Universidade de São Paulo
Instituto de Astronomia, Geofísica e Ciências Atmosféricas
Departamento de Astronomia

Giovanna Souza Rodrigues Costa

Investigação da Origem Física de Glitches nos Detectores de Ondas Gravitacionais LIGO

São Paulo
2026

Giovanna Souza Rodrigues Costa

Investigação da Origem Física de Glitches nos Detectores de Ondas Gravitacionais LIGO

Trabalho de conclusão de curso apresentado ao Instituto de Astronomia, Geofísica Ciências Atmosféricas da Universidade de São Paulo como requisito parcial para a obtenção do título de Bacharel em Astronomia

Área de Concentração: Astronomia Multimensageira

Orientadores:

Prof.(ª) Dr.(ª) Jane Gregório-Hetem (IAG/USP)

Prof. Dr. Odylio Denys Aguiar (INPE/MCTI)

Dr. (ª) Tábata Aira Ferreira (INPE)

Msc. Júlio Cesar Martins (INPE)

São Paulo

2026

Dedico este trabalho àqueles que vieram antes de mim. Este é o resultado dos caminhos que começaram muito antes do meu próprio nascimento. Espero honrar cada renúncia, cada força e cada escolha que me permitiram chegar até aqui. Que os meus sonhos sejam sempre uma extensão e uma celebração da grandeza de onde eu vim.

Agradecimentos

Ao Instituto de Astronomia, Geofísica e Ciências Atmosféricas (IAG-USP) e ao Instituto Nacional de Pesquisas Espaciais (INPE), por ajudarem na minha formação e fornecerem a infraestrutura e o ambiente necessários para o desenvolvimento deste trabalho. À *LIGO Scientific Collaboration* (LSC), pela oportunidade e pelo privilégio de fazer parte dessa colaboração tão especial e por me permitir contribuir com a minha pesquisa para uma questão tao importante para a astronomia.

Aos meus orientadores, Odylio Denys Aguiar, Tábata Aira Ferreira e Julio Cesar Martins. Obrigada pelos ensinamentos, paciência, orientação e, acima de tudo, pela confiança. Sem vocês, este trabalho não seria possível, e eu certamente não estaria tão próxima de me tornar a cientista que espero ser. Agradeço também a professora Jane Gregório Hetem pelo apoio e orientação na redação deste trabalho.

Aos meus pais, obrigada por acreditarem em um caminho que muitas vezes pode parecer abstrato ou distante do cotidiano, por respeitarem as ausências que precisei cumprir para chegar aonde estou hoje. Escolho essa carreira por mim, mas espero honra-los em cada uma das minhas conquistas que virão, devolvendo ao mundo tudo o que puder descobrir.

Aos meus amigos, André, Bruno, Gustavo, Eduarda, Hevelyn, Isabela, Isabella, Isabelle, João, Karoline, Lavínia, Miguel, Nicolas, Paulo, Taísa e Yasmmin. Ter o carinho e a parceria de vocês foi um dos meus maiores privilégios.

À Giovana, por acreditar em mim e me incentivar sempre, sua companhia tornou até os mais simples momentos em belas histórias.

Agradeço à FAPESP pelo financiamento desta pesquisa.

A todos que, de forma direta ou indireta, tornaram este caminho não apenas possível, mas profundamente significativo.

Esta tese/dissertação foi escrita em LaTeX com a classe IAGTESE, para teses e dissertações do IAG.

"*Que nada nos limite, que nada nos defina, que nada nos sujeite. Que a liberdade seja a nossa própria substância.*"

Simone de Beauvoir

"*O tempo passa e a gente aprende com os erros, com a dor e com a escassez. Mas a mente de quem sonha alto não tem fronteira, não tem CEP e não aceita limites. Nós somos o reflexo daquilo que insistimos em construir, mesmo quando ninguém acreditava que seria possível.*"

Racionais Mc's

Resumo

A detecção de ondas gravitacionais pelos interferômetros terrestres do LIGO é frequentemente prejudicada pela presença de perturbações transitórias não periódicas, conhecidas como *glitches*, que podem mascarar ou imitar sinais astrofísicos reais. Este trabalho tem como objetivo investigar a origem física e a dinâmica temporal de diferentes morfologias de *glitches* nos detectores de Hanford (H1) e Livingston (L1) durante a terceira (O3) e a quarta (O4) corridas observacionais. Para isso, utilizou-se a base de dados de classificação *Gravity Spy*. A metodologia baseou-se na análise da ocorrência temporal desses transientes através da distribuição estatística de Poisson, quantificando os desvios por meio do coeficiente de determinação R^2.

Analisou-se a evolução temporal comparativa através do cálculo das taxas de ocorrência horárias e diárias. Os resultados demonstram que as dinâmicas de ocorrência variam de acordo com a morfologia. Observou-se também que o aprimoramento da sensibilidade dos detectores na corrida O4 tornou a rede altamente suscetível a ruídos sísmicos, de modo que a análise temporal provou que, embora essas classes compartilhem parte da origem física, elas respondem a dinâmicas ocasionalmente independentes.

A forte correlação temporal e física entre os *glitches Extremely Loud* e *Koi Fish* sustentam a hipótese de que o *Extremely Loud* atua em partes como uma versão saturada e de alta energia (alta Relação Sinal-Ruído) do *Koi Fish*, possuindo assim a mesma origem física. Por fim, o estudo identificou limitações e erros sistemáticos nas classificações do *Gravity Spy*, de modo que a caracterização fenomenológica detalhada desenvolvida neste estudo fornece indicativos relevantes para o desenvolvimento de novas metodologias de classificação e estratégias de mitigação de ruídos, passos fundamentais para elevar a confiabilidade estatística na era da astronomia de ondas gravitacionais.

Abstract

The detection of gravitational waves by LIGO's ground-based interferometers is frequently hindered by the presence of non-periodic transient disturbances, known as *glitches*, which can mask or mimic real astrophysical signals. This work aims to investigate the physical origin and temporal dynamics of different *glitch* morphologies in the Hanford (H1) and Livingston (L1) detectors during the third (O3) and fourth (O4) observational runs. To this end, the *Gravity Spy* classification database was utilized. The methodology was based on analyzing the temporal occurrence of these transients through a Poisson statistical distribution, quantifying deviations via the coefficient of determination R^2.

The comparative temporal evolution was analyzed by calculating hourly and daily occurrence rates. The results demonstrate that the occurrence dynamics vary according to morphology. It was also observed that the enhancement of detector sensitivity in the O4 run rendered the network highly susceptible to seismic noise, such that temporal analysis proved that, although these classes share a partial physical origin, they occasionally respond to independent dynamics.

The strong temporal and physical correlation between *Extremely Loud* and *Koi Fish* *glitches* supports the hypothesis that the *Extremely Loud* acts, in part, as a saturated, high-energy (high Signal-to-Noise Ratio, or SNR) version of the *Koi Fish*, thus sharing the same physical origin. Finally, the study identified limitations and systematic errors in the *Gravity Spy* classifications, showing that the detailed phenomenological characterization developed in this study provides relevant insights for the development of new classification methodologies and noise mitigation strategies, which are fundamental steps to elevate statistical reliability in the era of gravitational-wave astronomy.

Lista de Figuras

1.1 Evolução do número cumulativo de detecções e candidatos a eventos de ondas gravitacionais 21
1.2 Sinal GW170817 22

2.1 Verificação de consistência da hipótese de Poisson 27

3.1 *Glitches* do tipo *Extremely Loud* durante a O4a em Hanford. 31
3.2 *Glitches* do tipo *Low Frequency Burst* durante a O4b em Livingston. . . . 31
3.3 Espectro da densidade de ruído de deformação (*strain noise*) dos detectores LIGO (Hanford e Livingston) comparando as corridas observacionais O3 e O4. 34
3.4 Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Livingston durante a corrida observacional O4a. 36
3.5 Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Hanford ($H1$) durante a corrida observacional O4b. 36
3.6 Comparação morfológica entre *glitches* no plano tempo-frequência. Respectivamente, um exemplo de *Koi Fish* e um exemplo de *Extremely Loud*. . . 37
3.7 Comparação temporal das taxas médias diárias para os glitches das classes Extremely Loud e Koi Fish no detector ($H1$) durante a O3a. 38
3.8 Distribuição temporal em taxa de ocorrência das classes *Extremely Loud* e *Koi Fish* tratadas como uma única categoria de *glitch*. 39
3.9 Espectrogramas de *Low Frequency Lines* rotulados incorretamente pelo algoritmo como *Scattered Light*. 40

3.10 Evolução temporal e alteração no comportamento de ocorrência dos glitches *Scattered Light* contaminados com *Low Frequency Lines* 41
3.11 Espectrogramas da classe *Low Frequency Burst* evidenciando desvios estruturais. 41
3.12 Comparação morfológica entre um ruído *Whistle* e duas das classificações incorretas geradas pelo algoritmo. 42

A.1 Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Livingston ($L1$) durante a corrida observacional O4b. 51
A.2 Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Hanford ($H1$) durante a corrida observacional O4a. 52
A.3 Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Livingston ($L1$) durante a corrida observacional O3a. 52
A.4 Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Livingston ($L1$) durante a corrida observacional O3b. 52
A.5 Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Hanford ($H1$) durante a corrida observacional O3a. 53
A.6 Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Hanford ($H1$) durante a corrida observacional O3b. 53

Lista de Tabelas

3.1 Resumo das caracterizações das morfologias de *glitches* entre os detectores e os períodos da corrida observacional O4. 32

3.2 Resumo comparativo unificado das morfologias de *glitches* entre os detectores e os períodos das corridas observacionais O3 e O4. 33

3.3 Taxas médias horárias (glitches/hora) e diárias (glitches/dia) para as classes de *Low Frequency Burst* e *Low Frequency Lines* durante a O4. 34

B.1 Resumo da caracterização das morfologias de *glitches* entre os detectores e os períodos da campanha observacional O3 55

Sumário

1. Introdução 19
1.1 *Glitches* 21

2. Dados e Metodologia 25
2.1 Modelo Estatístico 26
2.2 Teste de Qualidade de Ajuste de χ^2 26
2.3 Coeficiente de Determinação (R^2) 28
2.4 Análise de Taxas Temporais e Validação de Morfologias 28

3. Análise 31
3.1 Caracterização dos *glitches* 31
3.2 Sensibilidade dos detectores LIGO 33
3.3 Scattered Light, Low Frequency Lines e Low frequency Burst 35
3.4 Extremely Loud e Koi Fish 37
3.5 Limitações e Classificações Incorretas no Gravity Spy 39

4. Conclusões 43

Referências 45

Apêndice 49

A. Distribuição temporal comparativa das classes Scattered Light, Low Frequency Lines e Low Frequency Burst 51

B. Caracterização do comportamento dos glitches durante a corrida O3 55

Capítulo 1

Introdução

As ondas gravitacionais surgem como uma consequência direta da Teoria da Relatividade Geral (Einstein, 1916, 1918), que descreve a gravidade como uma deformação do espaço-tempo causada por concentrações de energia e matéria. Quando há alteração na distribuição de matéria por aceleração, perturbações são geradas na estrutura do espaço-tempo e se propagam pelo Universo à velocidade da luz, sendo mais intensas no caso de objetos compactos, como estrelas de nêutrons ou buracos negros, devido à grande quantidade de massa concentrada em um pequeno volume. Essas perturbações, as quais chamamos de ondas gravitacionais, carregam informações da fonte astrofísica e tornam possíveis os estudos dos fenômenos mais energéticos e extremos do cosmos. Devido à baixa amplitude característica dessas deformações e sua pequena interação com a matéria, a sua detecção representa um dos avanços mais significativos da astrofísica instrumental, com consequências significativas para a área de altas energias e astronomia multimensageira.

Desde a primeira detecção realizada pelo *Laser Interferometer Gravitational-Wave Observatory* (LIGO) em setembro de 2015 (Abbott et al., 2016b), múltiplos eventos cósmicos extremos foram observados, em que o encerramento do quarto intervalo de medição, chamado de "quarta corrida observacional" (O4), elevou o número total de detecções a 300 candidatos de ondas gravitacionais (LIGO Scientific Collaboration, 2024; Abbott et al., 2023, 2024; Collaboration et al., 2026). Essas observações oferecem muitas oportunidades para investigar fenômenos astrofísicos de altas energias, seja testando os limites de validade da Teoria da Relatividade Geral (Abbott et al., 2019) em regimes de campos gravitacionais intensos ou simplesmente ampliando a compreensão a respeito da dinâmica e evolução de sistemas compactos no universo.

Para viabilizar a captura desses sinais, a infraestrutura instrumental do LIGO baseia-

se no princípio de funcionamento de um interferômetro de Michelson (Rudenko e Sazhin, 1980), adaptado com cavidades ópticas de Fabry-Perot para amplificar a sensibilidade do sistema (Aasi et al., 2015). Cada observatório é composto por dois tubos perpendiculares em vácuo com 4 km de extensão, por onde um feixe de laser é dividido e viaja até ser refletido por espelhos de teste suspensos. A passagem de uma onda gravitacional altera o comprimento relativo desses braços, gerando uma variação infinitesimal na fase do laser, que se manifesta como um padrão de interferência no fotodetector (Abbott et al., 2016a). Para melhorar a confiabilidade nas observações, o experimento opera de forma síncrona com dois detectores geograficamente separados nos Estados Unidos: o observatório de Hanford (H1), no estado de Washington, e o de Livingston (L1), na Louisiana. É válido ainda ressaltar que os observatórios Virgo (Itália) e Kagra (Japão) também compartilham dados observacionais das corridas (Abbott et al., 2020).

Entretanto, a busca pela captura e interpretação das ondas gravitacionais encontra-se repleta de desafios. Os detectores interferométricos terrestres se destacam na função de detectar as mínimas distorções no espaço-tempo ocasionadas pela passagem das ondas gravitacionais. Durante as corridas observacionais, os interferômetros acabam por medir não apenas oscilações do espaço-tempo causadas por fontes astrofísicas, mas também captam uma gama de perturbações transitórias, comumente denominadas *glitches* (Glanzer et al., 2023) (Ashton, 2023). Essas anomalias, caracterizadas por sua natureza não periódica e frequentemente imprevisível, constituem um desafio significativo para a detecção e interpretação confiável dos sinais emitidos por ondas gravitacionais, já que podem mascarar ou até mesmo imitar sinais reais de ondas gravitacionais, reduzindo a confiança estatística das detecções e aumentando a taxa de alarmes falsos (Buikema et al., 2020). Esses fenômenos já foram amplamente investigados e caracterizados em diversos estudos na literatura (Arnaud, 2022), com o objetivo de compreender seus mecanismos e desenvolver estratégias de mitigação.

Sendo assim, considerando a transição para a quarta corrida observacional, subdividida nos períodos O4a e O4b, foi possível ampliar significativamente o horizonte científico da astrofísica de ondas gravitacionais, vide Figura 1.1. O aumento da sensibilidade dos detectores resultou em um acréscimo massivo no volume de dados e também no número de eventos astrofísicos candidatos (Capote et al., 2025). Da mesma forma, essa maior capacidade de observação de eventos tornou os instrumentos ainda mais suscetíveis a no-

vas dinâmicas de interferências ambientais, antropogênicas e instrumentais. A análise e a caracterização detalhada das populações de *glitches* na era da O4 e sua comparação com o comportamento observado na O3 são, portanto, fundamentais para investigar a origem e a distribuição desses *glitches*, de modo a minimizar seu impacto nos dados observacionais.

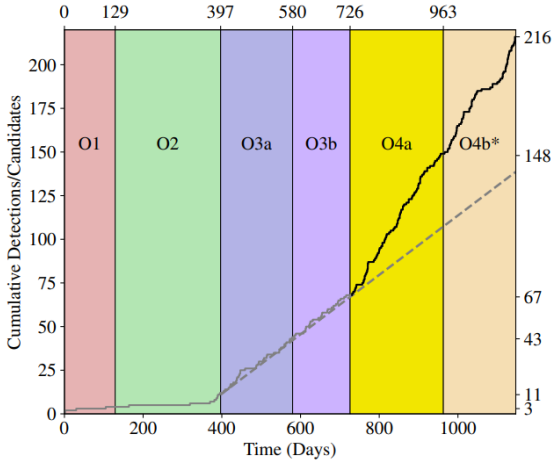

Figura 1.1: Evolução do número cumulativo de detecções e candidatos a eventos de ondas gravitacionais ao longo das corridas observacionais da colaboração LIGO-Virgo-KAGRA. O eixo horizontal indica o tempo em dias e os marcadores no topo mostram os dias exatos de transição entre os períodos com o eixo vertical direito destacando o total acumulado de eventos ao final de cada corrida.

1.1 *Glitches*

Durante o evento da primeira coalescência da binária de estrelas de nêutrons, ocorreu um caso real do impacto de *glitches*. A Figura 1.2 mostra o sinal *chirp* característico de uma coalescência, no qual a frequência aumenta com o tempo, porém com um *glitch* significativo no final. Neste caso, o *glitch* foi observado apenas no interferômetro de Livingston e não teve uma causa conhecida (Abbott et al., 2017).

Os *glitches* manifestam-se como perturbações súbitas no espectrograma tempo–frequência, apresentando uma ampla variedade de padrões morfológicos com base em suas características observadas nesse espaço, em que maioria dessas classes ainda carece de uma identificação física bem estabelecida, o que torna o processo de mitigação e remoção desses ruídos muito complexo.

É importante ainda reconhecer que os *glitches* são ruídos causados por fontes físicas (aparelhos de ar condicionado, motores ou outras fontes de ruído mecânico, espalhamento de luz, fontes elétricas e/ou magnéticas, dilatações térmicas, movimentos sísmicos, ventos,

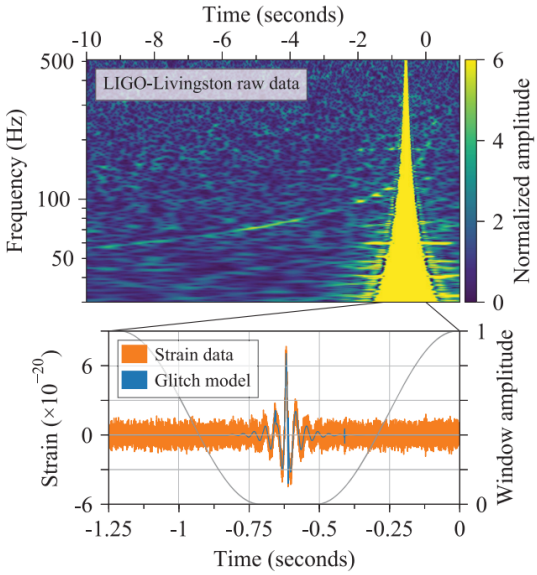

Figura 1.2: Na parte superior, é ilustrado o sinal do evento GW170817 contaminado por um *glitch* que foi reconstruído via *wavelet* (parte inferior).

variações gravitacionais), que fazem com que os espelhos dos detectores se movimentem, que partículas das paredes dos tubos de vácuo caiam e passem na frente do feixe laser, ou que outros processos mecânicos e/ou eletrônicos ocorram, resultando na indicação de sinal no detector de fótons.

Na primeira parte deste trabalho, foi caracterizada a distribuição temporal dos *glitches* observados na O3 de acordo com uma distribuição de Poisson. Essa distribuição modela a probabilidade de ocorrência de eventos discretos e estatisticamente independentes dentro de um intervalo fixo, sob a hipótese de uma taxa média constante. Uma distribuição que é amplamente utilizada para descrever processos como a detecção de fótons, o decaimento radioativo e o *shot noise*, este último particularmente relevante em sistemas ópticos e eletrônicos, como os interferômetros de ondas gravitacionais utilizados pelo LIGO (Abadie et al., 2011).

Em contraste, muitos fenômenos de ruído ambiental ou instrumental desviam-se do modelo poissoniano. O ruído térmico, por exemplo, resulta de flutuações termodinâmicas nos materiais e é tipicamente modelado como ruído branco gaussiano.

É importante ressaltar que o LIGO passa por constantes modificações, adaptações e atualizações de *hardware* tanto no intervalo entre as corridas observacionais quanto em paradas técnicas programadas, visando ao aprimoramento contínuo dos interferômetros (Aasi et al., 2015). Essa constante evolução instrumental motivou a aplicação da presente

análise estatística de forma segregada em diferentes blocos e períodos (como as subdivisões O3a, O3b, O4a e O4b), permitindo verificar se essas alterações específicas influenciaram diretamente a distribuição temporal ou taxa de geração de *glitches*.

Apesar dessas mudanças instrumentais ao longo do tempo, nota-se que certos *glitches* aparecem repetidamente com o mesmo formato nos gráficos de tempo-frequência (espectrogramas). Embora muitos desses eventos ocorram de forma estocástica e sem uma periodicidade fixa, a repetição de suas características permite agrupá-los sistematicamente. Esse comportamento é o que torna possível classificar e dar nomes a esses ruídos, definindo as categorias já bem estabelecidas na literatura, como *Blip*, *Scattered Light*, *Tomte* e *Koi Fish*.

Dessa forma, o presente estudo analisa e estuda a caracterização poissoniana dos *glitches* registrados durante a terceira e a quarta corrida observacional dos interferômetros LIGO (Davis et al., 2021). A partir dessas relações, a análise da evolução temporal das taxas horárias e diárias desses *glitches* ao longo das corridas e interferômetros, oferece através de métodos estatísticos e computacionais um subsídio fenomenológico útil para investigar suas origens físicas e auxiliar em estratégias de mitigação e aperfeiçoamento da observação de sinais em ondas gravitacionais.

A partir disso, avalia-se a evolução temporal das taxas horárias e diárias desses ruídos ao longo dos diferentes períodos e detectores (Hanford e Livingston). O objetivo é usar métodos estatísticos e computacionais para identificar padrões no comportamento desses *glitches*, de modo a obter fundamentos a respeito de suas origens físicas. Com esses resultados, busca-se contribuir para o desenvolvimento de estratégias de mitigação desses ruídos, melhorando a capacidade dos interferômetros de identificar sinais reais de ondas gravitacionais.

Além de mapear o comportamento dos transientes, as investigações feitas neste trabalho também permitiram identificar falhas de classificação cometidas pelo *Gravity Spy*, a principal ferramenta automatizada de classificação de *glitches* do LIGO. Essa verificação revelou que o algoritmo frequentemente confunde ou classifica de forma errada certos ruídos, às vezes por semelhanças na morfologia de outra classe ou por falhas do próprio *pipeline*, que está sujeito às limitações de usar apenas redes neurais supervisionadas e rígidas. Sendo assim, é possível que a caracterização de *glitches* no LIGO se beneficie de métodos alternativos, como algoritmos de aprendizado não supervisionado que agrupem os transientes

por diferentes fatores de semelhança, sem depender de rótulos pré-definidos.

Capítulo 2

Dados e Metodologia

É essencial caracterizar os interferômetros do LIGO e monitorar seu comportamento em relação à ocorrência de *glitches*. Como a classificação desses eventos pode fornecer indicativos importantes sobre suas possíveis origens, especialmente quando são observadas similaridades morfológicas entre eles, é necessário utilizar uma base de dados confiável de classificação de morfologias de *glitches*. Neste trabalho, foi utilizada a base de dados desenvolvida pelo projeto Gravity Spy (Glanzer et al., 2023; Zevin et al., 2024), que emprega um modelo de aprendizado de máquina baseado em espectrogramas tempo–frequência para categorizar ruídos transientes em 23 classes morfológicas predefinidas.

O *Gravity Spy* é uma das principais bases de dados do LIGO na análise de *glitches*. Ele utiliza como entrada os dados gerados pelo *pipeline* Omicron (Robinet et al., 2020), que identifica *glitches* ao detectar excesso de potência em relação ao ruído de fundo em um espectrograma multirresolução. Embora o Omicron, por si só, possa abranger uma faixa de frequência mais ampla e caracterize eventos por meio de transformadas-Q, o Gravity Spy restringe sua análise a *triggers* do Omicron no intervalo entre 10 e 2048 Hz. Para este estudo, foram utilizados tanto os dados abertos e publicamente disponíveis fornecidos pelo *Gravity Spy* quanto dados fechados e restritos aos membros da Colaboração Científica LIGO (*LIGO Scientific Collaboration* - LSC).

A análise tratou separadamente os diferentes segmentos de corrida para evitar a mistura de condições experimentais distintas que pudessem comprometer a consistência e distribuição dos resultados estatísticos.

2.1 Modelo Estatístico

Para modelar a ocorrência temporal dos *glitches*, assumiu-se um processo estocástico governado pela distribuição de Poisson Durran (1970). Assim, o número médio de ruídos em um intervalo de tempo (*chunk*) de duração t_0 é dado por $\mu = \lambda t_0$, onde λ representa a taxa média de ocorrência (razão entre o total de eventos e o tempo efetivo de observação). A probabilidade $P(r)$ de registrar exatamente r *glitches* em um único *chunk*, bem como a frequência esperada $N(r)$ de intervalos contendo essa contagem dentro de um total de G *chunks* analisados, são descritas por:

$$P(r) = \frac{e^{-\mu}\mu^r}{r!} \quad \text{e} \quad N(r) = G \cdot P(r). \tag{2.1}$$

Visando à consistência estatística frente ao ciclo de trabalho (*duty cycle*) do LIGO, o algoritmo desenvolvido em Python filtrou estritamente os períodos de operação do interferômetro. Os segmentos de observação contínua foram separados em múltiplos inteiros de t_0, de modo que fossem descartados os trechos observacionais menores que t_0 e intervalos em que o detector se encontrava *offline*, preservando a estacionariedade do modelo.

Por fim, adotou-se o limite mínimo de $N_{\text{tot}} \geq 400$ eventos por morfologia como critério de corte para a confiabilidade da análise. Esse limiar baseia-se na incerteza relativa associada à estatística de contagem de Poisson ($1/\sqrt{N_{\text{tot}}}$), garantindo que o erro relativo amostral permaneça abaixo de 5% e minimizando flutuações e instabilidades nas métricas descritivas para as classes com baixa taxa de ocorrência.

2.2 Teste de Qualidade de Ajuste de χ^2

Utilizou-se o teste estatístico do χ^2 (qui-quadrado) para verificar se as ocorrências de *glitches* são compatíveis com o modelo de Poisson.

Contudo, o teste rejeitou a hipótese de Poisson para todas as morfologias analisadas. Essa rejeição sistemática deve-se, majoritariamente, ao alto volume de segmentos temporais analisados, o que torna o teste excessivamente sensível a desvios estatísticos mínimos (Rice, 2007). Além disso, limitações inerentes ao algoritmo de classificação *Gravity Spy* indicam que cerca de 10% dos *glitches* na amostra podem ter sido classificados incorretamente devido a incertezas na forma do ruído (Seção 3.5).

Diante da incapacidade do teste χ^2 em confirmar a natureza poissoniana, buscou-se

quantificar o grau de desvio de cada classe por meio do coeficiente de determinação R^2. Esta métrica mede o quanto a distribuição real de cada morfologia se afasta de um cenário ideal simulado.

Para garantir que as rejeições encontradas nos dados reais resultam de características físicas do detector (ou ambientais) e não de erros computacionais no algoritmo de análise, realizou-se a mesma análise também utilizando dados simulados (sintéticos), gerados perfeitamente sob o modelo de Poisson. Foram geradas 1000 realizações sintéticas preservando a taxa média (λ), o tamanho do intervalo t_0 e o número total de segmentos G usados na análise real da morfologia *Blip*.

As conclusões obtidas dessa simulação de teste estão apresentados na Figura 2.1. O painel esquerdo mostra que, quando os dados realmente seguem a hipótese de Poisson, os valores de R^2 ficam muito próximos de 1, o que confirma que essa métrica funciona bem para avaliar o ajuste geral. Ao mesmo tempo, o painel direito mostra que o χ^2 calculado para essas amostras simuladas acompanha a distribuição teórica esperada, que está indicada pela linha vertical tracejada, ficando concentrado ao redor do número de graus de liberdade. Além disso, cerca de 5% das simulações geram $p - values$ abaixo de 0,05, que é exatamente o comportamento padrão esperado para um teste estatístico ideal. Esse comportamento prova que tanto o código quanto a metodologia são adequados. Portanto, qualquer falha de ajuste ou desvio encontrado nos dados reais do LIGO reflete apenas o comportamento físico real dos dados.

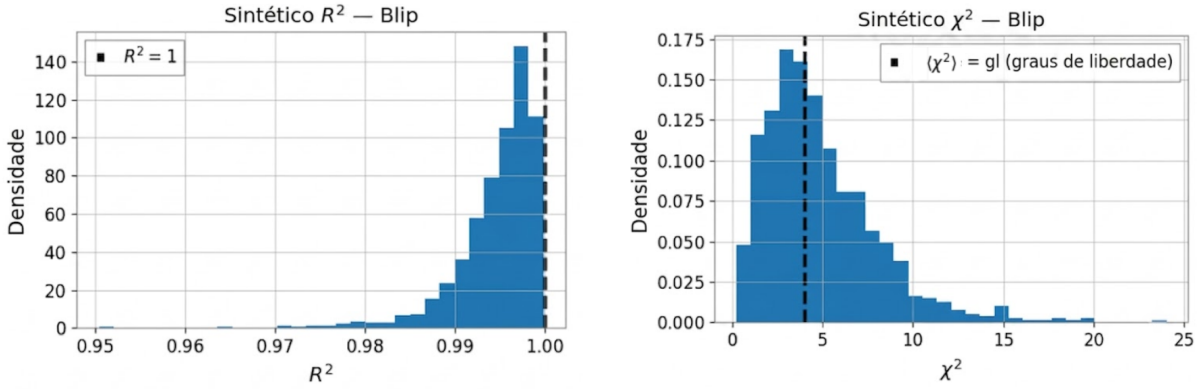

Figura 2.1: Verificação de consistência da hipótese de Poisson usando 1000 realizações sintéticas da morfologia Blip. Painel esquerdo: distribuição dos valores de R^2 obtidos a partir de conjuntos de dados gerados sob a suposição de Poisson, preservando a mesma taxa λ inferida, tamanho do intervalo t_0 e número total de segmentos G da análise real. Painel direito: distribuição dos valores de χ^2 computados usando os mesmos critérios de agrupamento de categorias adotados para os dados observados. A linha vertical tracejada indica o valor teórico esperado $\langle\chi^2\rangle = \text{gl}$ para um processo ideal de Poisson.

2.3 Coeficiente de Determinação (R^2)

Para avaliar a concordância global entre o modelo de Poisson e as contagens observadas, calculou-se o coeficiente de determinação R^2, definido como:

$$R^2 = 1 - \frac{\sum(N_{\text{obs}} - N_{\text{exp}})^2}{\sum(N_{\text{obs}} - \bar{N}_{\text{obs}})^2}, \tag{2.2}$$

onde N_{obs} é o número observado de *chunks* contendo *glitches*, N_{exp} é o número esperado correspondente calculado pelo modelo de Poisson e $\bar{N}_{\text{obs}}$ é a contagem média observada. Valores de R^2 próximos de 1 indicam uma forte concordância global entre as distribuições observada e esperada, enquanto valores baixos refletem discrepâncias crescentes.

O modelo foi validado utilizando conjuntos de dados reais e sintéticos. Os conjuntos sintéticos foram gerados por meio do gerador de números aleatórios poissonianos implementado na biblioteca *NumPy*, preservando o mesmo tempo total de observação, a mesma taxa λ, o mesmo tamanho de intervalo t_0 e o mesmo número total de *chunks* G da análise real.

Essas amostras simuladas funcionam como um caso de controle, validando o pipeline de análise sob condições em que a hipótese de Poisson é explicitamente obedecida. Para um processo puramente poissoniano, os valores esperados de R^2 se concentram fortemente próximos à unidade, confirmando que o pipeline estatístico está devidamente calibrado e não induz rejeições espúrias da hipótese nula.

2.4 Análise de Taxas Temporais e Validação de Morfologias

Para interpretar a dinâmica temporal dos *glitches* em diferentes escalas de magnitude, implementou-se o cálculo e a sobreposição das taxas de ocorrência horária e diária. A taxa horária foi computada contabilizando o número de eventos dentro de cada janela de tamanho $t_0 = 3600\,\text{s}$.

Complementarmente, a taxa média diária foi calculada a partir da média aritmética das taxas horárias contidas dentro de um intervalo de 86.400 segundos, com o ponto central fixado ao meio-dia (UTC) de cada dia operacional. Com esse processo é possível interpretar as tendências globais e de curto prazo de ocorrência de cada morfologia ao longo de toda a corrida observacional.

O *Gravity Spy* trabalha principalmente com aprendizado de máquina supervisionado,

ou seja, ele utiliza redes neurais convolucionais (CNNs) treinadas a partir de um conjunto de dados de referência que foi classificado por especialistas e voluntários do projeto. O aprendizado não supervisionado é usado em segundo plano, com o objetivo de agrupar e identificar novos formatos de ruído que ainda não foram catalogados. Assim, a cada *glitch* é atribuído um índice de probabilidade de acerto, denominado valor de confiança. Para evitar ao máximo que classificações incorretas e contaminações cruzadas distorcessem a análise estatística, o critério de seleção restringiu-se a eventos com confiança mínima de 90%.

Mesmo sendo muito preciso, o sistema do programa pode errar e gerar falsos positivos. Esses erros acontecem tanto por falhas de classificação da própria rede neural quanto por alarmes falsos gerados pelo pipeline *Omicron*, programa responsável por encontrar os transientes nos canais de dados brutos do detector.

Assim, para garantir que os resultados fossem confiáveis e que os surtos e flutuações vistos nos gráficos de taxa diária e horária fossem causadas pelo aumento real da ocorrência dos *glitches* ao invés de um erro de classificação, foi feita uma checagem manual dessas regiões em todas as classes. Essa verificação consistiu em analisar diretamente os espectrogramas de todos os picos de ocorrência encontrados nas taxas horárias e diárias. Esse processo serviu para confirmar se um surto inesperado era um aumento real daquele ruído no interferômetro ou se era apenas um erro de classificação do *Gravity Spy* ou do *Omicron*.

Capítulo 3

Análise

3.1 *Caracterização dos glitches*

Para a primeira etapa da análise, foram produzidos histogramas da distribuição temporal dos *glitches* da O4 em comparação com a frequência de ocorrência dos eventos observados, tendo como base o modelo estatístico de Poisson. As curvas observadas são representadas em verde, enquanto as distribuições de Poisson esperadas para cada morfologia de *glitch* são indicadas em vermelho. Como citado anteriormente na Seção 2.3, concordâncias ou discrepâncias entre os dados observados e o modelo estatístico proposto são avaliadas quantitativamente por meio do coeficiente de determinação R^2. As Figuras 3.1 e 3.2 abaixo apresentam, respectivamente, os histogramas de uma classe com comportamento semelhante ao de Poisson e de outra classe discrepante durante a O4.

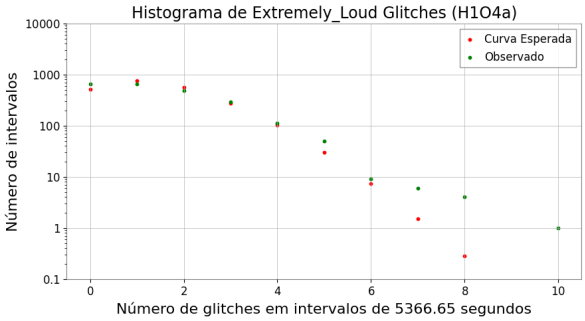

Figura 3.1: Glitches do tipo *Extremely Loud* em intervalos de 5.366,65 segundos durante a corrida O4a no detector H1.

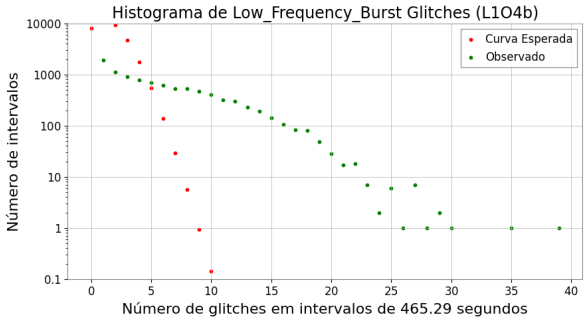

Figura 3.2: Glitches do tipo *Low Frequency Burst* em intervalos de 465,29 segundos durante a corrida O4b no detector L1.

Sendo assim, a Tabela 3.1 sintetiza a análise da distribuição de Poisson.

Tabela 3.1 - Resumo das morfologias de *glitches* entre os detectores e os períodos da corrida observacional O4, mostrando o coeficiente de determinação R^2 entre as distribuições de Poisson observadas e esperadas. A última coluna traz uma interpretação qualitativa do comportamento, baseada na concordância geral das distribuições reais com o modelo de Poisson.

Morfologia do Glitch	**L1 O4a** R^2	**L1 O4b** R^2	**H1 O4a** R^2	**H1 O4b** R^2	**Comportamento Fenomenológico**
Blip	0.980	0.841	0.857	0.964	Poissoniano
Blip Low Frequency	0.693	0.479	0.847	0.879	Dependente da corrida
Extremely Loud	0.921	0.974	0.950	0.985	Poissoniano
Fast Scattering	0.718	-	0.509	-	Fortes desvios
Koi Fish	0.961	0.906	0.949	0.952	Poissoniano
Low Frequency Burst	0.289	0.221	0.722	0.695	Dependente da corrida
Low Frequency Lines	0.291	0.536	0.733	0.636	Dependente da corrida
No Glitch	-	-	-	0.826	Poissoniano
Paired Doves	-	-	0.263	-	Fortes Desvios
Repeating Blips	-	-	-	0.996	Poissoniano
Scattered Light	0.450	0.446	0.618	0.813	Dependente da corrida
Tomte	0.766	0.847	0.640	0.774	Dependente da corrida

É válido ainda fazer algumas considerações ao comparar as classificações da Tabela 3.1 para a O4 com a Tabela B.1 desenvolvida para a O3 (disponível no Apêndice B), durante a primeira parte deste trabalho, notam-se algumas divergências comportamentais. Primeiramente, os *glitches* da classe *Scattered Light* mostraram uma mudança de comportamento importante. Na O3, essa classe apresentava desvios concretos do modelo de Poisson em ambos os observatórios, com valores de R^2 sempre abaixo de $0,490$. Já na O4, para o detector de Hanford, esse comportamento se aproxima do poissoniano. Porém, durante a O4b, a contagem de *glitches Scattered Light* foi consideravelmente menor do que nas demais corridas. Esse fator de baixa contagem aparentemente afetou outras classes durante a O3 e O4, como a *No Glitch* durante a O4, as *Low Frequency Lines* e *Low Frequency Burst* durante a O3 em L1 e a O4 em H1 e a *Fast Scattering* em Livingston durante a O4a

Por fim, a evolução na sensibilidade dos detectores e a reconfiguração dos subsistemas na O4 provocaram o desaparecimento ou o surgimento de populações específicas. Classes como *1400Ripples*, *Power Line*, *Air Compressor* e *Whistle*, que pontuavam com fortes desvios na O3, perderam volume estatístico e não tiveram amostragem suficiente para serem classificadas na O4. Em contrapartida, houve ainda o aumento considerável na contagem de *Paired Doves* em Hanford (H1 O4a), que apresentou fortes desvios da distribuição de Poisson.

Com base nessas condições, é possível realizar uma classificação ainda mais geral, utilizando o coeficiente de determinação como parâmetro e favorecendo, nos casos abstratos, o comportamento fenomenológico das corridas observacionais com maiores contagens de *glitches*. Essa relação é apresentada na Tabela 3.2.

Tabela 3.2 - Resumo comparativo unificado das morfologias de *glitches* entre os detectores de Livingston ($L1$) e Hanford ($H1$) para as corridas observacionais O3 e O4. Os valores indicam o coeficiente de determinação R^2 obtido em relação ao modelo ideal de Poisson. A última coluna apresenta uma classificação geral baseada no comportamento predominante nos regimes de maior contagem de eventos.

Morfologia do Glitch	**L1 O3a** R^2	**L1 O3b** R^2	**L1 O4a** R^2	**L1 O4b** R^2	**H1 O3a** R^2	**H1 O3b** R^2	**H1 O4a** R^2	**H1 O4b** R^2	**Classificação Geral**
1400Ripples	0.508	-	-	-	-	-	-	-	Fortes Desvios
Air Compressor	-	0.200	-	-	-	-	-	-	Fortes Desvios
Blip	0.940	0.932	0.980	0.841	0.909	0.887	0.857	0.964	Poissoniano
Blip Low Frequency	0.957	0.903	0.693	0.479	0.920	0.698	0.847	0.879	Dependente da Campanha
Extremely Loud	0.970	0.977	0.921	0.974	0.916	0.962	0.950	0.985	Poissoniano
Fast Scattering	0.269	0.277	0.718	-	-	0.322	0.509	-	Fortes Desvios
Koi Fish	0.953	0.952	0.961	0.906	0.922	0.963	0.949	0.952	Poissoniano
Low Frequency Burst	0.562	0.602	0.289	0.221	0.209	0.290	0.722	0.695	Dependente da Campanha
Low Frequency Lines	0.552	0.605	0.291	0.536	0.364	0.504	0.733	0.636	Dependente da Campanha
No Glitch	0.490	0.675	-	-	0.315	-	-	0.826	Fortes Desvios
Paired Doves	-	-	-	-	-	-	0.263	-	Fortes Desvios
Power Line	0.204	-	-	-	-	-	-	-	Fortes Desvios
Repeating Blips	-	-	-	-	0.850	-	-	0.996	Poissoniano
Scattered Light	0.370	0.481	0.450	0.446	0.279	0.236	0.618	0.813	Dependente da Campanha
Tomte	0.929	0.901	0.766	0.847	0.851	0.503	0.640	0.774	Dependente da Campanha
Whistle	0.623	0.379	-	-	0.294	0.415	-	-	Fortes Desvios

3.2 Sensibilidade dos detectores LIGO

A sensibilidade de um interferômetro terrestre de ondas gravitacionais é quantificada por meio da densidade espectral de amplitude do ruído de deformação (*strain noise*), expressa em unidades de $\mathrm{Hz}^{-1/2}$. A Figura 3.3 ilustra o comportamento da sensibilidade de observação em termos da frequência nos detectores de Hanford ($H1$) e Livingston ($L1$) ao longo das corridas observacionais O3 e O4. As curvas posicionadas em patamares inferiores no gráfico denotam um instrumento mais sensível e capaz de registrar eventos astrofísicos mais distantes ou de menor amplitude, ou seja, quanto mais para baixo estiverem as curvas no gráfico, melhor a sensibilidade do interferômetro.

A região central do gráfico, compreendida entre aproximadamente 50 Hz e 300 Hz, delimita a banda de maior sensibilidade física dos interferômetros. Nesse intervalo, o detector de Livingston ($L1$) durante a O4a e O4b atingiu os menores índices de ruído históricos da rede, aproximando-se do limiar de $3 \times 10^{-24}\,\mathrm{Hz}^{-1/2}$. Por fim, as estruturas pontiagudas e estreitas que quebram o perfil contínuo das curvas representam ruídos bem conhecidos, associados a harmônicos da rede elétrica norte-americana (60 Hz)(Buikema et al., 2020), linhas injetadas para calibração instrumental e modos de ressonância mecânica (*violin modes*) das fibras que suspendem os espelhos.

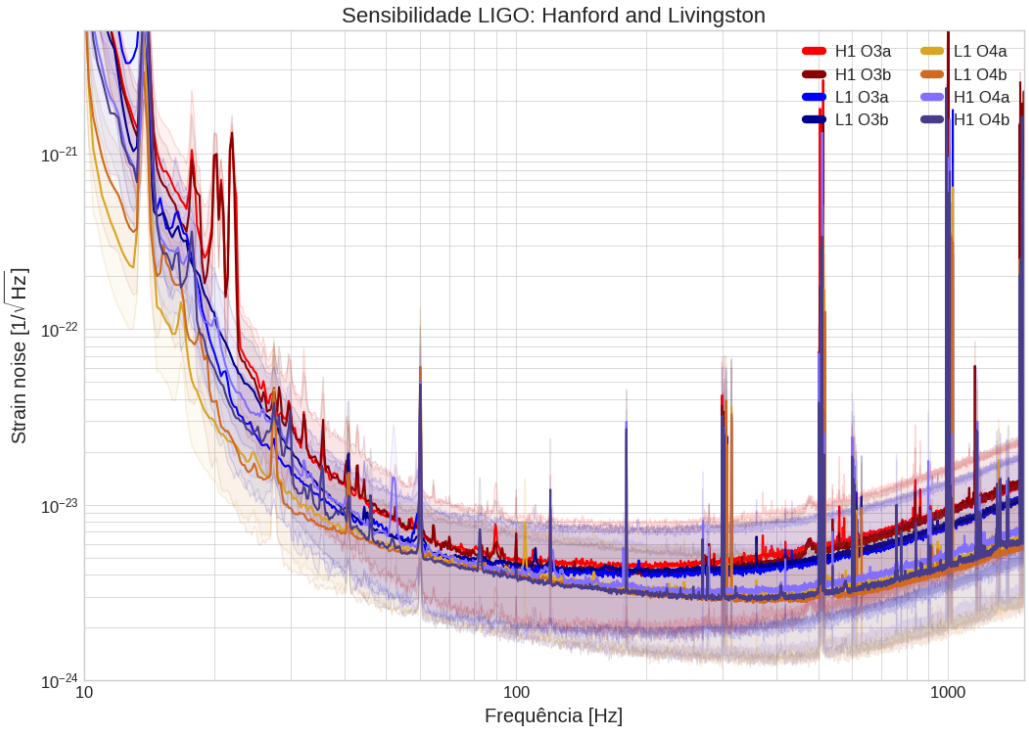

Figura 3.3: Espectro da densidade de ruído de deformação (*strain noise*) dos detectores LIGO (Hanford e Livingston) comparando as corridas observacionais O3 e O4.

É fundamental destacar que o ganho de sensibilidade obtido nas bandas de baixas frequências ao transicionar da O3 para a O4 gerou um impacto direto na população de *glitches* observada nas corridas observacionais. Ao rebaixar significativamente o piso de ruído de deformação nessa região do espectro, o instrumento tornou-se muito mais suscetível aos ruídos de baixa frequência.

Esse fenômeno justifica o aumento massivo na detecção de morfologias como *Scattered Light*, *Low Frequency Lines* e *Low Frequency Burst*, das quais as duas últimas passaram a registrar altas taxas de ocorrência diárias e horárias durante a corrida O4, descritas na Tabela 3.3.

Tabela 3.3 - Taxas médias horárias (glitches/hora) e diárias (glitches/dia) para as classes de *Low Frequency Burst* e *Low Frequency Lines* durante a O4.

Morfologia do Glitch	**L1 O4a**		**L1 O4b**		**H1 O4a**		**H1 O4b**	
	Horária	Diária	Horária	Diária	Horária	Diária	Horária	Diária
Low Frequency Burst	10,8972	12,3719	11,5714	12,4664	0,1781	0,2242	0,2195	0,2385
Low Frequency Lines	7,5824	8,2031	12,7883	13,1945	2,0526	2,0023	0,2611	0,2763

3.3 Scattered Light, Low Frequency Lines e Low frequency Burst

A correspondência física entre as morfologias de *Scattered Light*, *Low Frequency Lines* e *Low Frequency Burst* reside em sua forte associação com o ruído sísmico e microssísmico ambiental. O mecanismo gerador, por exemplo, decorre do movimento do solo (causado por atividade antropogênica ou por ondas oceânicas que colidem com a costa), o qual excita mecanicamente as bancadas ópticas e as paredes das câmaras de vácuo. Parte da luz do laser principal escapa do feixe ou é espalhada por imperfeições nos espelhos. Ao refletir em superfícies metálicas em vibração, essa luz espalhada adquire uma diferença de fase e, ao se recombinar com o feixe principal, pode gerar um glitch característico de *espalhamento da luz*. Esse processo sofre uma conversão ascendente de frequência, gerando os arcos e transientes característicos observados abaixo de 60 Hz (Longo et al., 2025). Uma vez que os surtos (*bursts*) e as linhas de baixa frequência compartilham desse mesmo canal de acoplamento mecânico e ambiental, eles respondem diretamente aos picos de atividade sísmica local.

Enquanto a infraestrutura *Gravity Spy* do LIGO segmenta esses ruídos em três categorias morfológicas distintas, a análise de dados do interferômetro europeu Virgo revela uma abordagem diferente devido à sobreposição dessas assinaturas. Estudos recentes focados no desenvolvimento do pipeline de classificação automatizada do Virgo, denominado *VIGI-Lant* (*Virgo Glitch Identification and Learning*), demonstraram que não é estatisticamente possível traçar uma distinção clara entre as classes *Scattered Light*, *Low Frequency Burst* e *Low Frequency Lines* nos dados do detector de Pisa (Fernandes et al., 2026), tratando-as como tipos de um mesmo fenômeno de ruído de baixa frequência.

Ao realizar a análise temporal dessas três classes, apresentada nas Figuras 3.4 e 3.5, nota-se que estas morfologias operam tanto de forma independente quanto em períodos temporais onde é possível notar comportamentos semelhantes entre as classes.

No cenário de alta contagem da corrida L1O4a (Figura 3.4), a correlação das fontes fica evidente na região de meados de 21 de dezembro. Nessa época, ambas as classes atingiram pontos de máxima taxa de ocorrência da corrida.

Além disso, outra assinatura incontestável de correlação local é registrada no início de setembro (02/09) e repetida em meados de outubro (12/10). Nesses instantes específicos, as curvas de *Low Frequency Burst* e *Low Frequency Lines* crescem em conjunto, subindo

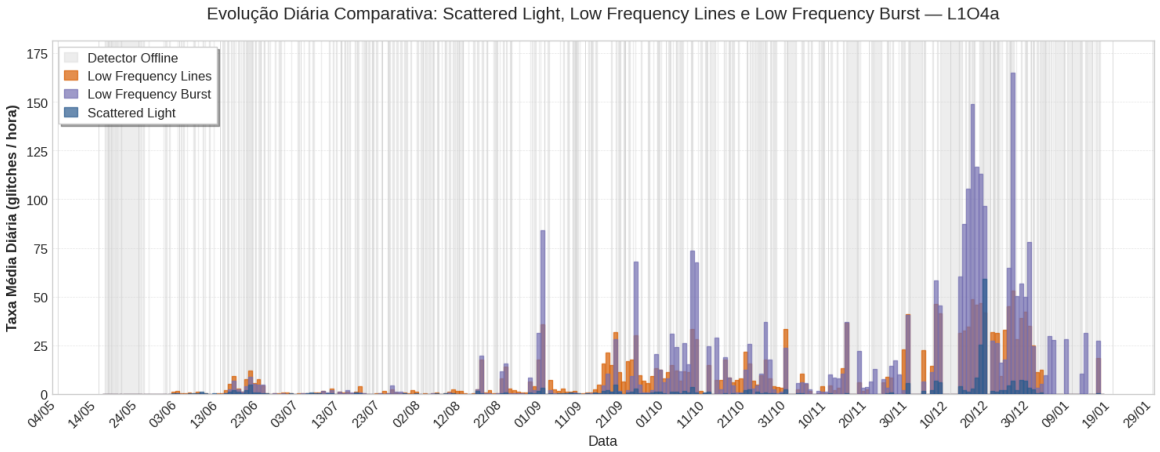

Figura 3.4: Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Livingston durante a corrida observacional O4a.

até patamares de 75 glitches/hora e 35 glitches/hora, respectivamente.

Essa dualidade entre independência e correlação torna-se ainda mais nítida no interferômetro de Hanford na corrida H1O4b (Figura 3.5).

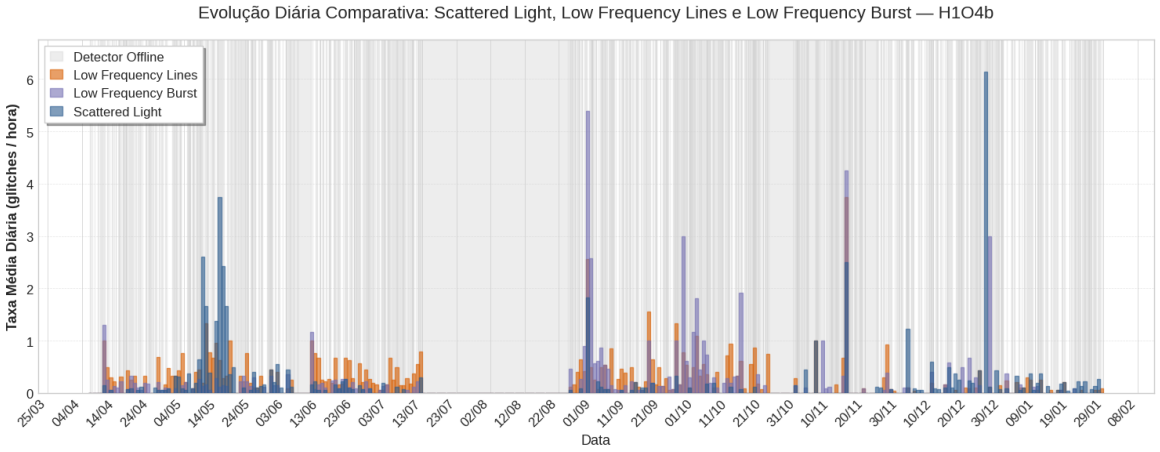

Figura 3.5: Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Hanford ($H1$) durante a corrida observacional O4b.

No dia 31 de dezembro, a classe *Scattered Light* apresenta o maior pico do período em taxa de ocorrência. Porém, esse comportamento não é imediatamente acompanhado pelos *Bursts* nem pelos *Low Frequency Lines*.

Portanto, com base na análise dos gráficos e considerando a classificação do *Gravity Spy*, essas três classes (*Scattered Light*, *Low Frequency Lines* e *Low Frequency Burst*) não apresentam um comportamento padrão que permita afirmar que elas vêm da mesma fonte física. O que os dados mostram é que elas provavelmente são causadas por tipos diferentes de sismos, de modo que ocasionalmente podem fomentar a ocorrência de mais de uma

classe. Além disso, os gráficos para as demais corridas encontram-se no Apêndice A.

3.4 Extremely Loud e Koi Fish

Outras duas classes que valem ser analisadas conjuntamente são as classes de *glitches* *Extremely Loud* e *Koi Fish*. Visualmente, um *Koi Fish* típico possui um centro bem definido em médias e altas frequências e traços laterais mais fracos em frequências mais baixas (Zevin et al., 2017). Por outro lado, estudos recentes avaliam se a classe *Extremely Loud* funciona mais como um indicador de excesso de energia do que como um formato de ruído isolado. Possivelmente, quando o evento que gera um *Koi Fish* é muito forte, ou seja, possui uma alta Relação Sinal-Ruído (SNR), o algoritmo da transformada Q espalha a energia nas regiões vizinhas do espectrograma (Chatterji et al., 2004). Esse efeito borra o gráfico de tempo-frequência e esconde as características mais fracas, transformando o formato do *glitch* em espectrogramas mais densos (Bahaadini et al., 2018). A Figura 3.6 demonstra exemplos do espectrograma dessas duas classes.

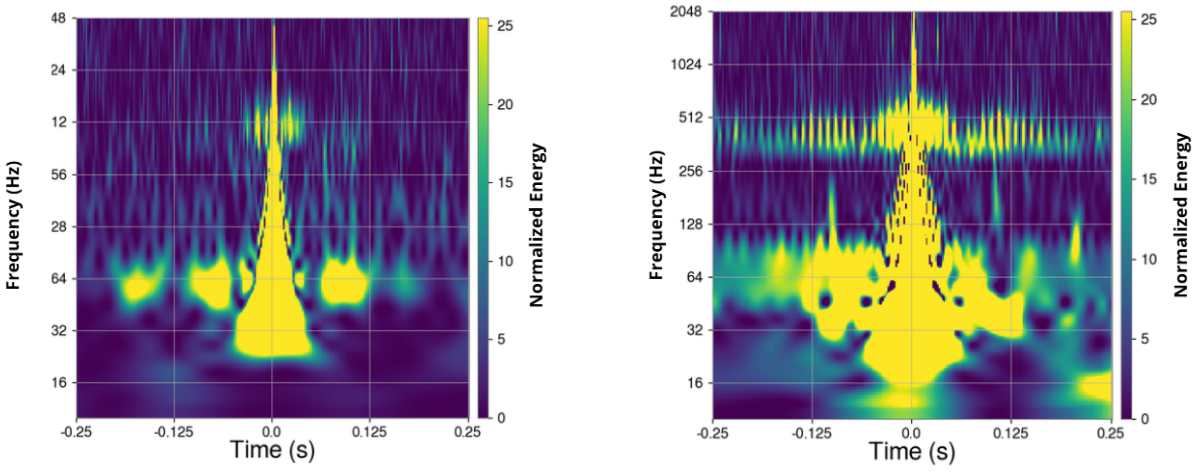

Figura 3.6: Comparação morfológica entre *glitches* no plano tempo-frequência. Respectivamente, um exemplo de *Koi Fish* e um exemplo de *Extremely Loud*.

Além dessa semelhança visual nos espectrogramas, a prova mais forte de que esses dois ruídos estão ligados aparece nas taxas diárias. A análise mostra que a distribuição dos eventos de cada uma dessas classes apresenta comportamentos semelhantes ao longo de toda a corrida observacional, vide Figura 3.7. Os dias com maior ocorrência e os períodos de taxa uniforme coincidem exatamente nas mesmas datas para ambas as populações.

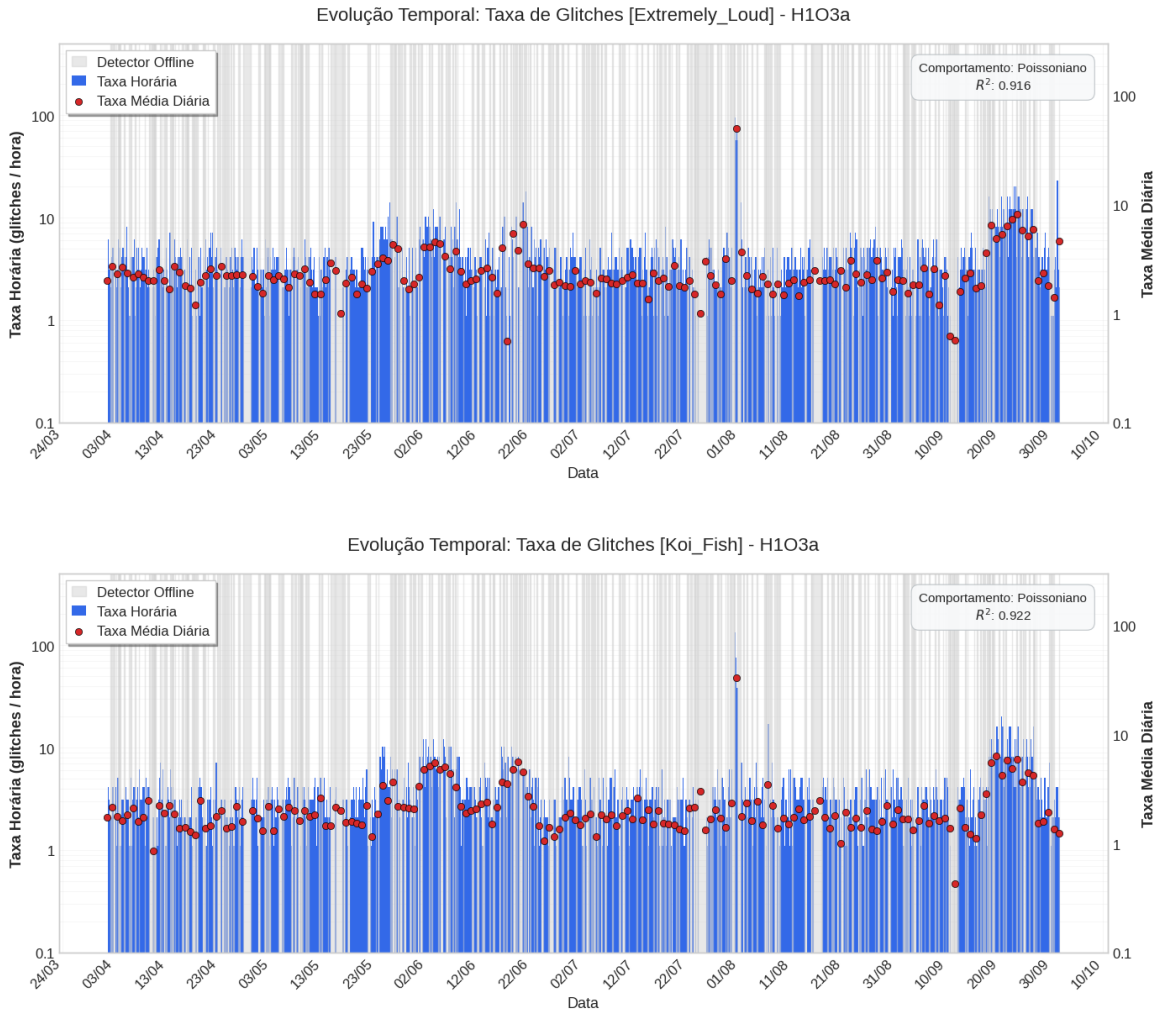

Figura 3.7: Comparação temporal das taxas médias diárias para os glitches das classes Extremely Loud (superior) e Koi Fish (inferior) no detector de Hanford ($H1$) durante a corrida observacional O4a.

Para testar a hipótese de que essas duas populações estão fisicamente ligadas, foi realizado um teste combinando os eventos das classes *Extremely Loud* e *Koi Fish* em uma única categoria unificada. Ao reanalisar a adequação ao modelo de Poisson para essa nova classe combinada, observou-se que o coeficiente de determinação (R^2) aumentou ou permaneceu igual em todas as corridas observacionais em relação aos resultados das classes separadas. Como ambas as classes já eram classificadas individualmente como poissonianas, a manutenção ou melhora desse ajuste estatístico mostra que a junção dos dados não gerou distorções no comportamento das classes, vide Figura 3.8.

Essa forte correlação temporal ganha ainda mais sustentação ao ser comparada com os estudos recentes de Ferreira e González (2025), que demonstraram que, ao normalizar a Relação Sinal-Ruído (SNR) dos dados brutos, as classes inicialmente classificadas como

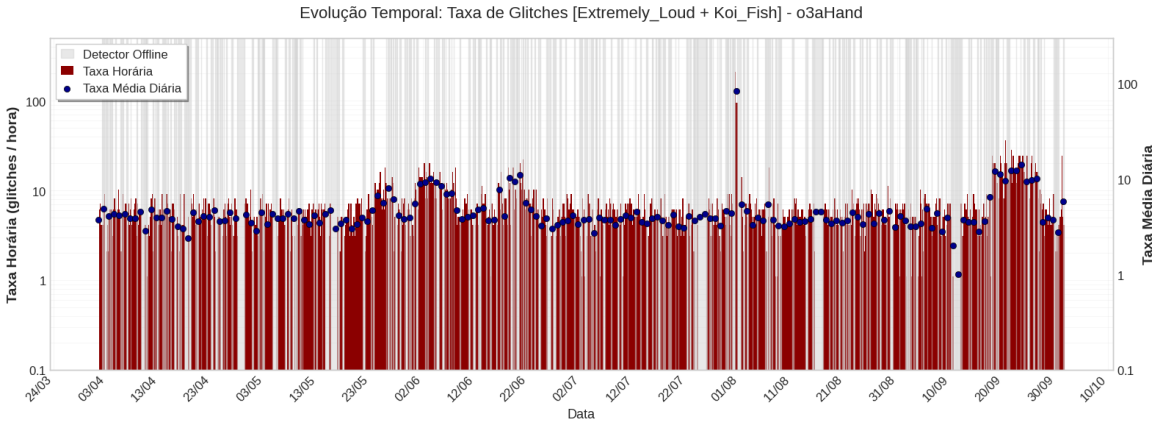

Figura 3.8: Distribuição temporal em taxa de ocorrência das classes *Extremely Loud* e *Koi Fish* tratadas como uma única categoria de *glitch*.

Koi Fish e *Extremely Loud* revelam semelhanças morfológicas. De acordo com as autoras, a sobreposição dessas duas populações nos algoritmos ocorre porque o *Extremely Loud* atua frequentemente como uma versão saturada de alta energia do próprio *Koi Fish*, o que corrobora os resultados de sincronia temporal encontrados neste trabalho.

Esse comportamento temporal pode indicar que o *Extremely Loud* e o *Koi Fish* compartilham a mesma origem física (Abbott et al., 2021). Portanto, os resultados sugerem que essas duas classes não devem ser tratadas como problemas independentes, mas que são, na verdade, duas manifestações do mesmo tipo de ruído, onde a classificação depende basicamente de quão forte foi a energia do evento (SNR).

3.5 Limitações e Classificações Incorretas no Gravity Spy

Apesar do sucesso do *Gravity Spy* na caracterização de transientes de ruído nos detectores da colaboração LIGO, foi realizada uma avaliação manual da classificação dos *glitches* para verificar a confiabilidade da principal ferramenta de caracterização de glitches do LIGO. Como consequência, foram observadas algumas divergências nas classificações. Os principais casos de classes afetadas por esses erros, bem como as limitações dos métodos de classificação, são analisados nesta seção.

O primeiro fator limitante do *Gravity Spy* que deve ser enfatizado reside na etapa de preparação dos dados de entrada. Para converter as séries temporais de deformação em espectrogramas utilizados por sua Rede Neural Convolucional (CNN), o algoritmo aplica a Transformada Q, que estabelece a relação de compromisso entre a frequência central (f_0)

e a largura de banda (δf) dos filtros aplicados:

$$Q = \frac{f_0}{\delta f}. \tag{3.1}$$

Como o *Gravity Spy* analisa os *glitches* processando a imagem sob um valor padronizado de Q que maximiza a Razão Sinal-Ruído (SNR) global do evento (Chatterji et al., 2004), o sistema perde a capacidade de registrar com nitidez transientes que possuem uma evolução espectral mais dinâmica.

Se as características físicas do ruído demandarem um *range* de Q diferente daquele fixado pelo algoritmo, o espectrograma resultante sofrerá distorções morfológicas, de modo que *glitches* rápidos de banda larga tendem a ter sua energia espalhada no espectrograma, enquanto sinais de banda estreita sofrem um alargamento temporal artificial (Zevin et al., 2017). Como a CNN toma decisões baseadas estritamente na geometria dessas imagens, essas distorções mascaram a assinatura real do ruído e induzem o modelo ao erro.

Durante a parte da O3b no detector de Livingston, identificou-se uma alteração no comportamento temporal do ruído de *Scattered Light*. Em virtude da mitigação de uma das causas físicas geradoras dessa classe, percebeu-se que, ao passo que deveria haver uma queda considerável na taxa de ocorrência dessa morfologia, o *Gravity Spy* classificou erroneamente a maioria desses eventos como *scattered Light*, quando, na realidade, a morfologia seria de *Low Frequency Lines* (Figura 3.9).

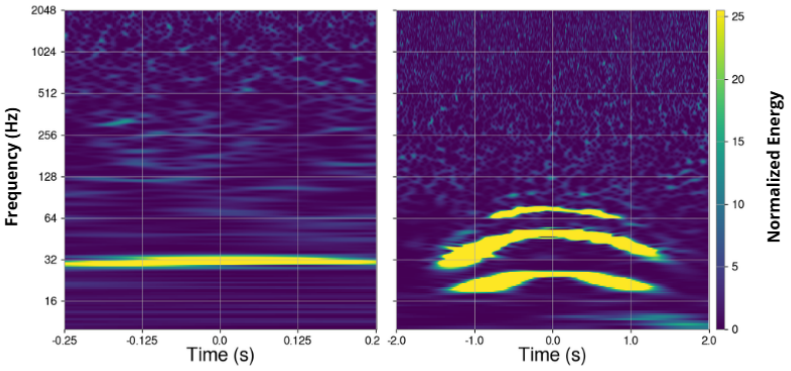

Figura 3.9: Espectrogramas tempo-frequência onde à esquerda, apresenta-se um *glitch* pertencente à classe *Low Frequency Lines* (com excesso de energia estacionário e de banda estreita concentrado em torno de 32 Hz) que foi rotulado incorretamente pelo algoritmo como *Scattered Light*. À direita, observa-se um *glitch* correto da classe *Scattered Light*, exibindo a morfologia característica de arcos paralelos e simétricos na região de baixas frequências.

Conforme ilustrado na Figura 3.10, essa contaminação cruzada teria afetado ao menos

metade das classificações que possuíam um limiar de confiança superior a 90% no período.

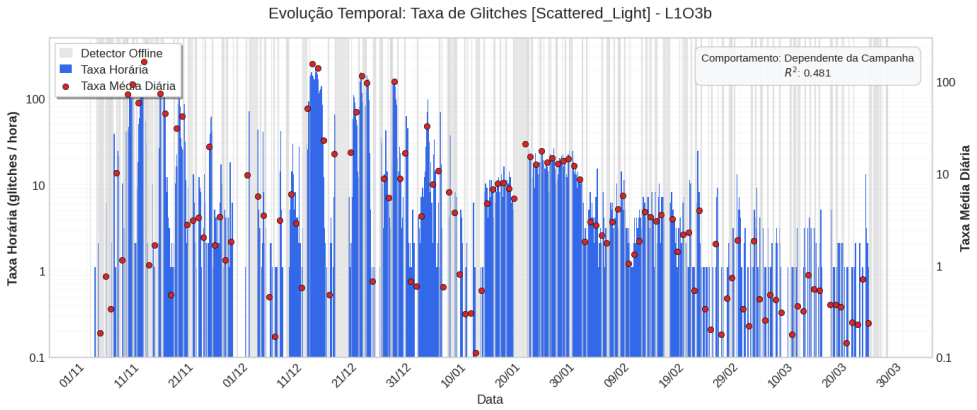

Figura 3.10: Evolução temporal e alteração no comportamento de ocorrência dos glitches de *Scattered Light* contaminados com *Low Frequency Lines* no detector L1 durante a O3b. Nota-se mudança comportamental nas taxas de ocorrência a partir de 20/01, época em que a uma das causas da classe *Scattered Light* foi mitigada.

Outra classe problemática nesta faixa espectral é a *Low Frequency Burst*. Foi observado que cerca de 40% das ocorrências atribuídas a essa categoria exibem anomalias. Como demonstrado na Figura 3.11, esses erros dividem-se entre sinais que apresentam um *burst* duplicado confinado no mesmo intervalo temporal ou eventos que manifestam outro sinal semelhante em frequências superiores.

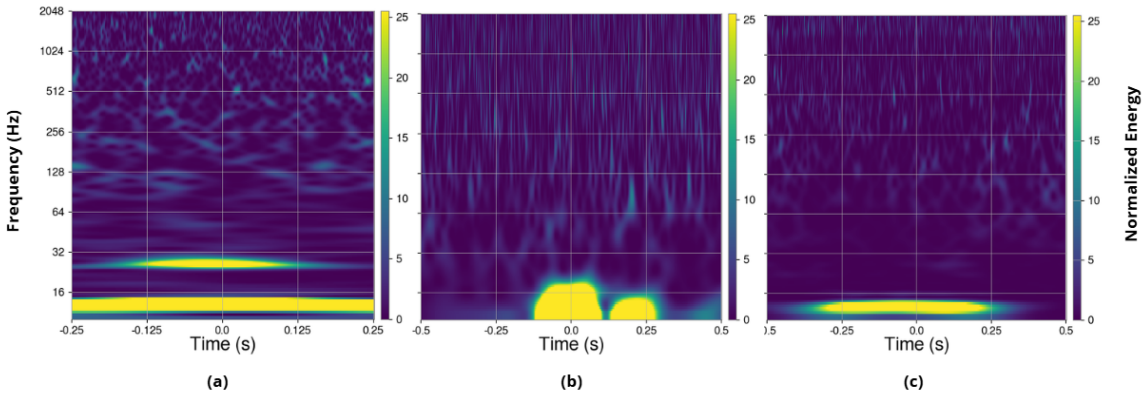

Figura 3.11: Espectrogramas da classe *Low Frequency Burst* evidenciando desvios estruturais: (a) transientes duplicados no mesmo sinal e (b) contrapartida em bandas de frequência superiores. (c) espectrograma de referência de um *Low Frequency Burst*.

Para contextualizar esse caso, é válido notar que o detector europeu *Virgo* aborda esse tipo de ruído de maneira distinta em seu *pipeline* de caracterização. O sistema do *Virgo* aplica um algoritmo de agrupamento temporal mais agressivo (Acernese et al., 2023), onde esses casos de sinais que ocorrem em rápida sucessão ou com réplicas em frequências um

pouco mais altas são interpretados como componentes de um único fenômeno físico, sendo direcionados para classes morfológicas dedicadas, como *Double Blips* ou *Crow*. No caso do LIGO, a ausência desse agrupamento prévio faz com que o *Gravity Spy* não separe esses fenômenos, e classifique como uma morfologia genérica de *Low Frequency Burst*, possivelmente comprometendo a consistência estatística dos dados.

No regime de frequências mais altas, a classe *Whistle* apresentou um desvio crítico durante um período anormal de alta taxa de ocorrência. A verificação demonstrou que cerca de 80% dos eventos catalogados pelo modelo nesse período não correspondiam à morfologia real de um *Whistle*. As discrepâncias entre a morfologia de um *whistle* e os *glitches* que contaminaram o período são apresentados na Figura 3.12.

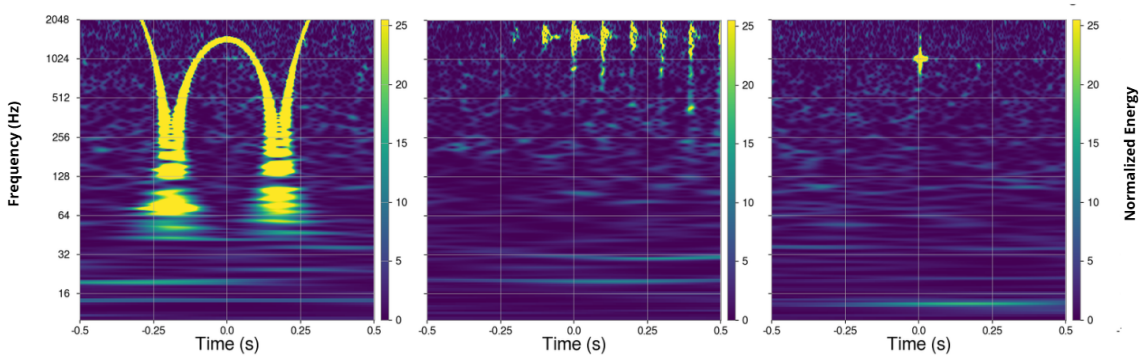

Figura 3.12: Comparação morfológica entre um ruído *Whistle* à esquerda e duas das classificações incorretas geradas pelo algoritmo durante o período de alta taxa de ocorrência à direita.

Para mitigar o impacto das classificações incorretas identificadas, os intervalos temporais mais contaminados do *Whistle* foram removidos do conjunto de dados, e a análise do coeficiente de determinação R^2 foi realizada novamente a partir dessa nova seleção. Neste caso em especial, o perfil da classificação fenomenológica dos *Whistles* permaneceu igual, continuando a apresentar valores de R^2 consistentemente abaixo de 0.5. A classe então mantém-se como um forte desvio ao modelo estatístico de Poisson, evidenciando que sua dinâmica de ocorrência temporal é de maneira geral governada por processos não estocásticos.

Capítulo 4

Conclusões

Este trabalho teve como objetivo investigar a origem física e o comportamento temporal de diferentes morfologias de *glitches* nos interferômetros de ondas gravitacionais LIGO (Hanford e Livingston) durante as corridas observacionais O3 e O4. Através da análise estatística baseada na distribuição de Poisson e da análise evolutiva das taxas de ocorrência diárias e horárias, foi possível avaliar ruídos poissonianos de *glitches* fomentados por processos ambientais e instrumentais não poissonianos.

A análise estatística revelou que as dinâmicas de ocorrência variam drasticamente entre as morfologias. Classes como *Blip*, *Extremely Loud* e *Koi Fish* apresentaram um coeficiente de determinação (R^2) consistentemente alinhado ao comportamento poissoniano, indicando que seus mecanismos geradores geram eventos de forma independente e aleatória. Em contrapartida, outras categorias como *Fast Scattering*, *Whistle* e *Paired Doves* exibiram fortes desvios da distribuição de Poisson, evidenciando causas físicas não independentes no tempo. Constatou-se também que o aprimoramento da sensibilidade dos detectores na faixa de baixas frequências, implementado para a corrida O4, tornou os instrumentos significativamente mais suscetíveis aos ruídos associados à movimentação sísmica, elevando de forma massiva a ocorrência das classes *Scattered Light*, *Low Frequency Lines* e *Low Frequency Burst*.

Concluiu-se com base na evolução temporal que as classes *Extremely Loud* e *Koi Fish* não apenas possuem taxas de ocorrência temporal quase idênticas, mas também mantêm ou melhoram sua aderência estatística quando tratadas como uma única categoria de *glitch*. Isso sugere fortemente que ambas partilham da mesma origem física, tornando o *Extremely Loud* com uma versão de alta energia (alta Relação Sinal-Ruído) do *Koi Fish*, cuja potência satura a visualização no espectrograma. Já em relação aos ruídos de baixas frequências

como *Scattered Light*, *Low Frequency Lines* e *Low Frequency Burst*, demonstrou-se que, apesar de compartilharem a mesma causa física associada aos sismos, não há uma fonte física única regendo todas simultaneamente. Elas respondem a estímulos distintos, intercalando dias de picos perfeitamente correlacionados com períodos de comportamento independente.

Além de caracterizar as famílias de ruídos, o estudo expôs severas limitações no *pipeline* de classificação automatizada atual, o *Gravity Spy*. O uso de uma Transformada Q otimizada para maximizar a Relação Sinal-Ruído global induz distorções visuais nos espectrogramas, fazendo com que o algoritmo frequentemente confunda e caracterize erroneamente classes de *glitches* com características semelhantes. O sistema também revelou a ausência de metodologias de classificação e separação de morfologias distintas na classe *Low Frequency Burst*, o que resulta em ao menos 3 morfologias diferentes classificadas como uma só, diferentemente do processo utilizado pelo Virgo. Além de registrar altos índices de erro na identificação de ruídos *Whistle*. Tais limitações indicam que a análise do LIGO pode se beneficiar futuramente da implementação de algoritmos de aprendizado de máquina não supervisionado, que não dependam exclusivamente de rótulos pré-definidos.

Em suma, a caracterização fenomenológica detalhada desenvolvida neste trabalho oferece um panorama robusto para a interpretação dos transientes no LIGO. A análise desenvolvida atua, portanto, como um subsídio relevante para viabilizar novas estratégias de mitigação. A superação dos desafios causados pelos *glitches* é indispensável não apenas para evitar a contaminação e taxas de falso alarme de sinais reais, mas para elevar o nível de confiança das detecções na era da astronomia multimensageira.

Referências Bibliográficas

Aasi J., et al., Advanced LIGO, Classical and Quantum Gravity, 2015, vol. 32, p. 074001

Abadie J., et al., A gravitational wave observatory operating beyond the quantum shot-noise limit, Nat. Phys., 2011, vol. 7, p. 962

Abbott B. P., et al., GW150914: The Advanced LIGO Detectors in the Era of First Gravitational-Wave Observations, Physical Review Letters, 2016a, vol. 116, p. 131103

Abbott B. P., et al., Observation of Gravitational Waves from a Binary Black Hole Merger, Phys. Rev. Lett., 2016b, vol. 116, p. 061102

Abbott B. P., et al., GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral, Physical Review Letters, 2017, vol. 119

Abbott B. P., et al., Tests of general relativity with the binary black hole signals from the LIGO–Virgo catalog GWTC-1, Phys. Rev. D, 2019, vol. 100, p. 104036

Abbott B. P., et al., A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave signals, Classical and Quantum Gravity, 2020, vol. 37, p. 055002

Abbott R., Abbott T., Abraham S., Acernese F., Ackley K., Adams C., Adhikari R., Adya V., Affeldt C., Aguiar O., et al., Characterization of transient noise in Advanced LIGO during the first half of the third observing run, Classical and Quantum Gravity, 2021, vol. 38, p. 134014

Abbott R., et al., GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo during the Second Part of the Third Observing Run, Phys. Rev. X, 2023, vol. 13, p. 041039

Abbott R., et al., GWTC-2.1: Deep extended catalog of compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run, Phys. Rev. D, 2024, vol. 109, p. 022001

Acernese F., et al., Virgo Detector Characterization and Data Quality: results from the O3 run, Classical and Quantum Gravity, 2023, vol. 40, p. 185006

Arnaud N., , 2022 LIGO-Virgo detector characterization and data quality (DetChar): from the O3 performance to the O4 preparation

Ashton G., Gaussian processes for glitch-robust gravitational-wave astronomy, Mon. Not. R. Astron. Soc., 2023, vol. 520, p. 2983

Bahaadini S., Noroozi V., Rohani N., Coughlin S., Zevin M., Smith J. R., Kalogera V., Katsaggelos A. K., Deep learning for gravity spy: Glitch classification and dataset manually normalized, Information Sciences, 2018, vol. 444, p. 172

Buikema A., et al., Sensitivity and performance of the Advanced LIGO detectors in the third observing run, Phys. Rev. D, 2020, vol. 102, p. 062003

Capote E., et al., Advanced LIGO detector performance in the fourth observing run, Phys. Rev. D, 2025, vol. 111, p. 062002

Chatterji S., Blackburn L., Martin M., Katsavounidis E., Multiresolution techniques for the detection of gravitational-wave bursts, Classical and Quantum Gravity, 2004, vol. 21, p. S1809

Collaboration T. L. S., the Virgo Collaboration the KAGRA Collaboration et al.,, 2026 GWTC-5.0: An Introduction to Version 5.0 of the Gravitational-Wave Transient Catalog

Davis D., et al., LIGO detector characterization in the second and third observing runs, Class. Quantum Grav., 2021, vol. 38, p. 135014

Durran J. H., SMP: Statistics and Probability. Cambridge University Press, 1970

Einstein A., Näherungsweise Integration der Feldgleichungen der Gravitation, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin (Erster Halbband), 1916, pp 688–696

Einstein A., Über Gravitationswellen, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin (Erster Halbband), 1918, pp 154–167

Fernandes T., Di Renzo F., Onofre A., Torres-Forné A., Font J. A., VIGILant: an automatic classification pipeline for glitches in the Virgo detector, arXiv preprint arXiv:2604.13687, 2026

Ferreira T. A., González G., Using t-SNE for characterizing glitches in LIGO detectors, Classical and Quantum Gravity, 2025, vol. 42, p. 105010

Glanzer J., et al., Data quality up to the third observing run of advanced LIGO: Gravity Spy glitch classifications, Class. Quantum Grav., 2023, vol. 40, p. 065004

LIGO Scientific Collaboration, 2024 LIGO Detects Dozens of Gravitational-Wave Candidates in First Months of New Observing Run

Longo A., Bianchi S., Valdes G., Arnaud N., Di Renzo F., Montani M., Plastino W., Identification and monitoring of scattered light noise sources in laser interferometers with adaptive algorithms, Journal of Applied Physics, 2025, vol. 138, p. 090701

Rice J. A., Mathematical Statistics and Data Analysis 3rd edn. Cengage Learning Belmont, CA, 2007

Robinet F., et al., Omicron: A tool to characterize transient noise in gravitational-wave detectors, SoftwareX, 2020, vol. 12, p. 100620

Rudenko V. N., Sazhin M. V., Laser interferometer as a gravitational wave detector, Soviet Journal of Quantum Electronics, 1980, vol. 10, p. 1366

Zevin M., Coughlin S., et al., Gravity Spy: integrating citizen science and machine learning for glitch classification in LIGO data, Classical and Quantum Gravity, 2017, vol. 34, p. 064002

Zevin M., et al., Gravity Spy: Lessons Learned and a Path Forward, Eur. Phys. J. Plus, 2024, vol. 139, p. 33

Apêndice

Apêndice A

Distribuição temporal comparativa das classes *Scattered Light*, *Low Frequency Lines* e *Low Frequency Burst*

Gráficos de evolução temporal comparativa para as classes *Scattered Light*, *Low Frequency Lines* e *Low Frequency Burst* ao longo das diferentes campanhas observacionais e detectores. É válido contextualizar que, durante a O3, a taxa de ocorrência de *Scattered Light* era muito superior à das outras duas classes e consideravelmente maior do que os níveis registrados mais tarde na O4. Essa grande quantidade de eventos na O3 aconteceu porque, no período, o interferômetro não era tão sensível para as frequências mais baixas, o que fez com que os interferômetros não fossem tão afetados pelos *Low Frequency Lines* e *Low Frequency Burst* quanto na O4.

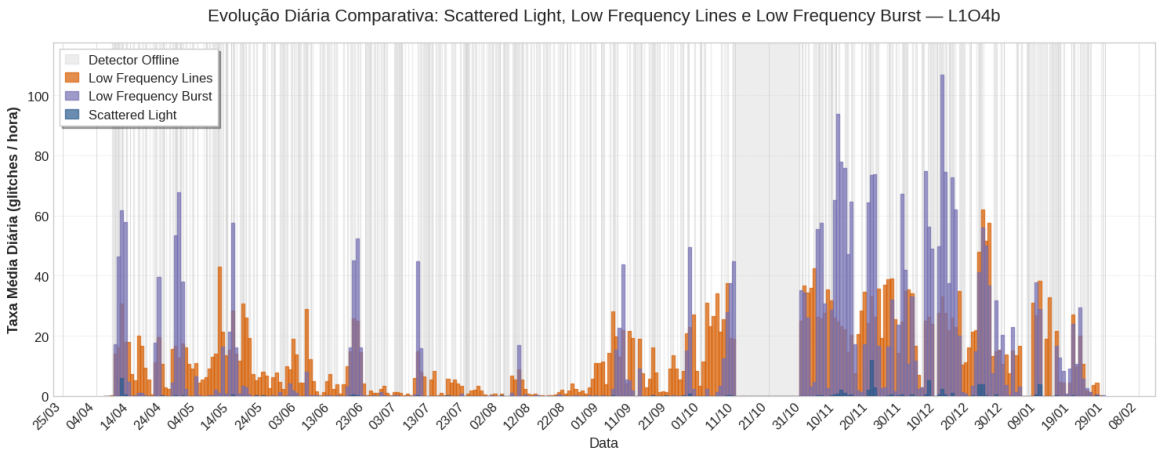

Figura A.1: Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Livingston ($L1$) durante a corrida observacional O4b.

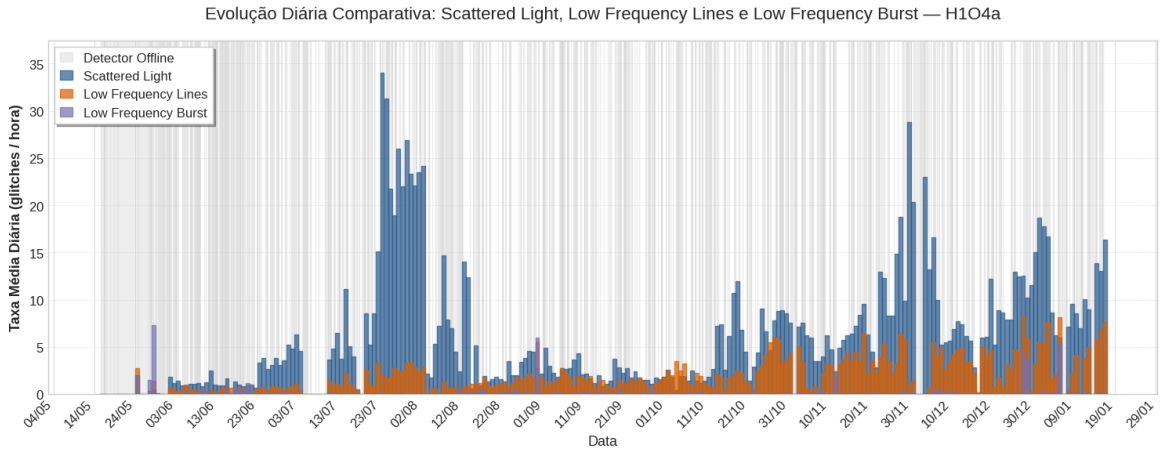

Figura A.2: Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Hanford ($H1$) durante a corrida observacional O4a.

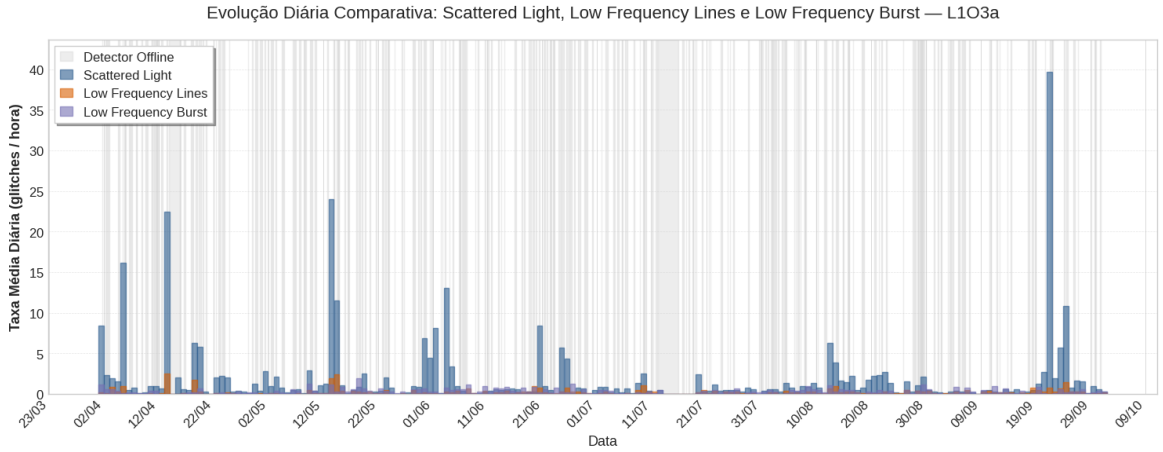

Figura A.3: Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Livingston ($L1$) durante a corrida observacional O3a.

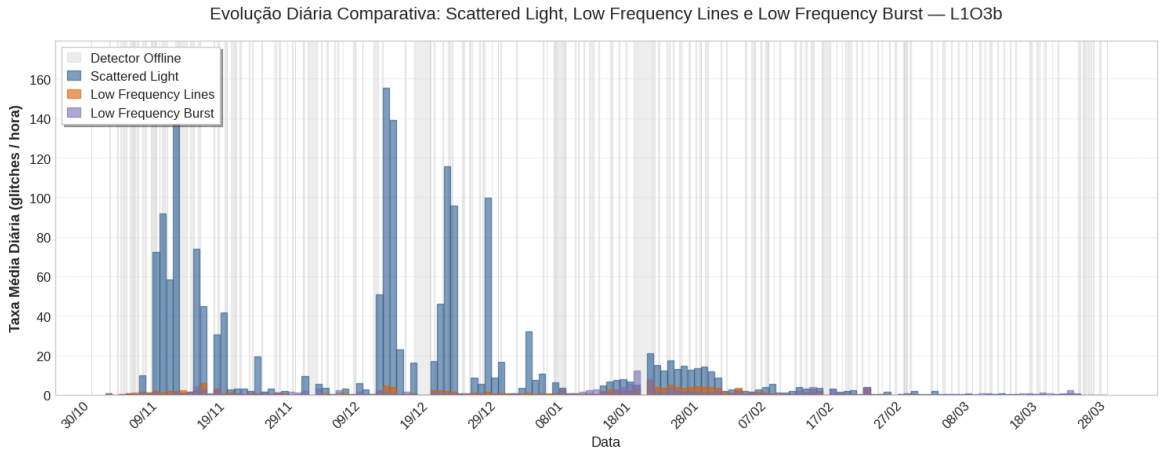

Figura A.4: Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Livingston ($L1$) durante a corrida observacional O3b.

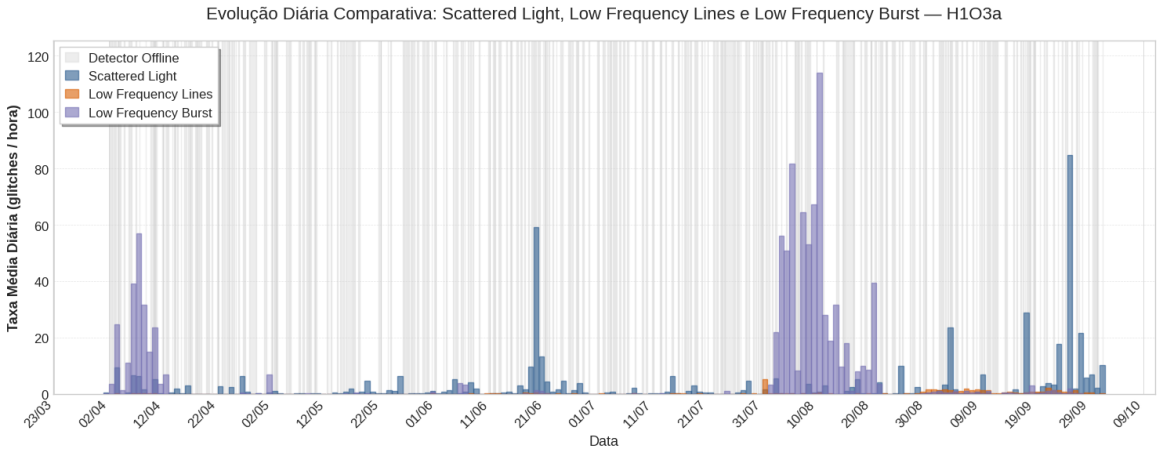

Figura A.5: Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Hanford ($H1$) durante a corrida observacional O3a.

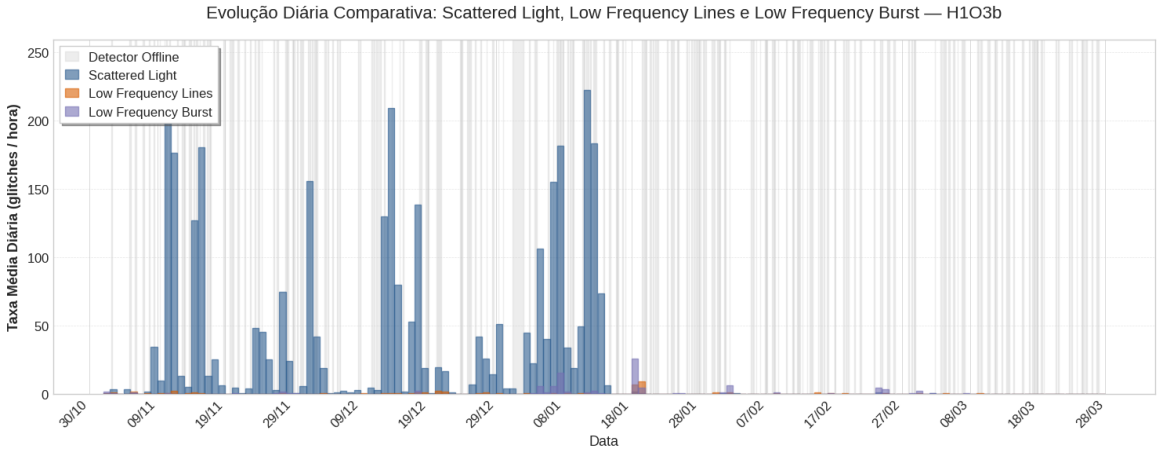

Figura A.6: Evolução temporal comparativa das taxas médias diárias de *glitches* para o detector de Hanford ($H1$) durante a corrida observacional O3b.

Apêndice B

Caracterização do comportamento dos *glitches* durante a corrida O3

Este apêndice traz a tabela de resumo das morfologias de *glitches* da campanha observacional O3, que foi obtida como resultado direto da primeira parte deste Trabalho de Graduação (TG1).

Tabela B.1 - Resumo das morfologias de *glitches* entre os detectores e os períodos da campanha observacional O3, mostrando o coeficiente de determinação R^2 entre as distribuições de Poisson observadas e esperadas. A última coluna traz uma interpretação qualitativa do comportamento, baseada na concordância geral das distribuições reais com o modelo de Poisson.

Morfologia do Glitch	**L1 O3a** R^2	**L1 O3b** R^2	**H1 O3a** R^2	**H1 O3b** R^2	**Comportamento Fenomenológico**
1400Ripples	0.508	—	—	—	Fortes Desvios
Air Compressor	—	0.200	—	—	Fortes Desvios
Blip	0.940	0.932	0.909	0.887	Poissoniano
Blip Low Frequency	0.957	0.903	0.920	0.698	Dependente da Campanha
Extremely Loud	0.970	0.977	0.916	0.962	Poissoniano
Fast Scattering	0.269	0.277	—	0.322	Fortes Desvios
Koi Fish	0.953	0.952	0.922	0.963	Poissoniano
Low Frequency Burst	0.562	0.602	0.209	0.290	Dependente da Campanha
Low Frequency Lines	0.552	0.605	0.364	0.504	Dependente da Campanha
No Glitch	0.490	0.675	0.315	—	Fortes Desvios
Power Line	0.204	—	—	—	Fortes Desvios
Repeating Blips	—	—	0.850	—	Poissoniano
Scattered Light	0.370	0.481	0.279	0.236	Fortes Desvios
Tomte	0.929	0.901	0.851	0.503	Dependente da Campanha
Whistle	0.623	0.379	0.294	0.415	Fortes Desvios