# Outline: continuous-wave detection validation steps (T060062-00-Z)

**I**. CW pipelines produce candidates that pass a number of tests.

**A**. The SNR is above a threshold set by a false alarm rate.

**B**. The candidate is not vetoed by coincidence test(s).

**1**. SNRs match in all IFOs within expected error.

**2**. Frequencies match in all IFOs within expected error.

**3**. Sky positions match in all IFOs within expected error.

**4**. Spindowns match in all IFOs within expected error.

**C**. The candidate is not vetoed by "goodness-of-fit" test(s).

**1**. These test might be applied before or after the coincidence test(s).

**2**. These tests have to undergo Monte Carlo simulations to set their false dismisal rates.

**3**. A chi-squared test in frequency-domain code has been implemented.

**4**. Other possible tests:

**a**. Test line width (instrument lines will be broadened by doppler demodulation).

**b**. Test that SNR grows as $\sqrt{T}$ on average.

**c**. Test SNR vs. sky position.

**d**. Time-domain code could test chi-squared value for parameters that minimize the posterior pdf.

**D**. The signal is not vetoed as a known instrument line.

**E**. Many candidates will survive this step.

**1**. Large false alarm rates will be used.

**2**. Very small (ideally zero) false dismisal rates will be used.

**3**. These rates are probably chosen to produce the approximate number of candidates that Step II can handle.

**4**. The exact value of these rates will be found by Monte Carlo simulations using software and hardware injections.

**II**. Follow-up studies are done on candidates that survive Step I.

**A**. A coherent search on a fine-grid parameter space surrounding the candidate's parameters is done on the same data.

**B**. Fine tune "goodness-of-fit" test(s).

**1**. Check that minimum $\chi^2$ or maximum likelihood value is consistent with a signal (i.e, that a CW model for the signal is not rejected based in this value).

**2**. Fine tune SNR vs. time tests.

**a**. Check that SNR grows as $\sqrt{T}$ on average (if not done in step I).

**b**. Check that SNR varies consistently with the diurnal antenna pattern.

**c**. Estimate parameters and perform chi-squared test of SNR vs. time using JKS equations for SNR.

**3**. Fine tune other "goodness-of-fit" test(s)?

**C**. Check that a joint coherent analysis using all IFOs is consistent.

**D**. Reproduce the results using data from a prior or subsequent run.

**E**. If any inconsistencies occur, check if a possible pulsar type "glitch" can account for it (i.e, does the data indicate the frequency changed discontinuously at some point, and can a better fit be found by modeling this).

**F**. Few candidates will survive this step.

**1**. This step should reduce the false alarm rate to a very small value.

**2**. The false dismisal rates should be kept as small as possible.

**3**. The exact value of these rates will be found by Monte Carlo simulations using software and hardware injections.

**III**. Candidates that survive Steps I and II should have very small false alarm rates and will be consistent with a real signal. Thus, it is time to find confidence intervals for $h_0$ or $A_1$, $A_2$, $A_3$, and $A_4$.

**A**. Predetermined unbiased approach(s) must be used to determine the confidence intervals.

**B**. Intervals for several levels of confidence could be found (e.g., 90%, 95%, 99.9%).

**C**. The method(s) should give $\sigma$'s for the estimated parameters.

**D**. Frequentist approach:

**1**. Parameters are estimated from minimizing chi-squared or maximizing the likelihood.

**2**. A fake signal with the parameter estimates is injected into the noise many times (at different frequencies). The parameters are re-estimated each time.

**3**. The $\sigma$'s of the parameters are found.

**4**. A boundary is drawn that contains $x$ percent of the estimates. The boundary would be determined by one of the following criteria:

**a**. A boundary of constant $\Delta\chi^2$ or constant likelihood ratio is used. (For example see Numerical Recipes and Feldman and Cousins)

b. A boundary that gives the central confidence interval is used. (In 1D this gives equal probability of finding a measurement below or above the acceptance interval).

c. A boundary based on the $\sigma$'s is used.

E. Bayesian approach: The method would be similar to the Frequentist approach, except the $\sigma$'s and confidence interval would be drawn from the posterior pdf.

F. How to handle the nuisance parameters.

1. Don't. Give the confidence ellipsoid for $A_1$, $A_2$, $A_3$, and $A_4$; display the result by projecting the ellipsoid onto each axis of this 4D parameter space. (For example see Numerical Recipes.)
2. Marginalize.
3. Use worst-case nuisance parameters.

G. Candidates survive this Step based on whether zero amplitude is not in the confidence interval(s) and/or on how many $\sigma$'s an estimated amplitude is from zero.

**IV**. Candidates that survive Steps I, II, and III will be "gold-plated" potential detections. Thus, it is time to rule out all other possibilities that could produce such a signal.

A. Review the validation of the software again.

1. Have any new bugs turned up?
2. Are any new validation tests or additional Monte Carlo simulations indicated?

B. Independent code should verify the result (this may already been done as a part of Step II).

1. If the frequency-domain code found the candidate use the time-domain code to verify this and vice versa.
2. Incoherent methods not already applied to this candidate might be run as further validation.

C. Check key results using independent SFTs.

D. Check the raw frames if the candidate is found in RDS data, and vice versa.

E. Check elogs for problems with excitations, DAQ corruption of data, etc....

F. Understand periodicities that can occur in the DAQ system that may have not already been vetoed.

**G**. Check excitation channels (make sure no accidental injection was done).

**H**. Check PEM and other channels for environmental causes.

**I**. Check frequencies of computer monitors and other electronics that might not already have been vetoed.

**J**. Check if up/down conversion can happen in the electronics and get into GW channel?

**K**. Check for other harmonics. Is there a signal at f/4, f/2, 2f, 4f or at ratios of the harmonics of the $r$-modes? Thus, can we determine if the signal is due to spin, precession, or a mode? (This may be very hard to do.)

**L**. Check if the parameters make astrophysical sense. (If not then this could be something really new, but do we require greater confidence in that case?)

**M**. Is there a known astronomical object associated with the candidate (e.g., pulsar, x-ray source, etc...). (If not this is not a problem; if so can we think of further consistency checks with astrophysical EM data for the source?)

**N**. If a signifant problem is found, we may need to adjust the pipeline in Steps I, II, and III and repeat Monte Carlo simulations.

1. Do we know how to do this without introducing bias?
2. How much of this to do we have to decide upon a priori?

**V**. If a candidate survives Steps I, II, III, and IV, should we seek corroboration?

**A**. Ask for astronomical data to seek EM counterpart?

**B**. Ask for data from other GW detectors?