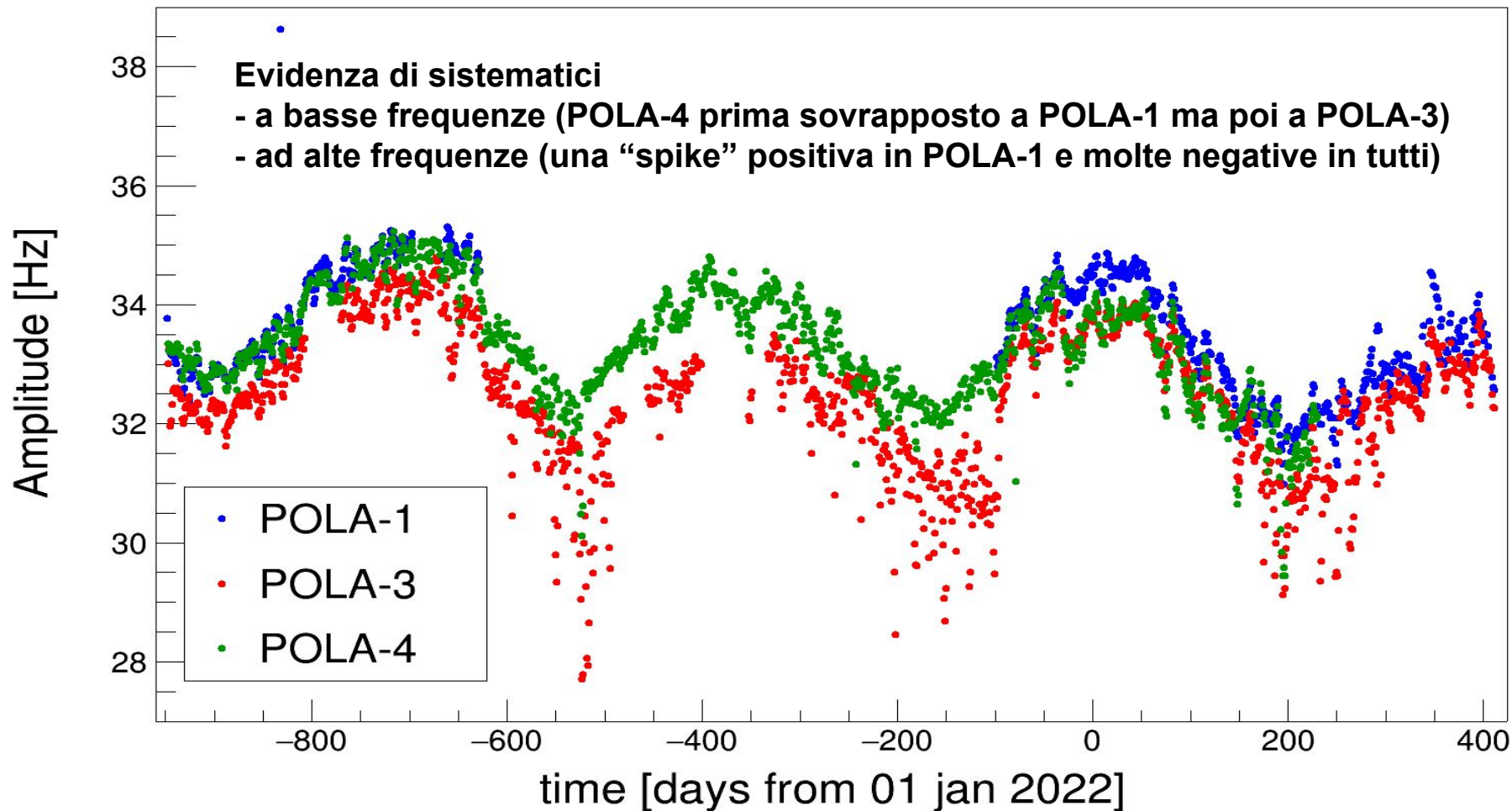


Test su ricerca di periodicità nelle time series di EEE



GOALS:

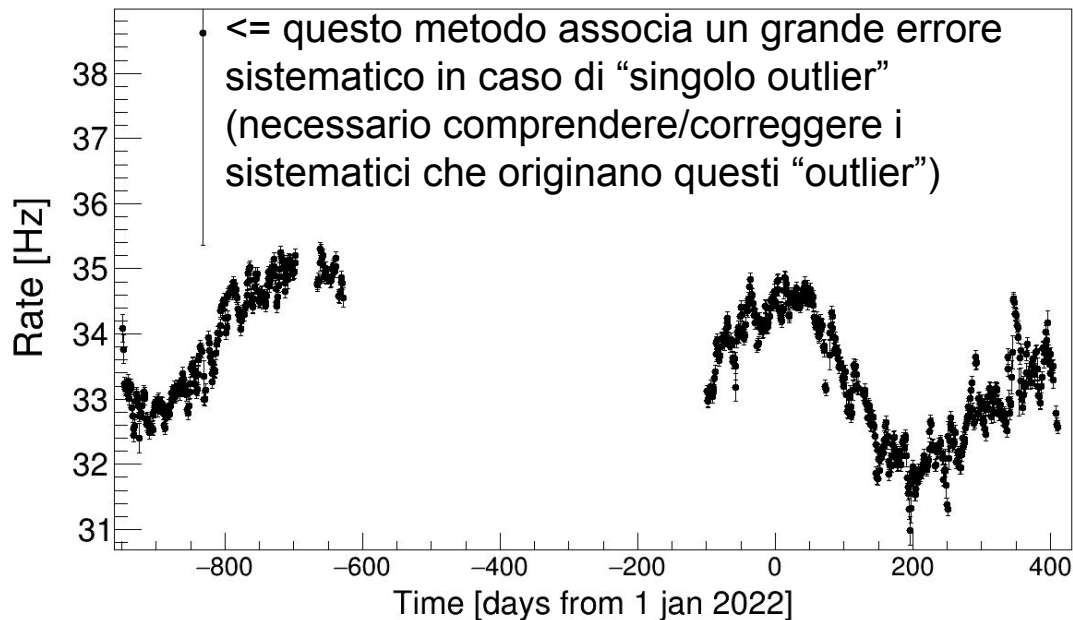
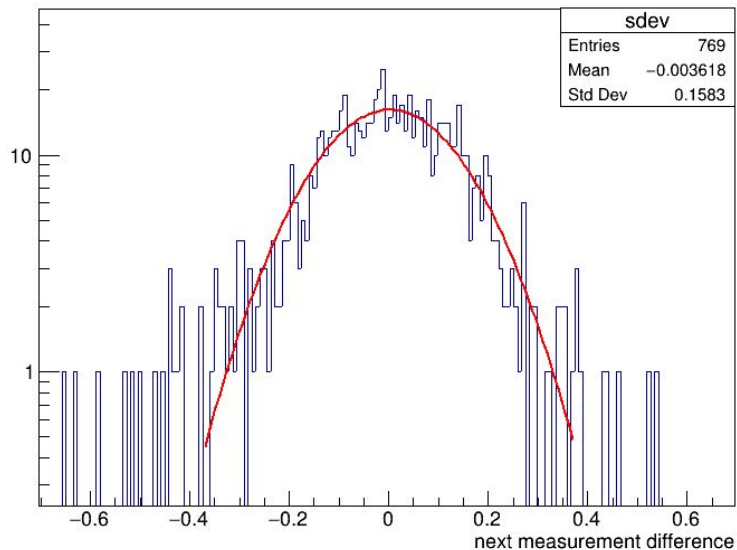
- 1) Gestire time series con “buchi” e/o campionamenti non uniformi**
- 2) Tenere in conto delle incertezze sperimentali associate ai campioni della time serie**
- 3) Valutare significatività statistica delle eventuali periodicità osservate**

Problema: definire l'incertezza statistica/sistemica della misura

Soluzione proposta:

- 1) Costruire istogramma della differenza $\text{rate}[i]-\text{rate}[i-1]$ (è una sovrastima dell'errore statistico)
- 2) Definire $\text{err}_{\text{stat}} = \sigma/\sqrt{2}$ dal fit gaussiano del bulk
- 3) Associare al punto $\text{err}_{\text{tot}}(i) = \text{Max}(\text{err}_{\text{stat}}, \text{Min}(|\text{rate}[i]-\text{rate}[i-1]|, |\text{rate}[i]-\text{rate}[i+1]|))$

Example: POLA-01

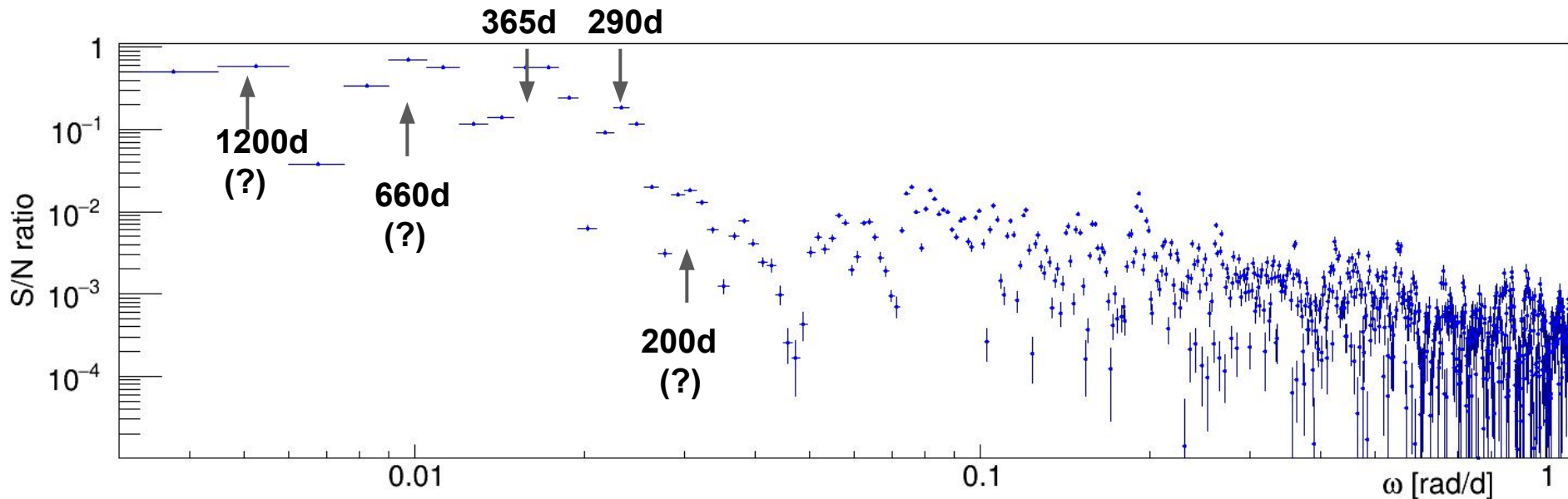


Ricerca di periodicità nella rate.

Metodo1: periodogramma Lomb-Scargle

Vantaggio: metodo “standard” ben noto (<https://arxiv.org/pdf/0901.2573.pdf>)

Dati: POLA-01 tutto il periodo

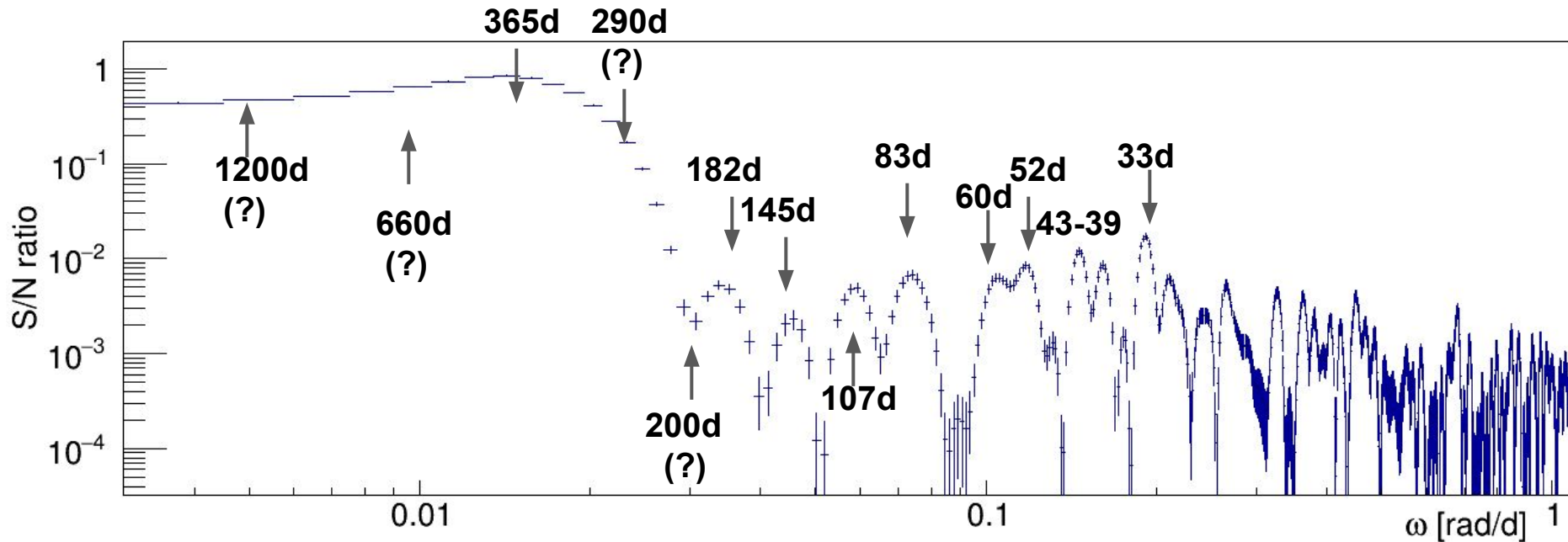


Problema: quali picchi sono significativi?

Metodo1: periodogramma Lomb-Scargle

Vantaggio: metodo "standard" ben noto (<https://arxiv.org/pdf/0901.2573.pdf>)

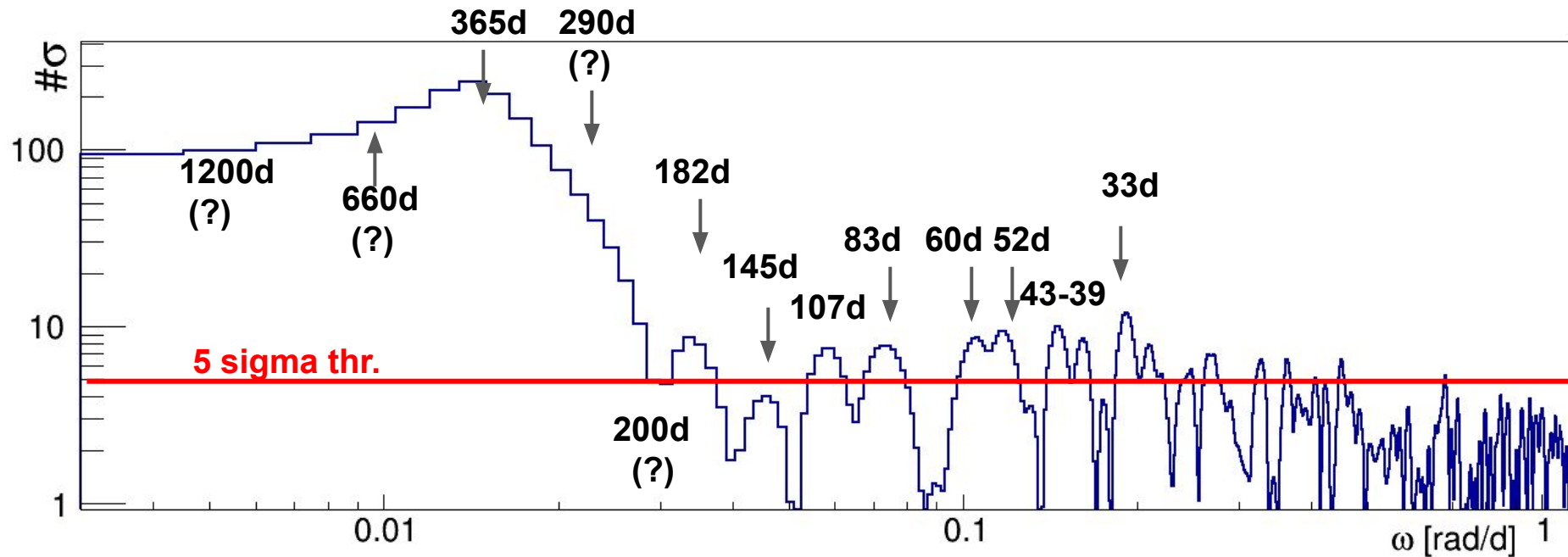
Dati: POLA-01 solo seconda tranche "dati curati da gaps ecc..."



Problema: i picchi sono riproducibili? => sistematici?

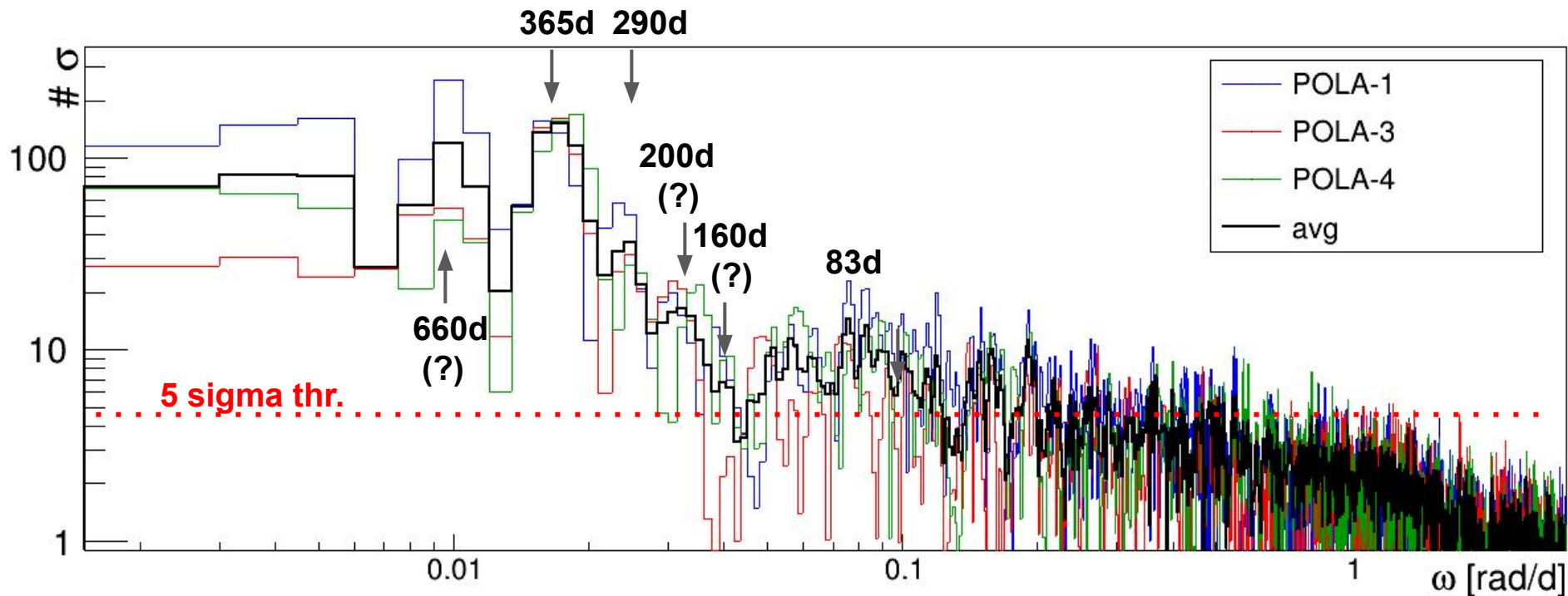
Valutazione **errore statistico**: resampling dei dati variando le misure secondo entro l'errore
50 pseudo-esperimenti => 50 periodogrammi da cui estrarre RMS => $val/RMS = \# \sigma$

Dati: POLA-01 solo seconda tranche "dati curati da gaps ecc..."



Problema: i picchi statisticamente significativi sono anche riproducibili? => sistematici?

Confronto POLA-1 POLA-3 POLA-4



Problema: i picchi statisticamente significativi sono anche riproducibili? => sistematici?

Probema0: percepito un trend pluriennale (drift detector o modulazione con periodo 11y?)

Problema1: non è chiaro come estrarre l'incertezza sulla frequenza (larghezza del picco?)

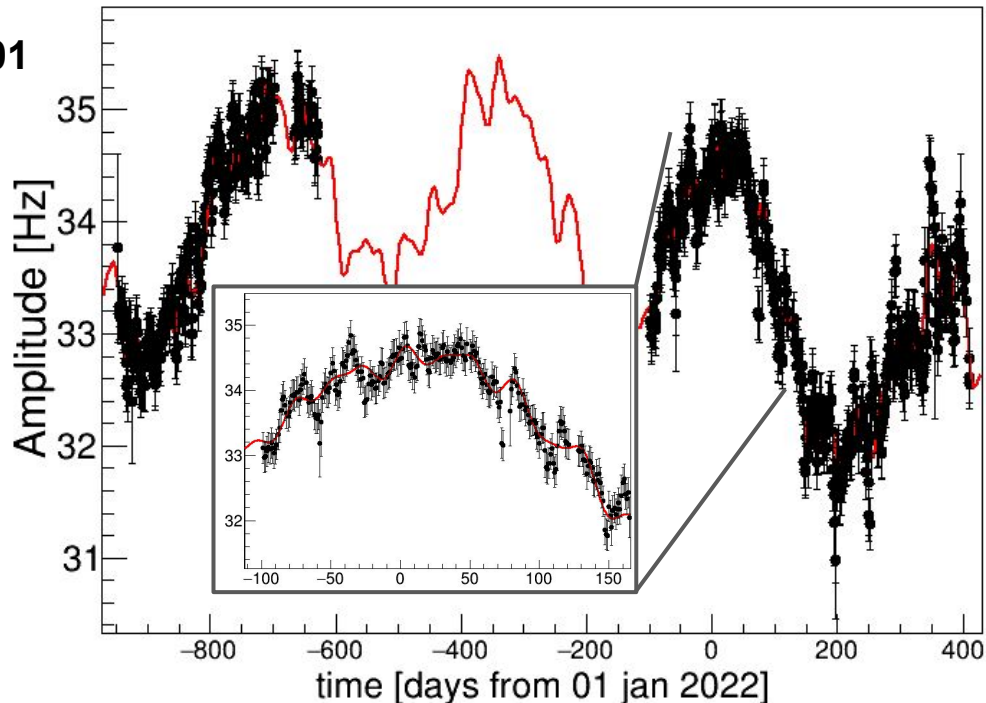
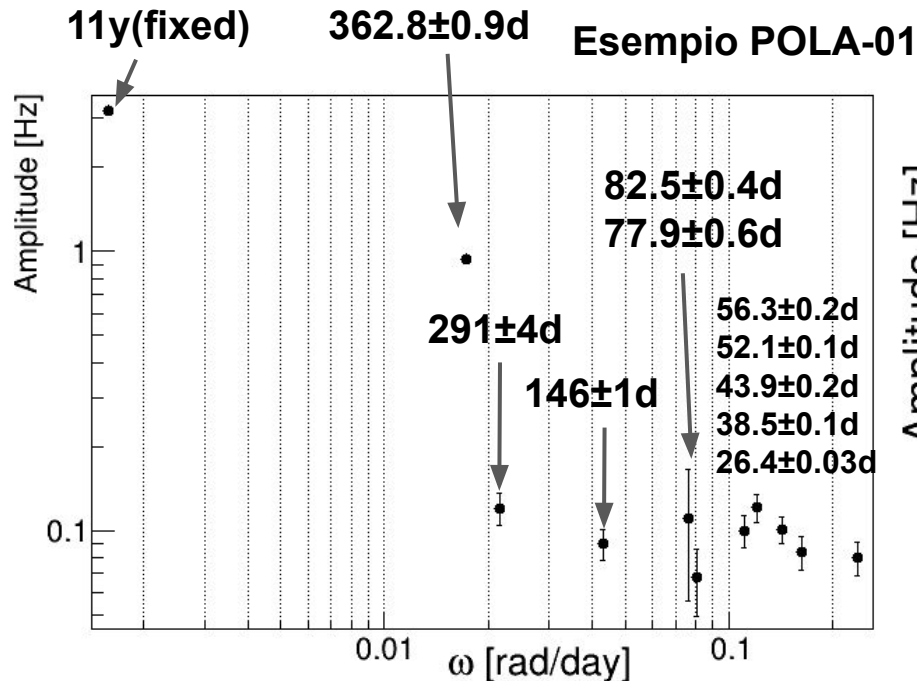
Problema2: “binning in omega”: tanti “picchi” vicini subito sotto soglia sono significativi?

Problema3: come interpretare il periodogramma in funzione delle ampiezze delle componenti?

Problema4: informazione complementare sulla fase?

Metodo2: Fit di N sinusoidi “brute-force”

Vantaggio1: applicazione immediata della statistica. N limitato con metodo del likelihood-ratio



Problema: non è facile far convergere un fit con 10 sinusoidi (30 param.) => molti minimi locali

Problema: serve rinormalizzare errore (x 2.4) per ottenere un ChiQuadro ragionevole

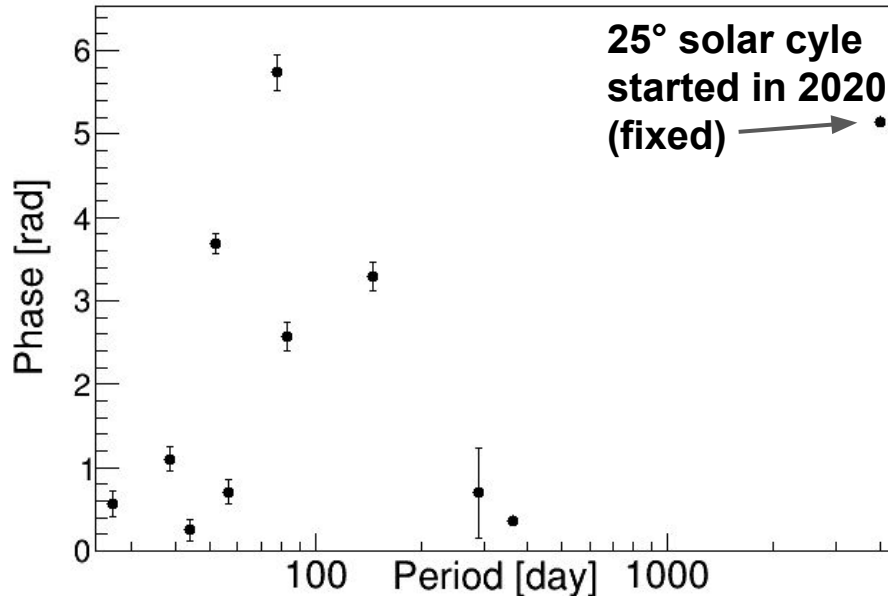
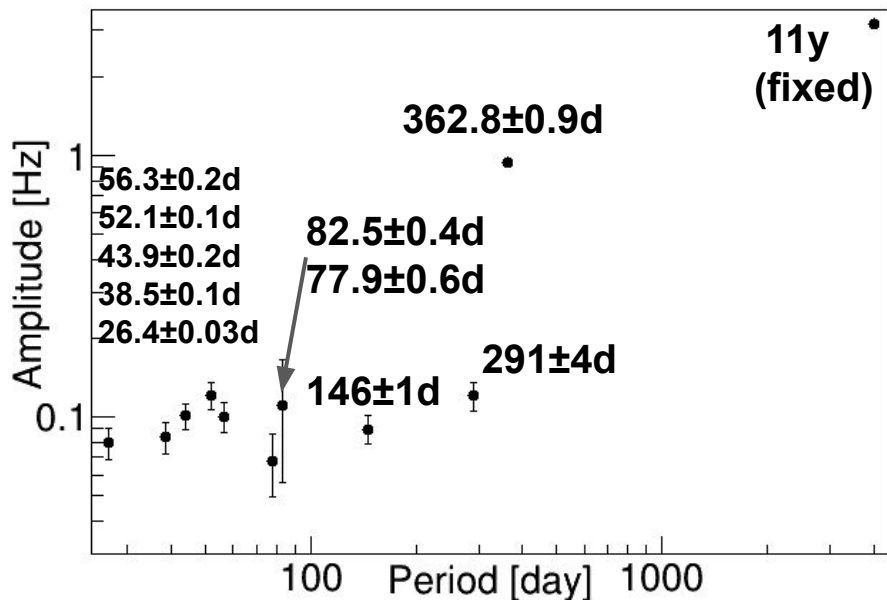
Problema: comunque gli errori ottenuti sono poco credibili rispetto alla riproducibilità

Osservata modulazione annuale di circa il 3%.

Misura complementare della fase:

Modulazione annuale con massimo il 20 ± 1 Gennaio

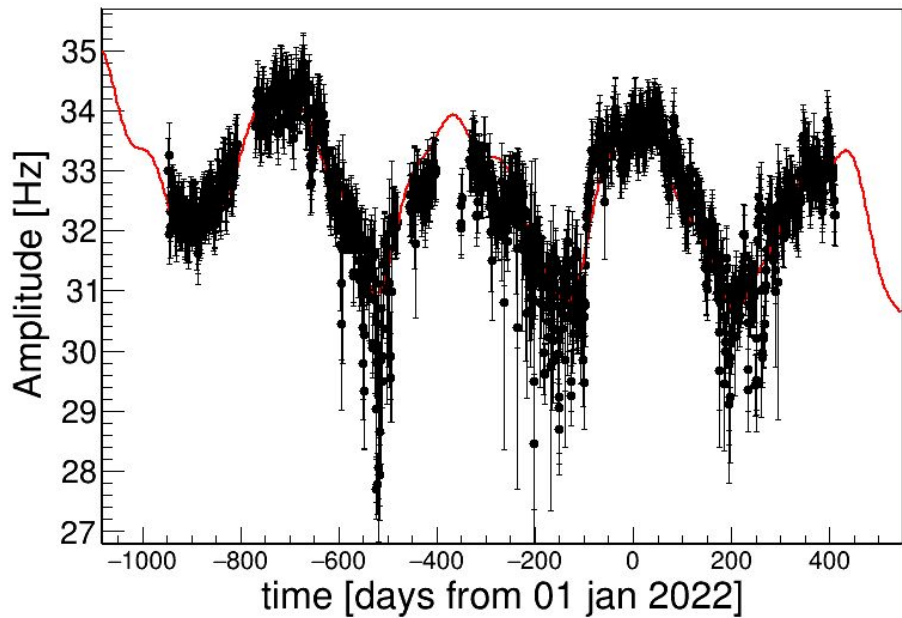
Esempio POLA-01



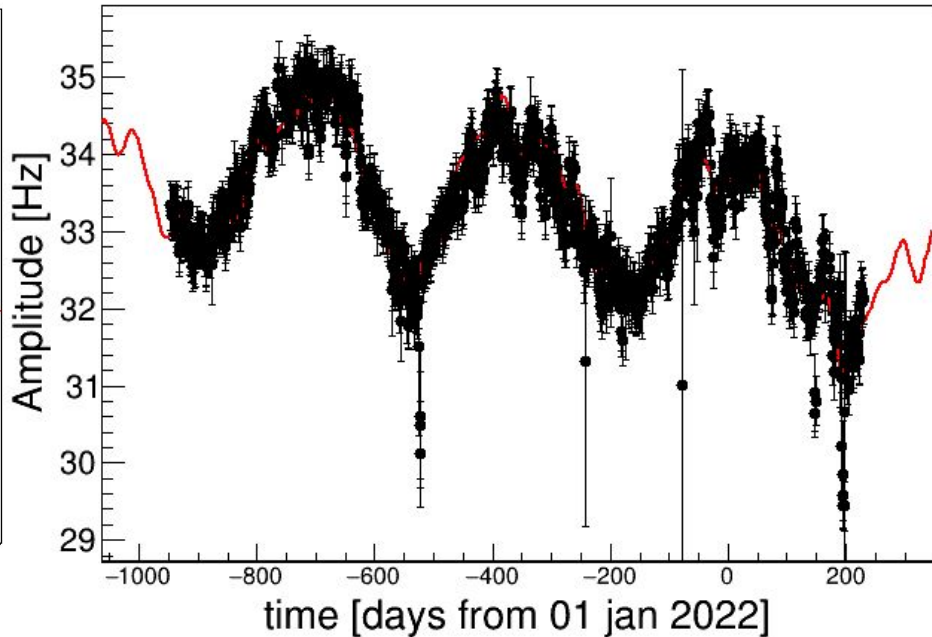
Ho provato a fissare la componente long-term con periodo 11y e fase 5.14 in maniera da avere il massimo di questa componente all'inizio del 2020 (inizio del 25° ciclo solare)

Possibile indicazione di una modulazione del 10% della rate dovuta al ciclo solare? Dubbia.

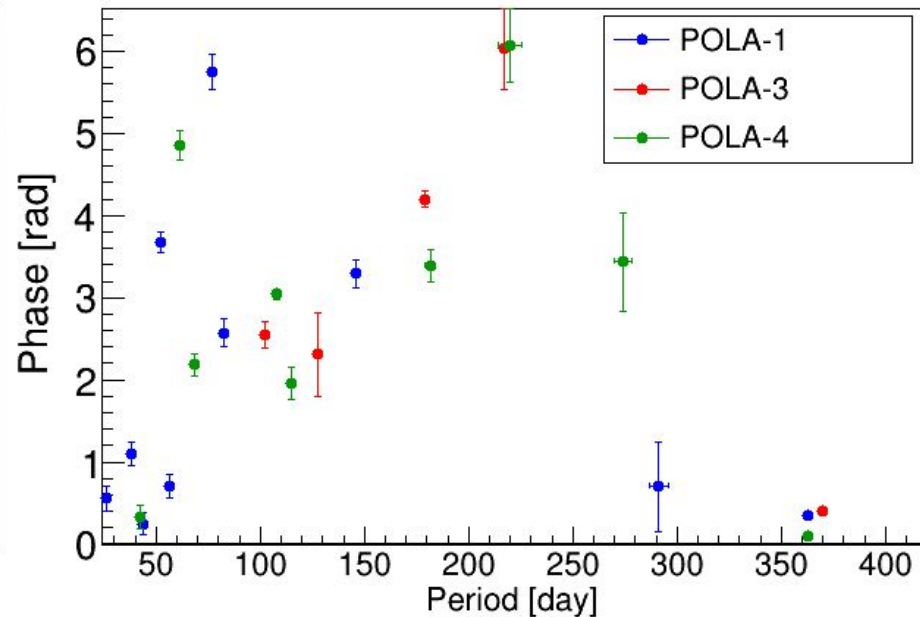
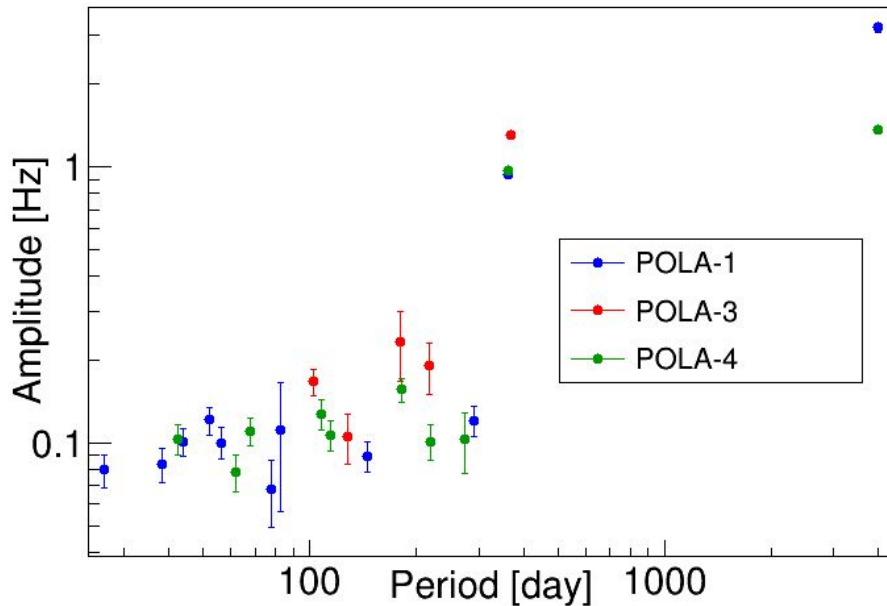
POLA-03



POLA-04



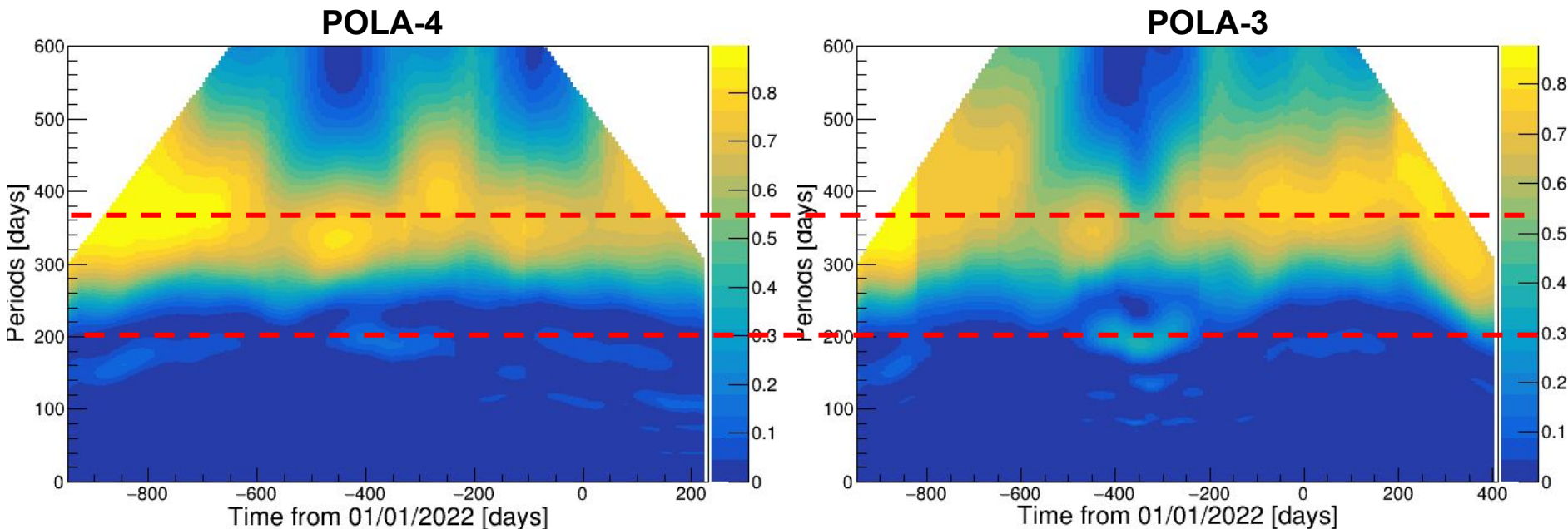
Confronto POLA-1 POLA-3 POLA-4



- Solo la Modulazione annuale è confermata contemporaneamente da tutti i POLA
- Evidenza di una sottostima notevole delle incertezze quotate nei fit results.
- Possibile mitigazione: su può provare a fittare contemporaneamente i tre datasets
- Da verificare se la freq/fase delle oscillazioni sub-annuali possa dipendere dal periodo dell'anno (cancellazione di parti con freq.fasi diverse è possibile)

Periodogramma Lomb-Scargle “TIME-DEPENDENT” (finestra 600d per “basse” frequenze)

Ogni “slice” verticale al tempo T è un periodogramma costruito con i dati [T-300, T+300] (quindi le slices in finestre di 600d non sono indipendenti)

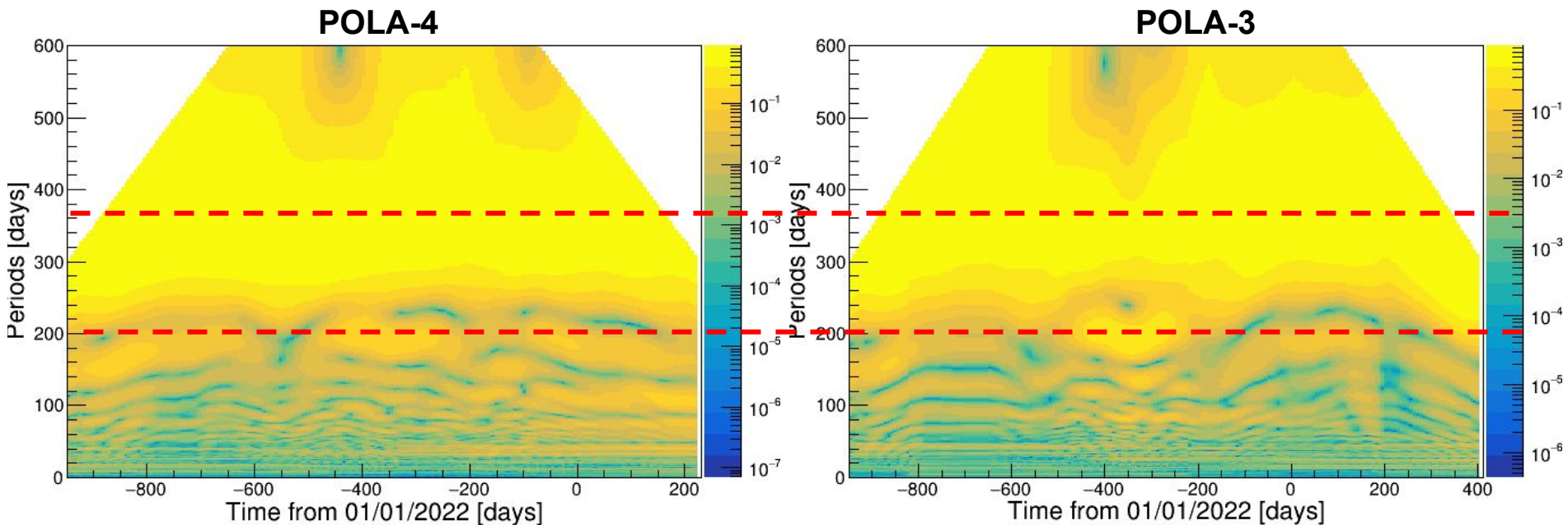


Effetto principale di modulazione annuale + possibile “banda” 200d (ma costante?)

interessante studiare cosa succede a frequenze più “alte” ...

Periodogramma Lomb-Scargle “TIME-DEPENDENT” (finestra 600d per “basse” frequenze)

Ogni “slice” verticale al tempo T è un periodogramma costruito con i dati [T-300, T+300] (quindi le slices in finestre di 600d non sono indipendenti)

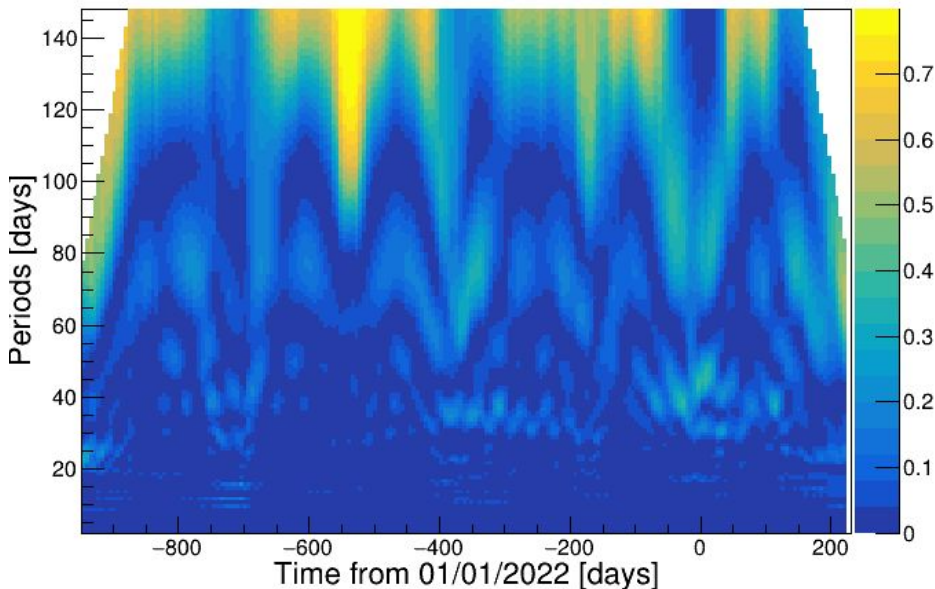


Effetto principale di modulazione annuale + struttura a bande che driftano oscillano in frequenza ma diverso comportamento (periodo/fase) tra POLA-3 e POLA-4 (?)
interessante studiare cosa succede a frequenze più “alte” ...

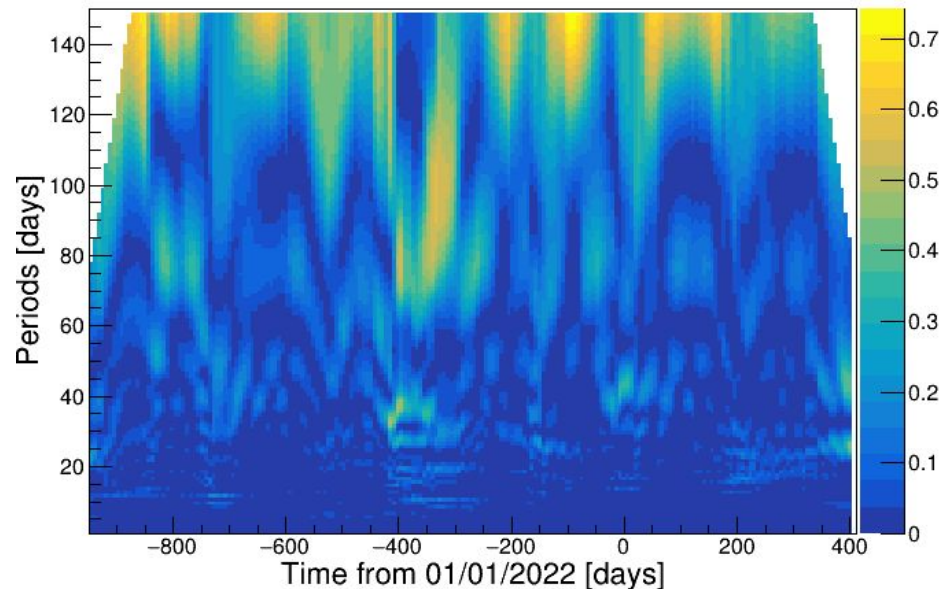
Periodogramma Lomb-Scargle “TIME-DEPENDENT” (finestra 150d per “medie” frequenze)

Ogni “slice” verticale al tempo T è un periodogramma costruito con i dati [T-75, T+75] (quindi le slices in finestre di 150d non sono indipendenti)

POLA-4



POLA-3

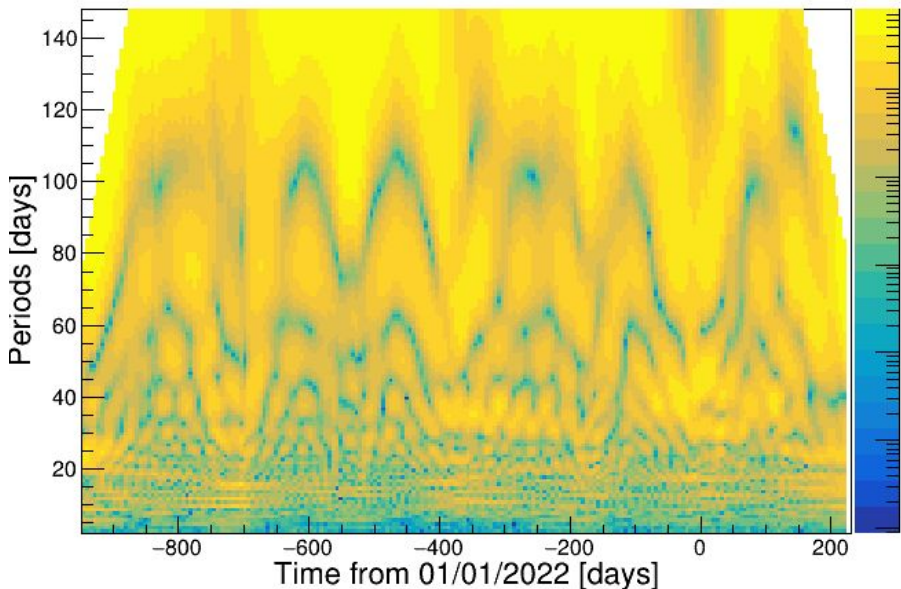


Struttura a bande con periodi oscillanti a T=80, 50, 35, 20 ... da approfondire

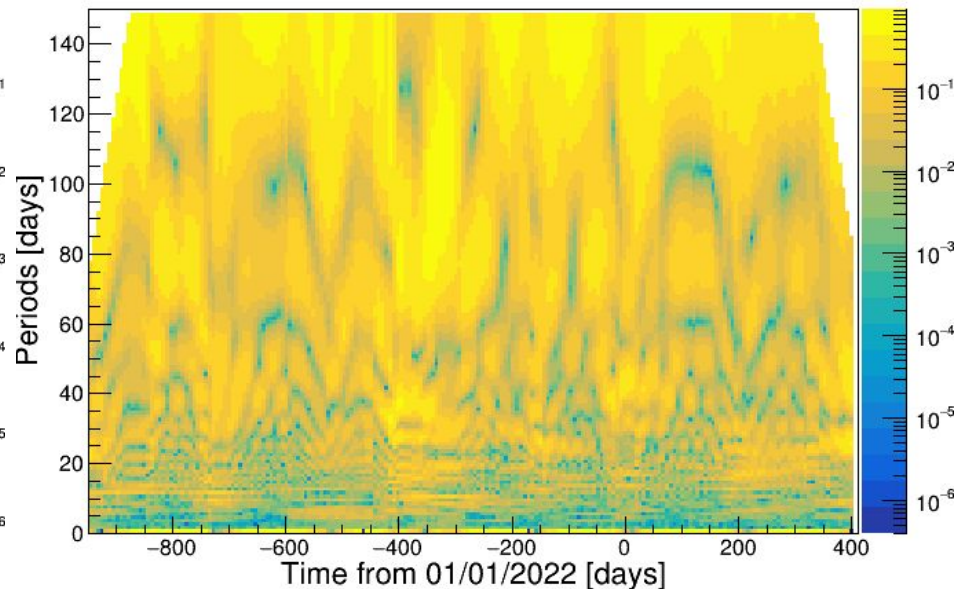
Periodogramma Lomb-Scargle “TIME-DEPENDENT” (finestra 150d per “medie” frequenze)

Ogni “slice” verticale al tempo T è un periodogramma costruito con i dati [T-75, T+75] (quindi le slices in finestre di 150d non sono indipendenti)

POLA-4



POLA-3



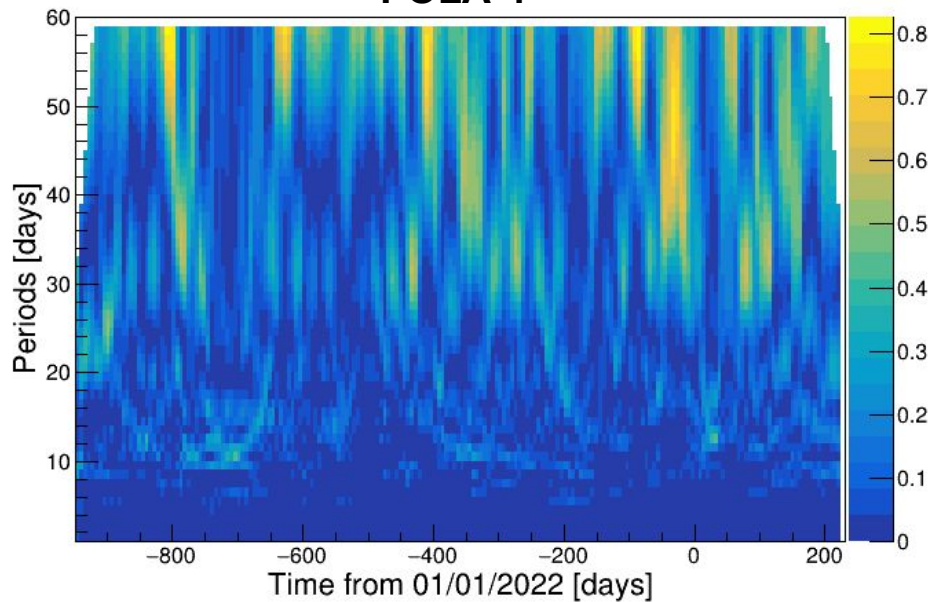
Complessa struttura di bande driftanti e periodiche

Struttura a bande non driftanti con periodi sotto 20d ... da approfondire

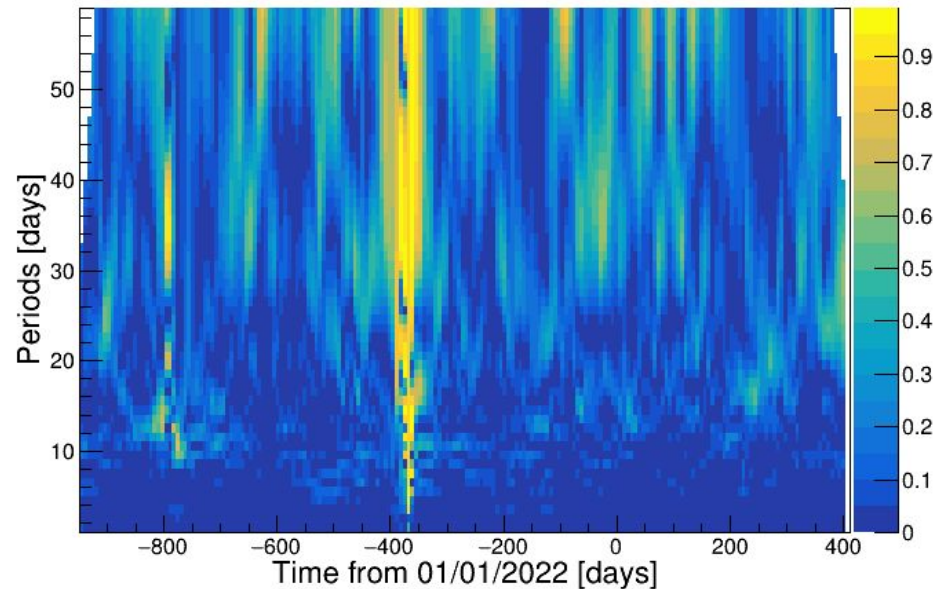
Periodogramma Lomb-Scargle “TIME-DEPENDENT” (finestra 60d per “alte” frequenze)

Ogni “slice” verticale al tempo T è un periodogramma costruito con i dati [T-30, T+30] (quindi le slices in finestre di 60d non sono indipendenti)

POLA-4



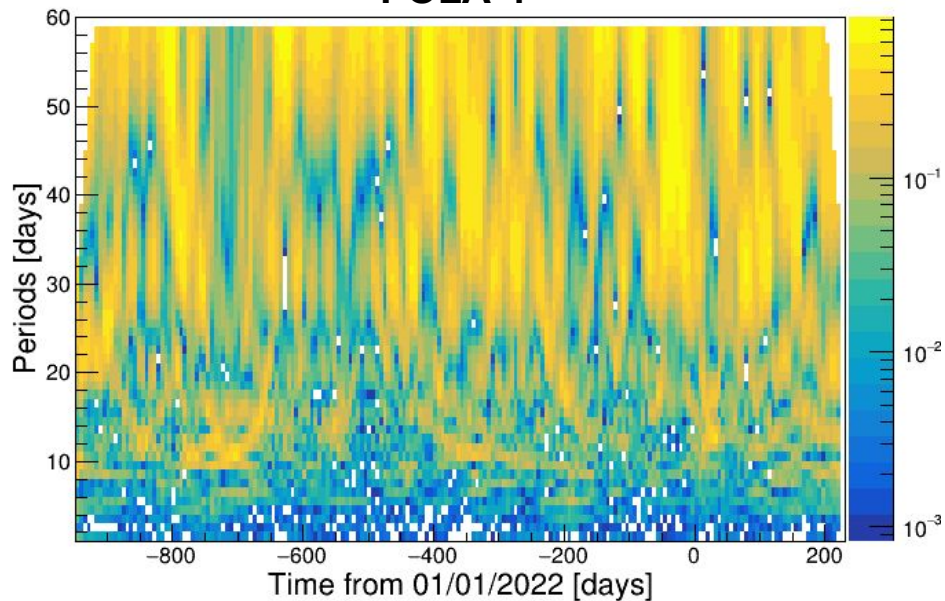
POLA-3



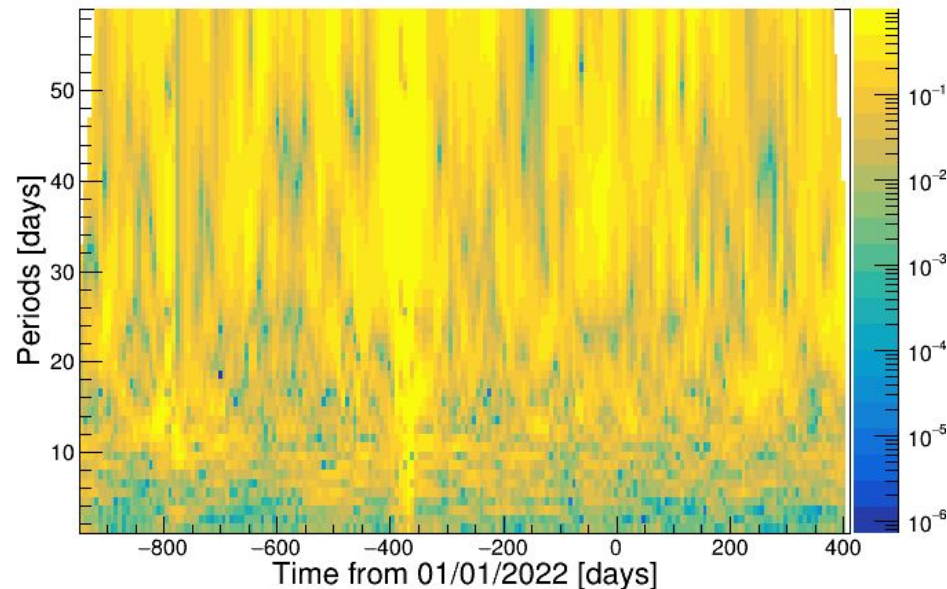
Periodogramma Lomb-Scargle “TIME-DEPENDENT” (finestra 60d per “alte” frequenze)

Ogni “slice” verticale al tempo T è un periodogramma costruito con i dati [T-30, T+30] (quindi le slices in finestre di 60d non sono indipendenti)

POLA-4



POLA-3



SUMMARY:

- 1) Sviluppate due approcci alternativi (“Lomb-Scargle” & “Brute-force-fit”)**
- 2) Hints di presenza di sotto-componenti**
- 3) Hints di presenza di sotto-componenti con frequenza “time dependent”**

TO DO LIST:

W1) produrre le time series delle rates "finalizzate" e quantificare errori statistici e sistematici

W2) produrre delle time series delle pressioni/temperature misurate sul detector ma anche locali dell'atmosfera sopra il detector, se esistono.

W3) comprendere come (e se) usare le pressioni e temperature per correggere le rates ovvero definire una strategia di correzione ed un errore sistematico associato a questa (rischio di sovra-correggere, ovvero correlazione potrebbe non indicare per forza causalità)

W4) studiare il contenuto in frequenza delle rates "corrette"

W5) valutare la possibile presenza di segnali non esattamente periodici con drift in frequenza