided by its length (*i.e.,* the mean). Guided by the results from analysing my simulated data, I find that using a cutoff probability of 0.1 to define a flare gave the best compromise between true and false positives. While this may seem low, I find that most non-flaring lightcurves have extremely low flare probabilities.

3.6 Results

In this section I first present the retrieval rates for classifying flares and non-flares in the case of each of my classes of simulated lightcurves (§3.6.1). For each class of lightcurve, true positive rates were calculated as the fraction of known flares with $P_{\mathrm{Flare}} > 0.1$. Similarly, the true negative rate is the fraction of control lightcurves with $P_{\mathrm{Flare}} < 0.1$. Afterwards, I analyse all of my unadulterated (*i.e.,* without injected flares) ZTF lightcurves to see which, if any, are flagged as containing flares; the results of this "blind" analysis are presented in §3.6.2.

3.6.1 Retrieval rates for simulated lightcurves

The confusion matrices for my perfect simulated lightcurves with injected Gaussian and gamma flares are shown in Fig. 3.10. The true positive rates are similar for both types of flare (91 and 92 per cent respectively), but the false positive rate is slightly higher for gamma flares (11 per cent compared to 7 per cent). Though the simulated flare parameters must be selected arbitrarily due to the rarity of AGN flares, I investigated the change in retrieval rates of simulated flares with specific properties. I found that the retrieval rates of the GP analysis decrease as the duration of the flare increases, and the amplitude of the flare decreases. For example, 95 per cent of flares with magnitude greater than 1.5 and 99 per cent of flares with duration less than 500 days are successfully detected by the GP analysis. Fig. 3.11 demonstrates the retrieval rate as a function of simulated flare amplitude, showing that the lowest amplitude flares are most difficult to detect by the GP analysis. The GP analysis is clearly struggling to distinguish simulated flares with an amplitude of one magnitude or less from the underlying variability.

As shown in Fig. 3.12, the retrieval rate reduces significantly when the lightcurves were sub-sampled to match ZTF cadence. There is little difference between the true positive and false positive rates of Gaussian and gamma flares, with the Gaussian flares

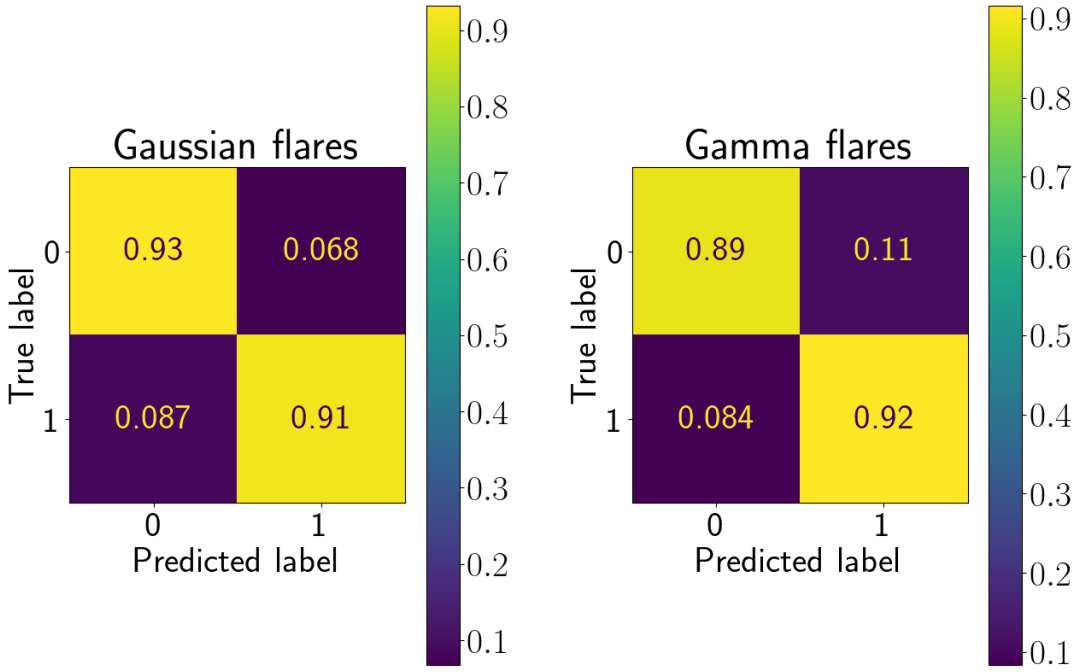

Figure 3.10: Confusion matrices for simulated lightcurves with a sampling of 10 days, in the case of injected Gaussian flares (left) and injected gamma flares (right). Note that zero and one refer to "non-flare" and "flare" respectively. The true positive rate is similar, shown in the bottom right panel (91 and 92 per cent respectively), but the false positive rate shown in the top right panel is slightly higher for gamma flares (11 per cent compared to 7 per cent). Figure from McLaughlin et al. (2024).

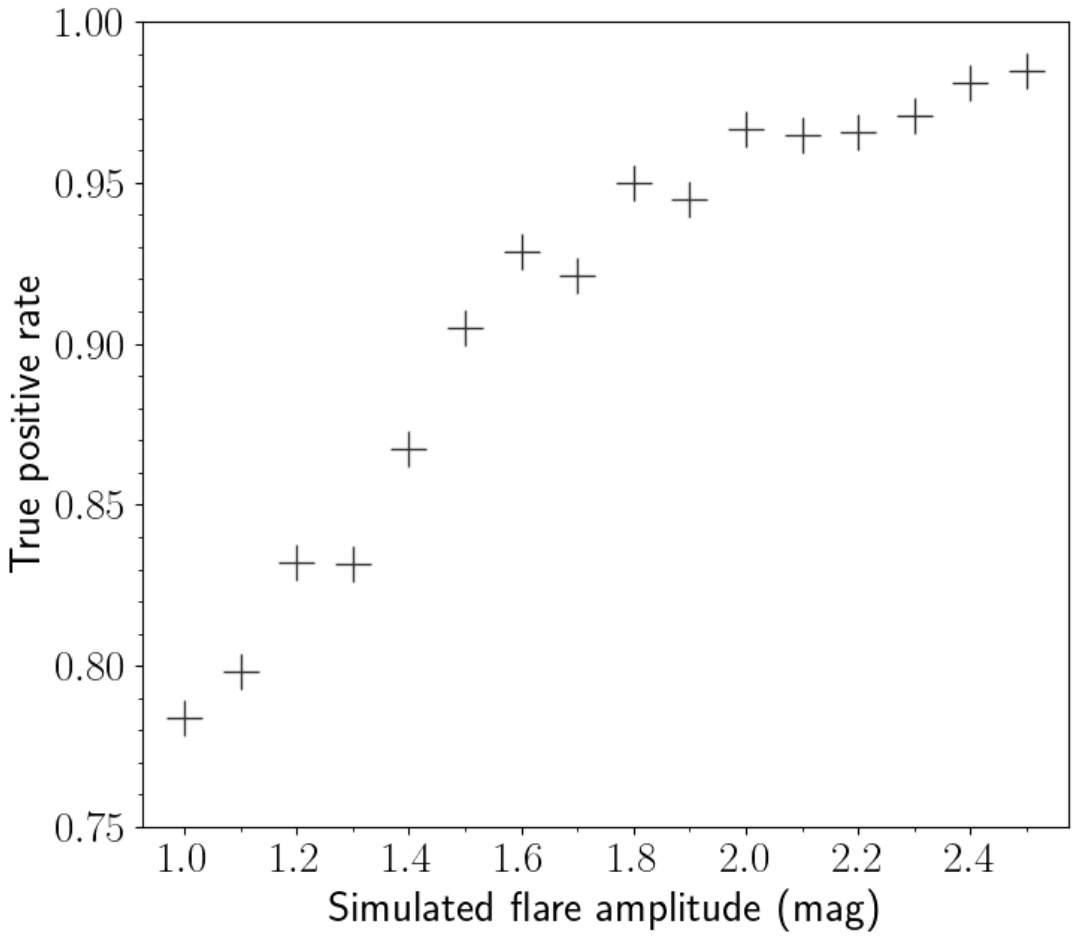

Figure 3.11: Retrieval rate (*i.e.,* true positive rate) of the GP analysis as a function of simulated flare amplitude for the perfect, simulated lightcurves. It is clear that the retrieval rate of the GP analysis decreases as the simulated flare amplitude decreases, since the hyperparameters of these lightcurves are more likely to reside in the overlapping region between flares and non-flares. Figure from McLaughlin et al. (2024).

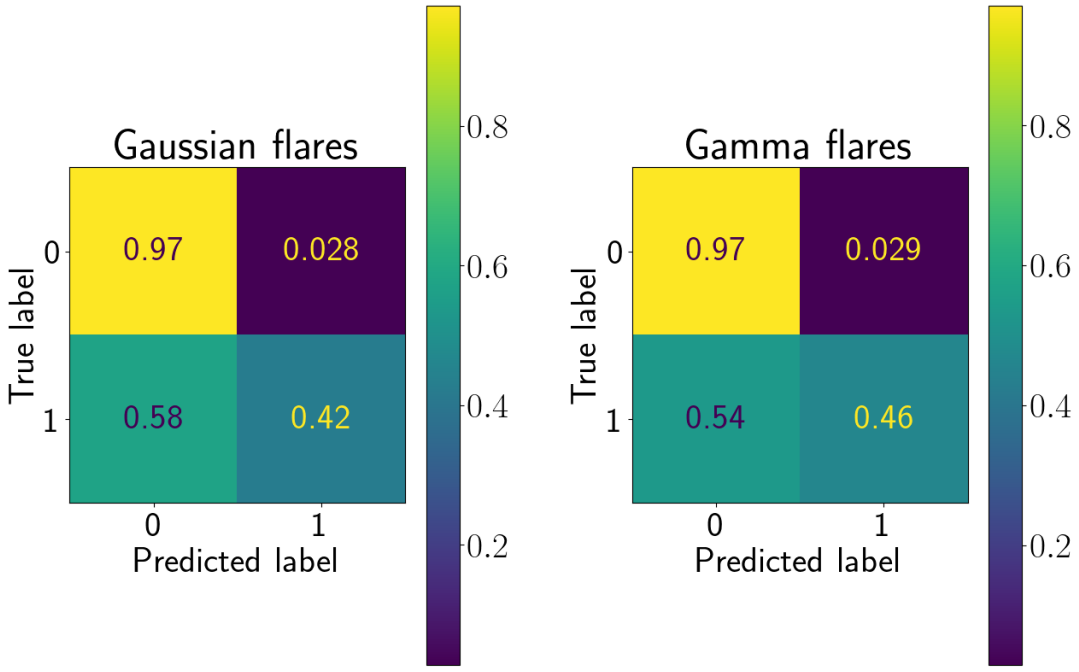

Figure 3.12: Confusion matrices for sub-sampled simulated lightcurves, in the case of injected Gaussian flares (left) and injected gamma flares (right). Note that zero and one refer to "non-flare" and "flare" respectively. The bottom right panel shows the true positive rate, which is 42 and 46 per cent for Gaussian and gamma flares respectively. The top right panel shows the false positive rate, which is 2.8 per cent for Gaussian flares and 2.9 per cent for gamma flares. Figure from McLaughlin et al. (2024).

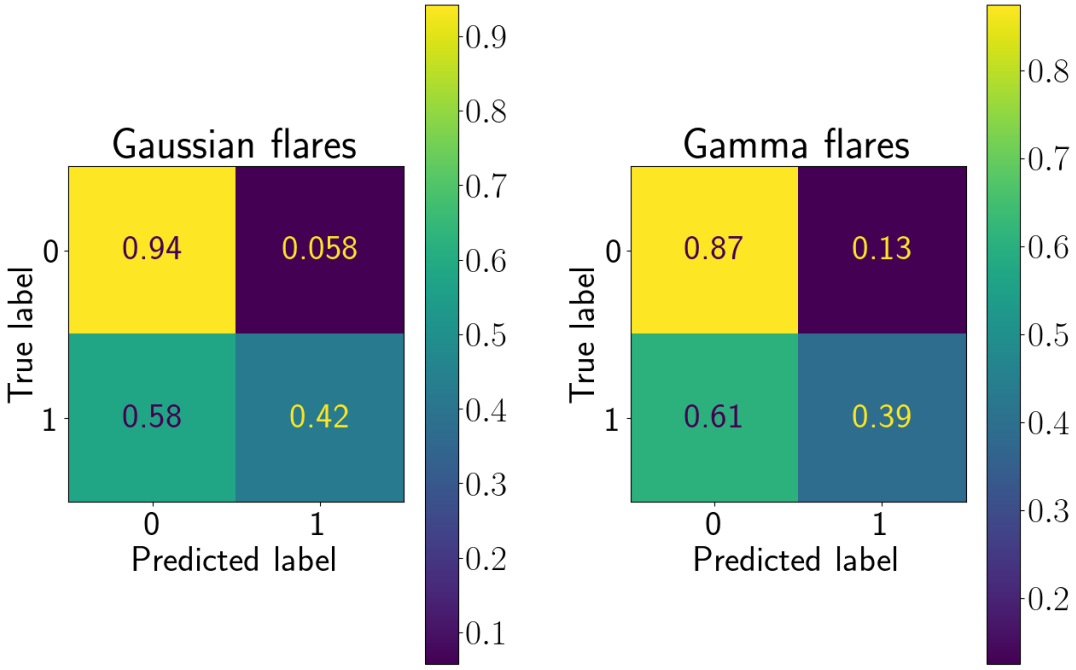

Figure 3.13: Confusion matrices for simulated and sub-sampled lightcurves with added outliers, in the case of injected Gaussian flares (left) and injected gamma flares (right). Note that zero and one refer to "non-flare" and "flare" respectively. The bottom right panel shows the true positive rate, which is 42 and 39 per cent for Gaussian and gamma flares respectively. The top right panel shows the false positive rate, which is 5.8 per cent for Gaussian flares and 13 per cent for gamma flares. Figure from McLaughlin et al. (2024).

having a true positive rate of 42 per cent and a false positive rate of 2.8 per cent. The gamma flares have a true positive rate of 46 per cent and a false positive rate of 2.9 per cent. As such, while the purity of the retrieved sample is relatively high (*i.e.,* low false positives), the completeness is low (*i.e.,* less than 50 per cent).

Fig. 3.13 shows the confusion matrices for simulated lightcurves with added outliers in the case of both Gaussian and gamma flares. This shows similar results as those found for sub-sampled lightcurves without outliers. In the case of Gaussian flares, the GP analysis is able to classify 94 per cent of non-flaring lightcurves correctly (a 6 per cent

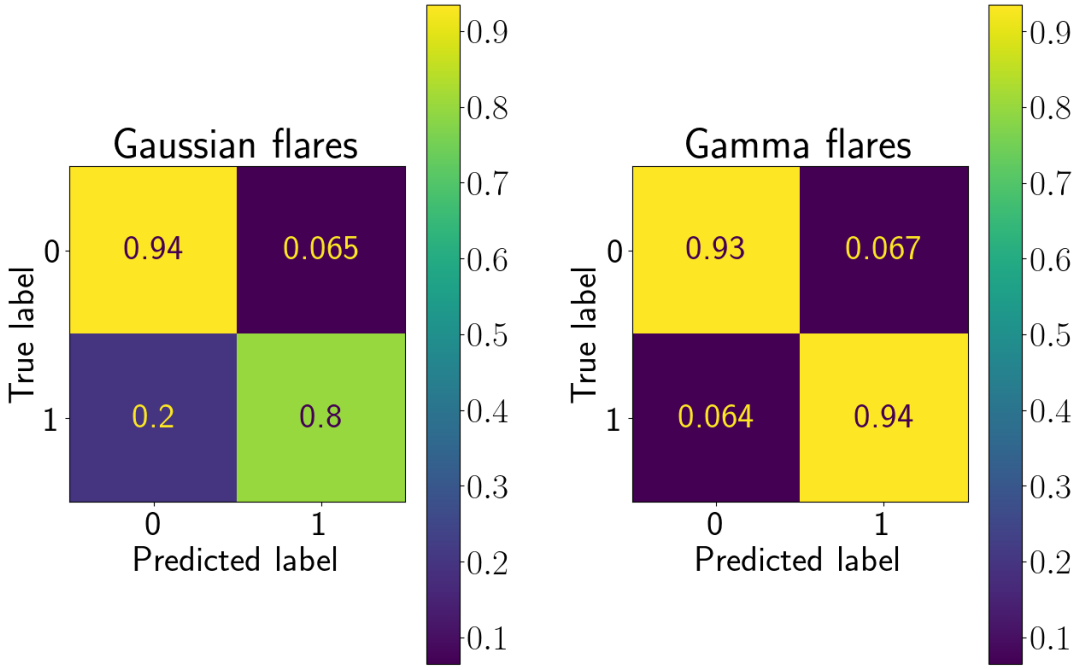

Figure 3.14: Confusion matrices for ZTF lightcurves with injected flares, in the case of injected Gaussian flares (left) and injected gamma flares (right). Note that zero and one refer to "non-flare" and "flare" respectively. The bottom right panel shows the true positive rate, which is 80 and 94 per cent for Gaussian and gamma flares respectively. The top right panel shows the false positive rate, which is 6.5 per cent for Gaussian flares and 6.7 per cent for gamma flares. Figure from McLaughlin et al. (2024).

false positive rate), but only 42 per cent of flaring lightcurves were classified correctly. The true positive rate of the gamma flares is slightly lower at 39 per cent with a higher false positive rate of 13 per cent.

Despite the GP analysis struggling to identify flares in sub-sampled lightcurves with or without systematic outliers, better results are achieved with the ZTF lightcurves with injected flares. For these lightcurves, the GP analysis was more effective at classifying flares and non-flares than with the simulated sub-sampled lightcurves; remarkably, this is in spite of using the ZTF cadence to sub-sample the simulated lightcurves. The confusion

matrices for ZTF lightcurves with both Gaussian and gamma injected flares are shown in Fig. 3.14. In the case of injected Gaussian flares, the GP analysis has an 80 per cent true positive rate and a 6.5 per cent false positive rate, compared with injected gamma flares with true and false positive rates of 94 per cent and 7 per cent respectively.

3.6.2 ZTF flares

The final step I took in testing the efficacy of GPs in detecting AGN flares was to perform the analysis on unadulterated AGN lightcurves. For this, 9035 ZTF lightcurves (§3.3.5) were analysed using a GP to determine if any would be flagged as containing flares or extreme variability.

I initially invoked a probabilistic cut-off of 0.1 for a lightcurve to be classified as a flare by the Gaussian Process. This cutoff resulted in a total of 257 flare candidates. On inspection, I found that a considerable number of these candidates were poorly sampled or had large gaps in their lightcurves. For example, 117 lightcurves contained fewer than 30 data points and 154 had gaps in their lightcurves lasting over 150 days. It is therefore feasible that some of these lightcurves may, indeed, contain (unsampled) flares, but I do not select them for visualisation purposes.[8]. It should also be noted that there were a number of lightcurves (117) that were assigned a high probability of containing a flare but were located in the far-left-hand side of the hyperparameter distribution and these were removed from selection due to having low values of σ and hence low amplitude values. Applying the above selections simultaneously and ignoring lightcurves with poor GP fits by visual inspection resulted in a sample of 27 flare candidates, which are shown as orange points in Fig. 3.15 and whose lightcurves are shown in Figs. 3.16 and 3.18-3.23.

The lightcurves of four examples chosen from the 27 identified flare candidates are shown in Fig. 3.16. In each of these four plots, I also include the lightcurves of 100 randomly-selected AGN that were not flagged as containing flares by the analysis. These

[8]These, together with all my labelled lightcurves, are available upon request.

non-flare lightcurves were normalized by calculating the "relative" magnitude by subtracting the first magnitude value from each magnitude value in each lightcurve (meaning that the first magnitude value of every non-flare lightcurve is zero and therefore each consecutive data point represents the relative deviation from this zero). Then I added this relative magnitude to the mean magnitude value of the flare lightcurve. This results in a visualisation of the extent to which the non-flare lightcurves deviate from that of the flare candidate. Normalisation of the lightcurves was performed for the purpose of comparison against a common baseline and for ease of visualisation. By comparing the flare candidates to these non-flaring lightcurves, it is clear that the former show extreme variability. Most notably, they display longer-term, more systematic departures from their starting point relative to the comparison (non-flaring) lightcurves. The full sample of flare candidates is shown in §3.9. Note that my analysis is only able to detect extreme variability, and hence classifies "flares" as objects that are becoming either brighter or fainter, rather than just brighter. This is not necessarily a drawback, since if the mechanism behind AGN flares is caused by changes in accretion state, then GPs may be able to detect changing-look AGN, which can both rapidly brighten or dim as their broad emission lines appear or disappear (*e.g.,* LaMassa et al., 2015; Gezari et al., 2017; Yang et al., 2023).

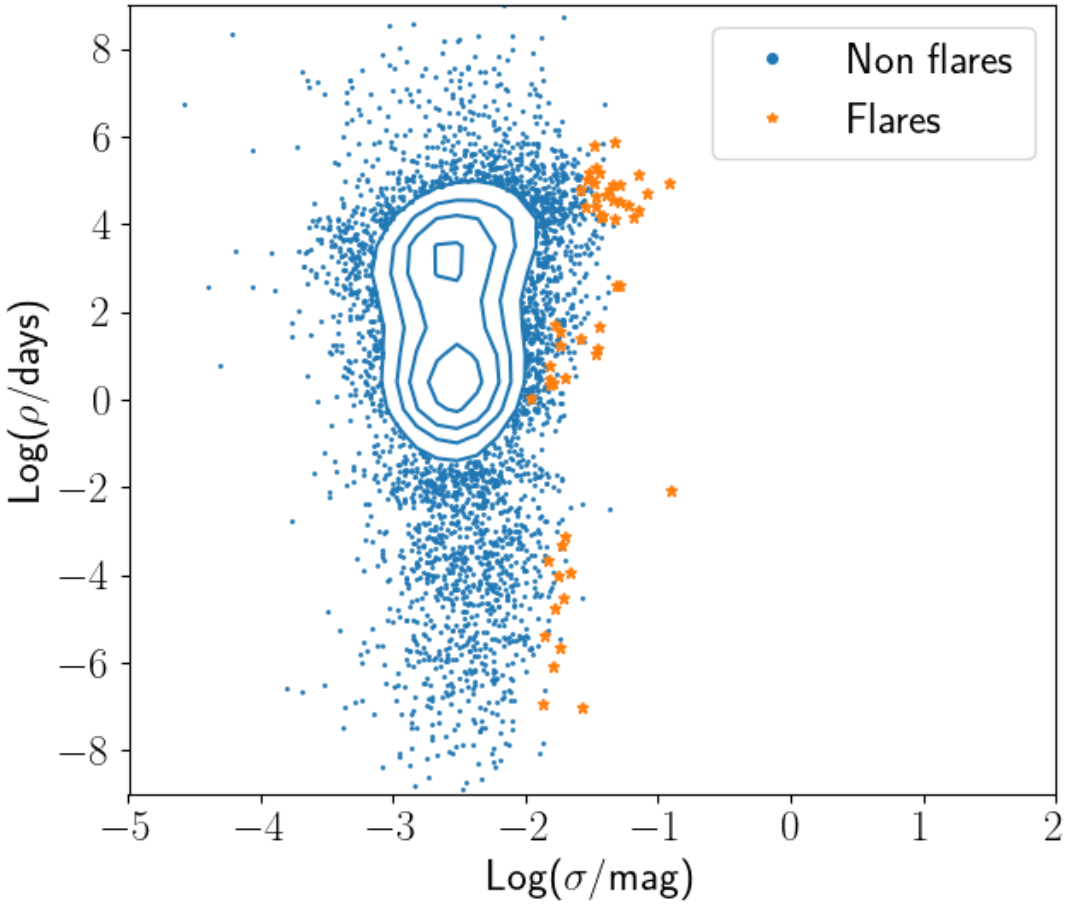

Figure 3.15: Distributions of hyperparameters for real ZTF lightcurves of Type 1 AGN. Lightcurves with a posterior probability greater than 0.1, the number of data points greater than 30, the maximum spacing between consecutive data points less than 150 days, and a sigma value of greater than -2 are shown in orange. This is the resulting flare candidate sample, before visual inspection to remove spurious GP fits. Figure from McLaughlin et al. (2024).

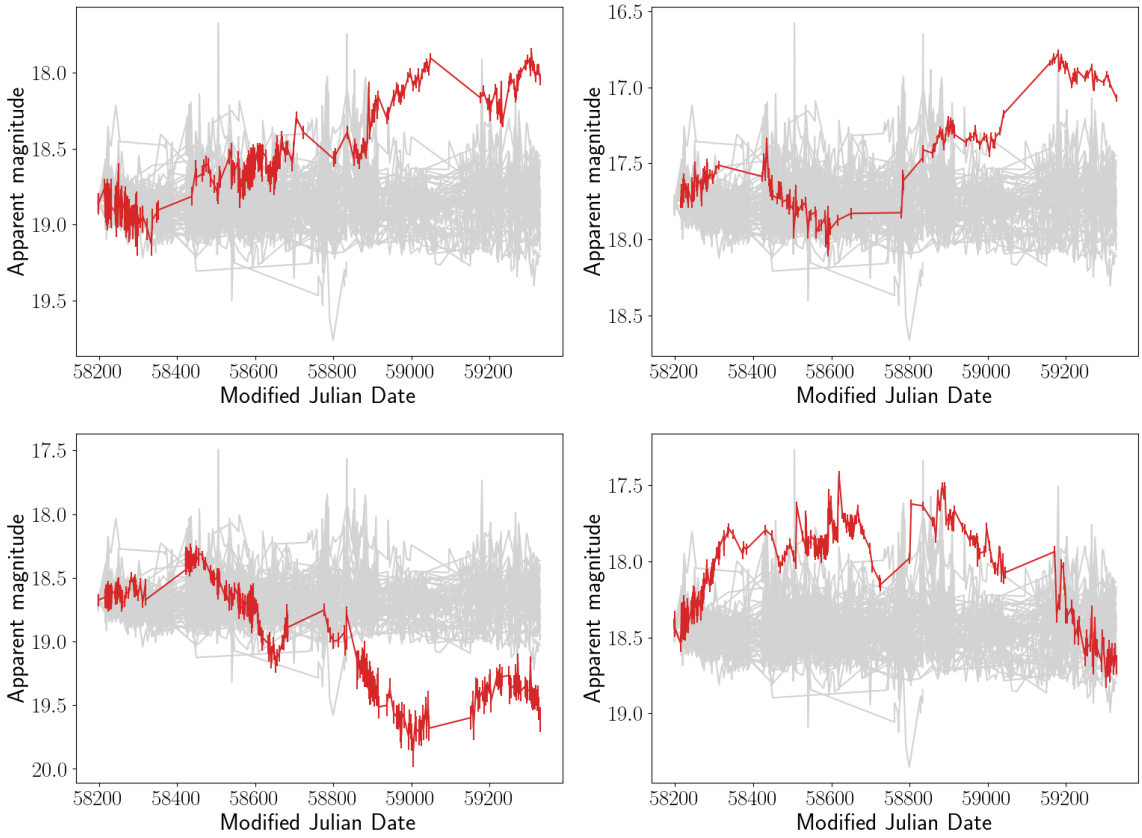

Figure 3.16: Four examples of ZTF lightcurves of flare candidates identified by the GP analysis. The red line shows the lightcurve of the flare candidate and the grey curves are a randomly-sampled selection of 100 lightcurves that were not flagged as flares by the GP analysis, demonstrating that they show extreme variability compared to the rest of the population. These lightcurves have been normalized for ease of visualization (see §3.6.2). The full sample of lightcurves is shown in the appendix. Figure adapted from McLaughlin et al. (2024).

3.7 Discussion

In §3.1 I described a specific problem associated with searching for flares in AGN lightcurves: namely, how does one detect the presence of a transient signal in data that is already intrinsically and stochastically variable? To solve this, it is necessary to quantify what constitutes a significant departure from this baseline variability in a statistical way. In this chapter, I have undertaken a feasibility study to determine whether Gaussian Processes (GPs) are an effective means to achieve this.

I find that GPs have the potential to correctly classify flaring and non-flaring simulated lightcurves with a high success rate, with regularly-sampled flare lightcurves being classified with a true positive rate of around 90 per cent for both Gaussian and gamma flares (see §3.6.1). However, in the case of simulated lightcurves that have been subsampled to mirror the cadence of real ZTF AGN lightcurves, this rate drops to around 40 to 45, depending on whether the injected flare is modelled as a Gaussian or gamma function (see §3.6.1). Similarly, the lightcurves with added outliers resulted in comparably low true positive rates (around 40 per cent; see §3.6.1). Despite this, when real ZTF lightcurves were injected with flares, the GP analysis successfully classified between 80-94 per cent per cent of the flaring lightcurves with false positive rates as low as 6.5 per cent (see §3.6.2). This false-positive rate is extremely promising, as when dealing with large amounts of data it is arguably far more important to have a low false-positive rate than a high true-positive rate, to ensure a high purity of the sample. To put this into context, I will demonstrate the expected purity of a sample of flare candidates selected using this method. Using the retrieval rate of the ZTF lightcurves with injected Gaussian flares (Fig. 3.14), the true positive rate is 80 per cent. This means that in a sample of 100 000 AGN lightcurves, where the true incidence of flares is 1 in 10 000 (*e.g.,* van Velzen, 2018), this technique will detect 8 of the total 10 real flares. The false positive rate in this case is 6.5 per cent, meaning that out of the total 100 000 lightcurves, 6500

will be false positives. With this in mind, it can be seen that although this GP analysis produces a contaminated sample of flare candidates, the considerably lower false positive rate means that it will be manageable to sift the real flare candidates using additional classification techniques; these techniques may include human inspection, citizen science, or the integrated use of human and machine classification (*e.g.,* Mantha et al., 2024). This is discussed further in Chapter 6.

It would be insightful to be able to place the above retrieval and contamination rates in the context of other methods of finding nuclear flaring events. However, with most studies focusing – quite reasonably – on the identification of new flares, rather than how many they may have missed, such success rates are difficult to quantify.[9] My results show that while GPs are not broadly robust against major outliers, they are still able to perform well when handling real data. It also suggests that my simulated outliers were "pessimistic", in that they gave the GP analysis a more difficult job than the real ZTF data. Furthermore, the retrieval rates for gamma and Gaussian flares are comparable, suggesting that the GP analysis is largely unaffected by the shape of the flare.

When I applied my GP analysis to 9035 real ZTF lightcurves of type 1 AGN, 27 flare candidates were identified (see §3.6.2 and Appendix 3.9). These lightcurves exhibit extreme variability when compared to 100 randomly sampled lightcurves that had not been flagged by the GP as flaring.

It is tempting to take the false-negative rates of the sub-sampled flares and the ZTF injected flares to estimate the number of real flares that I could be missing. However, since those false negative rates are based on simulated data, I cannot know what the actual false negative rates are for real flares.

I have shown that GPs are an effective way to detect extreme variability in simulated and real AGN lightcurves, especially in high-cadence data sets. In this chapter, whilst

[9]The systematic comparison of different methods of finding flares is beyond the scope of this thesis, though I discuss the possibilities of future work in this area in Chapter 6.

I have demonstrated this in the r-band only, it would be possible to modify the GP analysis to account for multiple bands. The use of GP analysis in lightcurve classification is not without caveats, however, as there are a number of limitations. As I have shown, a GP is not robust against extreme outliers. In addition, GPs optimise the kernel hyperparameters across the whole lightcurve, which favours the detection of longer-duration, larger-amplitude flares (and especially those that span a significant fraction of the lightcurve). To investigate the impact of these "average" hyperparameters, I sliced the sub-sampled, simulated flare lightcurves so that they contained only the flaring region of the lightcurve and repeated my analysis. The results are shown in Fig. 3.17. This shows that if a GP is somehow able to simultaneously "focus" on subsections of a lightcurve it would have a much higher success rate in terms of distinguishing between flares and non-flares. This demonstrates that GPs could be even more effective at flare classification if it were able to calculate a lightcurve's hyperparameters in a more localised way.

Furthermore, whilst they can quantify the probability of a lightcurve containing a flare, the GP analysis performed here cannot specify the location of the flare within the lightcurve. This is not necessarily a pitfall when searching for flares or extreme variability in archival data, but in the era of time-domain astronomy where surveys such as the LSST will detect potentially millions of transient sources per night, it warrants the ability to detect an AGN flare in real-time, ideally before it peaks. Such a requirement clearly demands an alteration of this method to enable the detection of flares as they happen.

These limitations highlighted above motivate the need to build on my techniques with the intention of localising flares within lightcurves. Two possible ways of achieving this are: tracking the posterior flare probability, $P(\sigma, \rho, \theta|y)$ as a function of time whilst feeding the GP new data, or calculating the posterior probability $P(y_{new}|y_{data})$ to determine whether new points in a lightcurve can be described by the current GP regime and flag them as a flare otherwise. However, these methods are potentially computa-

tionally intensive and so it is important to be able to devise a means of flare localisation in an efficient way. This may require more sophisticated techniques such as deep Gaussian Processes (Damianou & Lawrence, 2013) where the choice of kernel function will depend on training data. Other possibilities include change-point detection (Graham et al., 2023) or regime-switching models (Hamilton, 2010).

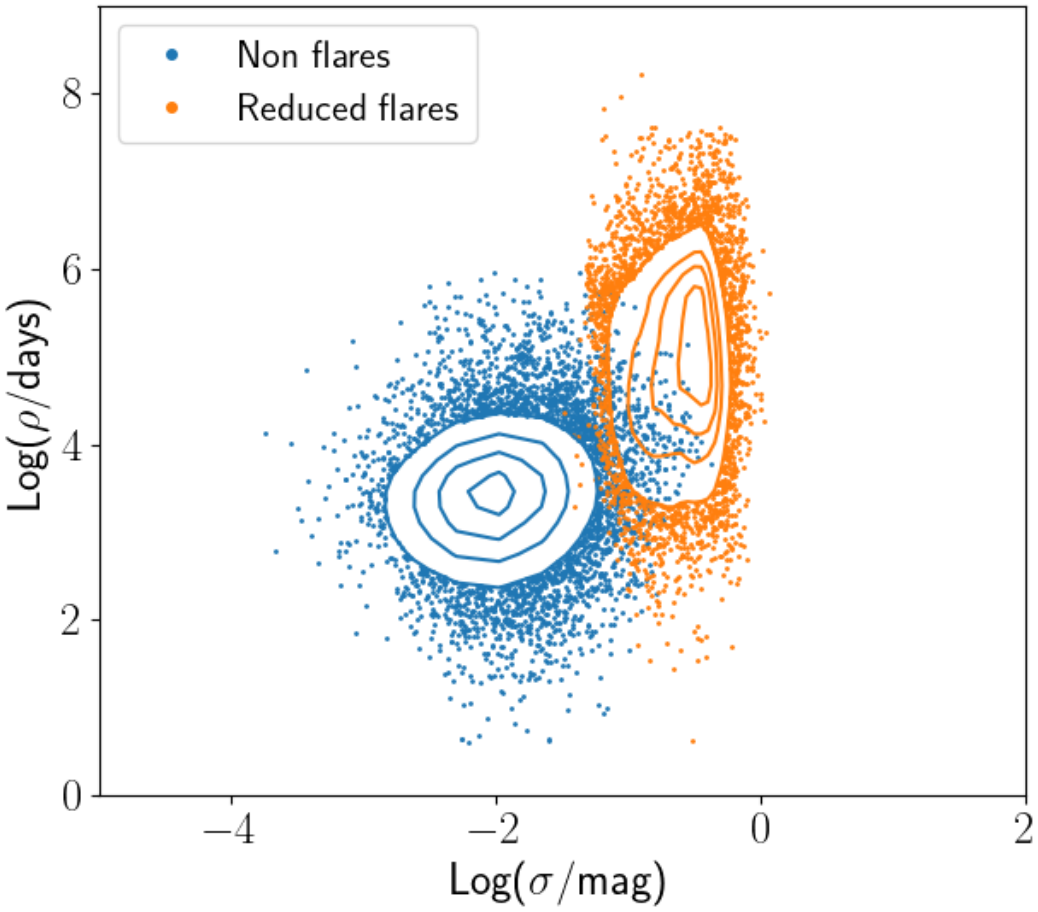

Figure 3.17: Distributions of hyperparameters for sub-sampled simulated lightcurves that have been reduced so that the flare lightcurve contains only the flaring region. The separation between distributions is much greater than in Fig. 3.2, highlighting that if one is able to localise sections of lightcurve it becomes more straightforward to distinguish between flares and non-flares. Figure from McLaughlin et al. (2024).

3.8 Conclusions

I have undertaken a feasibility study to investigate whether Gaussian Processes (GPs) are an effective means of identifying and classifying AGN flares in optical lightcurves. Using a combination of simulated and real AGN lightcurves, I used GP analysis to investigate how the distributions of kernel hyperparameters change after the injection of a simulated AGN flare into a lightcurve (§3.4). I then used these distributions as a basis to classify lightcurves in terms of whether they contain a flare or not, and calculate corresponding flare retrieval rates (§3.5). Throughout, I exploited five different classes of lightcurve, each more representative of real lightcurves than the last:

1. injected flares and simulated with a constant 10-day cadence;
2. injected flares and sub-sampled to match the cadence of real ZTF lightcurves;
3. as 2, but with added outliers;
4. real ZTF with injected flares;
5. real ZTF lightcurves.

In the case of 1, I find that the kernel hyperparameter distributions for flares and non-flares exist in different but partially overlapping regions of parameter space (§3.4). This means that the distributions can never be separated completely; however, since GPs are statistically robust, they can be used to distinguish between the distributions in a probabilistic way. In the cases of lightcurve classes 2-5, however, the hyperparameter distributions for flares and non-flares overlap significantly more than in the case of 1. Despite this, I have been able to demonstrate that GPs can be used as a classification tool for AGN flares, with varying degrees of success. The results can be summarised as follows:

- For simulated flares with a ten-day cadence and with injected flares, I find a true positive rate of 91-92 per cent and a false positive rate of 7-11 per cent for Gaussian and gamma flares respectively.

- When the lightcurves in 1 are sub-sampled to match the cadence of my sample of ZTF lightcurves (§3.3.5), the true positive rates reduce significantly to 42-46 per cent for Gaussian and gamma flares respectively, though the false positive rate is found to be approximately 3 per cent in each case.

- When outliers are added to these simulated lightcurves, the true positive rates remain similar to those found in 2, although the false positive rates increase to 6 and 13 per cent for Gaussian and gamma flares, respectively.

- When my sample of real AGN lightcurves are injected with simulated Gaussian and gamma flares, the results are more promising than in the cases of 2-5. I obtain true positive rates of 80 and 94 per cent for Gaussian and gamma flares, respectively, while the false positive rates remain similar as to that found for class 3 at approximately 7 per cent.

- Finally, I applied my GP analysis to the unadulterated sample of ZTF lightcurves to determine whether any real AGN lightcurves would be flagged as containing flares by my GP analysis. As shown in §3.6.2, the GP analysis classified 27 out of 9035 AGN lightcurves as containing flares or extreme variability. When compared with a randomly-selected sample of 100 lightcurves that were not flagged as flares, they indeed show greater levels of variability, particularly in the form of longer-term, systemic departures from their starting point.

Overall, I have demonstrated that GP analysis can be used to calculate the probability that an incoming AGN lightcurve contains a flare. I find that this is a promising method to detect flares in otherwise variable optical lightcurves, although it can be negatively

affected by extreme outliers and poorly sampled data. In order to keep up with the large amounts of data involved in future surveys such as the LSST (Ivezić et al., 2019), there is a growing requirement to be able to detect AGN flares and transients alike before they peak to enable for rapid follow-up. Therefore, since my GP analysis in this work is able to calculate the probability of an incoming lightcurve containing a flare but not the exact location of the flare within the lightcurve, there is a need to build on this GP technique to be able to localise a flare as it happens. As mentioned in §3.7, these techniques may be computationally intensive and so further feasibility studies are required to determine the most efficient way to achieve flare localisation within a lightcurve.

Finally, it is important to note that there have been more recent ZTF data releases since the download of the ALPAKA lightcurves from ZTF DR 6, which was acquired in 2021. To keep up to date with the most recent available data, I re-downloaded the most recent data release as of January 2024 (DR 20) and repeated the analysis described in this chapter. The results found using this updated data release are shown in the appendix (Appendix A).

3.9 ZTF lightcurves showing extreme variability

In this section I present the ZTF lightcurves of AGN that were classified as flares by the GP. Note that some objects are actually decreasing in brightness as the GP detects extreme variability in both directions.

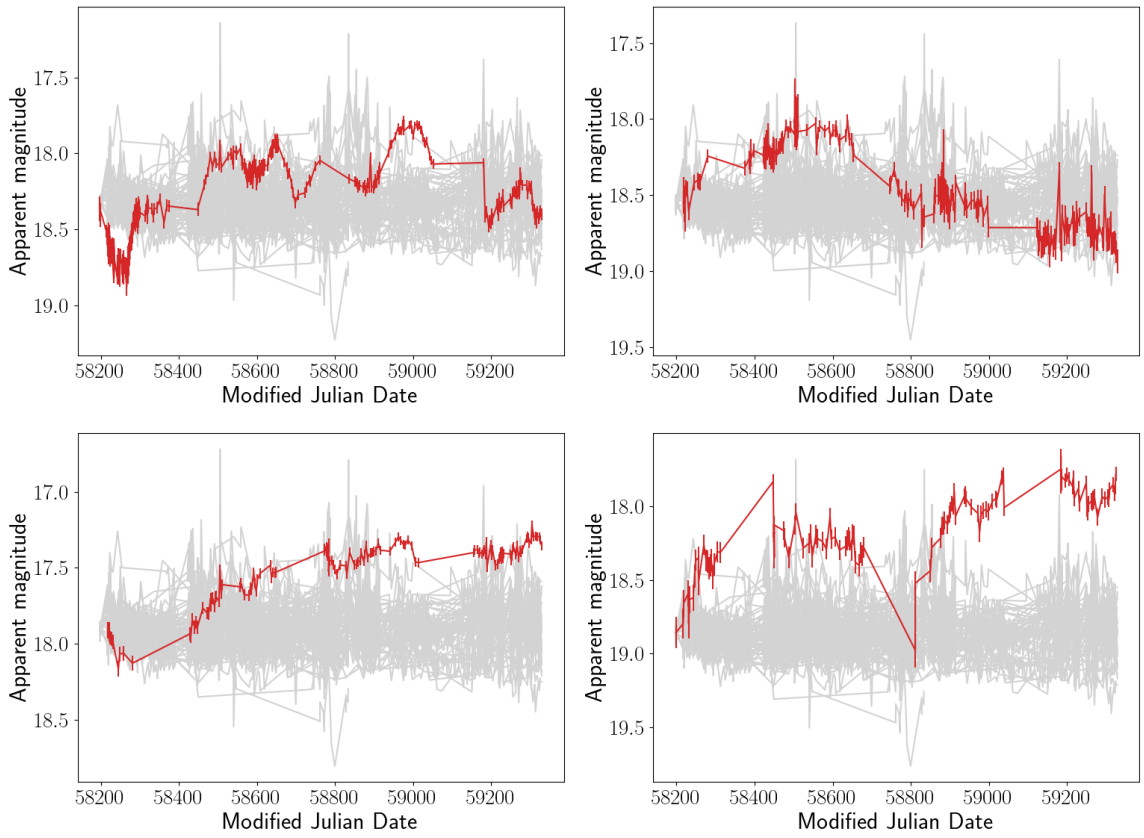

Figure 3.18: ZTF r-band lightcurves of the 5th–8th flare candidates out of the total 27 identified by the GP. The red line shows the lightcurve of the flare candidate and the grey curves are a randomly-sampled selection of 100 lightcurves that were not flagged as flares by the GP, demonstrating that they show extreme variability compared to the rest of the population. These lightcurves have been normalized for ease of visualization (see §3.6.2). Figure adapted from McLaughlin et al. (2024).

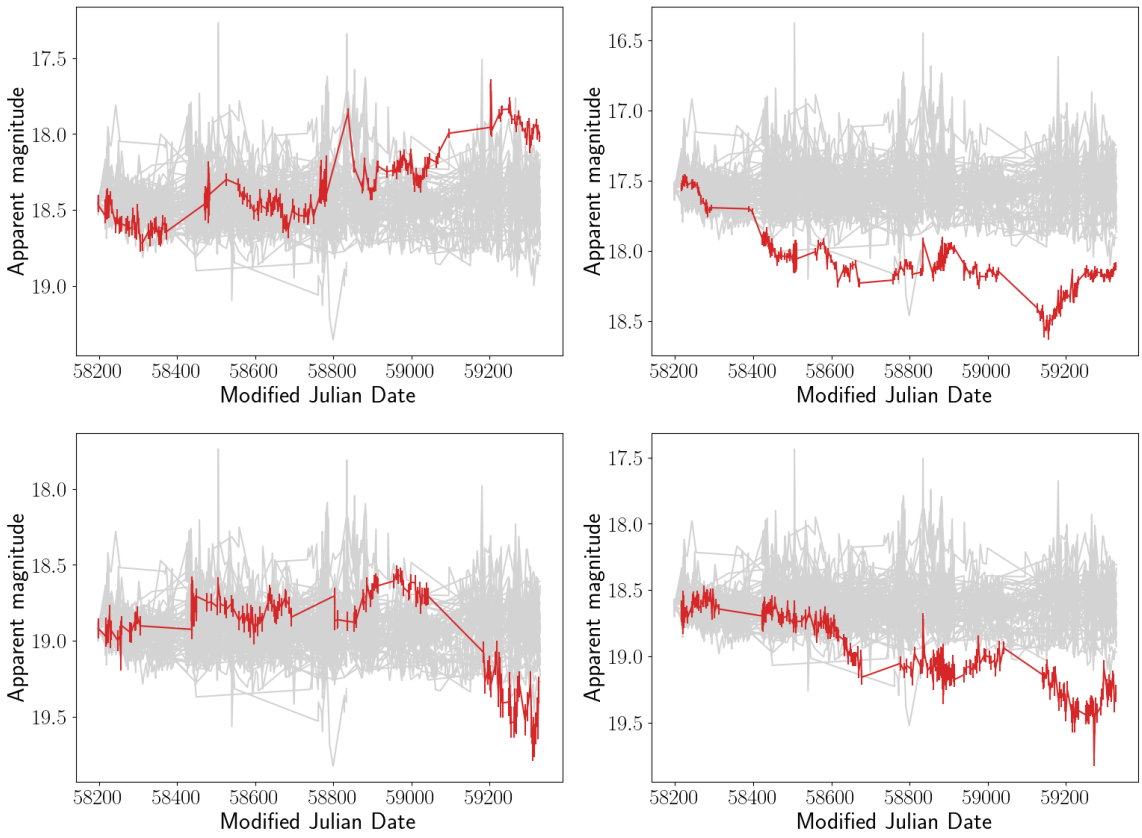

Figure 3.19: ZTF r-band lightcurves of the 9th–12th flare candidates out of the total 27 identified by the GP. The red line shows the lightcurve of the flare candidate and the grey curves are a randomly-sampled selection of 100 lightcurves that were not flagged as flares by the GP, demonstrating that they show extreme variability compared to the rest of the population. These lightcurves have been normalized for ease of visualization (see §3.6.2). Figure adapted from McLaughlin et al. (2024).

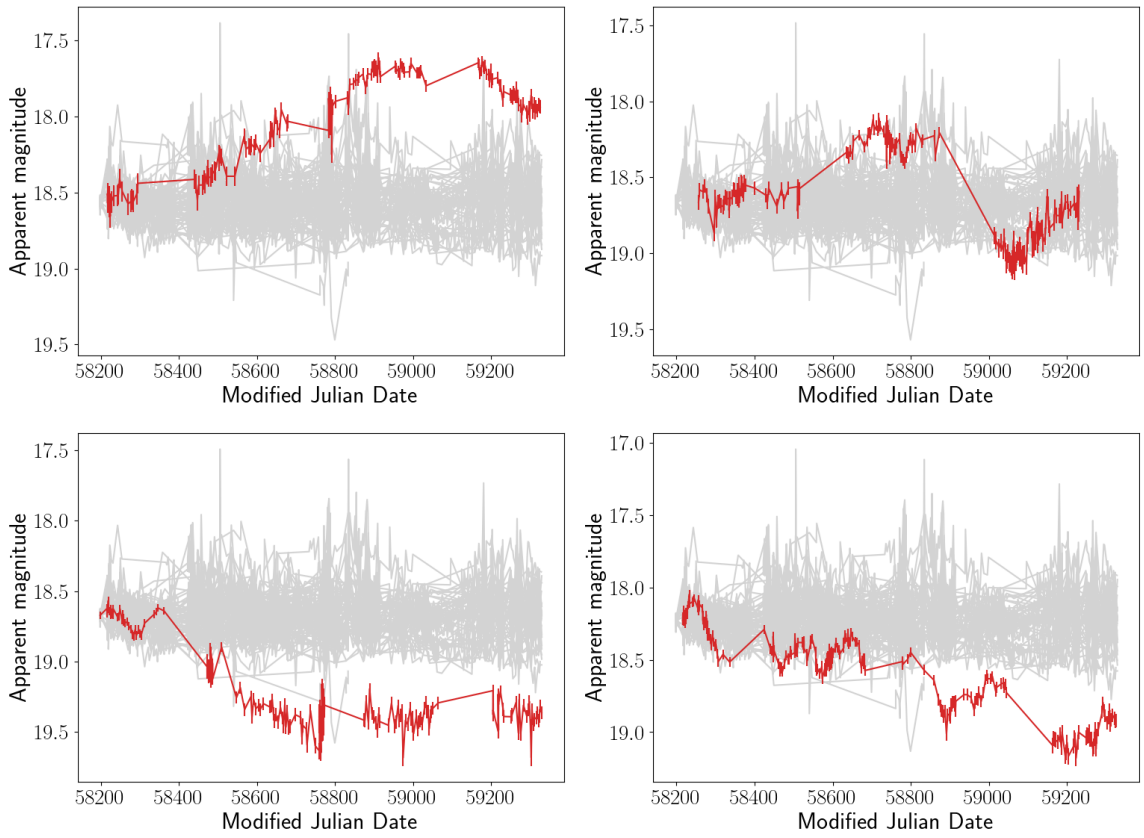

Figure 3.20: ZTF r-band lightcurves of the 13th–16th flare candidates out of the total 27 identified by the GP. The red line shows the lightcurve of the flare candidate and the grey curves are a randomly-sampled selection of 100 lightcurves that were not flagged as flares by the GP, demonstrating that they show extreme variability compared to the rest of the population. These lightcurves have been normalized for ease of visualization (see §3.6.2). Figure adapted from McLaughlin et al. (2024).

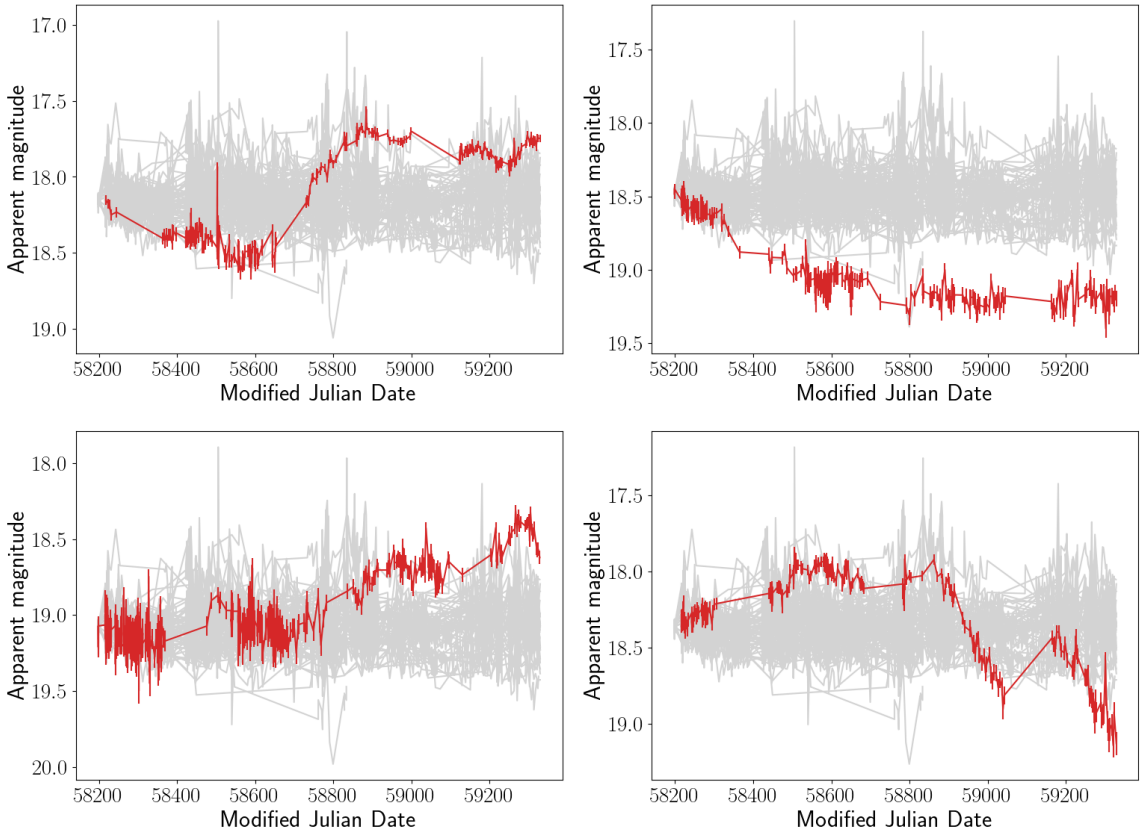

Figure 3.21: ZTF r-band lightcurves of the 17th-20th flare candidates out of the total 27 identified by the GP. The red line shows the lightcurve of the flare candidate and the grey curves are a randomly-sampled selection of 100 lightcurves that were not flagged as flares by the GP, demonstrating that they show extreme variability compared to the rest of the population. These lightcurves have been normalized for ease of visualization (see §3.6.2). Figure adapted from McLaughlin et al. (2024).

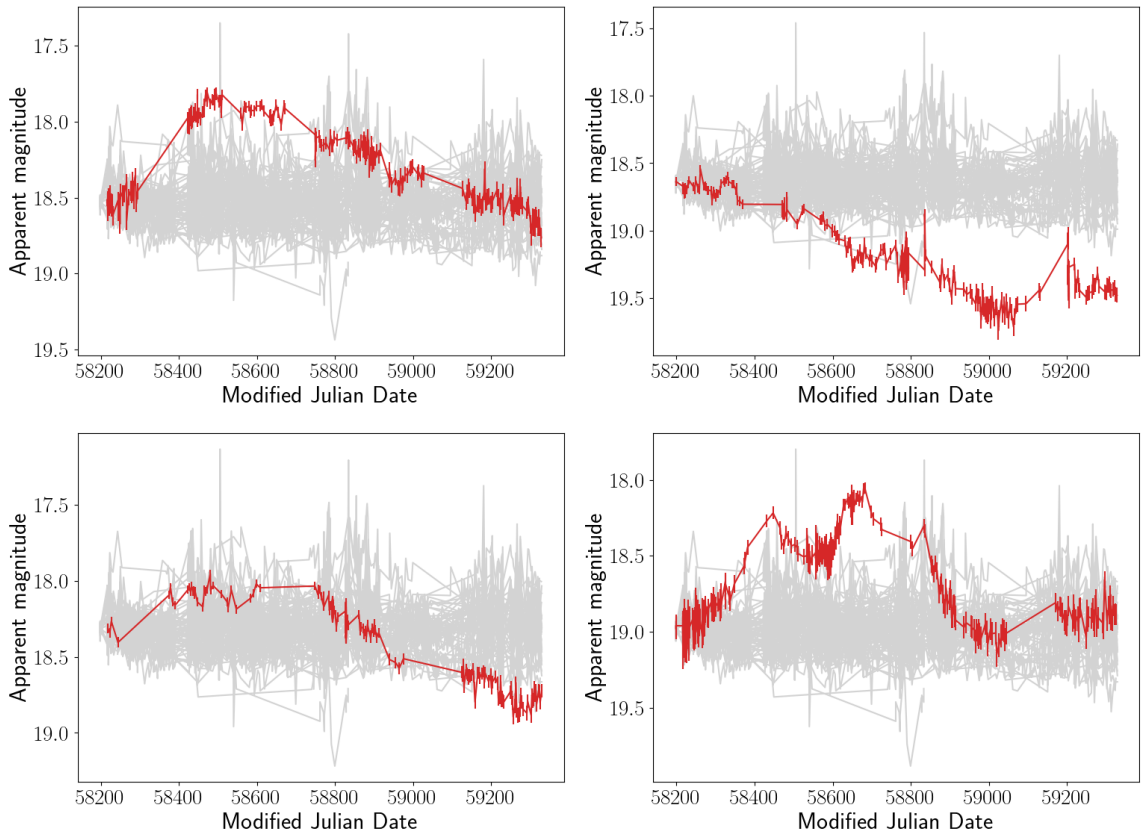

Figure 3.22: ZTF r-band lightcurves of the 21st-24th flare candidates out of the total 27 identified by the GP. The red line shows the lightcurve of the flare candidate and the grey curves are a randomly-sampled selection of 100 lightcurves that were not flagged as flares by the GP, demonstrating that they show extreme variability compared to the rest of the population. These lightcurves have been normalized for ease of visualization (see §3.6.2). Figure adapted from McLaughlin et al. (2024).

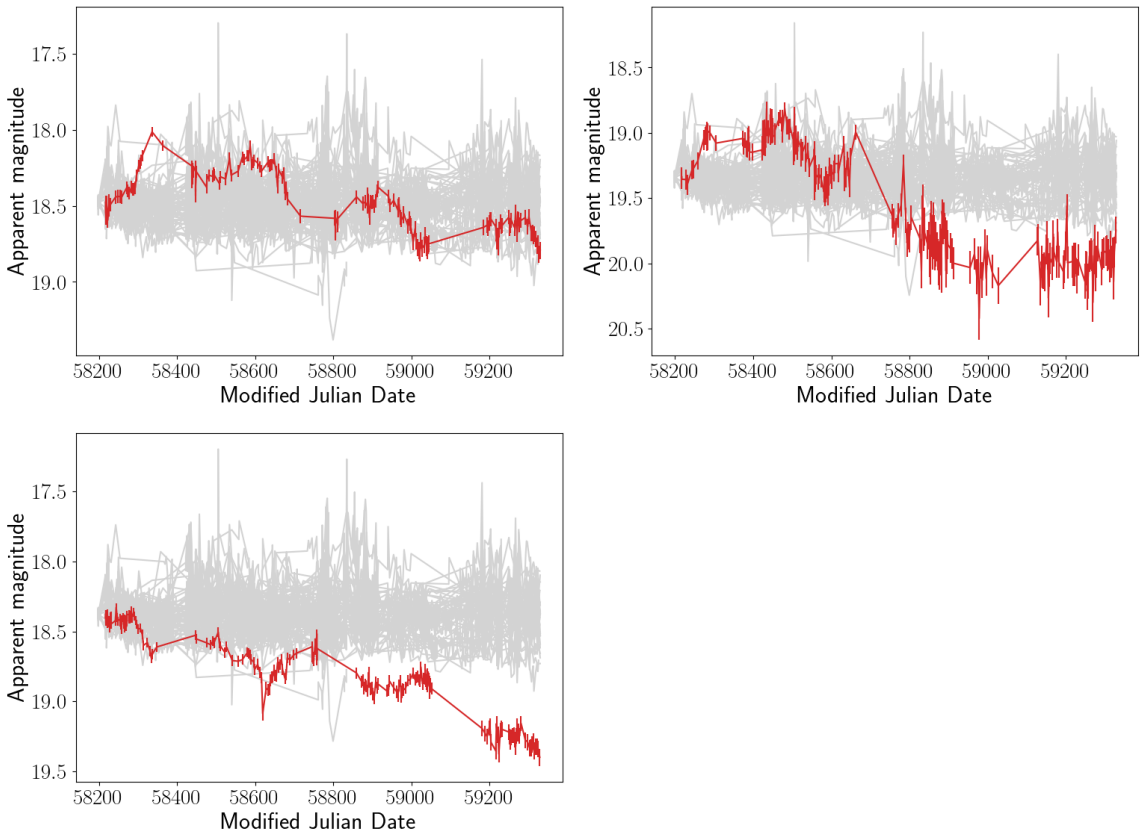

Figure 3.23: ZTF r-band lightcurves of the 25th–27th flare candidates out of the total 27 identified by the GP. The red line shows the lightcurve of the flare candidate and the grey curves are a randomly-sampled selection of 100 lightcurves that were not flagged as flares by the GP, demonstrating that they show extreme variability compared to the rest of the population. These lightcurves have been normalized for ease of visualization (see §3.6.2). Figure adapted from McLaughlin et al. (2024).

Chapter 4

Flare localisation within a lightcurve

4.1 Introduction

Up until this point in the thesis, my flare detection technique using Gaussian Processes (GPs: see Chapter 3) has looked at AGN lightcurves in their entirety and as such sought to determine whether a given entire AGN lightcurve differs statistically from the rest of the AGN population. This was achieved by fitting a GP to each AGN lightcurve individually and therefore obtaining the optimised hyperparameters for the entire lightcurve. These hyperparameters, which represent the overall variability of the lightcurve, have therefore been used as a summary statistic with which to compare a given AGN lightcurve against the rest of the population to determine whether the variability deviates significantly from the norm. In Chapter 3, I show that this technique is a viable means to distinguish lightcurves potentially containing AGN flares from the rest of the population. It is important to note here that this analysis provides the ability to predict that a given AGN lightcurve likely *contains* a flare, but not the ability to determine exactly which feature within that lightcurve has been identified as a potential flare (or indeed, the *location* of the flare within the lightcurve).

At face-value, flare detection in this way using a summary statistic of the entire

lightcurve is perfectly reasonable. If one's goal is simply to flag an object for follow-up after a GP fit to its lightcurve suggests that it is extremely variable compared to the rest of the AGN population, then this technique need not necessarily be developed further. However, there is justification for the community to aspire to detect AGN flares as they happen, to enable rapid follow-up observations for the purpose of understanding the extreme physics of the central black hole feeding. With this in mind, with the intention of developing a flare detection technique that is suitable for "real-time" AGN flare detection in advent of upcoming high-cadence surveys such as the Rubin Observatory Legacy Survey of Space and Time (LSST: Ivezić et al., 2019), there are still two key questions that one might like to address, for example:

1. Can we determine exactly when a flare actually occurs, i.e., its location within a lightcurve?

2. How early into a flare are we able to detect it?

Since one of the main science drivers of the LSST is "Exploring the Transient Optical Sky" and the survey expects an average of 10 million transient alerts per night (Ivezić et al., 2019), there is a need for rapid classification and follow-up by the community. With this in mind, I felt it important to alter my GP-based flare detection system to analyse AGN lightcurves in a more localised fashion in order to investigate the above questions. My research up to this point in the thesis has been able to determine if a given lightcurve is likely to *contain* a flare, but this chapter concerns itself with the ability to determine at *exactly what point within that lightcurve* does the flare occur and hereafter I refer to this as flare localisation.

To investigate flare localisation and prepare for upcoming surveys such as the LSST, in this chapter I outline how I have altered the GP flare detection from Chapter 3 to analyse both simulated and real AGN lightcurves in an ongoing fashion as more data is progressively added. The outline of this chapter is as follows: I describe the methodology

behind flare localisation in §4.2 before presenting the results of the analysis in §4.3. I discuss the findings in §4.4 and conclude the chapter in §4.5.

4.2 Flare localisation

In this section I discuss how I utilised a GP to parameterise the variability of an AGN lightcurve in an ongoing fashion, simulating the effect of receiving new observations from a survey in real-time. In this way, I step along the lightcurve and I re-evaluate the GP fit each time a new data point is added. I performed this technique on both simulated AGN lightcurves and the lightcurves of AGN from the ALPAKA catalogue (ALPAKA: Mullaney et al., 2013). In the latter case, the lightcurve data were obtained from the Zwicky Transient Facility Public Data Release 6 (ZTF: Masci et al., 2018; Bellm et al., 2019). Note that this dataset is the same as that used in Chapter 3 and the reader can refer to that chapter for the description of the AGN selection and data retrieval. First, I will provide a reminder for the reader about how GP regression works across an entire lightcurve and then I will describe how I alter this technique to achieve flare localisation in both simulated and real AGN lightcurves.

4.2.1 Rolling GP regression

In Chapter 3, a GP was used to produce a summary statistic of the variability of a given AGN lightcurve. As such, the GP was calculated across the entire lightcurve as shown in Fig. 4.1. The GP summarises the variability properties for the lightcurve (hyperparameters: see Chapter 3), which quantify how the AGN is varying over time. However, in the era of time-domain astronomy, and especially with the advent of high-cadence surveys such as the LSST, we are going to have the opportunity to detect transient events such as AGN flares in real-time. As a result, I sought to modify my GP analysis to simulate the effect of receiving data on an ongoing basis and also to determine

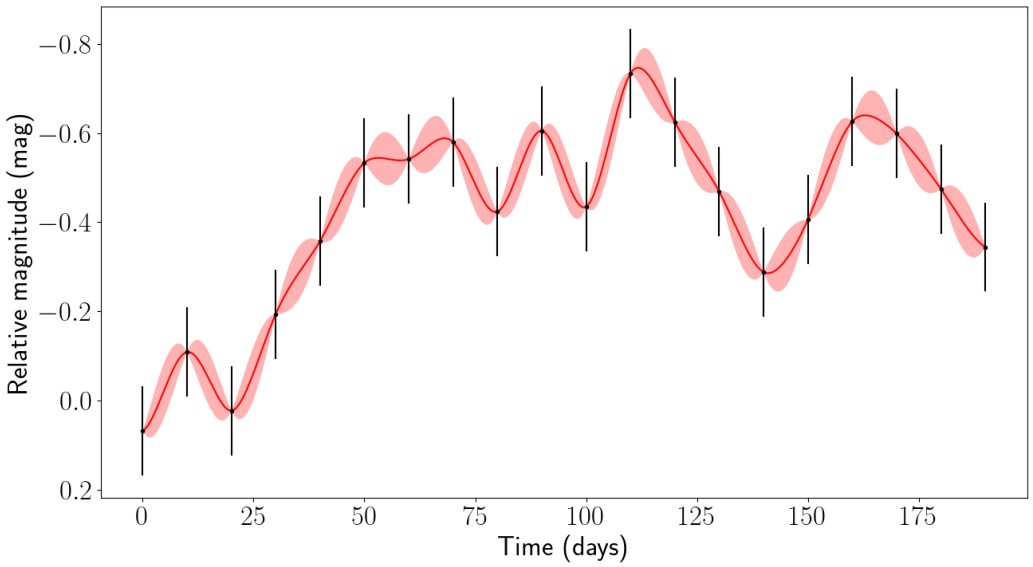

Figure 4.1: A Gaussian Process (GP) fit to a simulated AGN lightcurve. The red line shows the predicted mean distribution and the red shaded areas show the uncertainty in this mean.

whether it is possible to isolate the point at which a flare "goes off", *i.e.,* the moment at which the lightcurve deviates significantly from its normal behaviour.

One could imagine "stepping" through a lightcurve and fitting a GP each time a new data point is added, which in effect looks back through the lightcurve and re-evaluates the GP variability hyperparameters each time new data is received; this is shown in Fig. 4.2. This would allow for the possibility of identifying the point at which the GP hyperparameters deviate significantly from the region of parameter space where "normal" AGN lie and therefore signal that a flare may be occurring. In a similar vein to the analysis described in Chapter 3, for each iteration of the rolling GP, the optimised hyperparameters (when compared with the hyperparameters of the entire AGN sample in question) can be used to calculate the flare probability on an ongoing basis.

I used this rolling technique to investigate whether GPs can be used for flare localisation. As in Chapter 3, I first attempt to apply this approach to simulated AGN lightcurves to investigate the feasibility and then I test the approach on the real ZTF

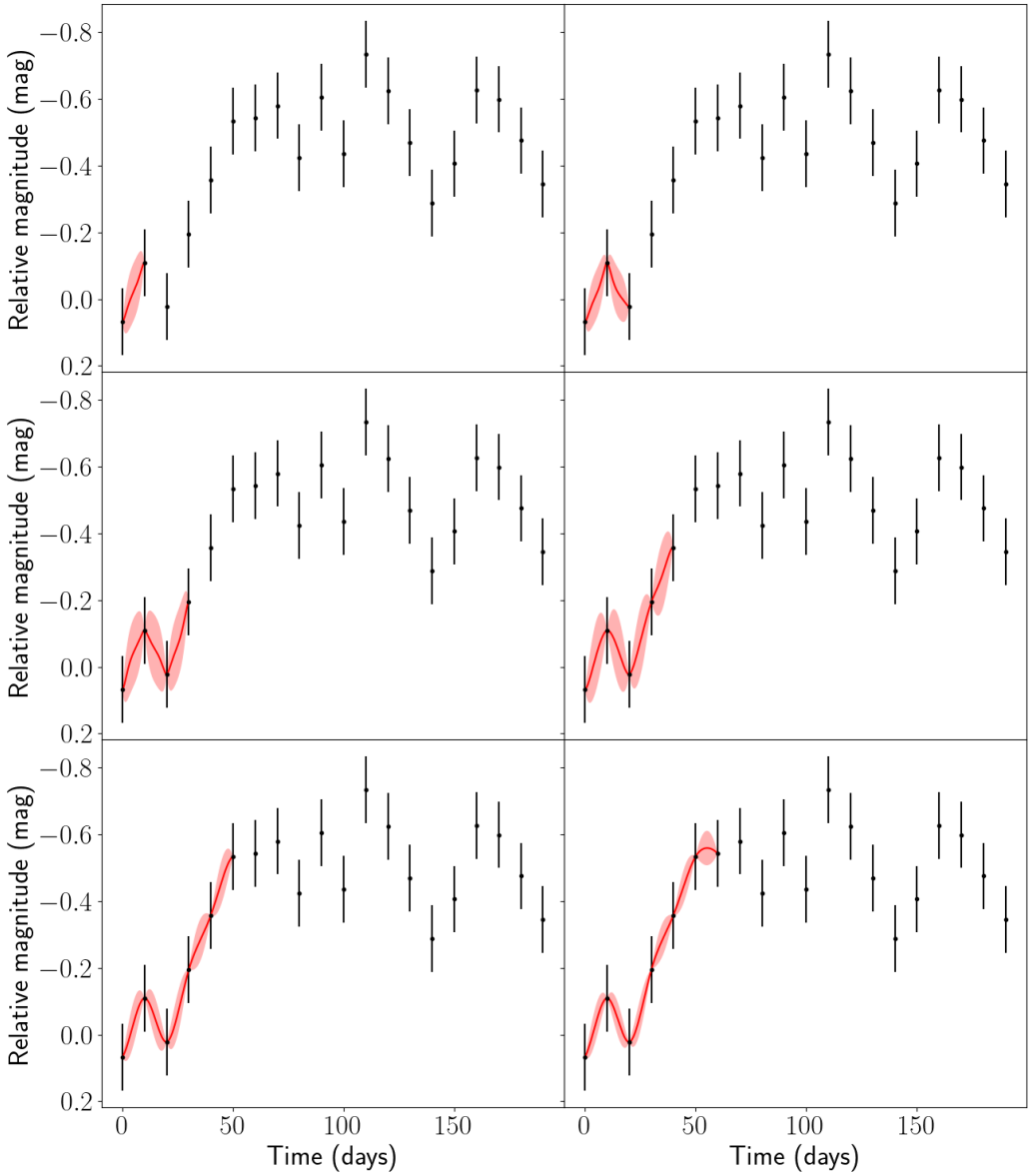

Figure 4.2: A rolling Gaussian Process fit to the first seven points in a simulated AGN lightcurve in succession. The red line shows the predicted mean distribution and the red shaded areas show the uncertainty in this mean.

lightcurves from the ALPAKA AGN catalogue (Mullaney et al. 2013, see Chapter 3). In Chapter 3 I used Bayesian hypothesis testing to determine the probability of a new lightcurve belonging to either the flare ($\theta = 1$) or non-flare ($\theta = 0$) populations. In that approach, the posterior probability of a lightcurve containing a flare or not is described (according to Bayes' theorem) as:

$$P(\sigma, \rho, \theta|y) \propto P(y|\sigma, \rho)P(\sigma, \rho|\theta)P(\theta), \tag{4.1}$$

where σ and ρ are the kernel hyperparameters, y is the data, and $P(\theta)$ is defined as a "hyper prior" (McLaughlin et al., 2024).

This is a statistically robust means of quantifying the probability that a given lightcurve contains a flare, but since this method utilises Markov Chain Monte Carlo (MCMC) sampling to sample the posterior probability distribution it is too computationally expensive to apply in a rolling fashion every time a new data point is added to a lightcurve. To circumvent this issue, the calculation of the flare probability can be simplified without loss of statistical rigour (provided the "probability" is viewed simply as a means to flag extreme variability rather than a true probability, see later). Under this name, the flare probability can be calculated as follows:

$$P(\theta_f|y) = \frac{P(\sigma, \rho|\theta_f)P(\theta_f)}{P(\sigma, \rho|\theta_f)P(\theta_f) + P(\sigma, \rho|\theta)P(\theta)}, \tag{4.2}$$

where θ_f and θ refer to the flare and non-flare regions of hyperparameter space respectively. This equation is simply a normalised ratio of the prior probability distributions of the non-flaring lightcurves and the flaring lightcurves. I refer to this ratio as a "probability" despite it not being a *true* probability; it is not statistically meaningful on its own as it is not "fully" Bayesian (*i.e.,* the full Bayes' Theorem in Eq. 3.5), but as long as one understands this nuance, it is reasonable to use this value as an alert or as a measure of the likelihood that a lightcurve is flaring. This is because it is a direct measure of

the discrepancy between the likelihood that the lightcurve belongs to the "normal" AGN population and the likelihood that it belongs elsewhere (*i.e.,* to the flaring region).[1]

Using this technique, I applied the rolling GP to both the sample of 10 000 simulated AGN lightcurves with and without injected flares and the 9035 real ZTF lightcurves from the ALPAKA catalogue (see Chapter 3 for details of these samples and their corresponding hyperparameter distributions). I realised that during early times where few data points in the lightcurve had elapsed, the GP was returning unphysical hyperparameters and on closer inspection the GP fit to the data was spurious. As a result, a minimum length for the rolling GP was chosen to be 30 data points to ensure that the resulting hyperparameters were reliable (*i.e.,* to minimise disruption due to irregular cadence and outliers and to provide enough data points to allow the GP fit to "settle" into a stable covariance). After the window of 30 data points is elapsed, the GP is re-evaluated after each new data point is fed into the lightcurve and the hyperparameters updated. Using Eq. 4.2, the flare probability is calculated after each new data point is added and therefore the flare probability can be tracked against time to determine when a flare may be occurring. In Section 4.3 I discuss the results of this technique in the case of both simulated and real AGN lightcurves.

4.3 Results

In this section I outline the results of the GP-based flare localisation technique described above. First, I show that it is possible to trace flare probability against time in the case of regularly-sampled simulated AGN lightcurves and that this technique can be used to isolate the point at which a flare occurs (§4.3.1). Secondly I show that despite the effects

[1]For further clarity, the "flare probability" $P(\theta_f|y)$ is referred to as a probability because it has been normalised to always lie between zero and one, which is essentially a means of finding the *dominant* probability (between that of the "normal" AGN region and the flaring region and the simulated lightcurve hyperparameters can be used to calibrate this). The use of the flare probability is suitable therefore as an alert but cannot be used in any further statistics and therefore the reader may wish to refer to it as a metric rather than a probability.

of irregular cadence and outliers that come with moving from simulated data to real data, this technique is still effective in real ZTF AGN lightcurves from the ALPAKA catalogue (§4.3.2) with and without injected simulated Gaussian flares. Finally I show how the rolling flare probability varies with time in the case of the 27 potential flares that were identified in Chapter 3.

4.3.1 Simulated lightcurves

Before presenting broader results from analysing the full sample of simulated lightcurves, I first present a singular example of the results of flare localisation to clarify the methodology described in §4.2.1. Fig. 4.3 shows how flare localisation allows for the tracking of flare probability against time. In this plot, a simulated AGN lightcurve with a sampling of 10 days and an injected flare at day 1100 is shown, with the rolling flare probability calculated using Eq. 4.2 also shown. It is clear that as more of the flare elapses, the flare probability increases. This is further illustrated in Fig. 4.4, which shows the path that the same lightcurve's hyperparameters take with respect to the hyperparameter distributions of simulated flares and non-flares from Chapter 3. This shows that as more data points in this lightcurve are included in the analysis, the hyperparameters move further into the flaring (orange) region.

Now that it has been established that flare localisation using this rolling GP technique is a viable means of tracking flare probability as new data points are received in real-time, it is important to ascertain some statistics about the full sample. This is especially important since Fig. 4.3 demonstrates that this approach works in an ideal case (*i.e.,* a simulated lightcurve with regular cadence) but is not necessarily representative of normal behaviour; key questions remain, such as how the full sample is distributed. Fig. 4.5 shows the distribution of the magnitude deviation from zero at which the flare probability exceeds 0.99. In other words, this demonstrates at what level of deviation does the flare

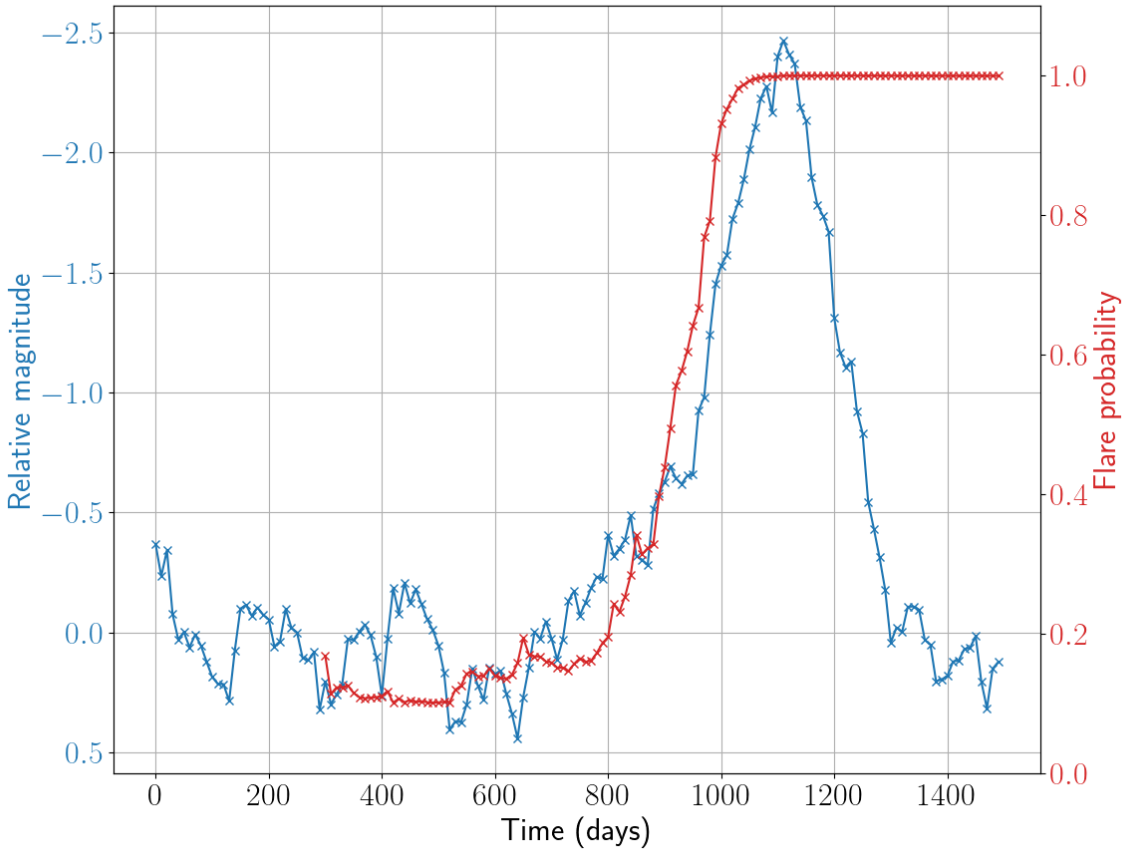

Figure 4.3: A simulated AGN lightcurve with an injected Gaussian flare at 1100 days in blue with its rolling flare probability shown alongside in red.

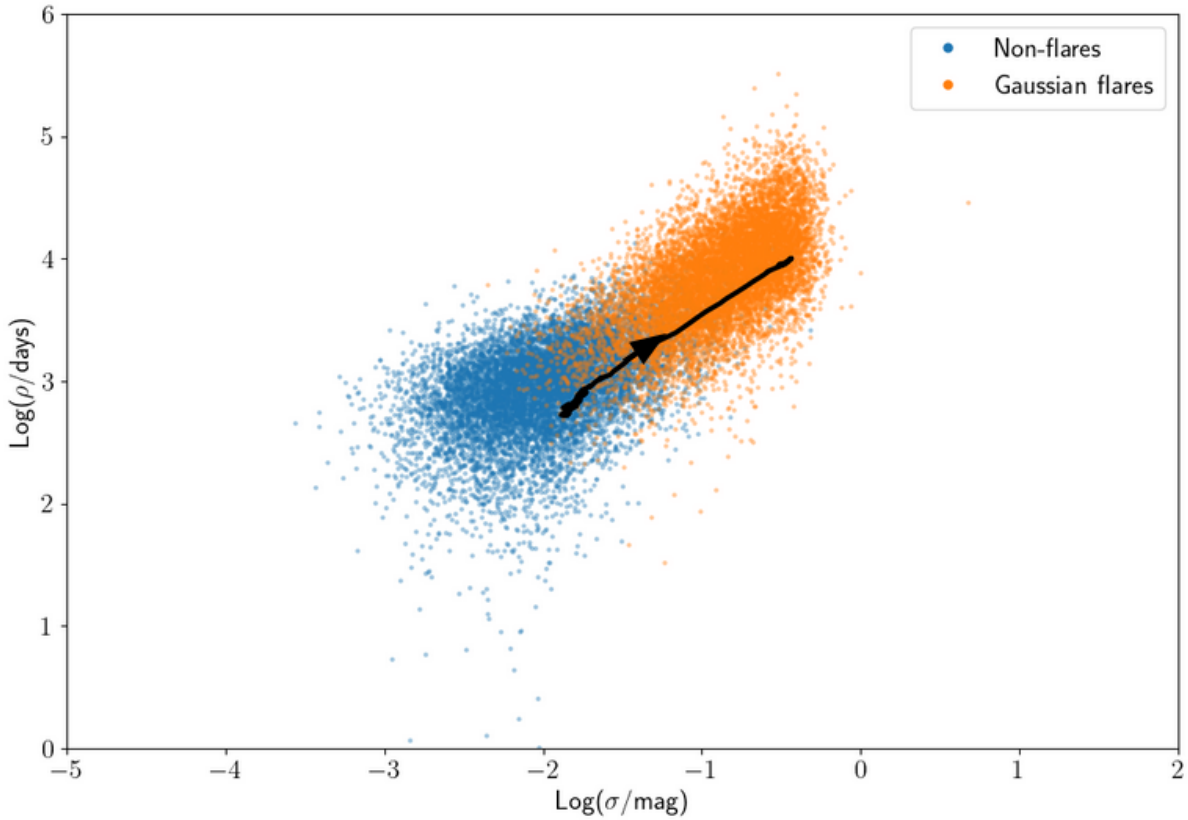

Figure 4.4: Distributions of flare and non-flare hyperparameters for well-sampled simulated lightcurves with injected Gaussian flares (see Chapter 3). The black line shows the path that the rolling hyperparameters of the simulated AGN lightcurve from Fig. 4.3 take in this parameter space and the arrow is indicating its direction, showing that as more data points in this lightcurve are elapsed, the hyperparameters move further into the flaring (orange) region.

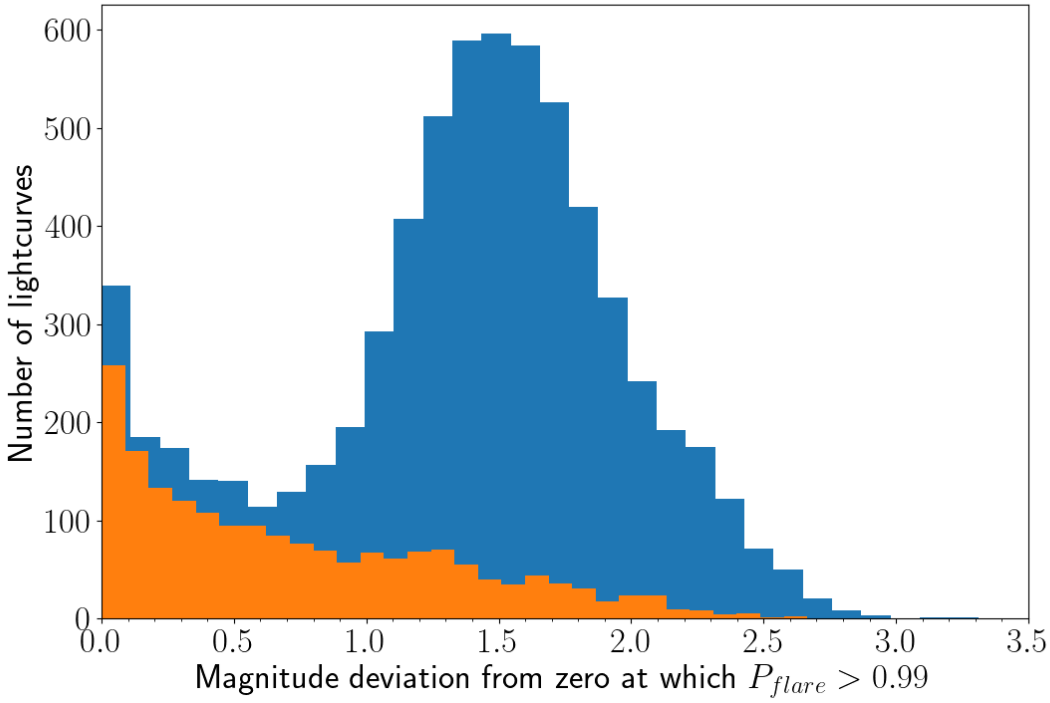

Figure 4.5: Distribution of the magnitude deviation from zero of each lightcurve where the flare probability exceeds 0.99. Lightcurves where the point at which $P_{flare} > 0.99$ does not coincide with the injected flare location ($\pm$ 300 days) are considered false positives and are shown in orange.

localisation technique predict a very high likelihood of flare activity. The distribution has a median value of 1.4, demonstrating that most simulated flares are successfully detected by the time their amplitude reaches this value. Note that there are lightcurves at very small magnitude deviations (< 0.5), which are false-positives in this case. Out of the 10 000 simulated lightcurves, the lightcurves with magnitude deviations below 0.5 represent 910 or 9.1 per cent.

Since I know the locations within the simulated lightcurves at which I injected Gaussian flares into, I investigated whether the flare localisation technique is effective at retrieving the locations of the flares. Figure 4.6 shows the distribution of the differences between the actual injected flare peak and the point in the lightcurve at which the flare probability exceeds 0.99. The majority of the differences are less than 1 day (8490),

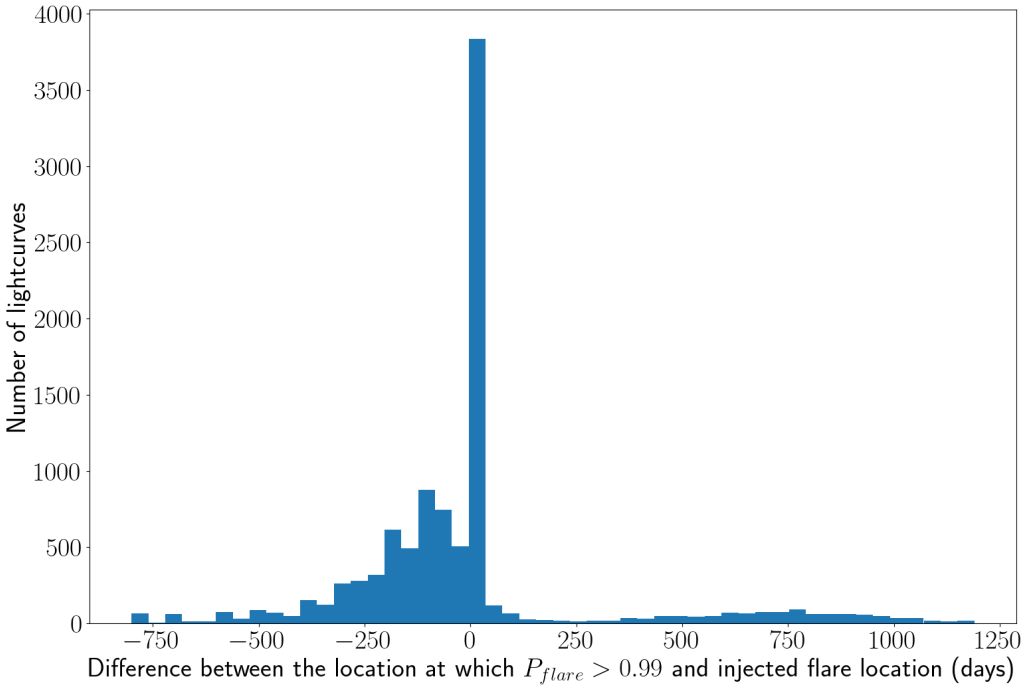

Figure 4.6: Distribution of the difference between the actual injected flare location and the point in the lightcurve at which the flare probability exceeds 0.99. Note that negative values indicate the number of days *before* the peak of the flare and therefore positive values indicate that the flare has been detected after its peak.

which corresponds to 85 per cent of the lightcurves being detected within one day of their peak.

4.3.2 ZTF lightcurves

My investigation into flare localisation has so far has been concerned with simulated AGN lightcurves with a 10-day cadence. Of course, in reality, the data from future surveys will be subject to observational effects such as irregular cadence and outliers (see Chapter 3). Therefore it is important for this technique to be tested on real data. This section outlines the results of the flare localisation technique on real ZTF lightcurves from the ALPAKA catalogue with injected Gaussian flares, and then I present the lightcurves and

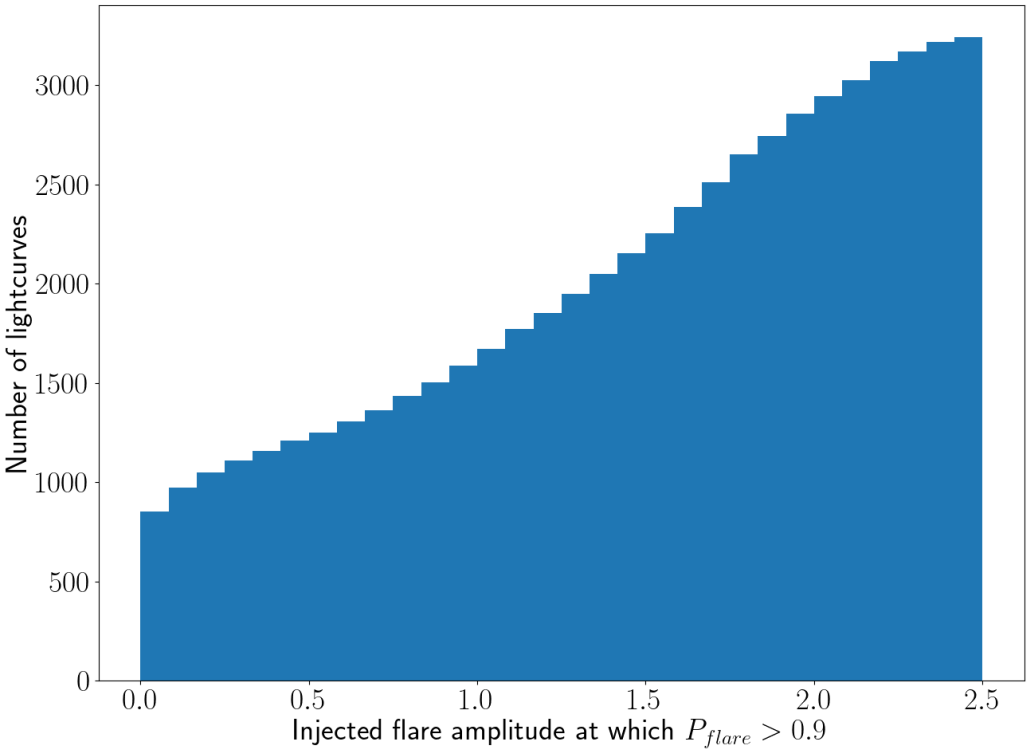

Figure 4.7: Cumulative histogram showing the injected flare amplitude at the point at which the flare probability exceeds 0.9.

corresponding rolling flare probabilities of the 27 flare candidates that were identified in Chapter 3.

Fig. 4.7 shows the cumulative distribution of the injected flare amplitude when the flare probability exceeds 0.9. Note that the threshold probability has been reduced from 0.99 in §4.3.1 to 0.9 to account for the fact that the GP is now being applied to real lightcurves with "flaws" (irregular cadence and outliers). When the threshold of 0.99 was used, many injected flares were not retrieved. Fig. 4.8 shows the same plot expressed as a fraction of injected flare height. More than half (64 per cent) of lightcurves with injected flares never achieved a flare probability of 0.9 and were therefore not detected using this threshold. Of the remaining 36 per cent of lightcurves with injected flares that were detected, 48 per cent were detected before the flare amplitude exceeded 1 magnitude

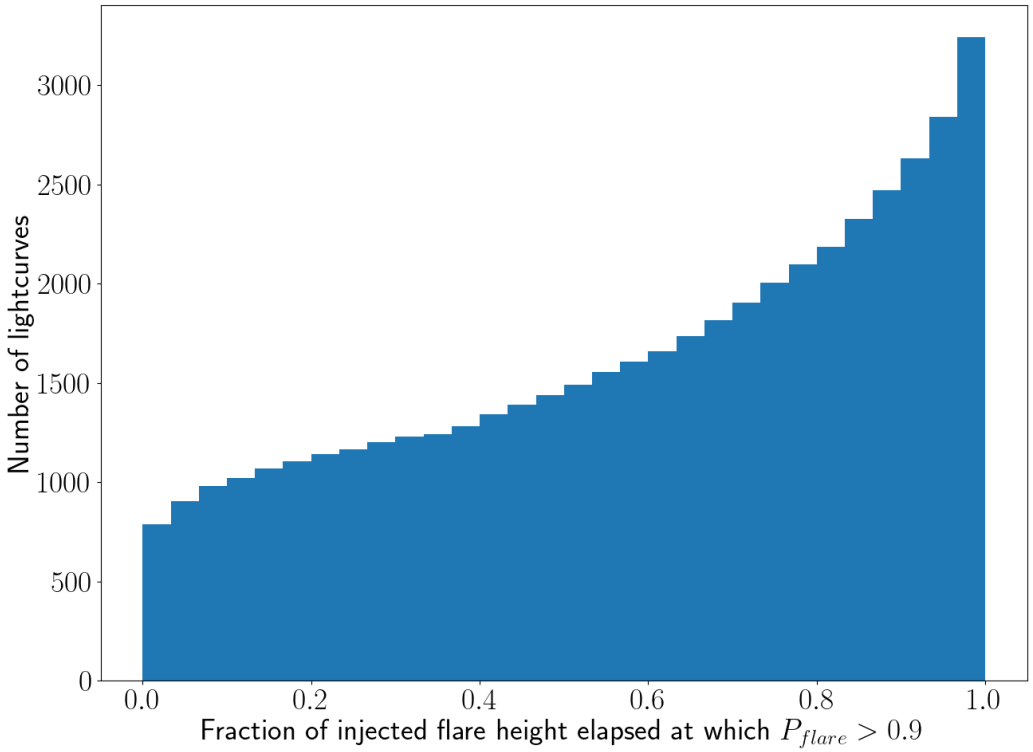

Figure 4.8: Cumulative histogram showing the fraction of injected flare amplitude at the point at which the flare probability exceeds 0.9.

and 88 per cent were detected before the flare amplitude exceeded 2 magnitudes. Since the injected flares were randomly assigned amplitudes between 1 and 2.5 magnitudes, Figure 4.8 expresses Figure 4.7 as a fraction of injected flare height. This shows that of the 36 per cent of injected flares that were detected, all of these flares were detected before their peak and 44 per cent of the flares were detected before the flare amplitude had reached half of its peak.

Finally, I present some flare localisation results from comparing the 27 flare candidates from Chapter 3 to a sample of randomly-selected ZTF lightcurves from the same parent sample that were not flagged as flares by the Chapter 3 analysis. Fig. 4.9 shows the distribution of the maximum flare probability achieved at any point within the lightcurve for both the 27 ZTF flare candidates from Chapter 3 and a randomly-selected

sample of 500 ZTF lightcurves that were not flagged as flares by the same analysis. The distribution shows that a significant proportion of AGN lightcurves never achieve a probability greater than 0.2, which can be used as a "cutoff" probability below which one can be confident that flare activity is not occurring in those lightcurves. Importantly, there is also a significant number of non-flaring ZTF lightcurves that achieve a maximum flare probability close to 1, representing false positives in the sample. To further investigate these false positives, Fig. 4.10 shows the distributions of the number of data points elapsed before the chosen cutoff probability 0.2. In this plot, it is clear that the vast majority of false positives result in a high flare probability very early in the lightcurve and therefore the GP fit at this stage can be spurious due to low numbers of data points. These spurious detections can be weeded out by placing a threshold on the number of data points in the lightcurve required to constitute a flare detection.

To finalise my flare localisation results, in Figs. 4.11, 4.12, 4.13 and 4.14 I present the lightcurves of the 27 ZTF flare candidates from Chapter 3 along with their corresponding rolling flare probabilities from the localisation analysis.

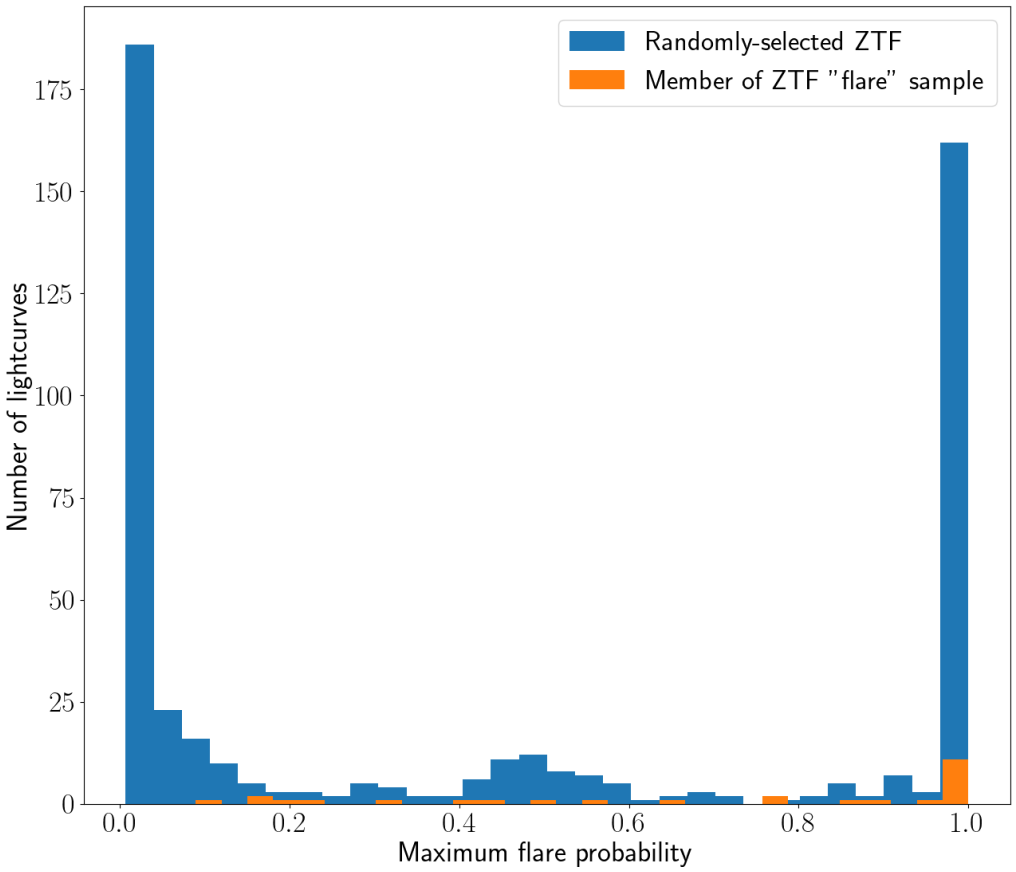

Figure 4.9: Distribution of the maximum flare probability achieved at any point in the lightcurve. The orange bars are members of the ZTF flare sample identified in Chapter 3 and the blue bars are a sample of 500 randomly-selected ZTF lightcurves that were not flagged as flares by previous GP analysis.

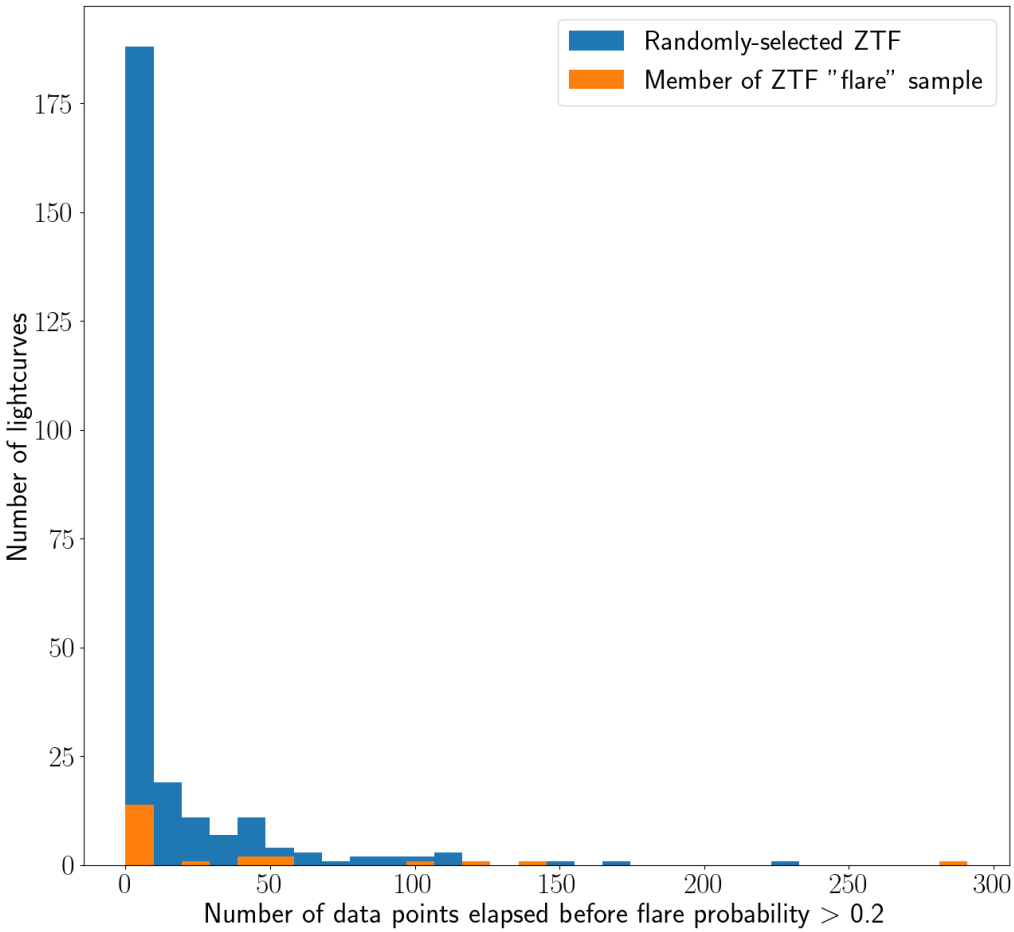

Figure 4.10: Of the lightcurves from Fig. 4.9 that achieve a flare probability exceeding 0.2, this is a plot of the distributions of the number of data points elapsed before the flare probability exceeds 0.2. The orange bars are members of the ZTF flare sample identified in Chapter 3 and the blue bars are a sample of 500 randomly-selected ZTF lightcurves that were not flagged as flares by previous GP analysis.

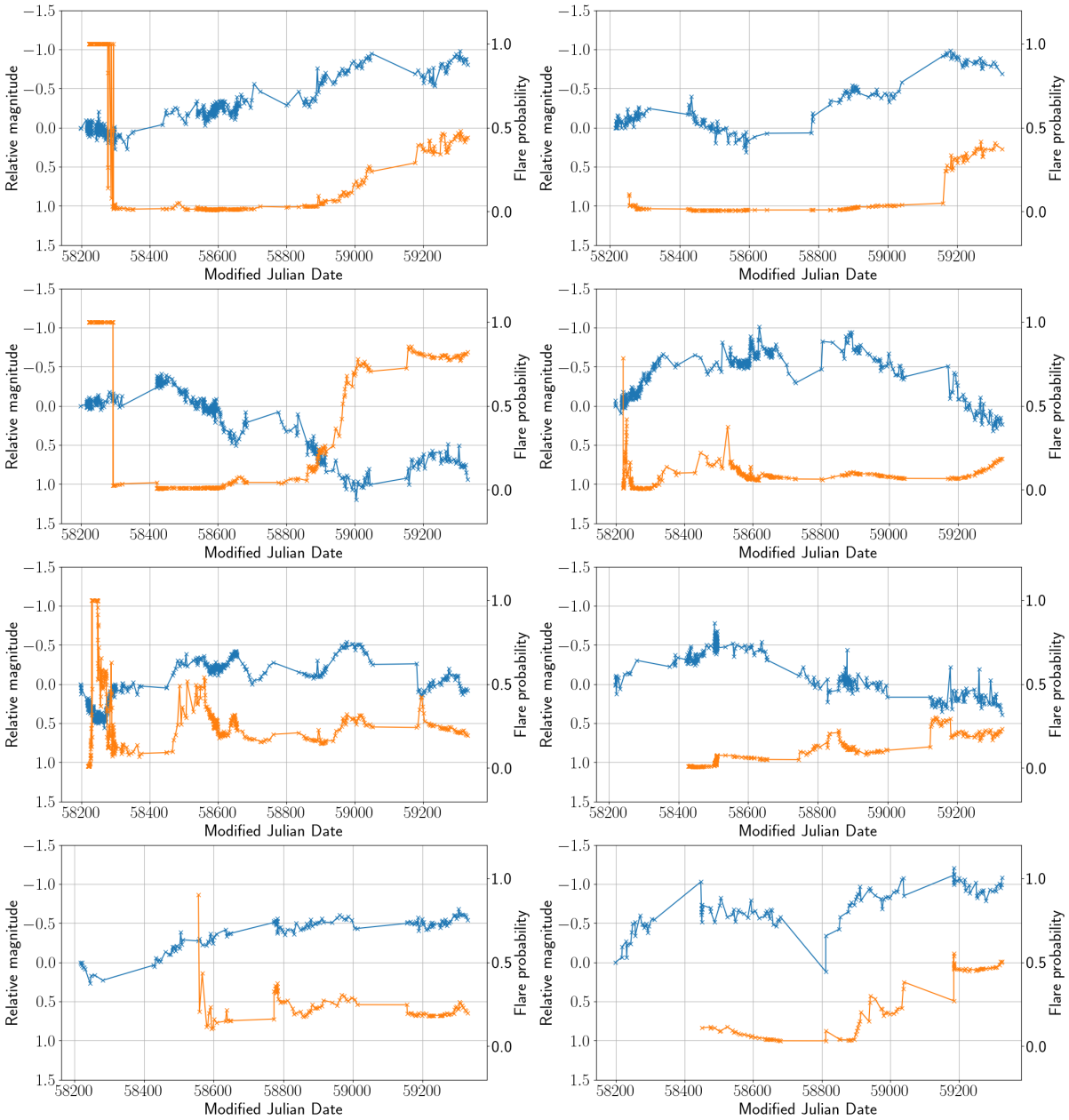

Figure 4.11: Optical r-band lightcurves of the first eight out of the total 27 flare candidates identified in Chapter 3. The lightcurve is shown by the blue points and the orange points show the corresponding rolling flare probability calculated by the flare localisation analysis.

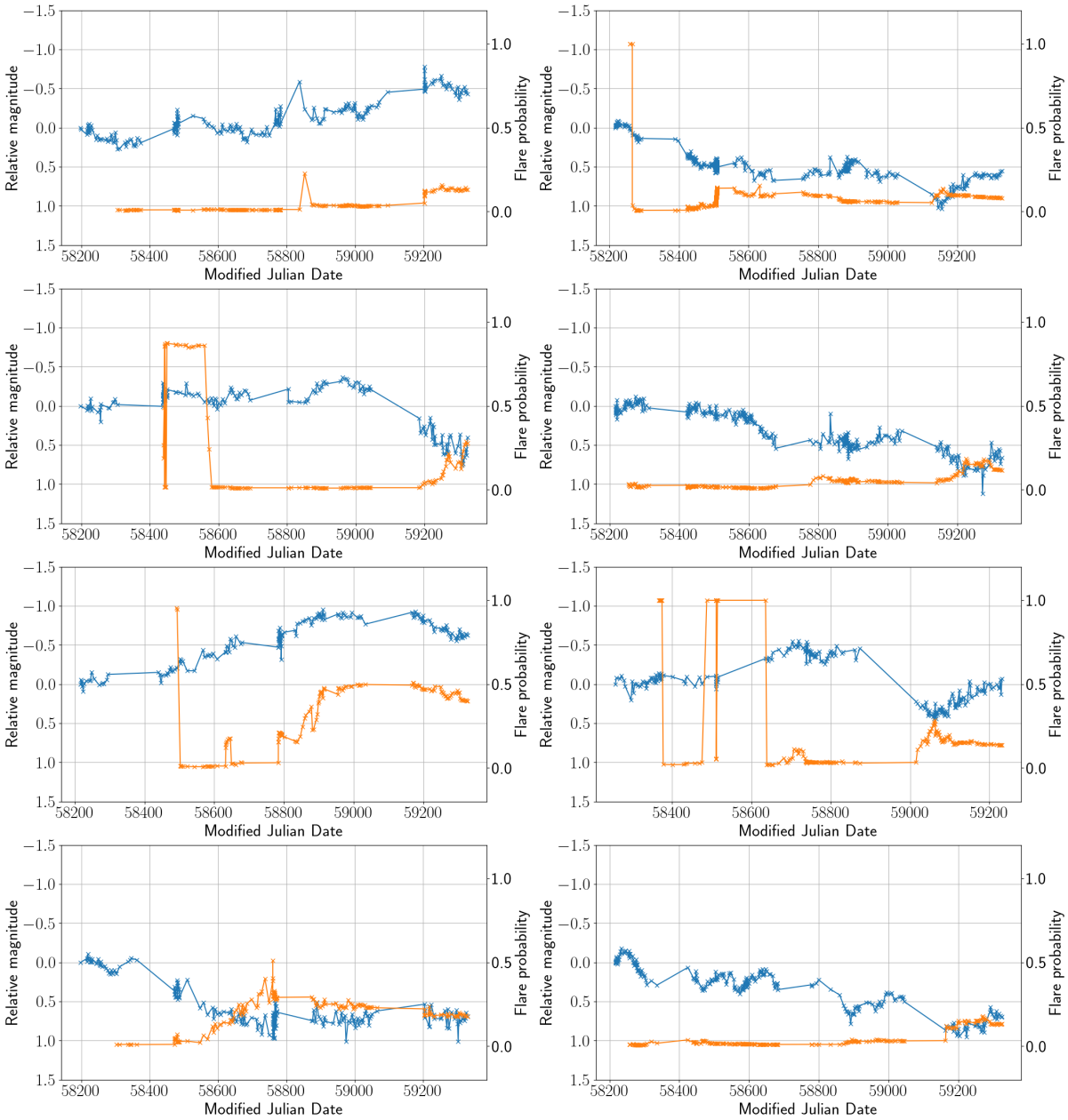

Figure 4.12: Optical r-band lightcurves of the 9th–16th out of the total 27 flare candidates identified in Chapter 3. The lightcurve is shown by the blue points and the orange points show the corresponding rolling flare probability calculated by the flare localisation analysis.

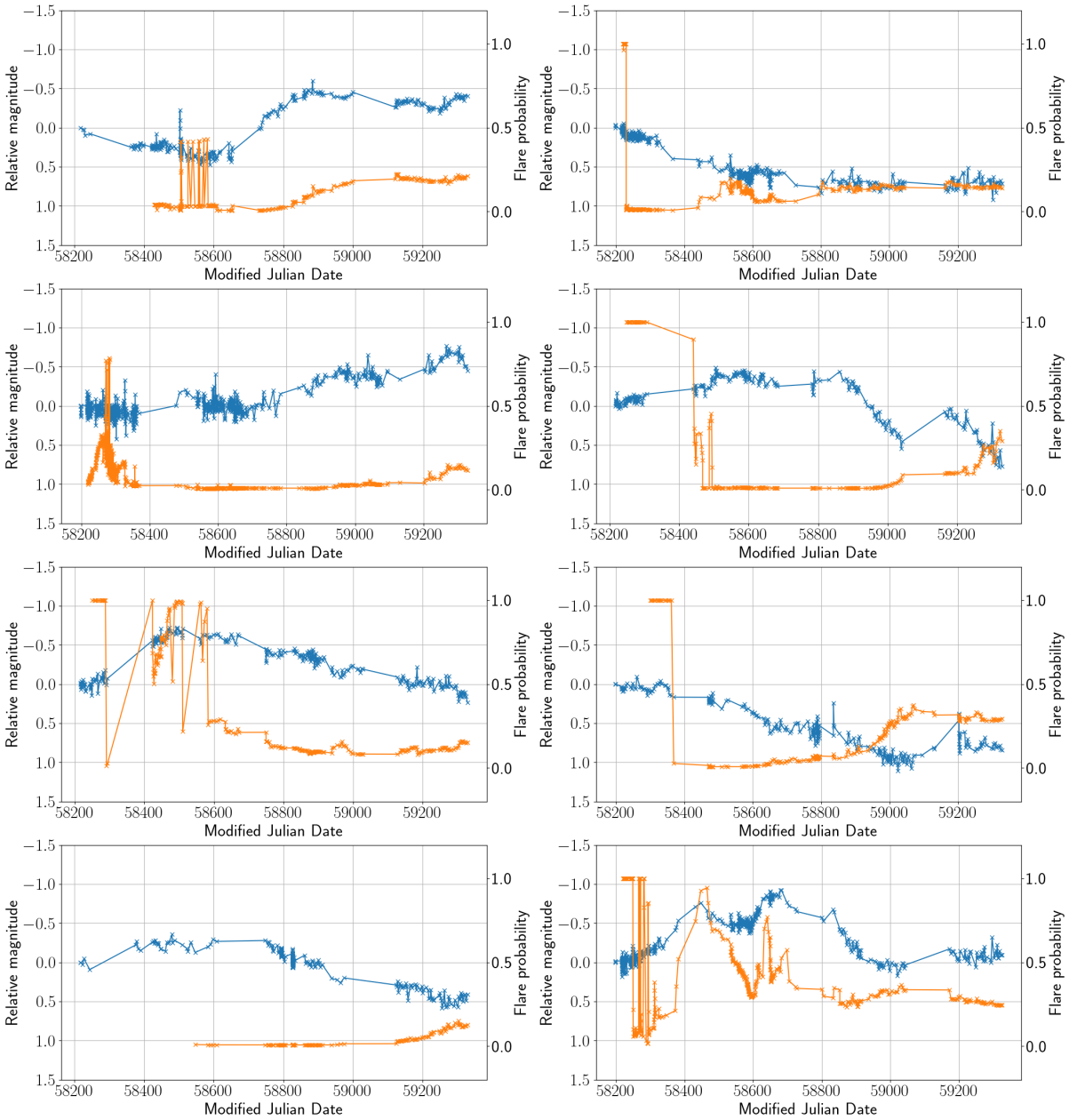

Figure 4.13: Optical r-band lightcurves of the 17th–24th out of the total 27 flare candidates identified in Chapter 3. The lightcurve is shown by the blue points and the orange points show the corresponding rolling flare probability calculated by the flare localisation analysis.

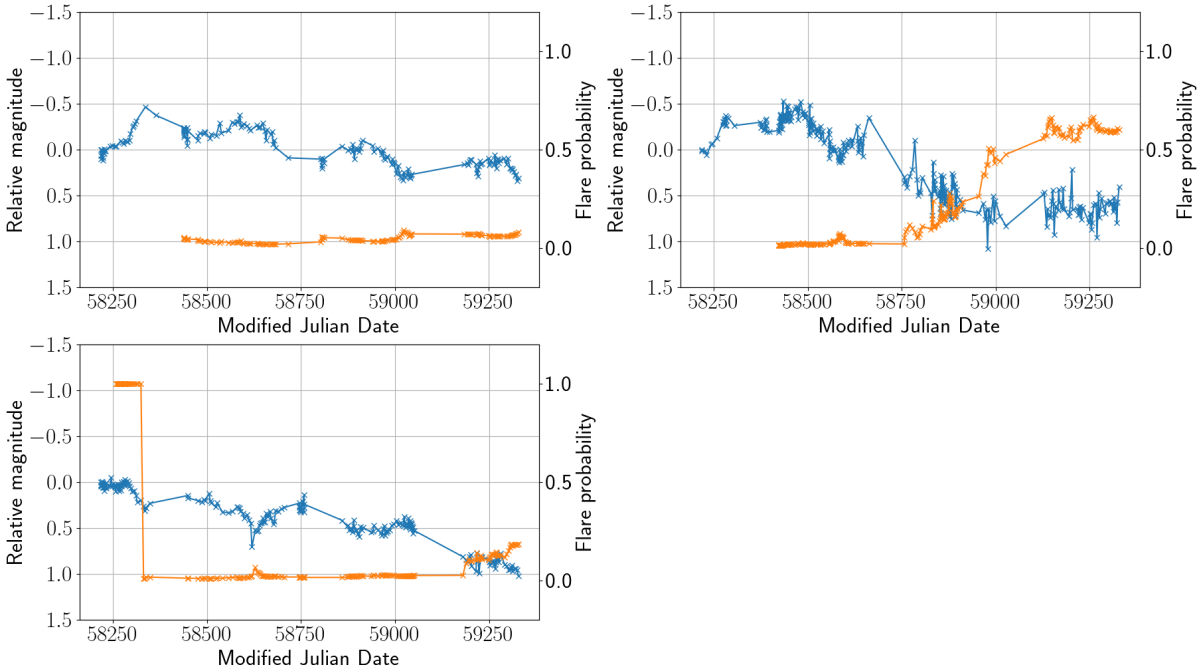

Figure 4.14: Optical r-band lightcurves of the 25th–27th out of the total 27 flare candidates identified in Chapter 3. The lightcurve is shown by the blue points and the orange points show the corresponding rolling flare probability calculated by the flare localisation analysis.

4.4 Discussion

In this chapter I have shown that flare localisation using GPs is a viable means of detecting flares in real-time in both simulated and real AGN lightcurves, which is crucial in the coming era of time-domain astronomy.

I have shown that in well-sampled simulated AGN lightcurves with injected flares, the flares can be retrieved by tracking the ongoing flare probability against time (Fig. 4.3). This is supported by the fact that an injected flare can cause the lightcurve's hyperparameters to move through parameter space from the "normal" population to the flare region (Fig. 4.4).

From analysing the results of the localisation technique on the full sample of simulated lightcurves with injected flares, it is clear that this method is still subject to a number of false positives (between 5-10 per cent), though this is due to the fact that at early

times with few data points the GP fit can be unreliable. Despite this, the majority of injected flares (55 per cent) were detected by the time the flare amplitude had exceeded 1.5 magnitudes (Fig. 4.5) and the vast majority of injected flares (85 per cent) were detected before their peak (Fig. 4.6). This suggests that flare detection in real-time as data comes in is indeed viable using this technique, provided that enough data points have elapsed prior to the flaring event to result in a trustworthy GP fit[2].

Figs. 4.7 and 4.8, although they display the results of real ZTF lightcurves with injected flares rather than simulated lightcurves, show that, as one would expect, greater amplitude flares are easier to detect and 44 per cent of injected flares (that have been successfully detected) are detected after the flare amplitude has reached half of its maximum height. However, more than half of the injected flares (64 per cent) were not detected using a probability threshold of 0.9. This suggests that a probability of 0.9 (already reduced from 0.99 to take into account the effects of cadence and outliers) is still too stringent for achieving a high flare retrieval rate. Of those flares that were detected, however, 100 per cent were detected before their peak (Fig. 4.8), which is promising if one considers the importance of rapid flare detection for the purpose of follow-up.

The key findings of this chapter are shown in Figs. 4.9 and 4.10. These results are important because they demonstrate two crucial points to consider when using this flare localisation technique: above what flare probability do you consider to be a detection, and what constitutes a *reliable* detection? Fig. 4.9 shows the distribution of maximum flare probability achieved in a sample of known flares[3] compared with a randomly-selected sample of non-flares. From this plot it is clear that a significant proportion of non-flares never exceed a probability of 0.2 and therefore this value can be considered a cut-off point above which is considered to be a flare detection to be flagged for follow-up. This

[2]Quantifying what constitutes a "good" or "bad" GP fit in order to systematically remove poor fits from ones analysis is not straightforward, even when it is obvious in individual cases by human inspection. This is a problem I consider worth investigation and I discuss ways in which this could be explored further in Chapter 6.

[3]*i.e.,* flare candidates identified from the GP analysis in Chapter 3.

probability cut-off can obviously be altered depending on whether one is more interested in the completeness or purity of one's sample. Fig. 4.9, however, shows a significant number of false positives (non-flare lightcurves that achieve a maximum flare probability close to one) and these are investigated in Fig. 4.10. From this plot it is clear that if one relies solely on flare probability to make a detection, this will result in a large number of spurious detections and therefore the second important factor to consider here is the reliability of the detection (in this case, the number of data points elapsed). With this in mind, Fig. 4.10 shows that any flare detections made in a lightcurve in which fewer than 80 data points have elapsed may be false-positives as a result of a bad GP fit.

In Figs. 4.11, 4.12, 4.13 and 4.14 I present the lightcurves of the 27 ZTF flare candidates from Chapter 3 along with their corresponding rolling flare probabilities from the localisation analysis. In many cases, it can be seen how the rolling GP is unstable at early times, resulting in an artificially high flare probability. This reinforces the findings shown in Figs. 4.9 and 4.10, which highlights the importance of taking into account the number of data points elapsed before one can "trust" a flare detection by the localisation analysis. After an unstable period in the beginning of the lightcurve, the flare probability stabilises once the GP has enough data points to generate a reliable fit to the data. Many of the lightcurves do exhibit an increase in flare probability as the lightcurve evolves, but at much lower values than the threshold of 0.9 placed in the case of Figs. 4.7 and 4.8, which confirms my suspicions that this threshold is too stringent in the presence of observational effects such as cadence and outliers. Thirteen of the 27 lightcurves achieve the aforementioned flare probability of 0.2, which was decided to be a suitable cutoff after considering Fig. 4.9.

4.5 Conclusions

The aim of this chapter has been to investigate the feasibility of adapting my GP-based flare detection technique from Chapter 3 – where my analysis resulted in the ability to determine whether or not a given AGN lightcurve contains a flare – to allow for the possibility of flare localisation whereby a flare can be detected in real-time and its location within the lightcurve identified.

In §4.2 I described how I applied the GP to individual lightcurves in a rolling fashion, constantly updating as new data points were added. By performing a straightforward calculation (Eq. 4.2) of the ratio between the likelihoods of the resultant hyperparameters belonging to either the "normal" AGN population or elsewhere, I was able to track the flare probability against time. I used this technique for well-sampled simulated lightcurves with injected flares (§4.3.1), real ZTF lightcurves with injected flares and real unadulterated ZTF lightcurves including the 27 flare candidates identified in Chapter §3 (§4.3.2). I summarise the main results as follows:

1. Flare localisation is highly effective at flare detection in well-sampled, simulated lightcurves and this is supported by the fact that one can witness the lightcurve's hyperparameters move through parameter space in real-time away from the "normal" AGN population (Fig. 4.4 in §4.3.1).

2. In real ZTF lightcurves with injected Gaussian flares, flare localisation is effective at detecting flares before their peak (§4.3.2).

3. Regarding the unadulterated set of ZTF lightcurves, Figs. 4.9 and 4.10 show that a cut-off probability of 0.2 is sufficient to successfully detect AGN flares in this sample but an important caveat is to consider the number of data points elapsed by that point to prevent spurious detections.

4. I present the lightcurves of the 27 ZTF flare candidates from Chapter 3 along

with their corresponding rolling flare probabilities and show that the flare localisation technique is successful at detecting "abnormal" behaviour in many of these lightcurves.

With these key findings in mind, in §4.1 I posed the following questions that I will now address:

1. How can we determine when a flare event actually takes place, *i.e.,* its location within a lightcurve?

2. How early into a flare are we able to detect it?

To answer question 1, after analysing the rolling probabilities of both non-flaring ZTF lightcurves and flare candidate ZTF lightcurves (Fig. 4.9), one must choose a probability cut-off above which a detection will be defined. This cut-off is subjective to the needs of the intended science case, *i.e.,* whether completeness or purity is preferred, and in this chapter I defined this to be 0.2 based on my sample.

Question 2 can be addressed by considering Figs. 4.7 and 4.10. In Fig. 4.7 it is clear that higher-amplitude flares are easier to detect but a significant number of flares can be retrieved before their peak. This is heavily caveated, however, by Fig. 4.10, which demonstrates the importance of taking into account the number of data points elapsed before detection and therefore the reliability of the GP fit, which I have determined from my sample to be after 80 data points.

As such, I have shown that GP analysis in a rolling fashion to achieve flare localisation is effective at detecting flares in real-time and is a viable option for transient detection in the advent of high-cadence surveys such as the LSST.

Chapter 5

Expanding the sample with the Million Quasars Catalogue

5.1 Introduction

Up to this point in the thesis, the data that has been used for the flare detection techniques in Chapters 2, 3 and 4 has been the ALPAKA catalogue (see Chapter 2); a sample of 9035 spectroscopically-selected Type 1 AGN from SDSS (ALPAKA: Mullaney et al., 2013). This sample has been large enough to investigate the feasibility of using Gaussian Processes for flare detection and to unveil potential flare candidates (see Chapter 3). Taking stock of the work preceding this current chapter, I have demonstrated that the following is true for a sample of AGN lightcurves of order the size of the ALPAKA catalogue.

- To disentangle the typical stochastic variability of AGN from a transient signal (i.e. flare), a statistically robust detection technique is required (Chapter 2).
- Gaussian Processes are a feasible means of determining if a given lightcurve *contains* a flare (Chapter 3).

- Gaussian Process analysis is also capable of identifying exactly when a flare occurs within its lightcurve (Chapter 4).

The ALPAKA sample is not, however, comparable to the sheer amount of data that will be made available by LSST in the coming years, with estimates predicting over 10 million transient alerts per night, and certainly millions of observable AGN (Ivezić et al., 2019, see Chapter 1). If one considers these unprecedented amounts of data, combined with the fact that flares are generally estimated to be extremely rare (for example, van Velzen 2018 place an estimate of the flare rate to be approximately $10^{-4}-10^{-5}\mathrm{gal}^{-1}\mathrm{yr}^{-1}$), it is clear that my GP analysis would benefit from a much larger sample size of the order one million. This larger sample size would (a) allow for an investigation as to whether GP analysis (as in Chapter 3) can handle this volume of data without being too computationally expensive, (b) provide a much larger sample of objects to increase the chances of identifying potential flares and (c) provide the added benefit, depending on the properties of the sample, of probing different regions of parameter space and hence AGN properties. For this purpose, I chose to perform the GP analysis outlined in Chapter 3 on the Million Quasars Catalogue (MILLIQUAS: Flesch, 2023), a sample of 907 144 Type 1 quasars and AGN which I describe in more detail in §5.1.1.

The outline of this chapter is as follows: in §5.2 I describe the methodology used to analyse MILLIQUAS with my GP flare detection technique, before describing the results in §5.3. Then I discuss my findings in §5.4 and compare the results with those found analysing the ALPAKA catlogue in Chapter 3 and present concluding remarks in §5.5.

5.1.1 The Million Quasars Catalogue

The Million Quasars Catalogue developed by Flesch (2023) is a compendium of 907 144 Type 1 quasars and AGN which is available online[1]. 66 026 QSO candidates are also

[1]https://heasarc.gsfc.nasa.gov/W3Browse/all/milliquas.html

included by radio or X-ray association. Blazars and Type 2 objects are also included, bringing the total count to 1 021 800 objects. The sample is heterogeneous (*i.e.,* an accumulation of most AGN detections in the literature up to June 2023) and is made up of optical, spectroscopically selected objects, mostly selected based on broad lines. A significant proportion of these AGN are from SDSS.

5.2 Methodology

To analyse the MILLIQUAS catalogue, their optical r-band lightcurves were downloaded from ZTF Data Release 23 which was made available in October 2024 (ZTF: Bellm et al., 2019; Graham et al., 2019). A total of 964 280 lightcurves of the above 1 021 800 were available and downloaded. Following this, Gaussian Process regression was used to determine the optimum variability parameters (hyperparameters) of each individual lightcurve. This technique is identical to that used in §3.4 in Chapter 3 but to remind the reader, I will briefly reiterate the method. The *prior* used in this case is a covariance matrix called a kernel; specifically, I use a Matérn-3/2 kernel which describes stochastic variability and is representative of the random nature of AGN variability (Griffiths et al., 2021). The covariance of this kernel is outlined in Equation 3.1

Exactly as described in §3.4 of Chapter 3, a GP with the above kernel is fit to each individual lightcurve in MILLIQUAS. The GP does this by calculating all of the possible functions that can both (a) describe the data while taking its error bars into account, and (b) be feasibly produced by the underlying kernel function (*i.e.,* functions that are representative of stochastic variability). From these calculations, a "best fit" mean function can be generated, which is determined to be the most likely fit given the data; the current hyperparameters of the kernel (σ and ρ from Eq. 3.1) that have generated that best fit function are returned. Each individual lightcurve therefore has its own unique set of hyperparameters that describe how the lightcurve is varying and

these hyperparameters can be thought of as a summary statistic of the data. For each of the 964 280 MILLIQUAS lightcurves downloaded, I used the GP with this Matérn-3/2 kernel to produce the optimised hyperparameters in each case and in the following section I describe the resulting hyperparameter distributions found.

5.3 Results

In this section I present the optimised hyperparameters returned by the GP for the MILLIQUAS catalogue and I compare them with the distributions of the ALPAKA catalogue. Figure 5.1 shows the optimised hyperparameters σ and ρ from Eq. 3.1 of the MILLIQUAS catalogue. There is significant variation across both axes, with the characteristic timescale $\log(\rho/\text{day})$ varying from 8 to -8. The figure is dominated by objects with $\log(\sigma/\text{mag})$ between -5 and 0. There is also a significant proportion of objects residing within the left hand side of the plot, indicating that these lightcurves show very low amplitudes of variability where $\log(\sigma/\text{mag}) <$ -6. Objects in the far right hand side of the plot beyond values of $\log(\sigma/\text{mag}) > 1$ indicate extreme variability.

Fig. 5.2 shows the hyperparameters of the MILLIQUAS catalogue with the hyperparameters of the ALPAKA catalogue for comparison. The distributions exist in similar regions of parameter space, the major differences being that the ALPAKA distribution sits slightly further left in the plot compared to the central MILLIQUAS distribution, indicating smaller amplitude variations and the MILLIQUAS distribution peaks at slightly longer timescales $\log(\rho/\text{day})$. Considering the MILLIQUAS sample contains quasars, which are known to be a more luminous subtype of AGN (Peterson & Burbidge, 1998), of the order 100 times brighter than typical AGN and with lower amplitudes of variability (Ulrich et al., 1997), this initially seems surprising, but it is important to note that MILLIQUAS does not comprise of only quasars and it also contains Type 1 AGN; the comparison is not necessarily valid. Also, it is possible that the sheer size of the

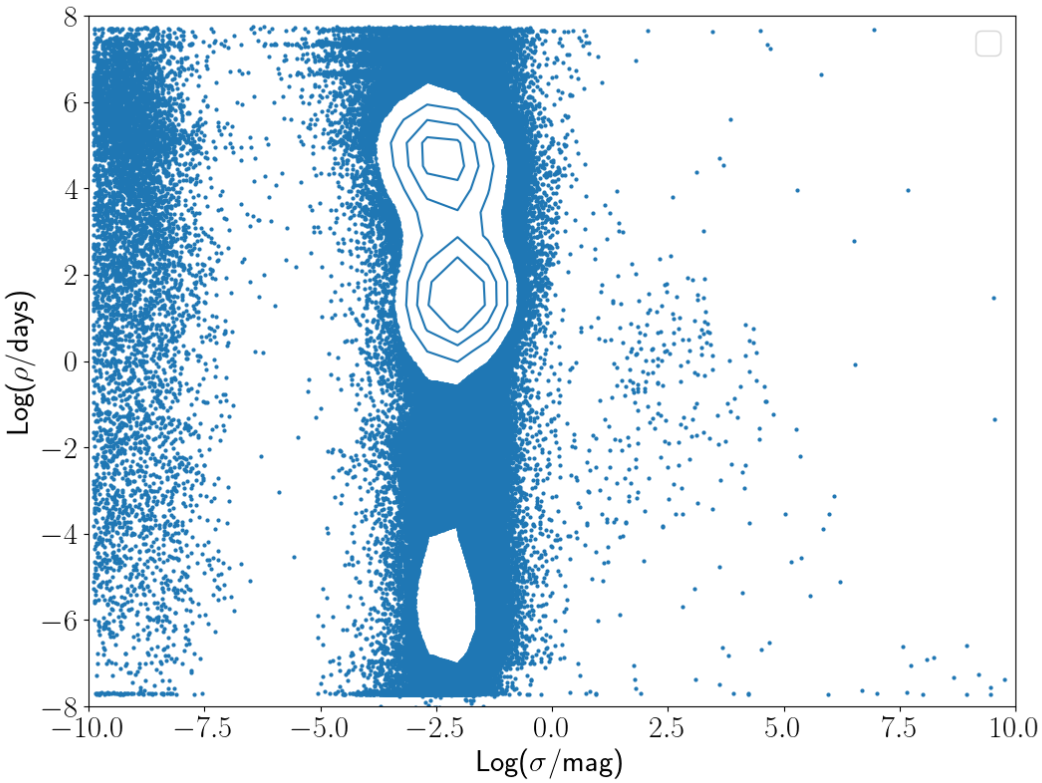

Figure 5.1: Distribution of the optimised hyperparameters σ and ρ for the Million Quasars Catalogue, excluding any objects with less than 30 data points. The contours are representative of the density of the data points.

MILLQUAS sample means that it covers the full range of possible parameter space, and that the ALPAKA sample merely subsamples that same parameter space. To investigate whether this is the case, I randomly downsampled MILLIQUAS to the size of ALPAKA to make for a fairer comparison and this is shown in Fig. 5.3. The difference in mean variability amplitude is still evident post-downsampling and is not simply a symptom of the large difference in sample sizes skewing the mean. The other notable difference in the two distributions is the presence of a number of MILLIQUAS objects in the far left side of the plot with $\log(\sigma/\mathrm{mag}) < -6$, which is indicative of very low variability amplitudes. This demonstrates that this population of objects is a feature of the MILLIQUAS distribution that is not present in the ALPAKA sample. On closer inspection this is a result of bad GP fits, often as a result of chaotic or extremely poor cadence (*e.g.,* lightcurves with very few data points). This indicates that the GP analysis is still very much susceptible to unreliable or sporadic cadence and this needs to be taken into account before trusting the resulting hyperparameter values, or corrected for as a data preprocessing step.

In Chapter 3, flare candidates were identified by combining the requirements that: (a) the lightcurve had a probability of containing a flare of above 0.1 as a result of Bayesian hypothesis testing (see §2.5 of Chapter 3) and (b) that the value of $\log(\sigma/\mathrm{mag})$ was greater than -2 (to ensure that the flare candidate is indeed in the region of parameter space where amplitudes are greater). Since in this current chapter I have not applied Bayesian hypothesis testing[2] to the MILLIQUAS dataset and instead obtained only the optimised hyperparameters of the lightcurves, a probabilistic cutoff for flare candidate selection is not possible. The initial flare selection criterion of $\log(\sigma/\mathrm{mag})$ was to be set the same as that in Chapter 3 as greater than -2, but this resulted in too many flare candidates to investigate (291 249 out of 514 333, corresponding to 57 per cent of the sample).

[2]The Bayesian hypothesis testing undertaken in Chapter 3 utilises Markov-Chain Monte Carlo and is therefore slow and computationally expensive. Later, in Chapter 4, I show that the technique is not needed and a more straightforward, faster calculation can be made as a proxy for flare probability.

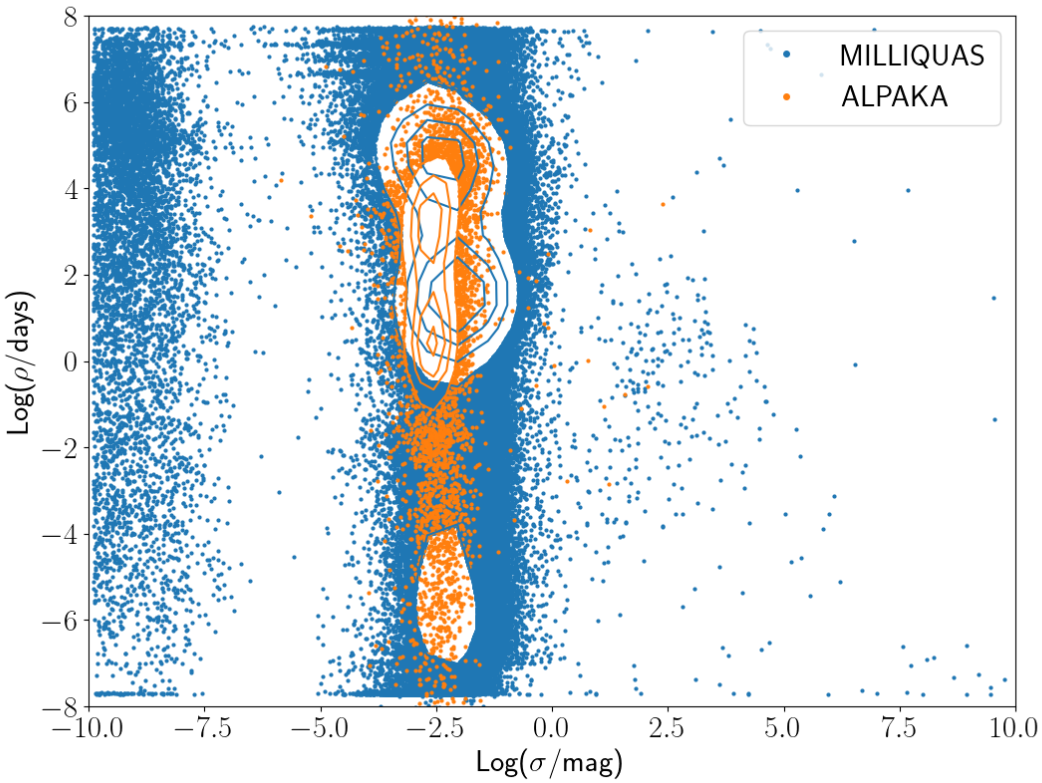

Figure 5.2: Distributions of the hyperparameters σ and ρ for the Million Quasars Catalogue (blue) and the ALPAKA catalogue (orange). The contours are representative of the density of the data points.

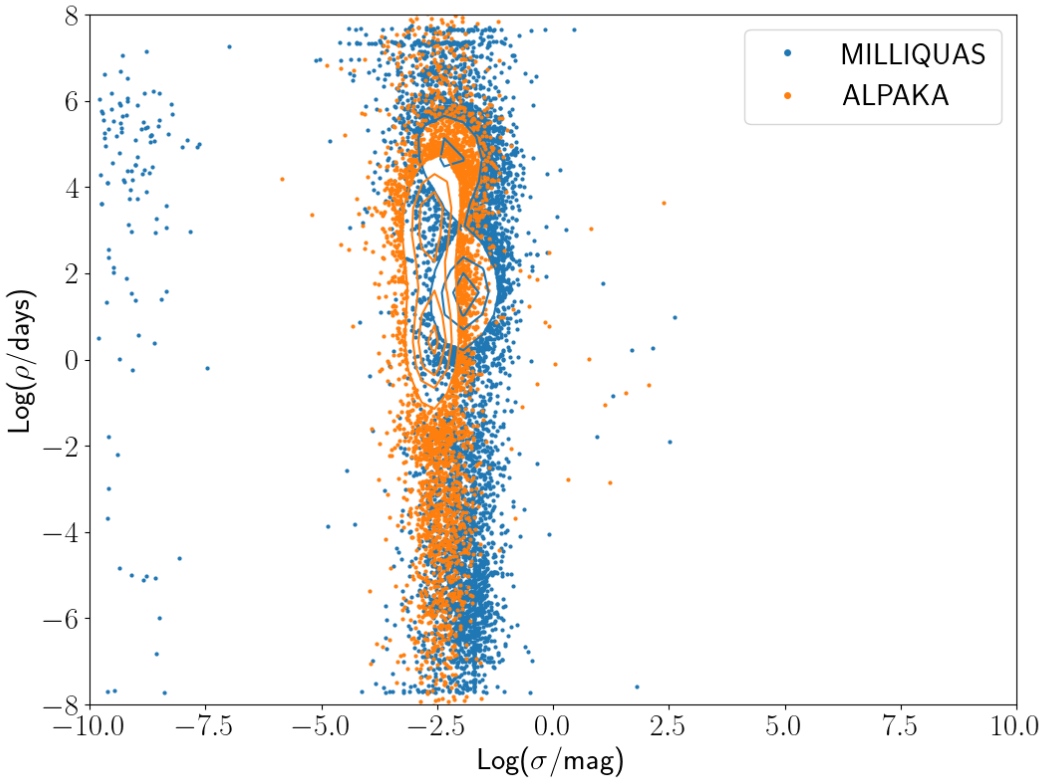

Figure 5.3: Distributions of the hyperparameters σ and ρ for the Million Quasars Catalogue (in blue, which has been randomly downsampled to the same size as the ALPAKA catalogue) and the ALPAKA catalogue (orange). The contours are representative of the density of the data points.

Instead, the criterion on $\log(\sigma/\mathrm{mag})$ was set to a value that resulted in a comparable fraction of flare candidates than that achieved using a cutoff of $\log(\sigma/\mathrm{mag}) > -2$ for ALPAKA (which is 11 per cent). The resulting choice for $\log(\sigma/\mathrm{mag})$ leading to 11 per cent of the MILLIQUAS sample to be defined as flare candidates is $\log(\sigma/\mathrm{mag}) > -1.3$. With this criterion, 58 100 potential flare candidates were identified. The lightcurves of a randomly-selected sample of 30 of 58 100 these objects is shown in Figs. 5.4, 5.5, 5.6, 5.7 and 5.8. This sample of 30 flare candidates is comparable in sample size to the 27 flare candidates identified by the analysis in Chapter 3. On inspection it is clear that whilst some of these lightcurves are indeed highly variable, there is a significant number of lightcurves that are not significantly variable by AGN standards (*i.e.,* a few tenths of a magnitude: Sesar et al., 2007) and a number of lightcurves contain potential outliers, which are likely due to observational effects rather than a flare, since they only affect one or two data points.

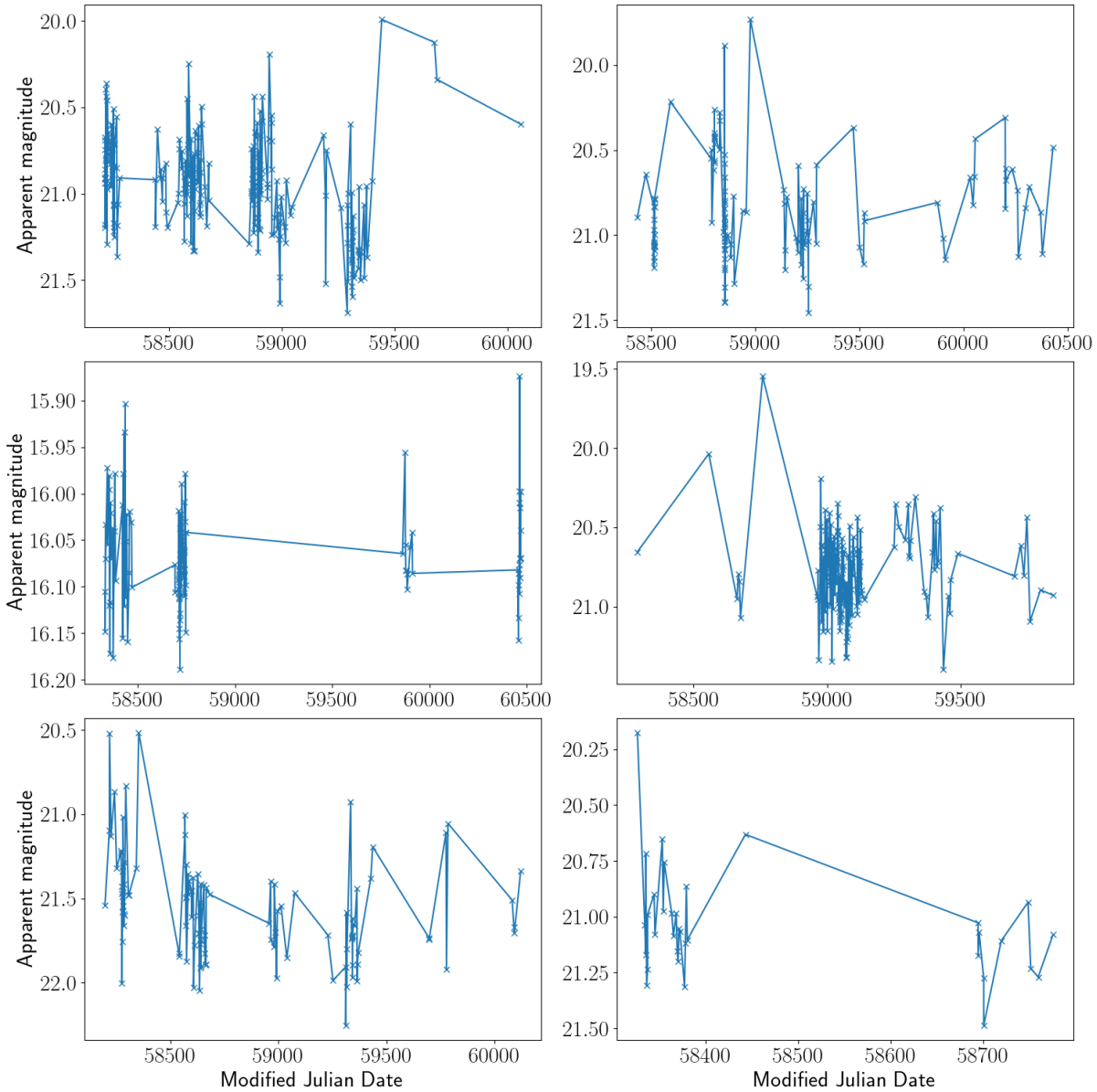

Figure 5.4: ZTF r-band lightcurves of the first six out of the total 30 flare candidates identified from the Million Quasars Catalogue using a cutoff of $\log(\sigma) > -1.3$.

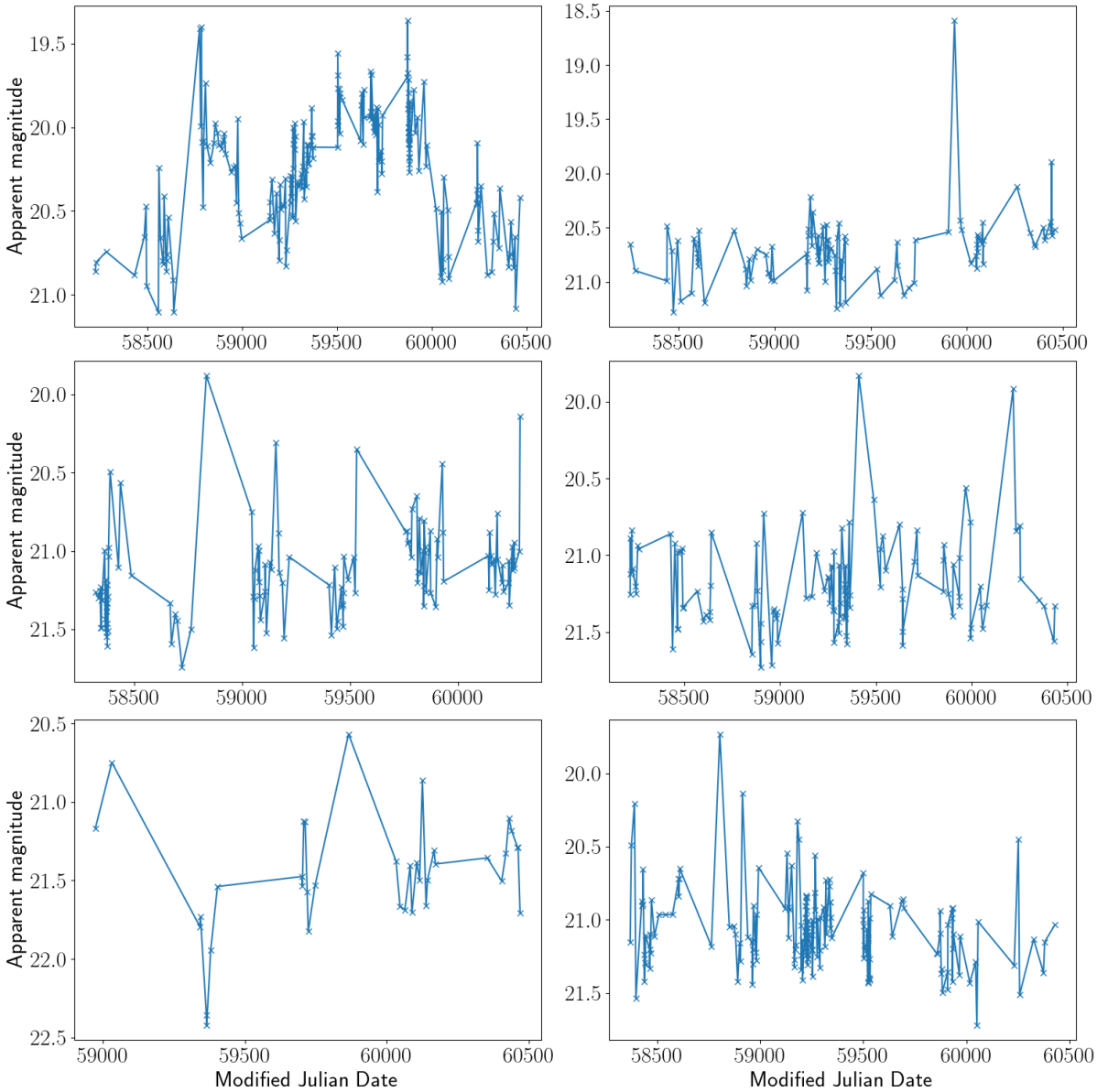

Figure 5.5: ZTF r-band lightcurves of the 7th–12th out of the total 30 flare candidates identified from the Million Quasars Catalogue using a cutoff of $\log(\sigma) > -1.3$.

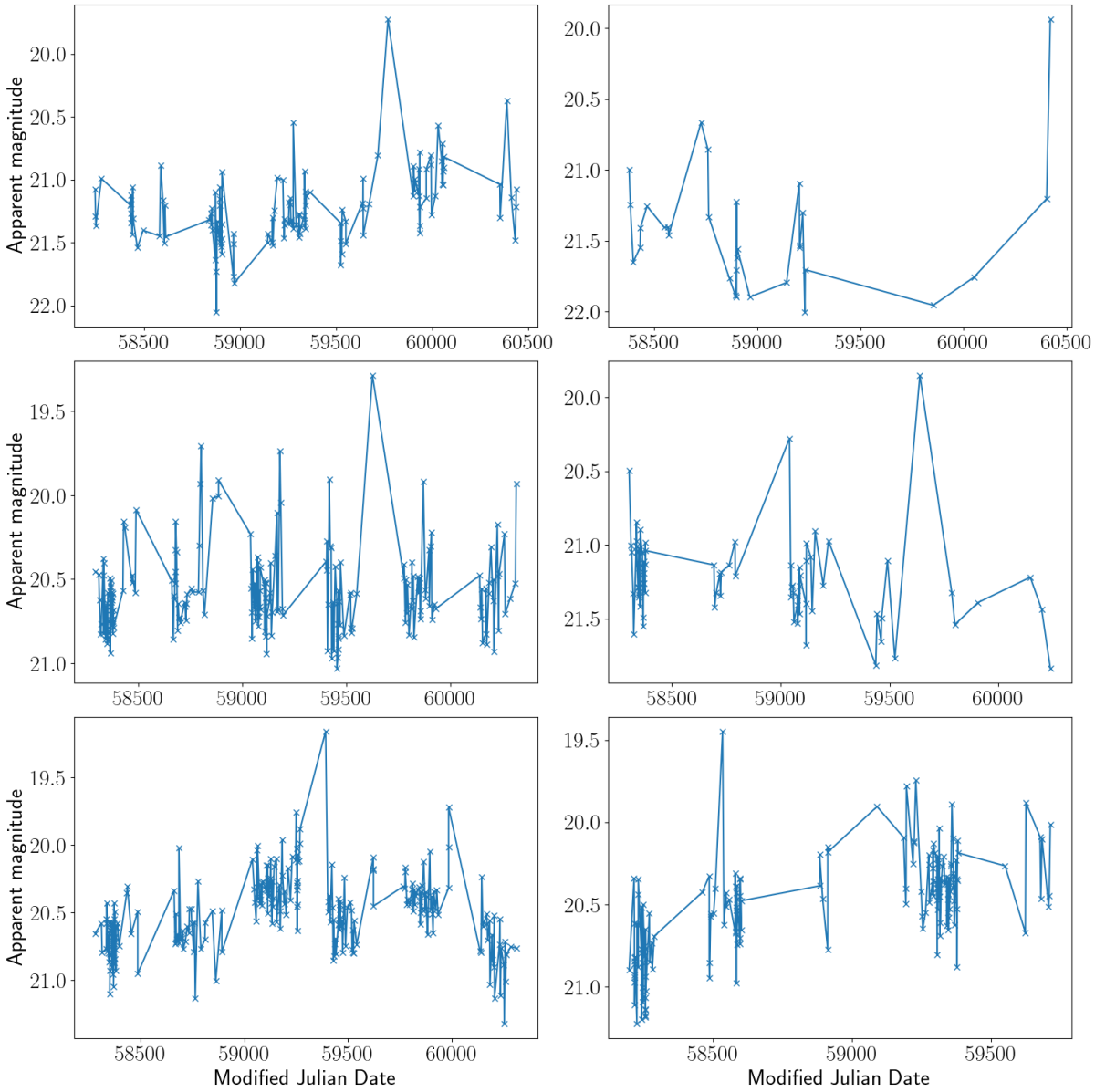

Figure 5.6: ZTF r-band lightcurves of the 13th–18th out of the total 30 flare candidates identified from the Million Quasars Catalogue using a cutoff of $\log(\sigma) > -1.3$.

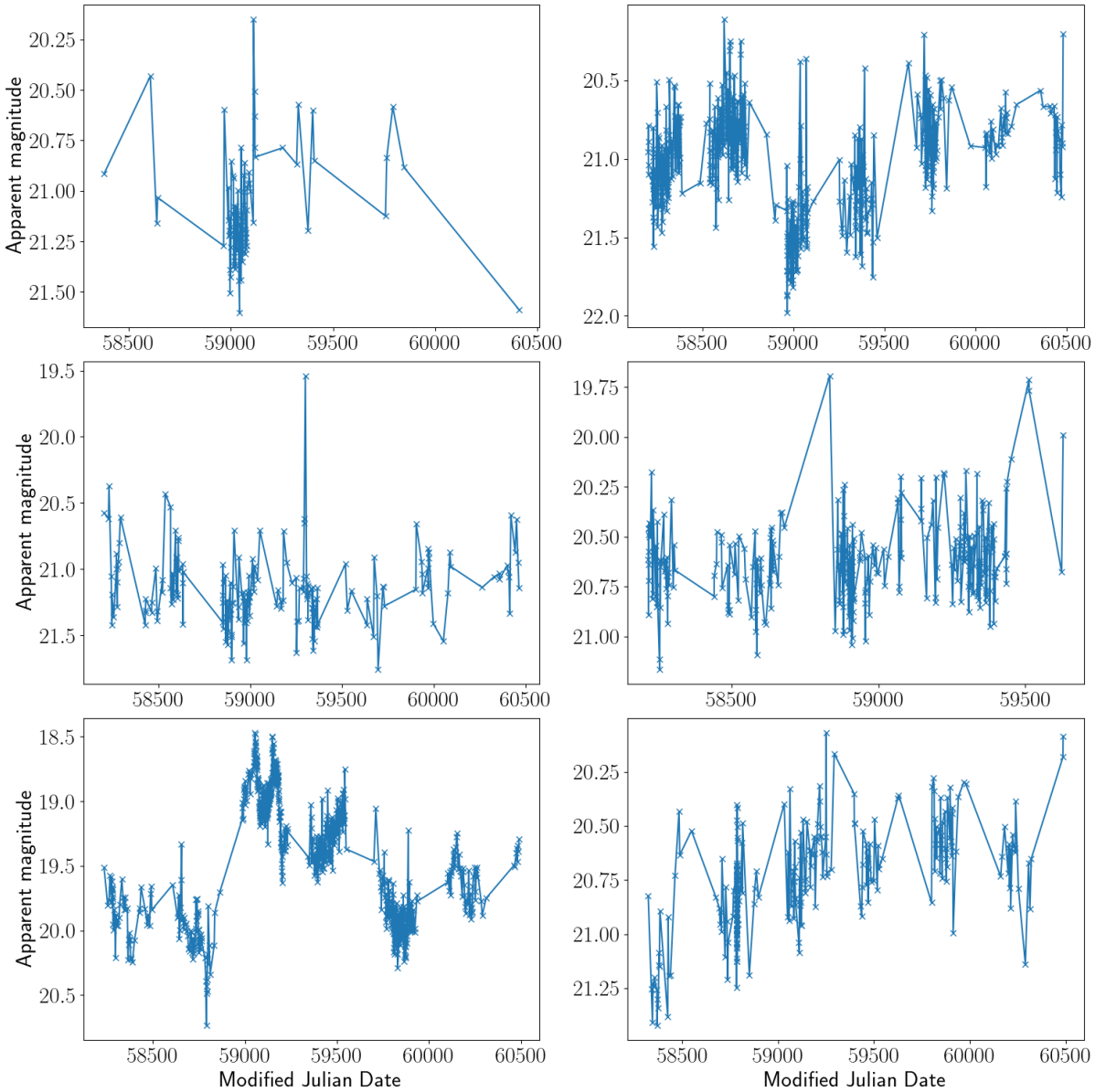

Figure 5.7: ZTF r-band lightcurves of the 19th–24th out of the total 30 flare candidates identified from the Million Quasars Catalogue using a cutoff of $\log(\sigma) > -1.3$.

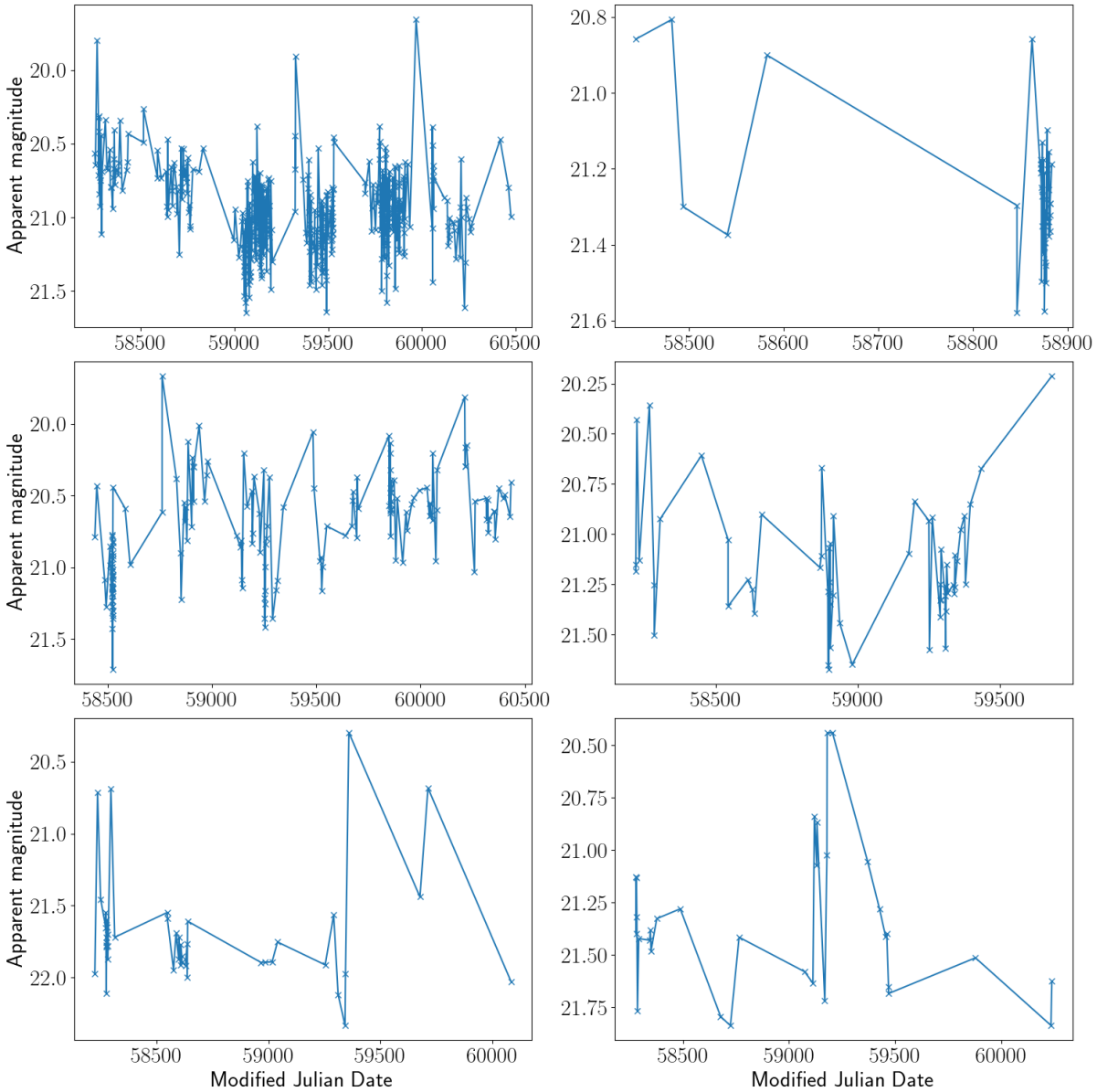

Figure 5.8: ZTF r-band lightcurves of the 25th–30th out of the total 30 flare candidates identified from the Million Quasars Catalogue using a cutoff of $\log(\sigma) > -1.3$.

5.4 Discussion

The hyperparameter disribution of the MILLIQUAS catalogue exists in a similar region of parameter space as the ALPAKA catalogue (Fig. 5.2). The key differences are that the mean of the central MILLIQUAS distribution is further to the right of that of the ALPAKA distribution, indicating that the amplitudes of variability are greater on average. Since MILLIQUAS is made up of both quasars and Type 1 AGN, conclusions can not necessarily be drawn about how this corresponds to the fact that quasars are generally less variable than other types of AGN (Ulrich et al., 1997). As shown in Fig. 5.3, this difference is not resolved by downsampling MILLIQUAS to the same size as ALPAKA to account for the stark difference in sample size, and this difference in mean is a product of the data. The second difference is the presence of a number of MILLIQUAS objects in the far left region of the hyperparameter distributions where $\log\sigma/\mathrm{mag} < -6$ indicating very low levels of variability. This was investigated and found to be a result of irregular or sparse cadence, leading to poor GP fits. This indicates that the GP analysis is still very much susceptible to unreliable or sporadic cadence and this needs to be taken into account before trusting the resulting hyperparameter values. It is curious that the mean amplitude of the MILLIQUAS distribution is greater than that of ALPAKA, given the understanding that quasars are less variable than standard AGN (*e.g.,* Kelly et al., 2009); this analysis could be enriched by removing the AGN from MILLIQUAS, and comparing a sample of pure MILLIQUAS quasars to ALPAKA to see if this is still the case and the implications.

The lightcurves of the 30 flare candidates (Figs. 5.4–5.8) overall do not show particularly compelling evidence of flare activity. There are a few examples of potential extreme variability, but the majority are typical of AGN variability and contain observational outliers, which have been flagged by the GP as extreme. Furthermore, the majority of these lightcurves are approaching the ZTF r-band magnitude limit (20.8:

see Bellm et al., 2019) and are potentially unreliable – this may explain why they have been erroneously flagged as flares by the GP analysis (see §1.4.1 for caveats of forced photometry). To put these findings into perspective, using the retrieval rates calculated in Chapter 3 where a false positive rate of 6.5 per cent was found (see Fig. 3.14), it is not surprising at this stage that a sample of 30 flare candidates out of 58 100 lightcurves did not contain any viable flares. This is reinforced by the estimated incidence of real flares (*e.g.,* $10^{-4} - 10^{-5}\mathrm{gal}^{-1}\mathrm{yr}^{-1}$: van Velzen, 2018). What these findings do suggest, however, is that more stringent quality control is required to ascertain a purer sample of flares. To investigate this, I placed a requirement on $\log(\sigma/\mathrm{mag})$ to be between $\pm$ 0.5, in an attempt to remove lightcurves with extreme hyperparameters as a result of bad GP fits. This resulted in a total of 844 candidates. Again, however, a significant proportion of these lightcurves were false positives due to outliers, cadence effects and bad GP fits. Through human inspection, I identified that 38 out of the total 844 lightcurves contained potential flare activity. I present the lightcurves of eight examples out of the total 38 that show potential extreme variability or flare activity in Figs. 5.9 and 5.10. This shows that there *are* identifiable flare candidates using this technique, but they are overshadowed by the sheer number of false positives if quality control is not applied *a priori*. This quality control should include the rejection of: spurious GP fits, poorly sampled lightcurves, and lightcurves close to the ZTF r-band magnitude limit of 20.8. The identification of 38 flare candidates out of a sample of 844 lightcurves corresponds to a purity of 4.5 per cent, which is a significant improvement from the rough estimate of the purity rate before quality control measures ($< 30/58\,100 \approx 0.05\%$).

It is clear, as above, that this technique is subject to the following limitations: poorly-sampled lightcurves (with few data points), sporadically-sampled lightcurves (with large gaps in the data), observational outliers, and unreliable data as a result of forced photometry upper limits. Each of these issues potentially result in poor GP fits and hence spurious hyperparameters. However, upon applying quality control measures, the purity

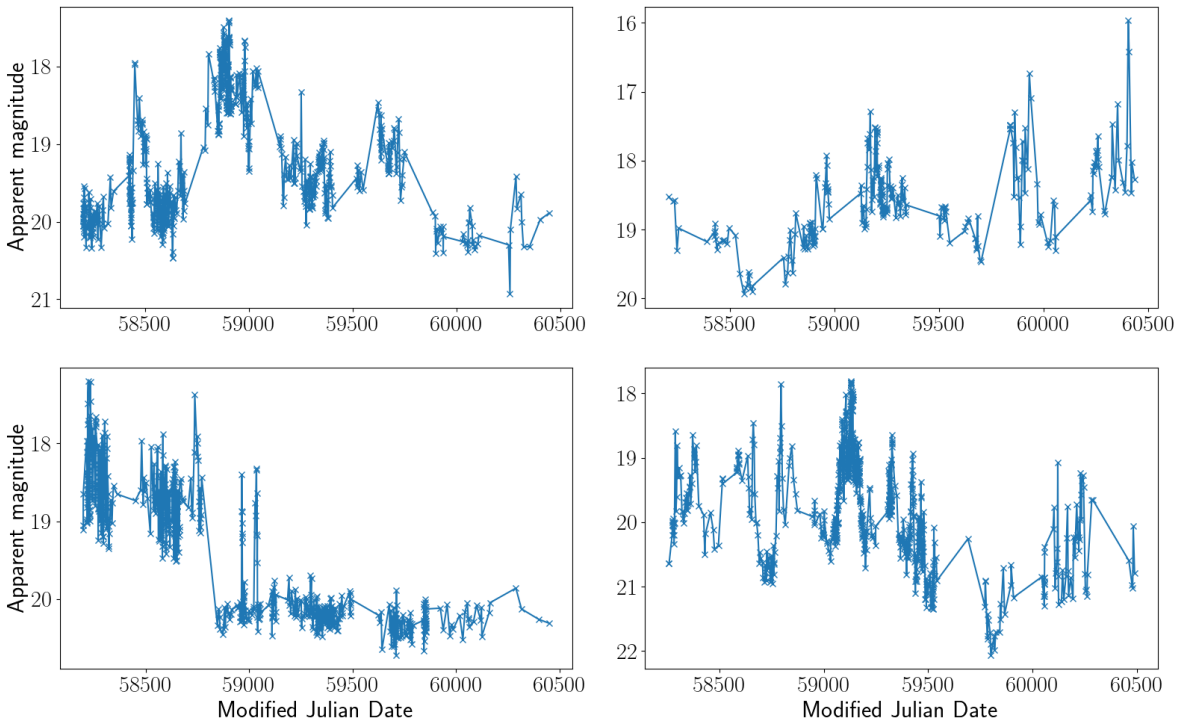

Figure 5.9: The first four examples of the ZTF r-band lightcurves of the flare candidates identified from the Million Quasars Catalogue setting a requirement on $\log(\sigma/\mathrm{mag})$ to lie between ± 0.5.

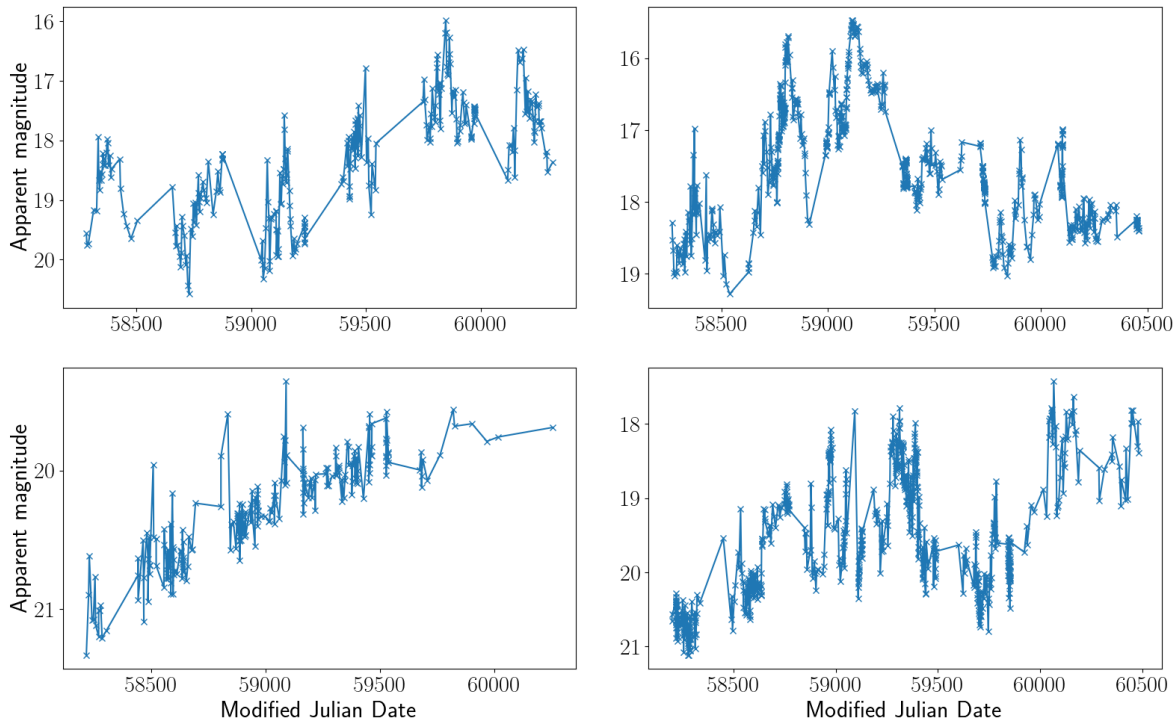

Figure 5.10: A second sample of four examples of the ZTF r-band lightcurves of the flare candidates identified from the Million Quasars Catalogue setting a requirement on $\log(\sigma/\mathrm{mag})$ to lie between ± 0.5.

of the final sample of flare candidates can be significantly improved. The visual inspection of the 844 lightcurves was feasible for one person to undertake in two hours, which means that human inspection is a viable tool for further flare classification. Indeed, this process could be scaled accordingly for proportionately larger sample sizes by utilising *e.g.,* citizen science. The true positive rate achieved (4.5 per cent) could be further improved by more stringent quality control measures as mentioned above, such as rejecting spurious GP fits, poorly sampled lightcurves, and lightcurves close to the ZTF r-band magnitude limit. In terms of improving the GP analysis to be resistant against outliers, this is less straightforward. Data smoothing or interpolation techniques have the potential to reduce the impact of individual outliers on the overall GP fit, but these techniques must be carefully applied to take into account the stochastic nature of AGN variability itself. Further suggestions on improving GP fits and finding ways to filter out poor fits, and other quality control measures are discussed in more detail in Chapter 6.

5.5 Conclusions

In this chapter I utilised the GP analysis from Chapter 3 on a much larger sample than previous (964 280 AGN lightcurves compared to 9035). I used the Million Quasars Catalogue (Flesch 2023: referred to as MILLIQUAS throughout this Chapter) to achieve this, whereby I fit a GP with a kernel function describing AGN variability (Matérn-3/2: see Eq. 1.4) to each individual lightcurve in the catalogue and analysed the resulting optimised hyperparemeters that summarise the lightcurve's variability.

I confirmed the findings of Chapter 3 in that I have shown that GP analysis in this capacity is a suitable means of encapsulating the variability of an AGN lightcurve in a statistically robust manner; to this end I found that the sample of Type 1 AGN from ALPAKA exist in a similar region of parameter space as the MILLIQUAS sample.

Upon identifying flare candidates, since in this chapter I did not utilise the Bayesian

hypothesis testing technique from Chapter 3 to assign a probability to each lightcurve, I placed a selection upon $\log(\sigma)$ to identify a similar proportion of flare candidates as was found in Chapter 3. This resulted in 58 100 potential flare candidates and it is therefore not feasible to inspect all of these; instead, I randomly selected a sample of 30 flare candidates from this sample to compare with those found in Chapter 3. This sample was found to be infiltrated by a large number of false-positives (the majority of the lightcurves not exhibiting flare activity or extreme variability), many as a result of outliers due to observational effects. GP analysis in this capacity is still susceptible to outliers and unpredictable cadence.

Chapter 6

Conclusions and future work

6.1 Thesis conclusions

In this thesis, I have demonstrated the importance of using a statistically robust technique for systematically detecting AGN flares in their already stochastically varying lightcurves. The thesis has described the development of a statistical means of flare detection that utilises Gaussian Processes to distinguish between flare activity and "normal" variability using a summary statistic of how each lightcurve is varying over time. This technique has been developed with the aim to be a viable means of sifting AGN flares from the large amounts of data that the community will face in the era of time-domain astronomy and upcoming surveys such as the LSST (Ivezić et al., 2019). In Chapter 2, I took a "first principles" approach to systematically detect AGN flares in their lightcurves by subtracting the Theil-Sen median and then identifying candidates with contiguous sets of points below this median. In Chapter 3, I made the case for changing approach to a statistically robust technique called a Gaussian Process and I embarked on a feasibility study to determine the viability of this method by using simulated lightcurves. Subsequently, I tested this technique on real AGN lightcurves from ZTF. In Chapter 4, I expanded on the technique from Chapter 3 by using GPs to analyse lightcurves in a

rolling fashion in order to identify the location within the lightcurve at which the flare is occurring. Finally, in Chapter 5 I used the original GP technique from Chapter 3 to analyse a much larger sample of AGN known as the Million Quasars Catalogue for the purpose of determining whether the technique is suitable for handling such volumes of data as will be required for surveys such as LSST.

In the following sections §6.1.1-6.1.4, I will provide a summary of the major findings in each chapter of the thesis. Then, in §6.1.5 I will comment on the progress made towards answering the open questions discussed in Chapter 1. Finally, in §6.2 I suggest possible avenues of future work that would be a valuable continuation of the work presented in the thesis.

6.1.1 A "brute force" approach to systematic flare detection

In Chapter 2, I used a brute-force approach building on the technique by Graham et al. (2017) to detect potential flares in AGN lightcurves from the ALPAKA catalogue. I took a straightforward approach, which involved calculating the Theil-Sen median of each lightcurve in sections along each lightcurve and subtracting this median from the overall flux. I then searched for sets of contiguous points above this median, which potentially indicate the presence of a flare or extreme variability. This first principles flare-finder algorithm identified a total of 16 flare candidates, 11 Type 1 AGN and 5 Type 2 AGN. Despite this, the technique was found to result in a significant number of false positives and is not statistically robust; in order to disentangle a flare from routine AGN variability a statistical approach is required.

6.1.2 Using Gaussian Processes to detect AGN flares

In Chapter 3 I undertook a feasibility study to determine whether the use of Gaussian Processes is viable for the detection of AGN flares in optical lightcurves. I used both

simulated and real AGN lightcurves fitted with a GP as a means to investigate how the distributions of resulting hyperparameters change as a result of each lightcurve's underlying variability (*e.g.,* upon injection of a simulated flare). I found from the analysis of simulated lightcurves with and without injected flares that their hyperparameters exist in different but partially overlapping regions of parameter space, indicating that whilst there will always be overlap between these distributions it is possible to differentiate between flares and non flares in this way. After discovering this, I was able to use the hyperparameters returned by the GP to distinguish between these distributions by assigning each lightcurve a probability of containing a flare based on its location in parameter space. I found that whilst observational effects are a hindrance to GP analysis in this capacity, *i.e.,* irregular cadences and outliers, the GP analysis was effective at distinguishing between real AGN lightcurves and the same lightcurves with injected flares with a true positive rate of 80 per cent and a false positive rate of only seven per cent. Using the same probability analysis, this technique resulted in 27 out of 9035 AGN lightcurves being flagged as potentially containing a flare. On balance, I found that the use of GPs is a suitable means of flare detection in AGN lightcurves, notwithstanding the fact that observational outliers and irregular cadence can reduce the ability for GPs to distinguish between typical and extreme variability.

6.1.3 Flare localisation within a lightcurve

Since the analysis involved in Chapter 3 enabled the determination of whether a given lightcurve *contains* a flare, it was not at that point able to calculate the *location* within the lightcurve at which the flare occurs. With this in mind, Chapter 4 sought to improve on this technique to achieve real-time flare detection, which I termed flare localisation. Instead of fitting a GP to each lightcurve in its entirety, I fitted a GP to each lightcurve in a rolling fashion, continuously adding data points one by one and updating the resultant

hyperparameters accordingly. This then enabled for the calculation of a proxy of the flare probability and I was able to trace the flare probability against time, identifying the point at which the lightcurve's GP hyperparameters transition from representing typical to extreme variability and thus the epoch of a flare. I found that this technique was highly effective at flare localisation within simulated lightcurves and is able to detect injected flares before they peak in real AGN lightcurves. I analysed the 27 flare candidates identified in Chapter 3 and found that flare localisation in this way was effective at identifying the point at which abnormal or extreme variability occurs.

6.1.4 Expanding the sample with the Million Quasars Catalogue

The real data involved in the preceding chapters of this thesis has been the ALPAKA catalogue, which is a sample of 9035 Type 1 AGN lightcurves. In the coming era of time domain astronomy especially with high-cadence surveys such as LSST planned to see first light in 2025, astronomers will have access to unprecedented volumes of data of order millions of objects per night (LSST: Ivezić et al., 2019). In Chapter 5 therefore, I applied the GP analysis from Chapter 3 to a much larger sample – the Million Quasars Catalogue (Flesch, 2023), which is a compendium of 907 144 Type 1 quasars and AGN to determine whether the analysis can handle such a sample size. I found that the analysis can certainly handle a sample size of order one million, without being too computationally expensive or time consuming. As in Chapter 3 I found that GPs are a viable means of encapsulating AGN variability as a summary statistic. Out of 58 100 potential flare candidates identified in MILLIQUAS, a sample of 30 was presented and it was found to be infiltrated by a large number of false-positives and lightcurves containing outliers. This demonstrated that the analysis is still susceptible (as was found in Chapter 3) to observational outliers and irregular cadence, which may be an issue when attempting to detect flares in large amounts of data.

6.1.5 The bigger picture

In Chapter 1, I posed the following questions, which I intended to address throughout the duration of the thesis:

1. Can systematic detection techniques distinguish between an AGN flare and typical AGN variability in optical lightcurves?

2. At what point can AGN flares be detected in lightcurves outside of the underlying variability?

3. Can we determine exactly when a flare actually occurs, *i.e.,* its location within a lightcurve?

4. How early into a flare are we able to detect it?

I will now comment and conclude on the progress made in each of these areas.

To address point one, I found that Gaussian Process analysis can distinguish between flare activity and typical AGN variability in optical lightcurves with a high success rate (see Chapter 3) however irregular cadences and observational outliers can reduce the efficacy of this technique.

Regarding point two, the point at which AGN flares can be detected outside of the underlying variability is not straightforward and requires a statistically robust technique. Chapter 3 concerns itself with investigating this question; I found that the ability to retrieve an injected flare from a simulated lightcurve is subject to the amplitude and duration of the flare. In general, flares with amplitudes lower than $\sim$ 1 magnitude become difficult to distinguish from the underlying variability though again this depends on the duration of such a flare. Indeed, in Chapter 2 I found that "normal" AGN variability with amplitudes between 0.5 and 1 mag is uncommon on timescales $\sim$ hundreds of days, but is still possible.

To address the third point, I showed in Chapter 4 that it is indeed possible to determine the point at which a flare occurs within a lightcurve. Using rolling GP analysis, I was able to track flare probability against time for a given lightcurve. From there, one must choose a probability cut-off above which a detection will be defined. This cut-off is subjective to the needs of the intended science case, *i.e.,* whether completeness or purity is preferred, and I defined this to be 0.2 based on my sample.

Finally, regarding point four, working out how early into a flare it can be detected is not straightforward. Since variability itself is stochastic, any selection based on magnitude change is biased. In Chapter 4, I showed that it is more meaningful to calculate a flare probability based on statistically robust GP hyperparameters. I subsequently found that a cut-off probability of 0.2 is sufficient to successfully detect AGN flares in the sample in question, but an important caveat is to consider the number of data points elapsed by that point to prevent spurious detections; I found that the minimum number of data points required to mitigate this was 80, which shows that detections made too early can be unreliable.

Overall, the work undertaken for this thesis has made significant progress in the field of AGN flare detection; this is particularly crucial in the era of high-cadence surveys such as the LSST. The study of AGN variability provides a window into the accretion processes occurring close to the supermassive black hole; processes that are poorly understood (Lawrence, 2018). Indeed, extreme variability such as AGN flare activity is both a potential problem for current accretion theory (*e.g.,* Lawrence, 2018; Wiseman et al., 2025) and also a means of studying the extreme behaviour of the supermassive black hole accretion disk (*e.g.,* Kelly et al., 2009). A much larger sample of AGN flare candidates is required, along with multi-wavelength follow-up observations to further understand the mechanisms behind these events and thus better constrain the physics of accretion.

6.2 Future work

The work undertaken and described in this thesis opens up a variety of important avenues to investigate and I summarise these below:

- **Making GP analysis resistant to outliers and irregular cadence.** Whilst I have shown that GP analysis can be successful at distinguishing AGN flares from typical AGN variability, observational outliers and irregular cadence in lightcurves can create false-positive detections. In this coming era of time domain astronomy whereby astronomers will have access to unprecedented amounts of data especially due to upcoming surveys such as LSST, it is important to reduce the false-positive rate of any transient detection algorithm as much as possible, to minimise the reliance on human inspection. It will be extremely beneficial if the GP analysis technique used in this thesis can be expanded upon or improved to render it resistant to such observational effects and reduce the number of false-positives. I suggest that the following is worth investigating to improve bad GP fits as a result of observational effects:

  - *Data preprocessing*: lightcurve smoothing or interpolation could be incorporated to minimise the effect of outliers and sparse cadence. However, it is important to note that any smoothing or outlier removal techniques must take into account the underlying AGN variability. Another potential option worth exploring is artificially increasing the error bars of the lightcurves to give the GP fit more flexibility. Alternatively, a more instrumentally realistic option would be to make an estimate of the added uncertainties due to the use of forced photometry (see §1.4.1), which would accomplish the same flexibility.

  - *Adjusting the GP itself*: switching to different kernels, potentially kernels that

result in smoother function realisations that are less susceptible to "spiky" variations like outliers.

- **Gaussian Process kernels.** The kernel (covariance function that encapsulates the variability of the lightcurve it is fitted to) that has been used in this research is the Matérn-3/2 kernel, but there are other choices of kernel that could reasonably describe the stochastic nature of variability. It would be valuable to investigate whether there are any differences in the ability to distinguish between AGN flares and typical AGN variability when different kernel functions are used. As well as the commonly used kernels such as the squared exponential kernel and the Matérn class of kernel (see Chapter 1 for explanations of kernels), there are also more sophisticated choices of kernel that invoke the use of neural networks, which may or may not improve the efficacy of flare detection and this is worth investigating. Furthermore, it will be useful to undertake an investigation into the range of kernel hyperparameters that circumscribe typical AGN variability. Indeed, this knowledge of the statistical parameters that encapsulate "normal" AGN variability can be utilised in more complex flare detection tools. On top of this, it would be interesting and highly informative to investigate if kernel hyperparameters change with AGN luminosity, host galaxy and black hole mass and redshift.

- **Comparison with other detection techniques.** Other flare and transient detection techniques are widely used, such as Machine Learning techniques (*e.g.,* Faisst et al., 2019; Wright et al., 2015; Sánchez-Sáez et al., 2021). It will be enlightening to compare the efficacy of these other techniques to the GP techniques discussed in this thesis.

- **Preparing for LSST.** Looking forward to first light of the Rubin Telescope planned for 2025, it will be beneficial to use the results found in this thesis to estimate the number of AGN flares one might expect to detect from LSST. Fur-

thermore, when regular data from LSST becomes available it will enable for an in-depth study of parameterising AGN variability, and this unprecedented sample size will provide detectable flare candidates. It will be highly valuable to undertake multiwavelength follow-up observations of flare candidates as they are found, and this will be integral for the study of the accretion processes occurring around supermassive black holes. It will then be possible to place estimates of the rate of AGN flaring events in context of the general population.

- **An end-to-end system to identify flare candidates in LSST.** Once LSST Data Release 1 becomes available (expected by January 2027), it will be possible to use the flare detection technique investigated in this thesis to systematically sift AGN flares from their optical lightcurves. To achieve this, I propose the following end-to-end system:

  1. Acquire a sample of the order ∼10-100 000 optical, r-band lightcurves from LSST Data Release 1. Limit the sample to contain no objects dimmer than 26.8 magnitudes (one magnitude brighter than the stacked r-band magnitude limit, see caveats of forced photometry in §1.4.1). Ensure that all lightcurves contain more than 80 epochs, which I found to be the minimum number of data points required for a trustworthy detection in §5.4.

  2. Use the GP analysis outlined in Chapter 3 to map the hyperparameter values of each lightcurve and assign each with a flare probability.

  3. Apply a threshold probability value of 0.2 (as determined in Chapter 4) to identify flare candidates.

  4. Flag the most promising flare candidates for follow-up observations by human inspection, which can be expedited using citizen science platforms (*e.g.,* Zooniverse: Fortson et al., 2018; Masters & Galaxy Zoo Team, 2020).

5. Request follow-up observations for the confirmation and classification of the detected flare candidates, to better understand the physical drivers of these events. Ideally, this would entail optical spectroscopy, which is a promising means of distinguishing between types of transient event including AGN flares (*e.g.,* Wiseman et al., 2025; Zabludoff et al., 2021). A good candidate for this purpose is the 4-metre Multi-Object Spectroscopic Telescope (4MOST: De Jong et al., 2019), which is a wide field spectroscopic survey instrument at the Visible and Infrared Survey Telescope for Astronomy (VISTA: Emerson et al., 2006; Dalton et al., 2006) telescope at Paranal, Chile.

Bibliography

Aigrain S., Foreman-Mackey D., 2022, Gaussian Process regression for astronomical time-series, doi:10.48550/ARXIV.2209.08940, https://arxiv.org/abs/2209.08940

Aigrain S., Pont F., Zucker S., 2012, Monthly Notices of the Royal Astronomical Society, 419, 3147

Alexander D. M., Hickox R. C., 2012, New Astron. Rev., 56, 93

Angus R., Morton T., Aigrain S., Foreman-Mackey D., Rajpaul V., 2017, Monthly Notices of the Royal Astronomical Society, 474, 2094

Antonucci R., 1993, ARA&A, 31, 473

Astropy Collaboration et al., 2013, A&A, 558, A33

Astropy Collaboration et al., 2018, AJ, 156, 123

Baade W., Minkowski R., 1954, ApJ, 119, 206

Begelman M. C., Silk J., 2017, MNRAS, 464, 2311

Bellm E., Kulkarni S., 2017, Nature Astronomy, 1

Bellm E. C., et al., 2019, PASP, 131, 018002

Bessiere P. S., Ramos Almeida C., Holden L. R., Tadhunter C. N., Canalizo G., 2024, Astronomy & Astrophysics, 689, A271

Blagorodnova N., Van Velzen S., Harrison D. L., Koposov S., Mattila S., Campbell H., Walton N. A., Wyrzykowski Ł., 2016, MNRAS, 455, 603

Blandford R., Meier D., Readhead A., 2019, ARA&A, 57, 467

Bruce A., et al., 2017, Monthly Notices of the Royal Astronomical Society, 467, 1259

Chan C.-H., Piran T., Krolik J. H., Saban D., 2019, The Astrophysical Journal, 881, 113

Collier S., Peterson B. M., 2001, ApJ, 555, 775

Creque-Sarbinowski C., Kamionkowski M., Zhou B., 2021, arXiv e-prints, p. arXiv:2110.13149

Crossfield I. J. M., et al., 2016, The Astrophysical Journal Supplement Series, 226, 7

Dalton G. B., et al., 2006, in McLean I. S., Iye M., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 6269, Ground-based and Airborne Instrumentation for Astronomy. p. 62690X, doi:10.1117/12.670018

Damianou A. C., Lawrence N. D., 2013, Deep Gaussian Processes (`arXiv:1211.0358`)

De Jong R. S., et al., 2019, Published in The Messenger vol. 175, pp. 3-11, March 2019.

Dekany R., et al., 2020, Publications of the Astronomical Society of the Pacific, 132, 038001

Dexter J., Begelman M. C., 2019, MNRAS, 483, L17

Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604

Drake A. J., et al., 2009, ApJ, 696, 870

Drake A. J., et al., 2011, ApJ, 735, 106

Emerson J., McPherson A., Sutherland W., 2006, The Messenger, 126, 41

Fabian A. C., 2012, ARA&A, 50, 455

Faisst A. L., Prakash A., Capak P. L., Lee B., 2019, The Astrophysical Journal Letters, 881, L9

Fath E. A., 1909, Lick Observatory Bulletin, 149, 71

Fausnaugh M. M., Peterson B. M., Starkey D. A., Horne K., t. A. S. C., 2017, Frontiers in Astronomy and Space Sciences, Volume 4 - 2017

Flesch E. W., 2023, The Open Journal of Astrophysics, 6

Foreman-Mackey D., Agol E., Ambikasaran S., Angus R., 2017, The Astronomical Journal, 154, 220

Fortson L., Wright D., Lintott C., Trouille L., 2018, arXiv e-prints, p. arXiv:1809.09738

Frederick S., et al., 2021, ApJ, 920, 56

Gezari S., 2021, Annual Review of Astronomy and Astrophysics, 59, 21

Gezari S., et al., 2017, ApJ, 835, 144

Graham M. J., Djorgovski S. G., Drake A. J., Stern D., Mahabal A. A., Glikman E., Larson S., Christensen E., 2017, MNRAS, 470, 4112

Graham M. J., et al., 2019, PASP, 131, 078001

Graham M. J., et al., 2023, ApJ, 942, 99

Greenstein J. L., Schmidt M., 1964, ApJ, 140, 1

Griffiths R.-R., et al., 2021, The Astrophysical Journal, 914, 144

Guo W.-J., Li Y.-R., Zhang Z.-X., Ho L. C., Wang J.-M., 2022, The Astrophysical Journal, 929, 19

Görtler J., Kehlbeck R., Deussen O., 2019, Distill

Hamilton J. D., 2010, Regime switching models. Palgrave Macmillan UK, London, pp 202–209, doi:10.1057/9780230280830_23, https://doi.org/10.1057/9780230280830_23

Hanbury Brown R., Jennison R. C., Gupta M. K. D., 1952, Nature, 170, 1061

Harrison C. M., 2016, in Observational Constraints on the Influence of Active Galactic Nuclei on the Evolution of Galaxies. https://api.semanticscholar.org/CorpusID:117874932

Harrison C. M., 2017, Nature Astronomy, 1, 0165

Hawkins M. R. S., 2007, Astron. Astrophys., 462, 581

Hawley J. F., Krolik J. H., 2001, ApJ, 548, 348

Ho L., Kormendy J., 2000, in Murdin P., ed., , Encyclopedia of Astronomy and Astrophysics. Institute of Physics Publishing, p. 2365, doi:10.1888/0333750888/2365

Hossein Nouri, Fatemeh Janiuk, Agnieszka 2024, A&A, 687, A184

Ivezić Ž., et al., 2019, ApJ, 873, 111

Kankare E., et al., 2017, Nature Astronomy, 1, 865

Kawaguchi T., Mineshige S., Machida M., Matsumoto R., Shibata K., 2000, PASJ, 52, L1

Kelly B. C., Bechtold J., Siemiginowska A., 2009, ApJ, 698, 895

King A., 2008, New Astron. Rev., 52, 253

Kormendy J., Ho L. C., 2013, ARA&A, 51, 511

Kovacevic N., et al., 2025, Comparing Optical Variability of Type 1 and Type 2 AGN from the BAT 9-Month Sample using ASAS-SN and TESS Surveys (arXiv:2504.08123), https://arxiv.org/abs/2504.08123

Kozlowski S., 2016, ApJ, 826, 118

Kozłowski S., 2017, Frontiers in Astronomy and Space Sciences, 4, 14

Kozłowski S., et al., 2010, ApJ, 708, 927

Kristian J., 1973, ApJ, 179, L61

LaMassa S. M., et al., 2015, The Astrophysical Journal, 800, 144

Lawrence A., 2018, Nature Astronomy, 2, 102

Lawrence A., et al., 2016, Monthly Notices of the Royal Astronomical Society, 463, 296

Liu T., et al., 2018, MNRAS, 479, 5022

Lodato G., Rossi E. M., 2010, MNRAS, 410, 359–367

Lynden-Bell D., 1978, Phys. Scr., 17, 185

MacLeod C. L., et al., 2010, ApJ, 721, 1014

MacLeod C. L., et al., 2012, ApJ, 753, 106

MacLeod C. L., et al., 2019, ApJ, 874, 8

Makrygianni L., et al., 2021, Publications of the Astronomical Society of Australia, 38

Mantha K., et al., 2024, Citizen Science: Theory and Practice, 9, 40

Masci F. J., et al., 2018, Publications of the Astronomical Society of the Pacific, 131, 018003

Masci F. J., et al., 2023, A New Forced Photometry Service for the Zwicky Transient Facility (arXiv:2305.16279), https://arxiv.org/abs/2305.16279

Mason R. E., 2015, Planet. Space Sci., 116, 97

Masters K. L., Galaxy Zoo Team 2020, in Valluri M., Sellwood J. A., eds, IAU Symposium Vol. 353, Galactic Dynamics in the Era of Large Surveys. pp 205–212 (arXiv:1910.08177), doi:10.1017/S1743921319008615

Matérn B., 1966, Spatial Variation; Stochastic Models and Their Application to Some Problems in Forest Surveys and Other Sampling Investigations. Stockholm. Statens Skogsforskningsinstitut. Meddelanden, University of Sweden, https://books.google.co.uk/books?id=HTWTwgEACAAJ

Matthews T. A., Sandage A. R., 1963, ApJ, 138, 30

Mattila S., Graham M. J., Kankare E., Kool E., Moriya T., Perez-Torres M., Wyrzykowski Ł., 2019, Proceedings of the International Astronomical Union, 339, 263

McHardy I. M., et al., 2016, Astronomische Nachrichten, 337, 500

McLaughlin S. A. J., Mullaney J. R., Littlefair S. P., 2024, Monthly Notices of the Royal Astronomical Society, 529, 2877–2892

Meusinger H., et al., 2010, A&A, 512, A1

Mullaney J. R., Alexander D. M., Fine S., Goulding A. D., Harrison C. M., Hickox R. C., 2013, Monthly Notices of the Royal Astronomical Society, 433, 622

Nathanail A., Fromm C. M., Porth O., Olivares H., Younsi Z., Mizuno Y., Rezzolla L., 2020, MNRAS, 495, 1549

Netzer H., 2015, ARA&A, 53, 365

Nicholson B. A., Aigrain S., 2022, Monthly Notices of the Royal Astronomical Society, 515, 5251

Osterbrock D. E., 1993, ApJ, 404, 551

Padovani P., et al., 2017, A&ARv, 25, 2

Payne A. V., et al., 2022, arXiv e-prints, p. arXiv:2206.11278

Peterson B. M., 2001, Variability of Active Galactic Nuclei. World Scientific, p. 3–68, doi:10.1142/9789812811318_0002, http://dx.doi.org/10.1142/9789812811318_0002

Peterson B., 2006, The Broad-Line Region in Active Galactic Nuclei. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 77–100, doi:10.1007/3-540-34621-X_3, https://doi.org/10.1007/3-540-34621-X_3

Peterson B. M., 2008, New Astron. Rev., 52, 240

Peterson B. M., Burbidge G., 1998, Physics Today, 51, 67

Peterson B. M., et al., 1999, ApJ, 510, 659

Portillo S. K. N., Speagle J. S., Finkbeiner D. P., 2020, The Astronomical Journal, 159, 165

Press W. H., Rybicki G. B., 1998, ApJ, 507, 108

Rasmussen C. E., Williams Christopher K., 2006, Gaussian processes for machine learning. Vol. 1, Springer

Rees M. J., 1984, ARA&A, 22, 471

Ricci C., Trakhtenbrot B., 2023, Nature Astronomy, 7, 1282

Ricci C., et al., 2025, Exploring the Extremes: A JWST Study of Hyper-Eddington AGN, JWST Proposal. Cycle 4, ID. #8457

Roelants P., 2019, Gaussian processes (1/3) - From scratch, https://peterroelants.github.io/posts/gaussian-process-tutorial/

Ryle M., Smith F. G., Elsmore B., 1950, MNRAS, 110, 508

Salpeter E. E., 1964, ApJ, 140, 796

Sandage A., 1964, ApJ, 139, 416

Sartori L. F., Trakhtenbrot B., Schawinski K., Caplar N., Treister E., Zhang C., 2019, The Astrophysical Journal, 883, 139

Schmidt M., 1963, Nature, 197, 1040

Seeger M., 2004, International Journal of Neural Systems, 14, 69

Sen P. K., 1968, Journal of the American statistical association, 63, 1379

Sesar B., et al., 2007, AJ, 134, 2236

Seyfert C. K., 1943, ApJ, 97, 28

Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337

Shields G. A., 1999, PASP, 111, 661

Smith H. J., Hoffleit D., 1963, Nature, 198, 650

Smith K. W., et al., 2019, Research Notes of the AAS, 3, 26

Somalwar J. J., Ravi V., Dong D., Graham M., Hallinan G., Law C., Lu W., Myers S. T., 2022, ApJ, 929, 184

Soraisam M. D., et al., 2020, ApJ, 892, 112

Sánchez-Sáez P., et al., 2021, The Astronomical Journal, 162, 206

Theil H., 1950, Indagationes mathematicae, 12, 173

Trakhtenbrot B., et al., 2019, Nature Astronomy, 3, 242

Ulrich M.-H., Maraschi L., Urry C. M., 1997, Annual Review of Astronomy and Astrophysics, 35, 445

Urry M. C., Padovani P., 1995, PASP, 107, 803

Villar V. A., et al., 2020, ApJ, 905, 94

Virtanen P., et al., 2020, Nature Methods, 17, 261

Wiseman P., et al., 2025, MNRAS, 537, 2024

Wright D. E., et al., 2015, Monthly Notices of the Royal Astronomical Society, 449, 451

Yang Q., et al., 2023, arXiv e-prints, p. arXiv:2303.06733

Yip C. W., et al., 2009, The Astronomical Journal, 137, 5120

Zabludoff A., et al., 2021, Space Science Reviews, 217

Zackay B., Ofek E. O., Gal-Yam A., 2016, ApJ, 830, 27

Zel'dovich Y. B., Novikov I. D., 1964, Soviet Physics Doklady, 9, 246

Zhang K., Wang T.-G., Gaskell C. M., Dong X.-B., 2013, ApJ, 762, 51

van Velzen S., 2018, ApJ, 852, 72

Appendix A

Appendix

A.1 GP analysis of the latest ZTF data release

The data used throughout this thesis has been the ALPAKA catalogue (see Chapter 2 for more details). These 9035 Type 1 AGN lightcurves were downloaded in August 2021 from the Zwicky Transient Facility Public Data Release 6, which was the most current data release at the time (ZTF: Masci et al., 2018; Bellm et al., 2019). Since then, more recent ZTF Data Releases have become available. In the interest of characterising AGN variability and indeed detecting rare AGN flares, longer baselines of lightcurves are imperative. With this in mind, I downloaded the updated ZTF lightcurves of the ALPAKA catalogue from ZTF Data Release 20 (made available in January 2024) to investigate whether the GP analysis from Chapter 3 yielded differing results, *e.g.,* (a) how the hyperparameter distributions have changed, if at all, (b) whether any lightcurves have changed position in hyperparameter space and (c) whether I can ascertain a preliminary estimate for the rate of flare activity or extreme variability in the population. I will briefly present the results of applying the same technique as used in Chapter 3 on the updated ZTF lightcurves and discuss the implications.

Figure A.1 shows the distributions of hyperparameters for the original sample used

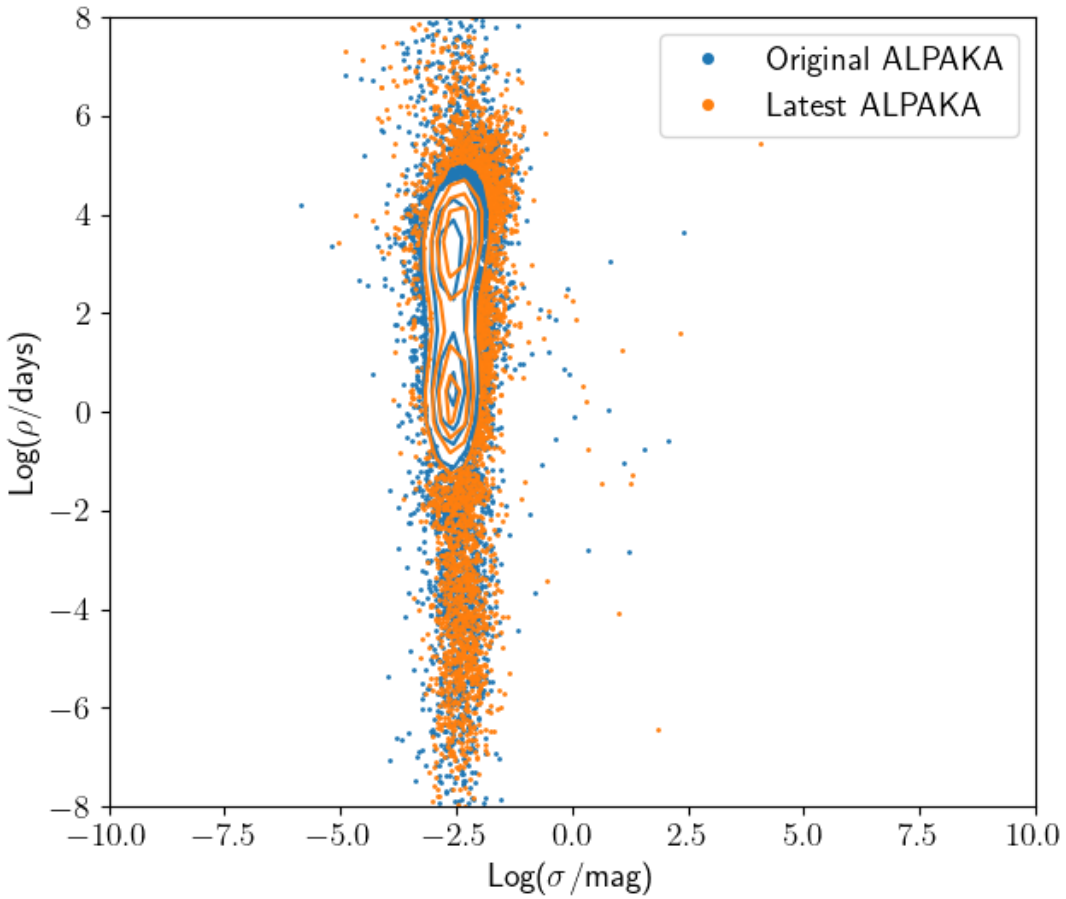

Figure A.1: Comparison of the distributions of hyperparameters for the original ZTF lightcurves of the ALPAKA sample and the updated ZTF lightcurves.

in Chapters 3 and 4 and the updated sample from the more recent data release. The distributions are virtually identical, which shows that although more data has been acquired for each lightcurve, statistically speaking, the underlying variability of the sample has largely remained the same.

To investigate whether any new flare candidates could be identified in the more recent data release, I applied the same selections on the data that were employed in Chapter 3. As a reminder, these were:

- A probabilistic cut-off of 0.1 for the flare probability as a result of the Bayesian hypothesis testing in Chapter 3.
- Lightcurves must contain greater than 30 data points.
- Lightcurves must not contain gaps in the data exceeding 150 days.

As a result, 21 new flare candidates were found that were not identified in the previous data release. A total of 16 of the 27 flare candidates identified in Chapter 3 were also retrieved. The remaining 11 lightcurves did not meet the above requirements for detection because the flare probability had decreased in the presence of new data. Lightcurves of the 21 new flare candidates are shown in Figures A.2-A.6. It is clear that these lightcurves exhibit extreme variability and those of particular interest are shown in red.

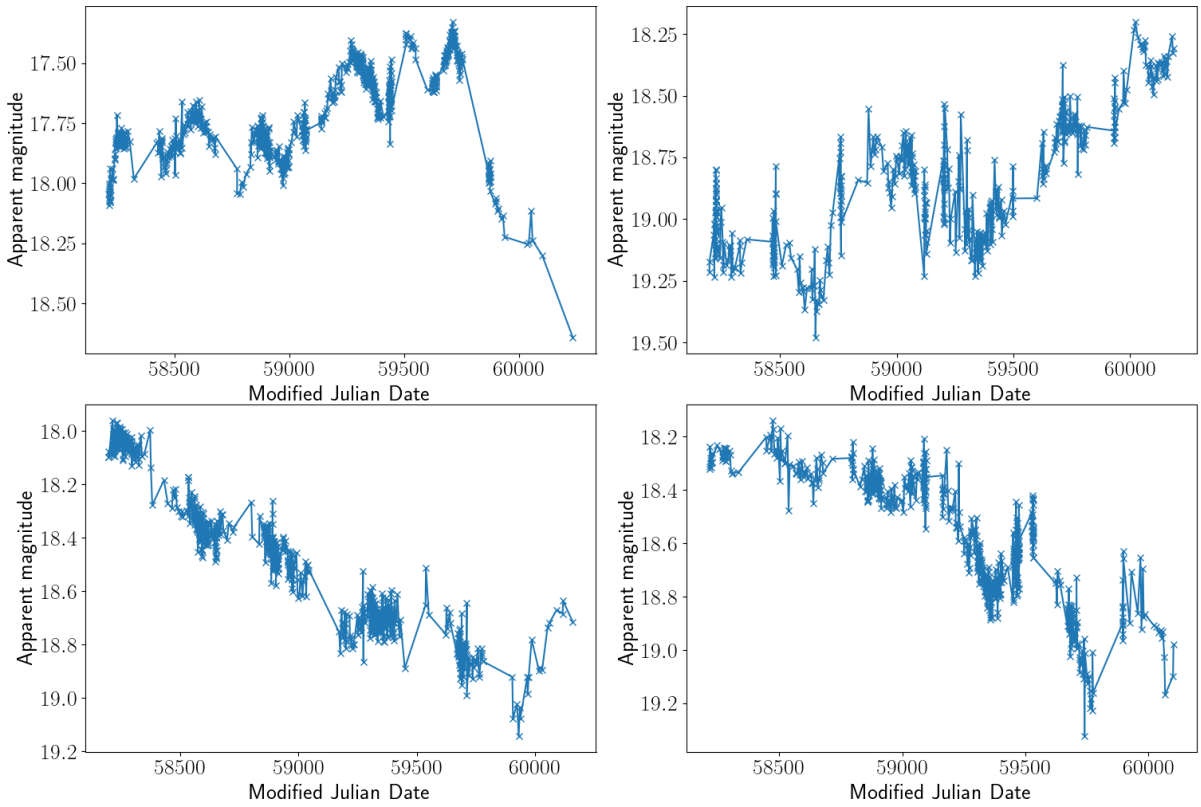

Figure A.2: ZTF r-band lightcurves of the first four out of the total 21 flare candidates identified from the updated ZTF data release.

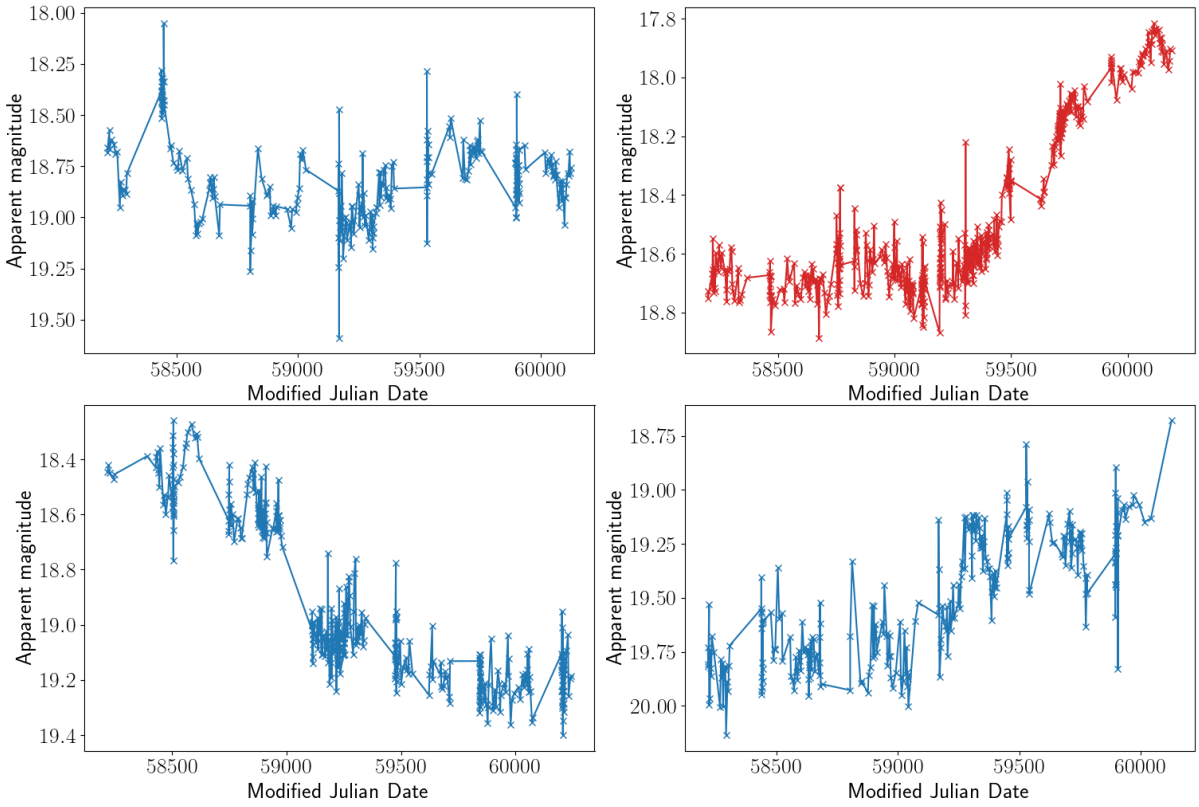

Figure A.3: ZTF r-band lightcurves of the 5th–8th out of the total 21 flare candidates identified from the updated ZTF data release. Lightcurves of particular interest for their extreme variability are indicated in red.

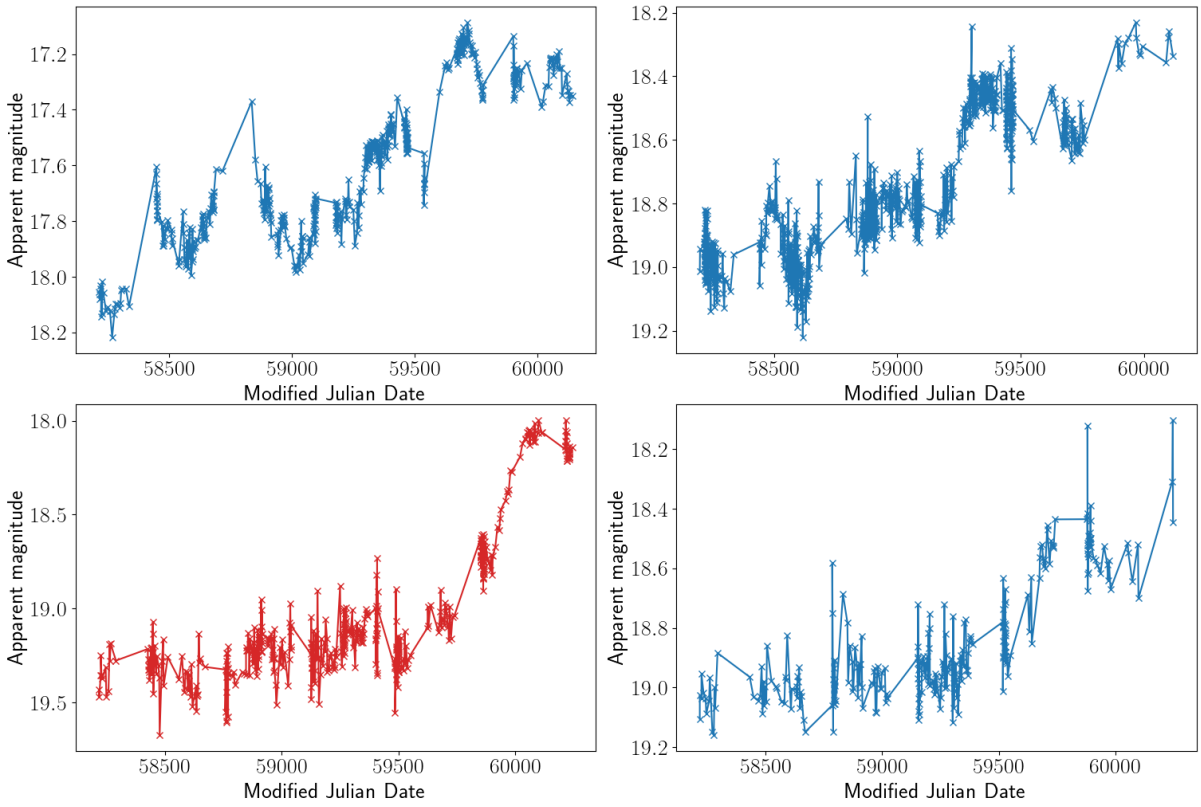

Figure A.4: ZTF r-band lightcurves of the 9th–12th out of the total 21 flare candidates identified from the updated ZTF data release. Lightcurves of particular interest for their extreme variability are indicated in red.

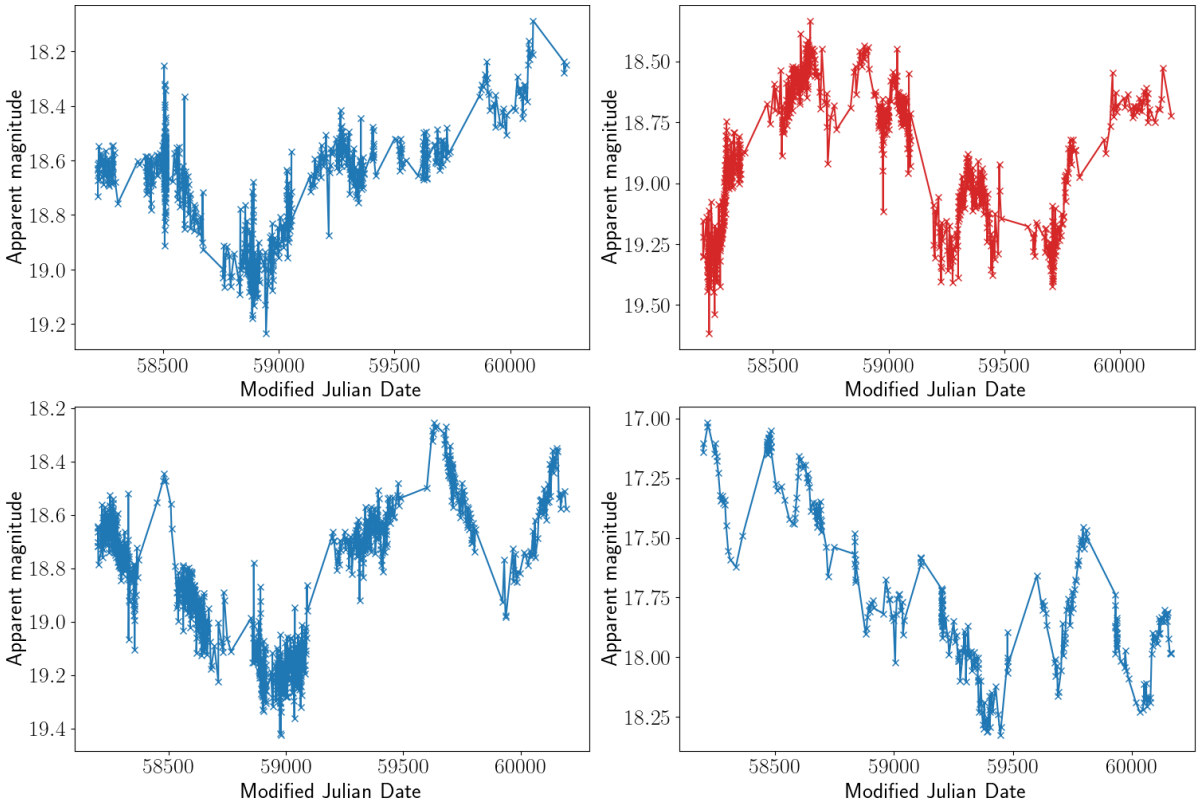

Figure A.5: ZTF r-band lightcurves of the 13th–16th out of the total 21 flare candidates identified from the updated ZTF data release. Lightcurves of particular interest for their extreme variability are indicated in red.

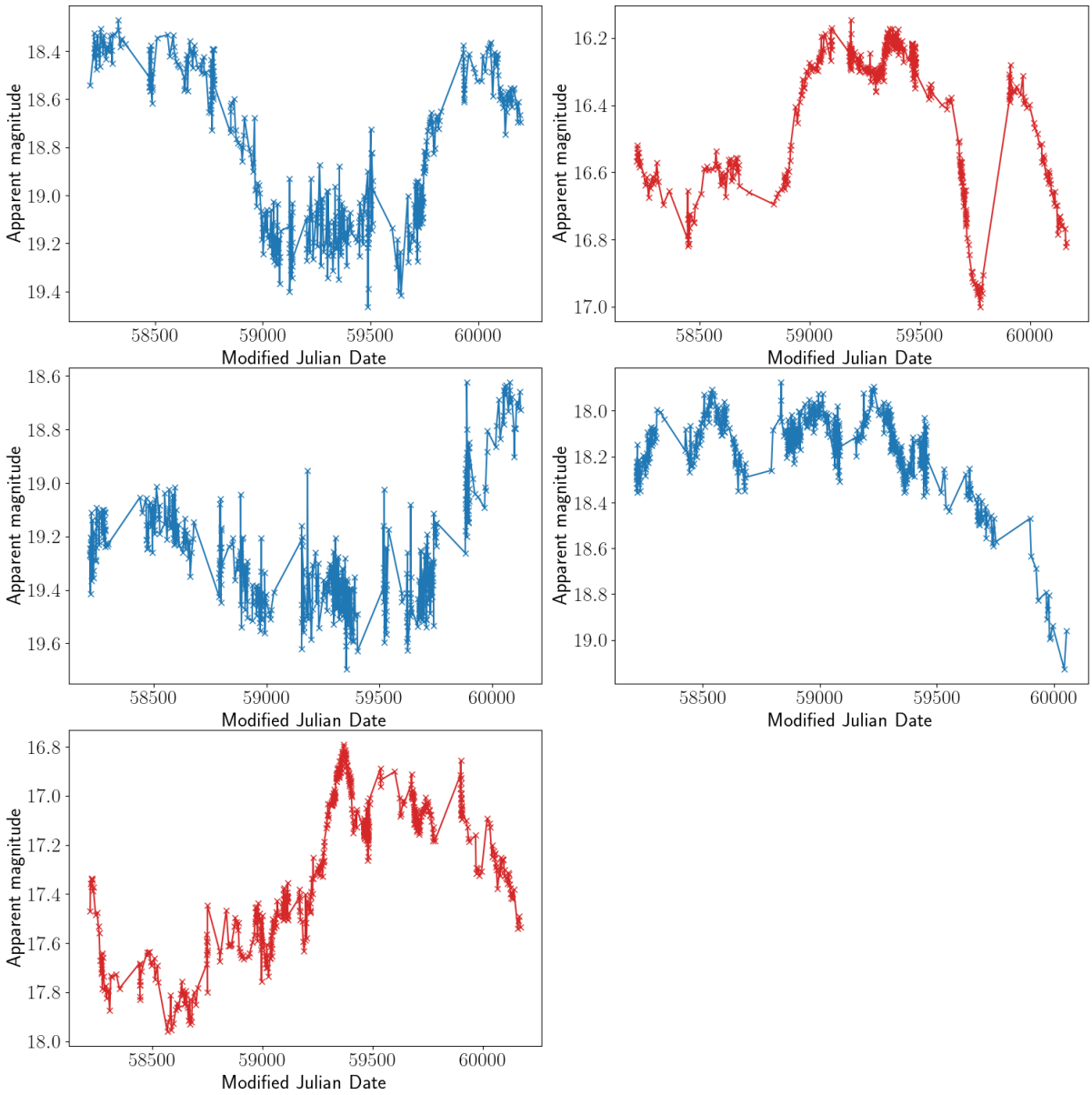

Figure A.6: ZTF r-band lightcurves of the 17th–21st out of the total 21 flare candidates identified from the updated ZTF data release. Lightcurves of particular interest for their extreme variability are indicated in red.