[RADIOCARBON, VOL 26, NO. 2, 1984, P 159-165]

Radiocarbon

1984

TIME SERIES ANALYSIS OF LOW LEVEL GAS COUNTING DATA

LAURI KAIHOLA*, HENRY POLACH**, and HANNU KOJOLA†

ABSTRACT. We demonstrate the feasibility of pulse time of arrival information for early detection of periodic events in low level counting. Time of arrival data allows us to apply time series analysis and serial correlation tests which, in graphic form, give the user an illustrative view of the parameters affecting the validity of counting statistics. The decision to discontinue counting can already be made on the basis of less than 100 counts from the time information alone if more than 10 of these are non-Poisson periodic counts. The analyses also serve as a means of quality control for low level counting, being directly based upon the interval distribution of the Poisson process.

INTRODUCTION

We are developing an ultra low level gas counter that is capable of counting simultaneously 14 different CH_4 samples contained in 10ml proportional detectors at 1 to 10 atm pressures (Polach *et al*, 1982; Kaihola *et al*, 1983). Like Currie *et al* (1983) we have found it useful to provide each pulse with time of arrival (TA) information in addition to pulse height (PH) and rise-time (RT) data. Time of arrival analyses are powerful in detecting very low count rate, periodic noise. Thus, we have included pulse time series programs in the counting software in order to enhance detection of spurious events that do not fit Poisson statistics. TA, PH, and RT data are stored on disk for the analyses to be carried out after the runs. In low-level counting the required storage capacity remains reasonable even during prolonged (several days') counting.

To suppress HV capacitor induced spurious events by electronic means we have encapsulated these components in resin enabling up to 7kV to be applied without noise. Thus rigidly held, the HV components settle very fast, producing only 10 to 20 spurious pulses when first turned on in contrast to thousands of counts when they remain exposed to the atmosphere. To suppress radio frequency and line noise, the pulses detected by an aerial are amplified and channeled into the anticoincidence (pulse inhibit) module.

TIME INTERVAL PROBABILITY DISTRIBUTION

Radioactive decay is a Poisson process when the source half-life is much longer than the observation time, *ie*, the decay probability is time-invariant. When the recording system dead time is small compared with the mean time interval between decays then also the counting process can

* Department of Physical Sciences, University of Turku, SF-20500 Turku 50, Finland

** Radiocarbon Dating Research, Australian National University, POB 4, Canberra 2600, Australia

† Wallac Oy, Research Department, POB 10, SF-20101 Turku 10, Finland

be considered to be a Poisson one. Both the conditions are very well met in ^{14}C and ^{3}H counting. The mean modern count rate is only 0.5cpm in our sample detector, for example.

Let λ be the mean rate of occurrence for the detector pulses. Then the probability of the number of events N_t observed in time interval of length t has a Poisson distribution of mean λt (fig 1),

$$\text{prob}\,\{N_t = n\} = \frac{(\lambda t)^n e^{-\lambda t}}{n!},\ (n = 0, 1, \ldots). \tag{1}$$

The probability for the time interval X from time origin to the next pulse to be $X > t$ is equivalent to the fact that no pulse is observed in time interval (0,t) and

$$\text{prob}\,\{X > t\} = \text{prob}\,\{N_t = 0\} = e^{-\lambda t}. \tag{2}$$

Further, prob {there will be a pulse in the time interval (t, t+dt)} = λdt. Therefore, the joint probability for no event in the interval (0,t) and one event in (t, t+dt) is the pulse time interval probability density function

$$f_X(t) = \lambda e^{-\lambda t},\ (t \geqslant 0). \tag{3}$$

The origin of the counting process can be freely selected (Cox & Lewis, 1966).

The pulse time interval distribution function $F_X(t)$ is prob {at least one event between (0,t)} or

$$F_X(t) = 1 - e^{-\lambda t} \tag{4}$$

(see fig 2). Of the number N_0 of pulse intervals there are $N = N_0(1-e^{-1}) = 0.63N_0$ intervals smaller than the mean interval length $1/\lambda$ and half of the intervals are shorter than 0.69λ. The density of decay intervals increases towards the shorter lengths according to the negative exponential gap distribution Eq (3).

The so-called survivor function $R(t) = 1-F_X(t)$ contains the same information as Eqs (3) and (4). On a logarithmic scale deviation of R(t) from linearity indicates non-Poisson component in the counting process, seen as a mismatch of the predicted and actual data.

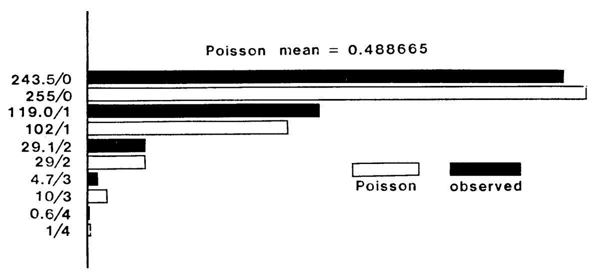

Fig 1. Distribution of counts in 200 s counting intervals. Total number of counts was 194 in 22.1 hr (397 intervals). The theoretical Poisson probability histogram is calculated using the observed mean = 0.49 c per interval. Bar length indicates number of intervals containing given number of counts.

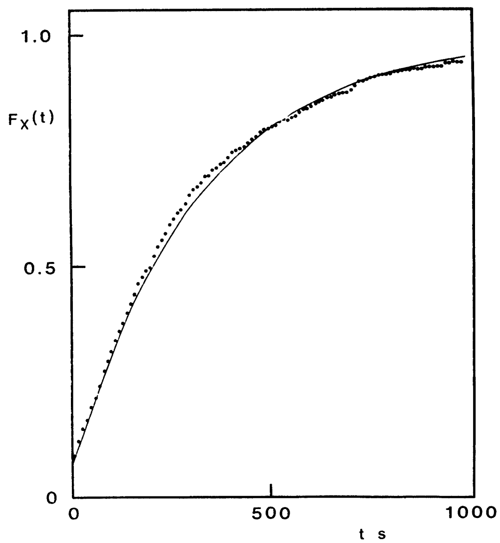

Fig 2. Pulse time interval distribution function (normal case, 373 c/33.8 hr; see Eq (4)).

PULSE TIME INTERVAL SPECTRUM

Non-random events can be tested by calculating the time intervals Δt_i from the pulse i to a later pulse i+j, $\Delta t_i = t_{i+j} - t_i < \tau$, $j = 1,2, \ldots$ with a reasonable upper limit for the time interval span. A histogram is constructed to visualize the frequency of time intervals between pulses in constant segment categories over τ (fig 3). Periodic events will be disclosed as non-randomly distributed peaks in the spectrum (*cf* fig 3, normal case, with fig 6, non-random).

SCATTER DIAGRAM

As a test of serial correlation between successive time intervals between pulses a scatter dot diagram is useful (Cox & Lewis, 1966) *ie,* we plot time interval $\Delta t_{i+1} = t_{i+1} - t_i$ as a function of $\Delta t_i = t_i - t_{i-1}$. A higher number of short intervals leads in a normal case to concentration of dots along to the axes, their equidensity curves being hyperbolic (fig 4).

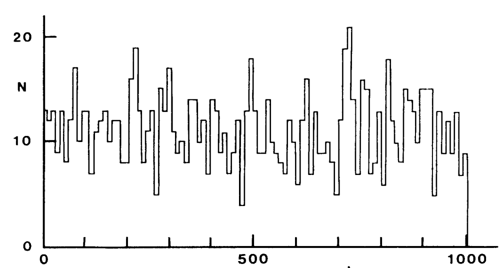

Fig 3. Pulse time interval spectrum for Poisson-distributed data of figure 2. Number of intervals of length from t to t+10 s are indicated by bars.

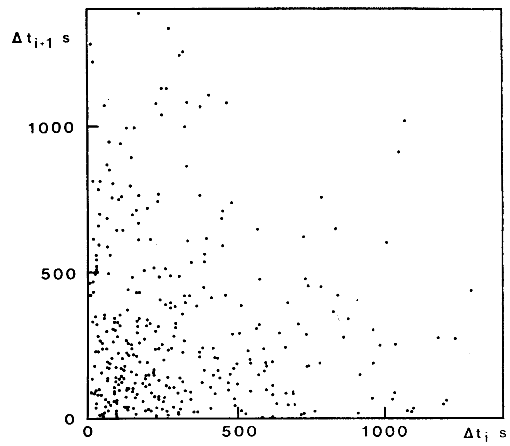

Fig 4. Scatter diagram for the data of figures 2 and 3.

CUMULATIVE PLOT OF COUNTING DATA

The accumulated number of pulses as a function of counting time is a rising line at a slope λ with some scatter around it (Cox & Lewis, 1966). It is convenient to turn the plot horizontal, *ie*, we plot

$$N = N_t - \dot{N}t. \tag{5}$$

This plot is produced after the data has been collected on the basis of *a posteriori* Poisson mean count rate $\dot{N}$. Deviations from horizontal indicate fluctuations in the count rate, *ie*, due to cosmic flux or anticoincidence shield efficiency variations (fig 5).

RESULTS

A background run of 167 counts in 7.6 hr was divided into 200-second periods. We observed 62 periods containing one pulse while the theoretical Poisson distribution based on the observed mean count rate predicted only 48.9 such periods. χ^2 test and index of dispersion showed excessive variance (table 1). The time interval spectrum showed a 400 s period with ± 100 s scatter in its length (fig 6). The same fact is dramatically illus-

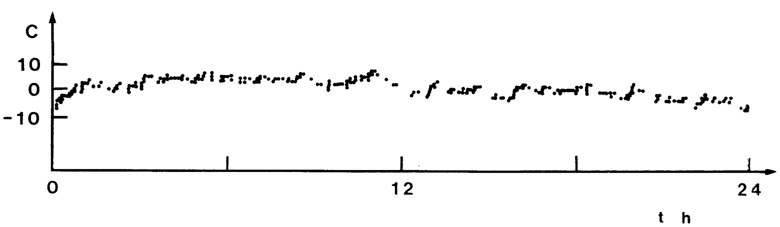

Fig 5. Horizontal cumulative plot for the first 24 hr of the data of figures 2, 3, and 4.

TABLE 1
A background run containing non-random counts

Total number of counts	= 167
Counting time	= 7.57 hr
Mean count rate ± σ	= 0.37 ± 0.10 cpm
Segmented in 200 s intervals:	
Number of intervals	= 136
Mean number of counts/interval	= 1.23 ± 1.36 c
Expected Poisson error/interval	= 1.11 c
χ^2	= 203.5
Index of dispersion = χ^2/degrees of freedom	= 1.51

trated in the time interval distribution function (fig 7). Because this is a background run, as a gross deviation from the predicted curve we can estimate the fraction of the extra counts to be some 30% of the total number.

Another run of 606 counts in 20.5 hr contained 60% of extra pulses. The horizontal cumulative plot is not time-dependent. The scatter diagram showed a clear correlation between successive time intervals (fig 8). The time interval spectrum and distribution function both show a periodicity of 270 seconds.

Time resolution of the above run was rounded from 10 ms to 1 s and 10 s without loss of the interval spectrum structure. The data file was also examined in sections of 100 seconds. No great change was observed beyond the first 100 s while the first section of data showed less prominent periodicity.

To test the detection level of the time series analysis in low-level counting, cyclic spurious counts were added into pure random data files. Fifty counts were added into 373 counts collected in 33.8 hr with 20 s and 40 s scatter around the mean 400 s interval length (see figs 9a, 9b). If <10% of spurious counts or >10% scatter in interval length are introduced, it is difficult to resolve any periodic structure in the interval spectrum.

The addition of artificial spurious periodic counts is a "rigorous" test of the method because we assume that the pulse height spectrum of the added counts does not differ from that of the true counts.

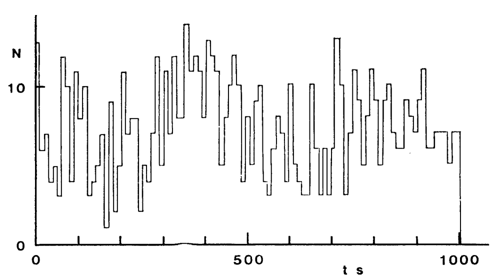

Fig 6. Periodic pulses in the data of table 1 indicating 400 s period and its harmonic at ca 800 s. Time window is 10 s.

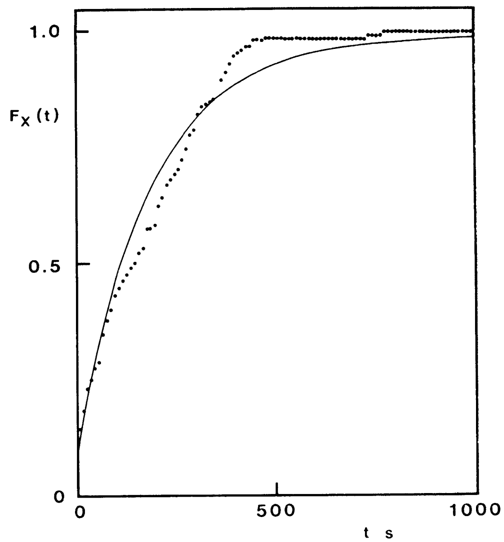

Fig 7. Time interval distribution function for the data of figure 6 indicating too many intervals of length 400 s. Theoretical curve is drawn on the basis of the experimental mean count rate.

CONCLUSIONS

The use of pulse time arrival information in statistical analyses of low-level counting data offers a possibility of identifying the existence of

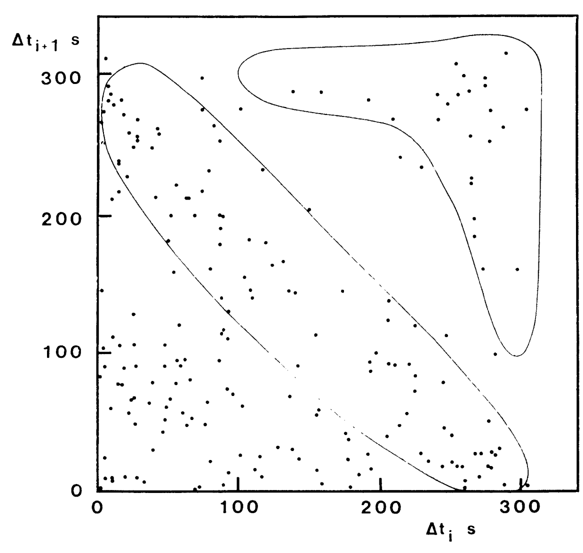

Fig 8. Scatter diagram for 200 intervals out of a run of 606 counts in 20.5 hr. Extra counts appear in encircled areas around the lines $\Delta t_{i+1} = 270$ s and $\Delta t_i = 270$ s. Grouping of dots around the descending line $\Delta t_{i+1} + \Delta t_i = 270$ s indicates cases where the 270 s periods are intercepted by single real events.

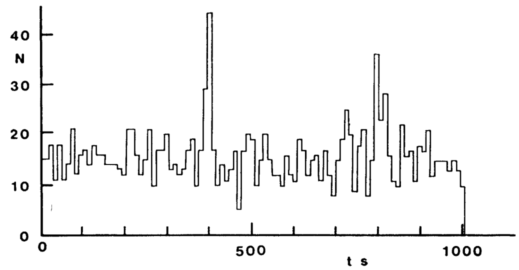

Fig 9a. The data of figure 2 made periodic by adding 50 pulses into the first 5.6 hr of the run. Equally distributed random scatter of ± 20 s was introduced into the mean period 400 s. Time window is 10 s.

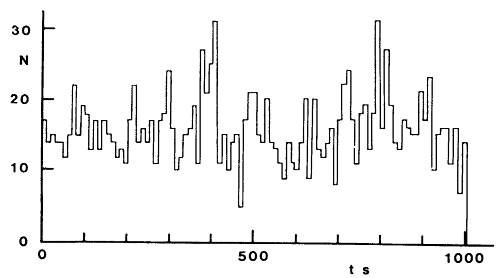

Fig 9b. Same as figure 9a except the scatter in the length of the period is ± 40 s.

periodic events at an early stage of counting. If the level of spurious counts reaches 10 out of 100 total counts, the spurious events can be detected. We have not used or planned to use time series analyses for tracing outliers, but only for establishing the possible existence of periodicity. The method is of particular advantage in examining very low frequency periodic phenomena, as higher frequency non-random events are detected by other means.

Time series analyses displayed in suitable graphic form permits the user to visually assess the validity of low count-rate data. Time-dependent effects, radon decay, anticoincidence shield efficiency fluctuations, and natural fluctuations in environmental radiation show up in cumulative plots that can be displayed while counting is in progress.

References

Cox, D R and Lewis, P A W, 1966, The statistical analysis of series of events: New York, John Wiley & Sons, p 1.

Currie, L A, Gerlach, R W, Klouda, G A, Ruegg, F C, and Tompkins, G B, 1983, Miniature signals and miniature counters: Accuracy assurance via microprocessors and multiparameter control techniques, *in* Stuiver, M and Kra, R S, eds, Internatl radiocarbon conf, 11th, Proc: Radiocarbon, v 25, no. 2, p 553-564.

Kaihola, L, Kojola, H, Polach, H, Soini, E, and Otlet, R, 1983, An evaluation of shielding characteristics applicable to mini-gas proportional counter-based ^{14}C dating systems, *in* Stuiver, M and Kra, R S, eds, Internatl radiocarbon conf, 11th, Proc: Radiocarbon, v 25, no. 2, p 577-580.

Polach, H, Soini, E, Kojola, H, Robertson, S, and Kaihola, L, 1982, Radiocarbon dating of milligram-size samples using gas proportional counters: an evaluation of precision and of design parameters, *in* Ambrose, W and Duerden, P, eds, Archaeometry: An Australasian perspective: Canberra, ANU Press, p 343-350.