The Research Council of Norway (NFR)

NORSAR

NORSAR Scientific Report No. 2-95/96

Semiannual Technical Summary

1 October 1995 - 31 March 1996

Kjeller, May 1996

7.3 Quality assessment of automatic onset times estimated by an autoregressive method

Introduction

In the previous semiannual report (Kværna, 1995), we described an experiment where we used an autoregressive method, denoted AR-AIC, for automatic estimation of phase onset times. In this report we will expand on the use of accompanying onset quality estimates as a tool to choose between onset times derived from different types of AR-AIC models, as well as for flagging onsets that have a high probability of being incorrect.

The human observation of a seismic phase is attributed to an amplitude increase and/or a change in the frequency content of the data. If the trace is properly filtered, an amplitude increase should be observable. For quality assessment of the automatically estimated onsets, we decided to derive additional signal parameters from the time domain data, filtered in the band that provides the highest SNR.To analyze the amplitude increase we found it convenient to create an envelope of the data from the filtered trace and its Hilbert transformed counterpart. The Hilbert envelope was gently smoothed with a lowpass filter. The procedure is illustrated in Fig. 7.3.1.

We defined the following set of measurements to be made on the envelope:

- $NOISE_{max}$ was taken to be the maximum of the envelope within a 3 second interval preceding the automatically estimated onset.
- $AMP_{0.5}$, $AMP_{1.0}$, $AMP_{2.0}$, $AMP_{3.0}$ and $AMP_{5.0}$ were the maxima of the envelope within 0.5, 1.0, 2.0, 3.0 and 5.0 seconds after the onset, respectively. The corresponding (quality) signal-to-noise ratios $QSNR_{0.5,...,5.0}$ were defined to be $AMP_{0.5,...,5.0}$ / $NOISE_{max}$.
- $T_{QSNR1.5}$ was the time from the onset to the point where QSNR exceeded 1.5.

Data

A database of 83 P-phases with SNR > 100 recorded at different GSETT-3 stations was created. The arrival times of each of the phases were picked manually and stored for reference. By successively reducing the SNR by adding scaled noise samples, the performance of the AR-AIC method and the associated quality measures were evaluated using the manually picked onsets as the reference.

AR-AIC models and quality metrics

For each of the down scaled signals, the AR-AIC method was applied with two different models as described by Kværna (1995):

- The first model, denoted $AR\text{-}AIC_{F+S}$, applies autoregressive coefficients derived both in a preceding noise interval and in a window within the signal.
- The second model, denoted $AR\text{-}AIC_F$, applies autoregressive coefficients derived only from the preceding noise interval.

Generally speaking, the overall accuracy of both manually and automatically estimated onsets depends on the SNR of the signal. It was therefore obvious to us that a quality metric should take into account this factor. To ensure that the SNR was measured in the vicinity of the actual onset we decided to use the envelope measurement $QSNR_{2.0}$, being the maximum QSNR-within 2 seconds of the onset. At the same time we wanted to include a factor that specifically contained information on a possible erroneous onset estimate. From experiments we found that the envelope measurement $T_{QSNR1.5}$, being the time from the onset to the point where QSNR exceeded 1.5, would yield low values for correct onsets and high values for both early and late onsets.

The working hypothesis was to compute the composite quality metric

$$QAIC = QSNR_{2.0} / T_{QSNR1.5}$$

for the onsets estimated by two different models of AR-AIC, and then from this quality metric to decide which one was the best.

The second working hypothesis was that once the best AR-AIC onset estimate was chosen, we could compare QAIC with the standard STA/LTA based SNR to identify onsets that had a high probability of being incorrect.

Results

Fig. 7.3.2a shows the time difference between $AR\text{-}AIC_{F+S}$ onsets and the corresponding manual pick of the unscaled signals, plotted against the standard SNR in the best frequency band. We can see that for SNR less than 5, the $AR\text{-}AIC_{F+S}$ onsets become random and unstable. We do currently not know if this is due to the method itself, or is an artifact of quantization problems introduced by the noise scaling or due to other small signals present in the scaled noise samples. However, we will in the following restrict our analysis to the cases where SNR exceeds 5.0.

As seen from Fig. 7.3.2a, one problem that arose with the $AR\text{-}AIC_{F+S}$ model, was that it sometimes estimated the onset too early even for large SNRs. When comparing to the $AR\text{-}AIC_F$ results shown in Fig. 7.3.2b, we find the number of early onsets to be much less. On the other hand, we found that in general the $AR\text{-}AIC_F$ onsets had a tendency of being late and that the $AR\text{-}AIC_{F+S}$ model should initially be preferred.

For phases with $SNR \geq 10$ we have in Fig. 7.3.3a plotted the composite quality metric of the $AR\text{-}AIC_{F+S}$ onsets versus the composite quality metric of $AR\text{-}AIC_F$ onsets, denoted $QAIC_{F+S}$ and $QAIC_F$, respectively. The cases where the $AR\text{-}AIC_F$ onsets are more than 0.2 seconds closer to the reference manual pick than the $AR\text{-}AIC_{F+S}$ onsets are emphasized by circles, being representative for the cases where $AR\text{-}AIC_F$ onsets should be preferred. It can be seen from this figure that we can, on the basis of comparing the quality metrics, come up with a general rule for when to use the onsets estimated by the $AR\text{-}AIC_F$ model instead of the $AR\text{-}AIC_{F+S}$ onsets. In fact, by slightly adapting the simple working hypothesis described above (i.e., selecting the onset with the highest QAIC value), we succeeded in making the correct choice in about 75% of the cases. Similar results for $5 \leq SNR < 10$ are shown in Fig. 7.3.3b.

By applying a somewhat more sophisticated selection method, it ought to be possible to improve these initial results. However, before concluding the details of the general selection rule, we plan to extend our database somewhat so that we can split the data set into two populations, i.e., one for learning and one for testing. It should also be noticed that the approach of comparing the quality metrics can easily be extended to cases where several different models or parametrizations of the AR-AIC method are run in parallel, and we plan to test such approaches as well.

After selection of the "best" AR-AIC model has been made in each case, the next step will be to assess the actual accuracy of the selected onset time. We note that even with optimized selection criteria there will be AR-AIC onsets that can be considered as “wrong”, and it will be important to identify these cases to avoid erroneous input to the subsequent event location process. In Fig. 7.3.4 we have plotted the QAIC metric (obtained by selecting the "best" model in each case) versus the standard SNR of the signal, and we have labelled with a “B” the onsets that are considered “bad”, i.e., onsets that are more than 0.3 seconds ahead of the reference manual pick or more than 2 seconds late. It can be seen that a majority of the “bad” onsets cluster in the lower left part of this plot, thus making it possible to design a rule for automatic flagging of the less reliable onset estimates. Developing such an algorithm will be a task for future work.

Conclusions

This study has demonstrated that the quality measurements made on the optimally filtered beam or single trace can be used both for selection of the best AR-AIC model as well as a tool for identifying onsets that have a high likelihood of being wrong. The data set should, however, be expanded before concluding on any final decision rules, and it is also our intention to further investigate the relation between the envelope quality measurements and the onset picking error. So far we have only utilized two of the envelope measurements, but with a larger data set we can through the use of neural networks or statistical analysis investigate the utility of the other measurements.

T. Kværna

References

Kværna, T., 1995. Automatic onset time estimation based on autoregressive processing. Semiannual Technical Summary, 1 April - 30 September 1995, NORSAR Sci. Rep. No. 1 95/96, Kjeller Norway.

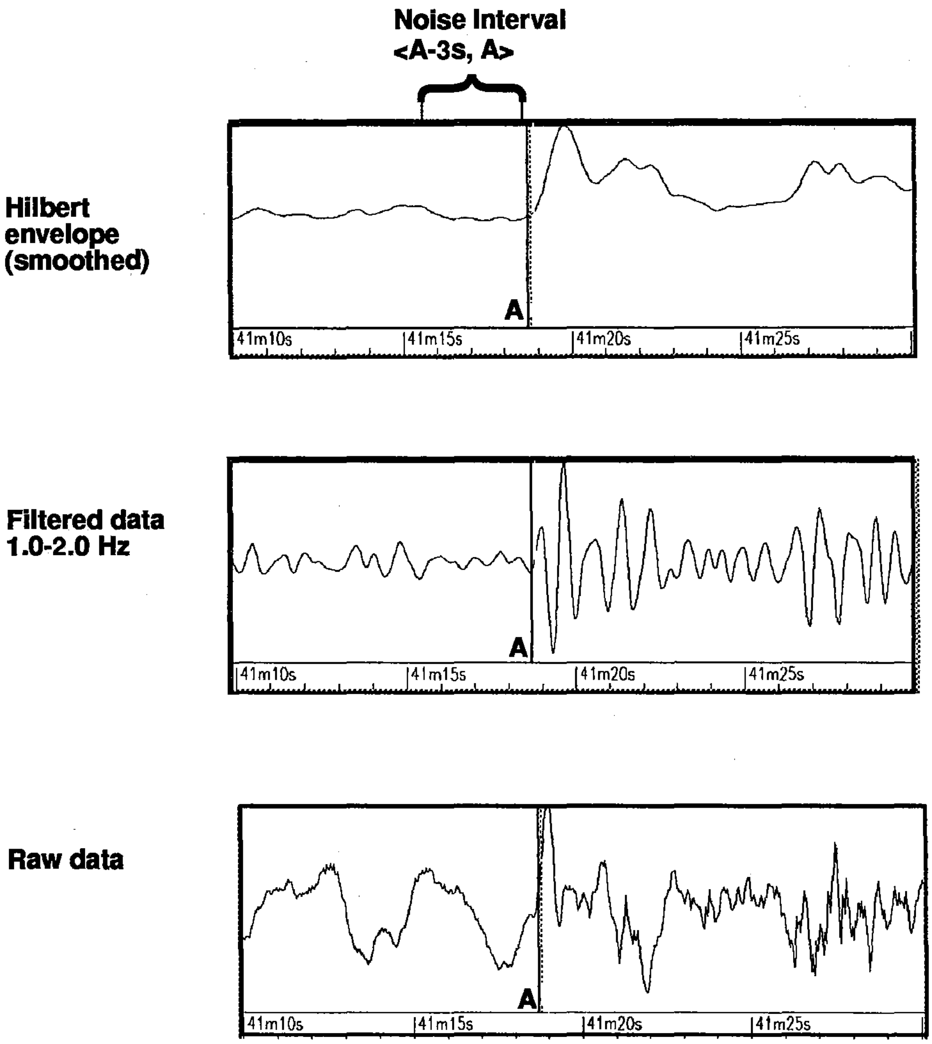

Fig. 7.3.1. Figure showing the raw data (lower panel), the data filtered in the best frequency band (middle panel) and the smoothed envelope (top panel) computed from the filtered time series and its Hilbert transformed counterpart. The 3 sec noise interval is indicated on the top panel.

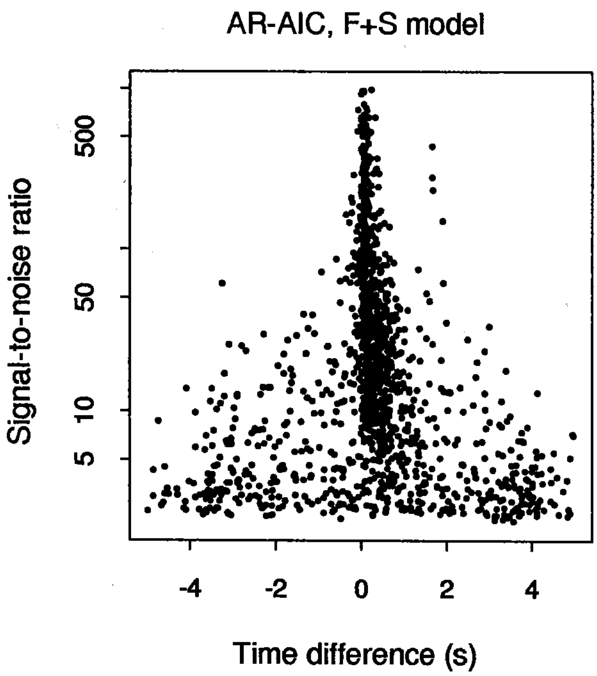

Fig. 7.3.2a. Time differences between AR-AIC$_{F+S}$ onsets and the reference manual picks plotted against the standard SNR of the signals.

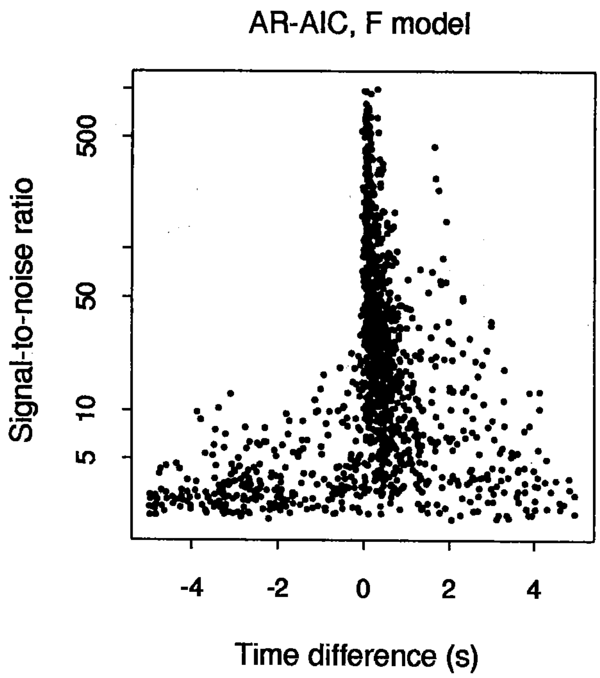

Fig. 7.3.2b. Time differences between AR-AIC$_F$ onsets and the reference manual picks plotted against the standard SNR of the signals.

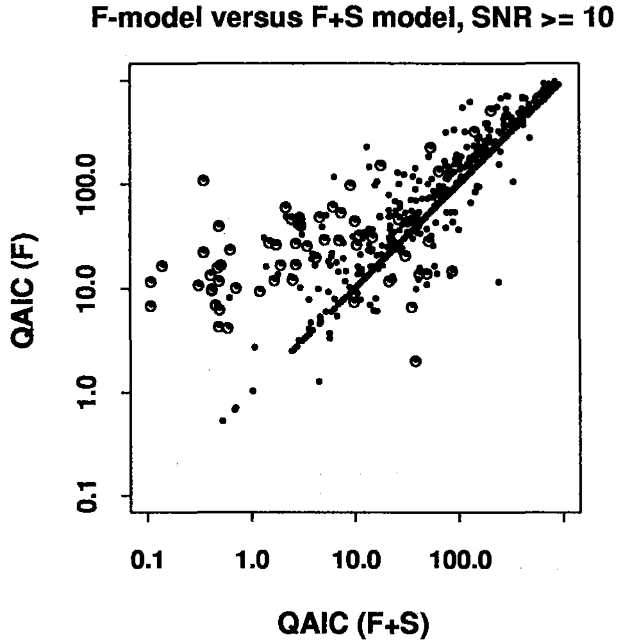

Fig. 7.3.3a. Onset quality metric for AR-AIC$_{F+S}$ plotted against the onset quality for AR-AIC$_F$ for phases with SNR >=10. The cases where the AR-AIC$_F$ onsets are more than 0.2 seconds closer to the reference manual pick than the AR-AIC$_{F+S}$ onsets are emphasized by circles.

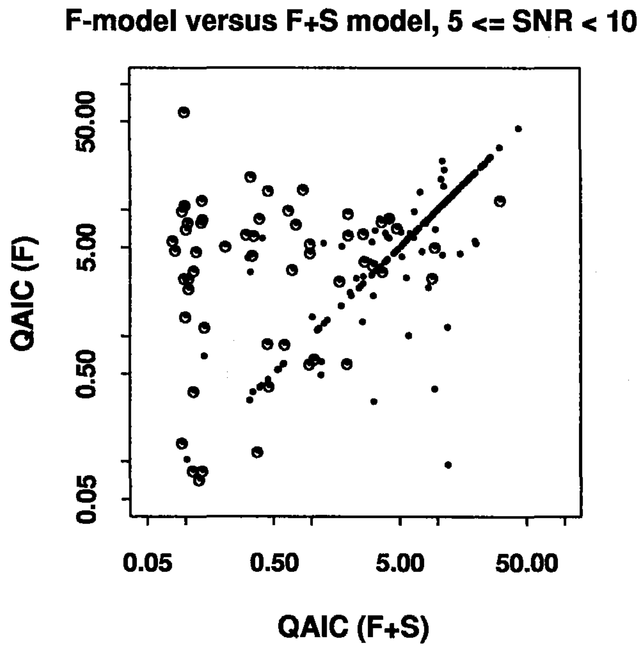

Fig. 7.3.3b. Onset quality metric for AR-AIC$_{F+S}$ plotted against the onset quality for AR-AIC$_F$ for phases with SNR between 5 and 10. The cases where the AR-AIC$_F$ onsets are more than 0.2 seconds closer to the reference manual pick than the AR-AIC$_{F+S}$ onsets are emphasized by circles.

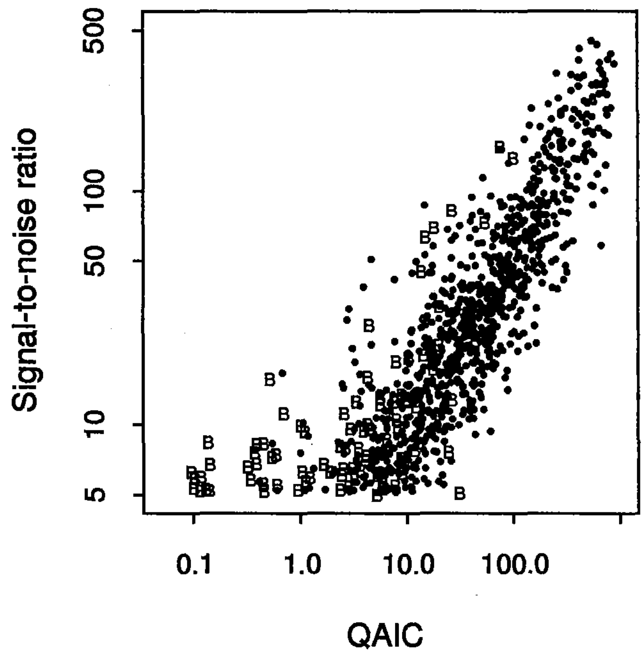

Fig. 7.3.4. QAIC metric plotted against the standard SNR of the signal. The "bad" onsets being more than 0.3 seconds ahead of the reference manual pick or more than 2 seconds late, are labelled "B".