CHAPTER 5. NEURAL NETWORK BASED REMOTE NOISE PREDICTION FILTER

5.1 General description of a remote noise prediction filter

Even though ultrasensitive sensors and micro-computers for in-field data processing have enabled collection of better data quality in an EM method, there are still limitations set by external noise. A further advance for improving data quality in low-frequency EM and TEM techniques has been achieved by the introduction of a remote-reference technique (Clarke et al., 1983; Gamble et al., 1979a, b; Goubau et al., 1984; San Filipo and Hohmann, 1983; Stephan and Strack, 1990; Wilt et al., 1983). The basic idea of the remote-reference technique in EM noise reduction is to find the earth transfer function between the EM fields at the local and remote receivers. The earth transfer function accounts for the lateral conductivity change between the receivers, the measurement errors of the differently calibrated receivers, and relative alignment errors between the local and remote receivers. During subsequent data collection, the EM noise can be subtracted in real time.

The remote-reference technique can be generally considered as a geometry where a local receiver measures EM signal and noise, while a remote receiver is simultaneously measuring solely electronics and EM noise, some of which may be correlated with the noise measured by the local receiver. In a simple model of a situation such as this, the total EM field recorded at the local receiver is given as follows:

$$l(t) = s(t) + n^{l}(t) + \varepsilon^{l}(t) \tag{5.1}$$

and the EM field measured at the remote receiver is given by:

$$r(t) = n^r(t) + \varepsilon^r(t) \,, \tag{5.2}$$

where $s(t)$ is the EM signal, $n^l(t)$ and $n^r(t)$ are the EM noise fields measured at the local and remote receivers, respectively, and $\varepsilon^l(t)$ and $\varepsilon^r(t)$ denote electronics noise at the local and remote reference sites which are not correlated to each other or with the EM signal.

Since $\varepsilon^l(t)$ and $\varepsilon^r(t)$ are not correlated, from Equations 5.1 and 5.2, the mathematical relationship between the EM noise fields at the local and remote receivers is given as:

$$n^l(t) = \int_0^\infty \alpha(\lambda) n^r(t+\tau) d\lambda \,, \tag{5.3}$$

whereby, assuming the receivers are identical, $\alpha(\lambda)$ is the impulse response of the earth transfer function and τ is the time shift between the local and remote time series. For a digitised time series, Equation 5.3 is written as:

$$n_i^l = \sum_{j=0}^{\infty} \alpha_j n_{i+\tau-j}^r, \quad i = 0, \ldots, k-1 \tag{5.4}$$

where k is the number of samples.

Theoretically, after removing the time shift between two time series, τ, calculated from the slope of the cross phase spectrum (Equation A2.14 of Appendix 2), Equation 5.4 is written as:

$$n_i^l = \sum_{j=0}^{\infty} \alpha_j n_{i-j}^r \,. \tag{5.5}$$

However, when the time shift, τ is not an integral multiple of the sample interval, a residual time shift remains. Therefore, Equation 5.5 is written as:

$$n_i^l = \sum_{j=0}^{\infty} \alpha_j n_{i+\beta-j}^r, \quad -\Delta t < \beta < +\Delta t \tag{5.6}$$

or

$$n^l(t) = A * n^r(t+\beta)\ , \tag{5.7}$$

where β is the residual time shift, Δt is the sample interval, A is the earth transfer function, and * represents a convolution operator.

For application of a neural network to a remote-reference technique, Equation 5.7 is rewritten as:

$$n^l(t) = F\left(n^r(t+\beta)\right), \tag{5.8}$$

where the function F represents a linear system. A neural network RNPF can be implemented by replacing the linear system by an artificial neural network with the hyperbolic tangent as an activation function.

5.2 Network model for a remote noise prediction filter

The prediction of a time series may be performed in a number of ways. In the time domain, an autoregressive moving average (ARMA) model and an artificial neural network may be fitted to the data. These then give empirical descriptions of the relationship between the time series that can be used to predict a local noise from a remote noise measurement.

A remote noise prediction filter (RNPF), based on an artificial neural network, is the technique that predicts a sferics pulse at the local site from the corresponding sferics pulse recorded at the remote receiver. A multi-layer feed-forward network is chosen for an RNPF so that it can find the impulse response function of the earth transfer function, α_j and accommodate any time shift, β, given by Equation 5.6. The abbreviation *(i+1)×h×i* represents the network as follows:

- The $i+1$ input PEs $\left(\beta, n_0^r, n_1^r, \ldots, n_{i-2}^r, n_{i-1}^r\right)$ are given by the residual time shift and the time series recorded at a remote receiver;

- The h hidden PEs are fully connected to the input PEs;
- The i output PEs are fully connected to the hidden PEs, producing the time series $(n_0^l, n_1^l, \ldots\ldots, n_{i-2}^l, n_{i-1}^l)$ recorded at a local receiver;
- Output and hidden PEs have adjustable biases;
- The connection weights can be positive, negative, or zero;
- The nonlinearities are located in the activation function of the hidden PEs for high-frequency sferics reduction, and both hidden and output PEs for background EM noise reduction. A hyperbolic tangent is used as an activation function.

5.3 Investigation of an optimal network configuration

In the investigation of an optimal network configuration for sferics reduction with an RNPF, the X-component time series of sferics pulses recorded in Darwin in December 1994 were used. These measurements were made with a receiver separation of 11 km (see Section 2.4.1).

A cross-validation method has been used for finding the optimal size of the network and when to stop training. As mentioned in Section 3.3, the whole available time series is divided into three sets: a training set, a validation set, and a prediction set. The *ANFE* for finding the optimal network size of an RNPF is defined as:

$$ANFE = \frac{1}{\sigma^2}\frac{1}{M}\frac{1}{N}\sum_{p=1}^{M}\sum_{i=1}^{N}\left(n_{pi}^l - \hat{n}_{pi}^l\right)^2, \tag{5.9}$$

where n_{pi}^l and $\hat{n}_{pi}^l$ are the true and predicted local time series, respectively, at the ith output PE in a given input pattern p, N is the number of the output PEs, M is the number of input patterns, and σ^2 is the variance of the true time series.

In order to investigate the dependence of performance of an RNPF on the number of input PEs, four networks, *4×20×3*, *6×20×5*, *11×20×10*, and *16×20×15*,

with a hyperbolic tangent as their activation function were trained for 1000 cycles. A training set for each network was made from 11 corresponding pairs of sferics pulses recorded at local and remote stations with a separation of 11 km in Darwin in December 1994. The number of training patterns for the four networks, *4×20×3*, *6×20×5*, *11×20×10*, and *16×20×15*, was 316, 190, 94, and 62 patterns, respectively. The initial learning and momentum rates were $2/n_{PE}$ and 0.6, respectively, where n_{PE} is the total number of the previous layer's PEs connected to the hidden or output layer. To avoid oscillations in the training of the network, these initial rates were reduced at 30, 120, 270, 480, and 750 cycles to 70% of the value used in the previous training stage.

Figure 5.1 shows *ANFE* values of each network as a function of the number of cycles. *ANFE* values were calculated using Equation 5.9 when each validation set was presented to the corresponding network every 10 training cycles. The validation set for each trained network were made from three pairs of sferics pulses which were not involved in the training set. The performance of an RNPF is best when the number of input PEs is six, (i.e., one PE for the residual time shift between two sferics pulses measured simultaneously at the local and remote stations and five PEs for the remote time series). The *ANFE* values of the *6×20×5* network are about 3% after 300 cycles (57,000 iterations with 190 training patterns). The *ANFE* values become larger as the number of input PEs increases or decreases from 6.

In order to investigate how the number of PEs at a hidden layer affects the network performance for high-frequency sferics reduction, the network configured as *6×h×5* with a hyperbolic tangent as an activation function was trained for 500 cycles (i.e., 95,000 iterations with 190 training patterns). The number of hidden PEs, *h*, varied from 5 to 30 in intervals of 5, and then to 50 in intervals of 10. The same learning schedule above was used during the training of the network. An *ANFE* value was calculated every 10 cycles during the training of the network. Figure 5.2 shows

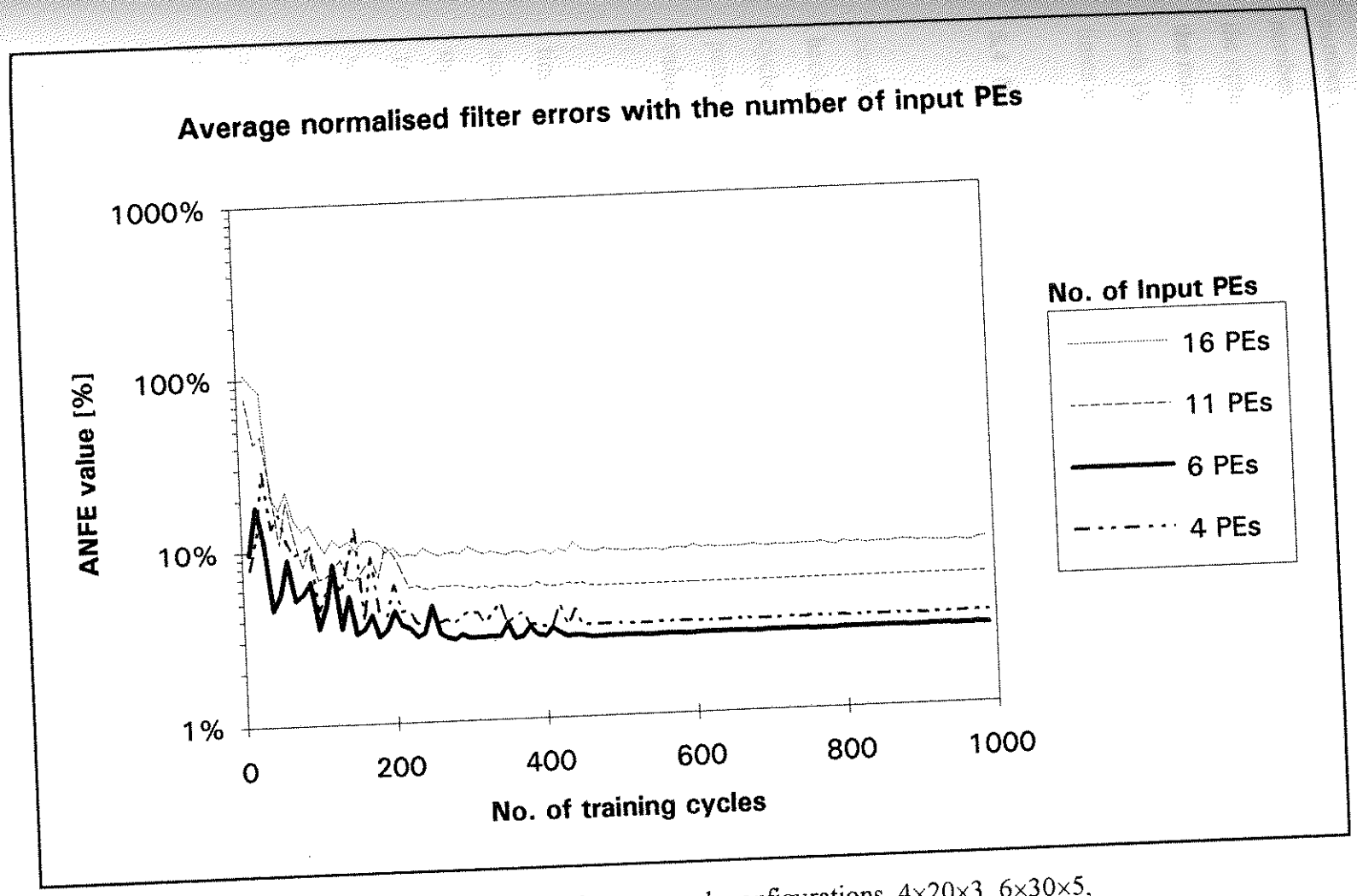

Figure 5.1 Comparison of four network configurations, 4×20×3, 6×30×5, 11×20×10, and 16×20×15, in terms of ANFE values as a function of the number of iterations when the corresponding validation sets was presented to each network.

the contour of *ANFE* values as a function of the number of hidden PEs and the number of training cycles. The *ANFE* values of the network gradually decrease as the numbers of hidden PEs and cycles increase up to 25 and 250, respectively. Therefore, the optimal network size for an RNPF for sferics reduction was chosen as $6 \times 25 \times 5$ and the number of cycles to stop training was chosen to be 250 (i.e., 47,500 iterations with 190 patterns). With this method, it takes less than 23 minutes for computing times including I/O operations on an Intel 486DX-66 MHz processor based PC.

5.4 Network performance for background EM noise reduction

For background EM noise reduction, the aforementioned network configuration used for sferics reduction was tested with the X component of background EM noise (VLF) measured in Darwin in December 1994. The number of cycles to stop training was determined as 170 (i.e., 34,000 iterations with 200 training patterns), because after 160 training cycles, the network became stable, and an *ANFE* value of about 23 % was obtained (Figure 5.3).

Figure 5.4 shows an example of an RNPF used to attenuate the X component time series of background EM noise recorded in Darwin during very low sferics activity period. This time series is a 4 ms block occurring 11 min after the time series used in the training of the network. The amplitude of background EM noise is about 10 times smaller than that of background EM noise in Ku-Ring-Gai National Park, NSW (see Figure 4.14). The dominant frequency of background EM noise is ~20 kHz (NWC). The ~20 kHz VLF noise is only a factor of two above the data acquisition system noise level (see Figure 4.13). The RNPF can remove only the VLF noise, as this is the only noise component that is correlated at the local and remote stations. The residual noise therefore still contains significant (random) noise compared with the measured noise, and a low *NRF* value of 2.2 is consequently obtained. The power spectrum of the residual noise (Figure 5.4) shows attenuation by 20 dB (i.e., by a

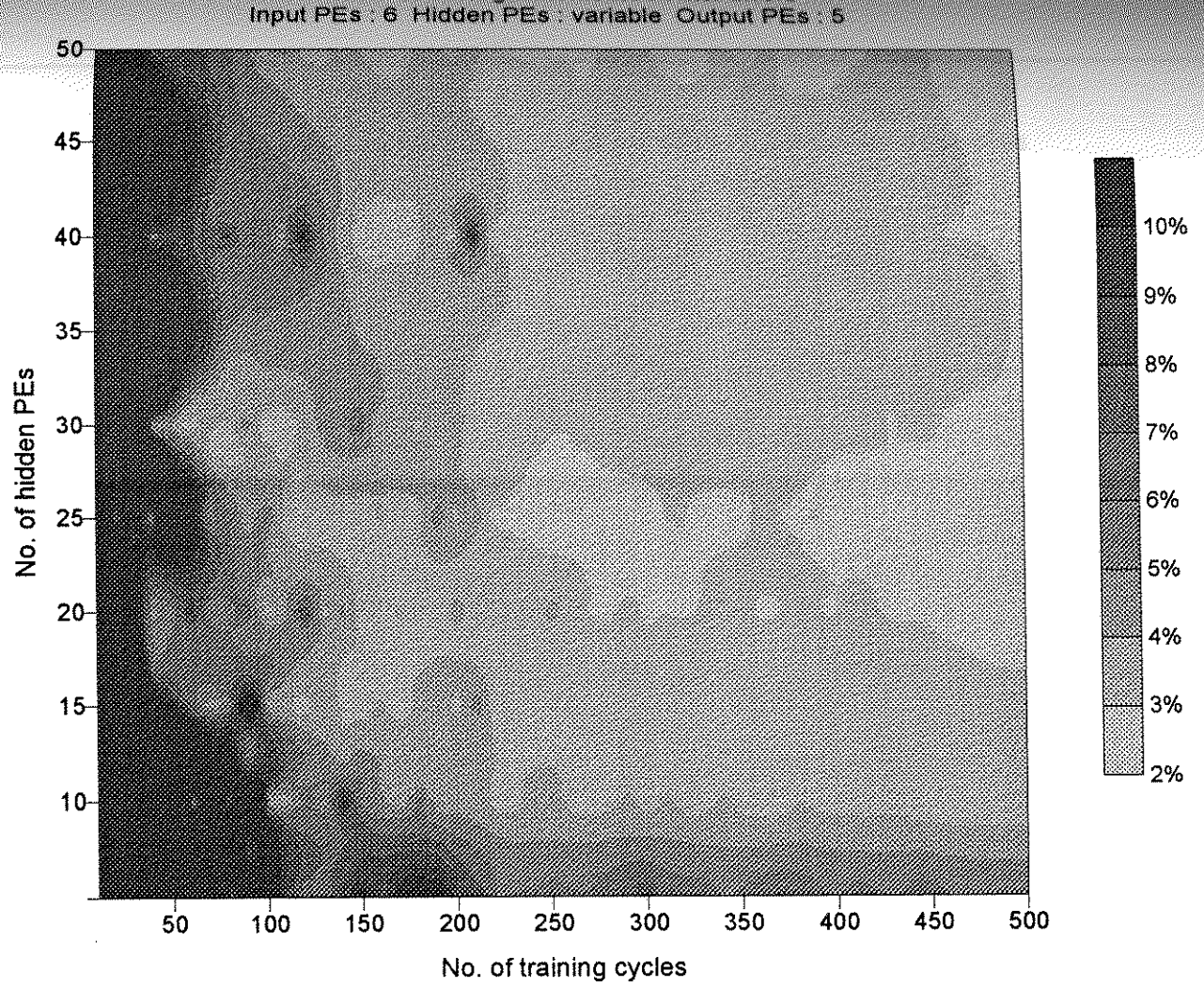

Figure 5.2 Distribution of ANFE values as a function of the hidden PEs and the number of iterations. An ANFE value was calculated every 10 cycles during training the 6×20×5 network.

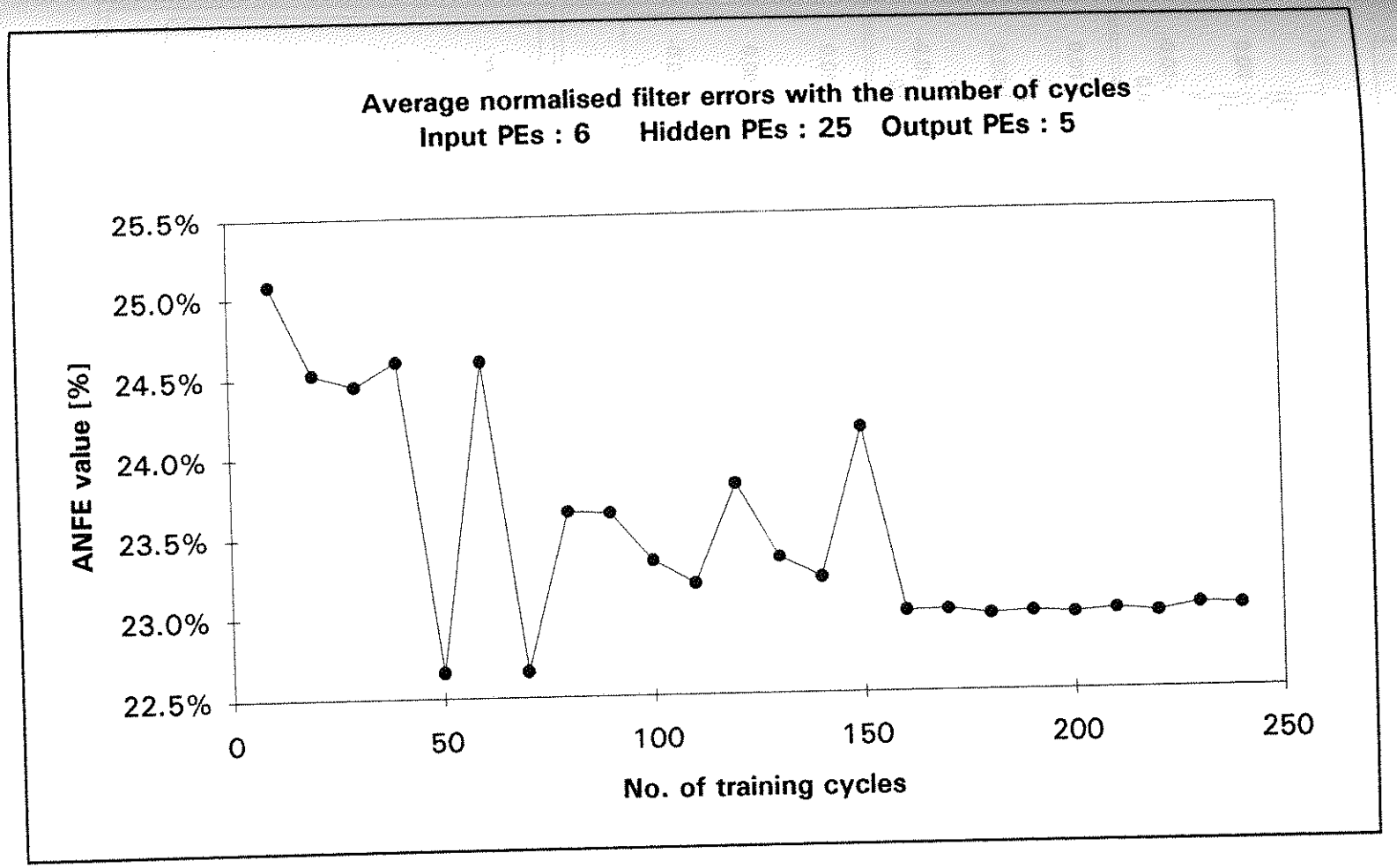

Figure 5.3 ANFE values as a function of the number of iterations when the 6×20×5 network was trained with the X component of background EM noise.

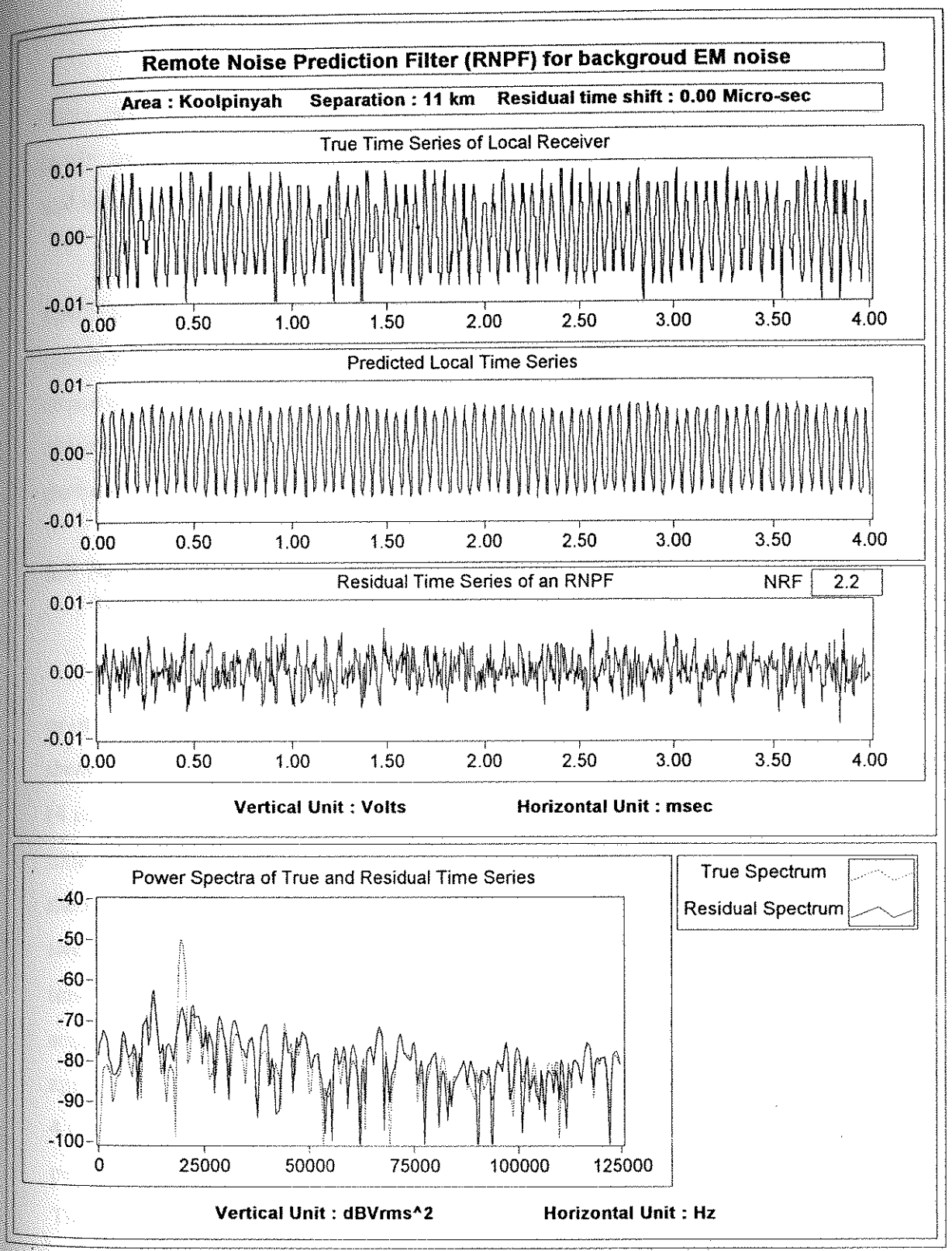

Figure 5.4 Example of performance of an RNPF used to attenuate the X component of background EM noise. The data were a 4 ms block of the time series occurring after the time series used in the training of the 6×20×5 network.

factor of 10 in amplitude) at the ~20 kHz VLF peak, and at other frequencies the power spectrum reflects the character of the random noise measured at both the local and remote receivers.

To test whether the performance of the RNPF is invariant with time, the RNPF was applied to other blocks of data. Figure 5.5 shows another example of results from use of the RNPF. The time series for the prediction set is a 4 ms block occurring 4 sec after the time series used in Figure 5.4. An *NRF* value of 2.2 is obtained by the RNPF. From these and other similar results, it is concluded that the performance of the RNPF is stationary.

5.5 Network performance for the reduction of the horizontal components of high-frequency sferics

Two time series of sferics pulses recorded simultaneously at two receivers must be corrected for any time shift between them before an RNPF can be applied to them. However, when this time shift is not an integral multiple of the sampling interval (4 μs in this case), a residual time shift remains. Figure 5.6 shows two X-component time series after correcting for a time shift of -36 μs, which is an integral multiple of the sampling interval. The residual time series produced by simple subtraction of the two series and by the application of an RNPF are also shown in this figure. A residual time shift of -1.34 μs still remains. The two time series were measured at 11 km separated stations. An *NRF* value of 8.5 was obtained with the neural network RNPF applied in this case, compared with a corresponding significantly lower *NRF* value of 4.0 obtained from simple subtraction of the two time series. This difference in performance is caused mainly by the residual time shift between the two time series. The simple subtraction is not accurate enough (without interpolation between sample points) to give substantial noise reduction. On the other hand, the neural network RNPF can be trained to accommodate any residual time shift. A comparison of the

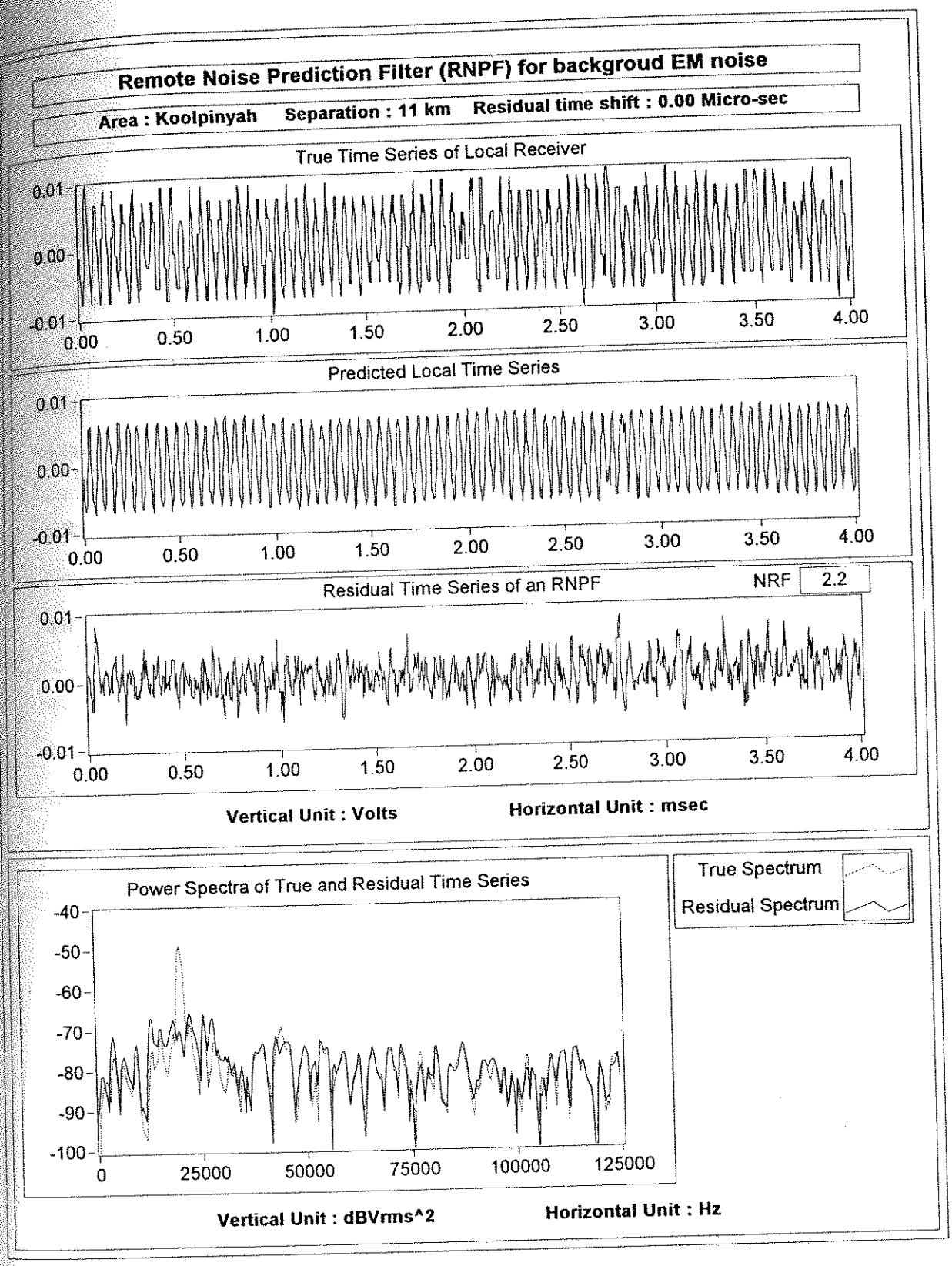

Figure 5.5 Example of stationarity test of an RNPF's performance for reduction of the X component of background EM noise. The data were a 8 ms block of the time series occurring after the time series shown in Figure 5.4.

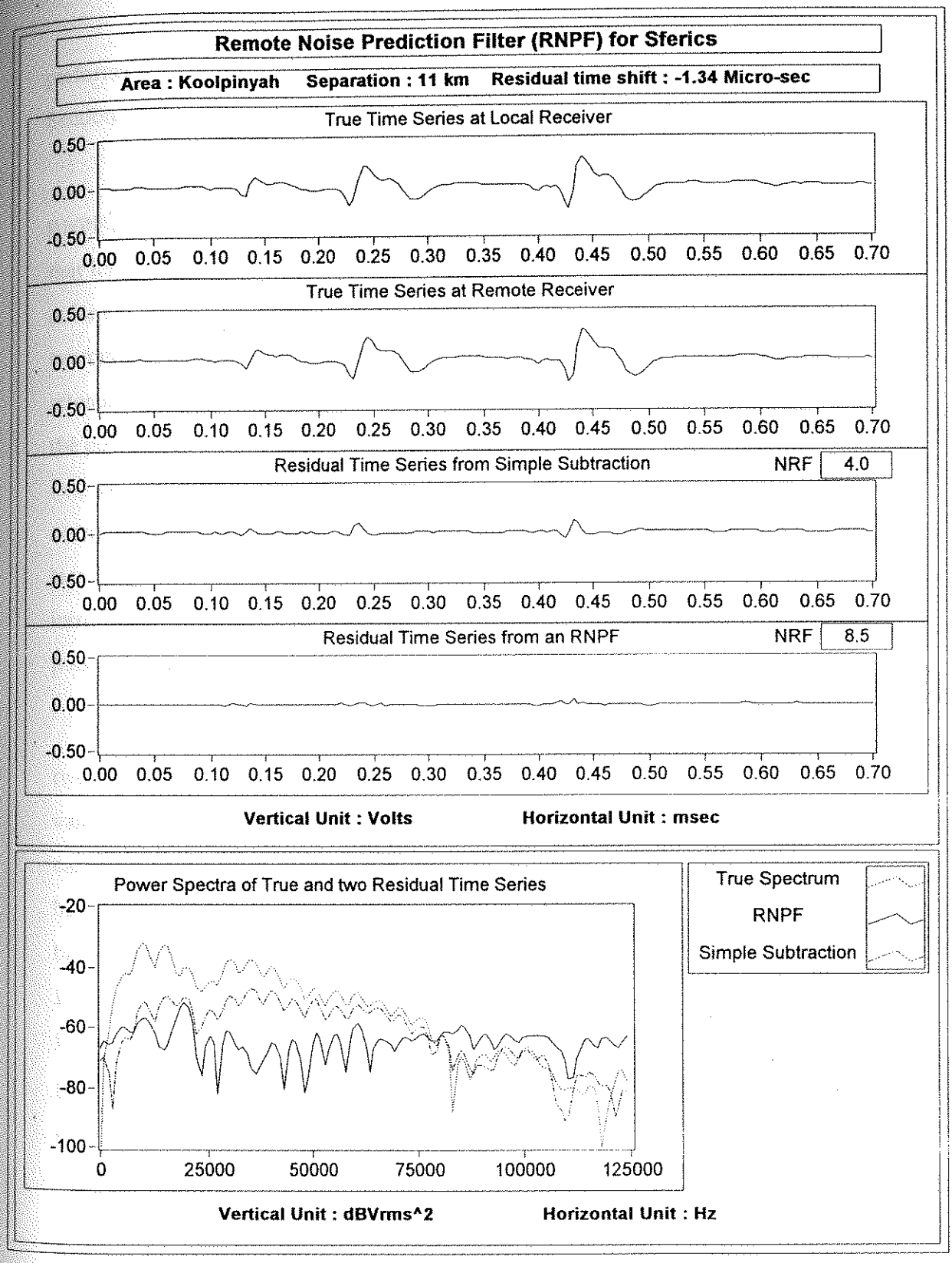

Figure 5.6 Example of performance of an RNPF used to reduce the X component of high-frequency sferics noise. Two sferics pulses were measured simultaneously at the local and remote receivers with a separation of 11 km and the residual time shift between the two sferics pulses is -1.34 μs.

performance of a neural network RNPF and simple subtraction is given in detail in Section 6.3.1 of Chapter 6.

The power spectra plotted in Figure 5.6 demonstrate that the neural network RNPF can effectively attenuate sferics noise. Examination of the power spectra shows that for amplitudes contributed by sferics in the frequency range of 5 to 50 kHz, the noise reduction is even greater than that indicated by an *NRF* value of 8.5. For example, at 10 kHz the RNPF reduces sferics noise by 25 dB, i.e., the amplitude is attenuated by a factor of 18. On the other hand, the VLF peak observed at ~20 kHz in the power spectrum is attenuated by 10 dB by both simple subtraction and the RNPF, i.e., the amplitude is attenuated by a factor of 3.2. When there is a time shift between corresponding sferics pulses, the attenuation of any accompanying VLF noise may be less than that of sferics because correction for the time shift of a given sferics pulse introduces an inappropriate time shift in the VLF signals accompanying the sferics pulse.

Results obtained for a residual time shift of +1.86 and +0.30 μs between the two X-component time series are presented in Figures 5.7 and 5.8, respectively. *NRF* values of 4.9 and 9.5 are obtained with the neural network RNPF applied in these two cases. The *NRF* value in Figure 5.7 is worse than those of Figures 5.6 and 5.8. This is because the amplitude of the sferics pulse in Figure 5.7 is smaller relative to background noise than the amplitude of the sferics shown in Figures 5.6 and 5.8. Inspection of the power spectra shown in Figures 5.7 and 5.8 shows that, at some frequencies where power is contributed by sferics pulse, the noise reduction is even greater than those indicated by *NRF* values of 4.9 and 9.5, respectively. For example, at ~9 kHz, the power is attenuated by 30 dB (i.e., a factor of 32 in amplitude) in Figure 5.7 and by 25 dB (i.e., a factor of 18 in amplitude) in Figure 5.8.

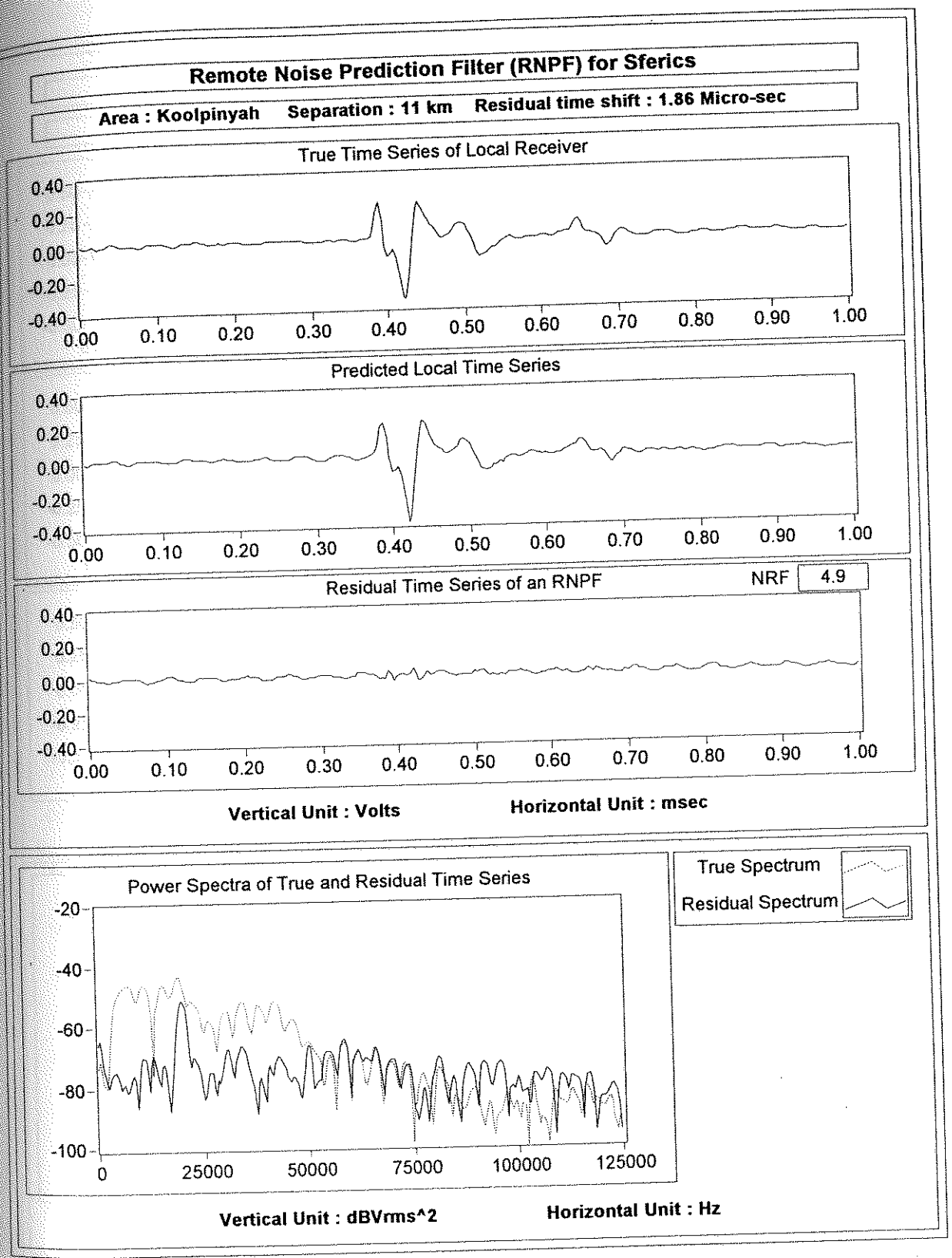

Figure 5.7 Example of performance of an RNPF used to reduce the X component of high-frequency sferics noise. Two sferics pulses were measured simultaneously at the local and remote receivers with a separation of 11 km and the residual time shift between the two sferics pulses is 1.86 μs.

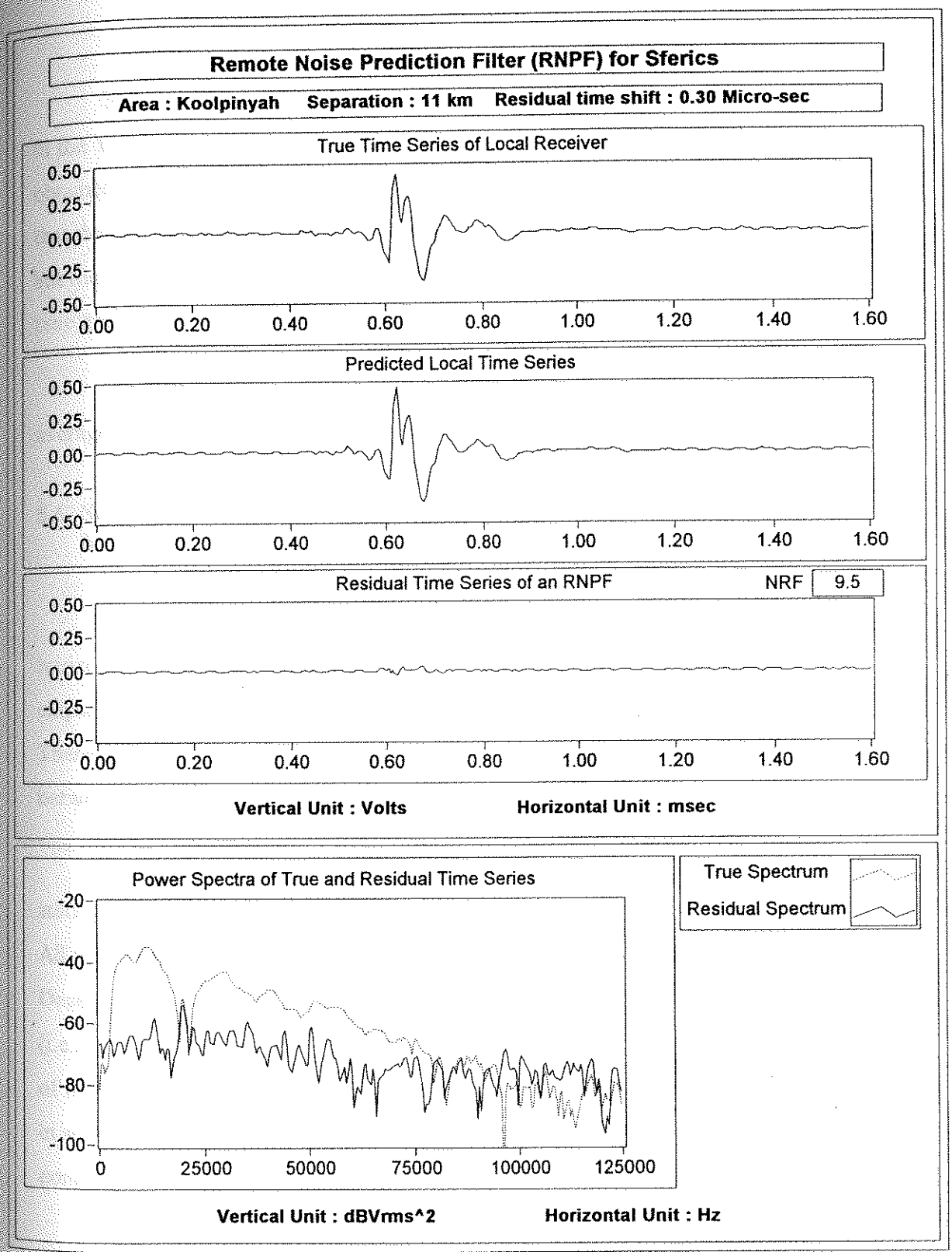

Figure 5.8 Example of performance of an RNPF used to reduce the X component of high-frequency sferics noise. Two sferics pulses were measured simultaneously at the local and remote receivers with a separation of 11 km and the residual time shift between the two sferics pulses is 0.3 μs.

Figures 5.9 and 5.10 show results when the neural network RNPF is applied to the Y-component time series of sferics noise. *NRF* values of 7.7 and 20.0 are obtained from the filter. This difference of the *NRF* values is mainly caused by the fact that the residual signals (the *NRF* denominators) are about the same in the two cases, but the peak-to-peak sferics amplitude (the *NRF* numerator) in Figure 5.10 is about 3 time greater than that in Figure 5.9.

A comparison of a neural network RNPF with an autoregressive moving average (ARMA) model is given in Section 6.3.2 of Chapter 6.

5.6 Network performance for the reduction of the vertical component of high-frequency sferics

For distant sferics and a uniformly conducting earth, the plane of the ellipse is horizontal. Local sferics and local conductivity inhomogeneities tilt the plane of the ellipse (Ward, 1967) and produce a vertical (Z) component of the sferics pulses. Therefore, this component is the one most affected by a ground conductivity change between two separated stations.

Figure 5.11 shows the effect of different ground conductivity at two separated stations at which simultaneous measurements of high-frequency sferics were made. All time series depicted in Figure 5.11 were measured at Koolpinyah near Darwin, in December 1994. For a separation of 0 km between two receivers, the amplitude ratio (i.e., 1.15) of the Z component of sferics is caused by differences in the characteristics of the two sensors. At 5 and 11 km separations, after compensating for the difference in sensor characteristics the amplitude ratios of 0.90 and 0.66 arise from the lateral conductivity change between the two stations.

The greatest effect of conductivity heterogeneity on the Z component measured at Koolpinyah is shown in the measurement with a separation of 8.9 km,

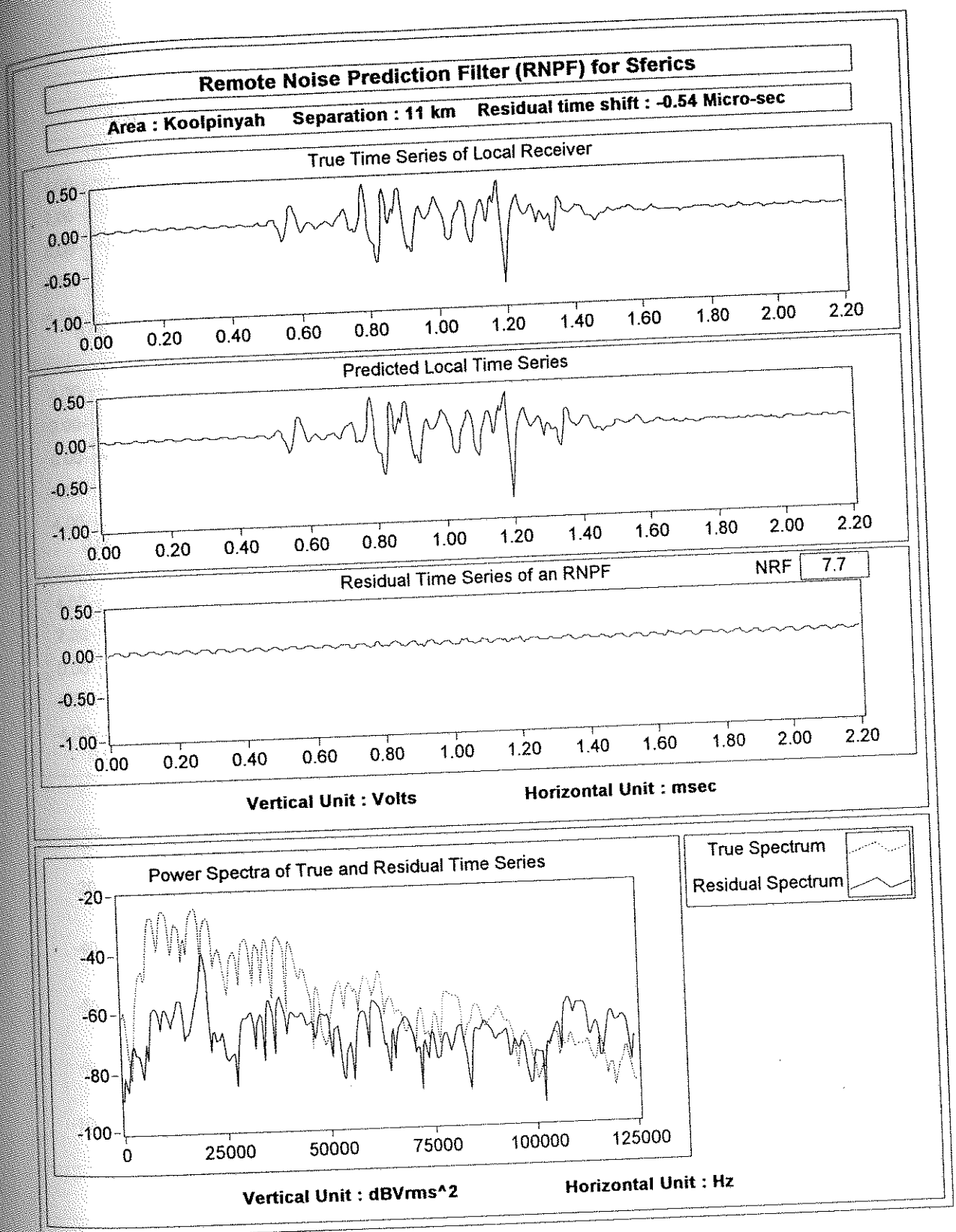

Figure 5.9 Example of performance of an RNPF used to reduce the Y component of high-frequency sferics noise. Two sferics pulses were measured simultaneously at the local and remote receivers with a separation of 11 km and the residual time shift between the two sferics pulses is -0.54 μs.

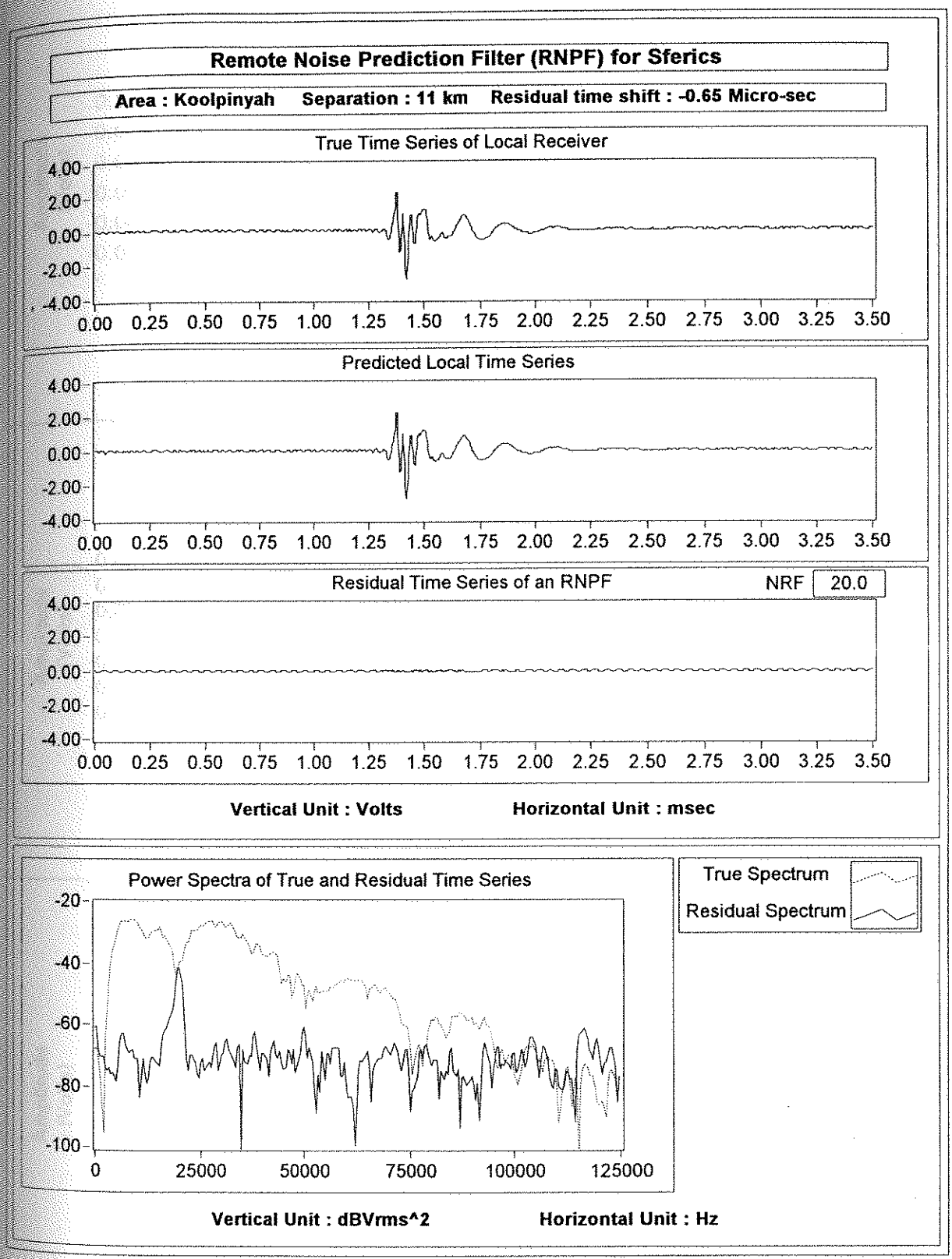

Figure 5.10 Example of performance of an RNPF used to reduce the Y component of high-frequency sferics noise. Two sferics pulses were measured simultaneously at the local and remote receivers with a separation of 11 km and the residual time shift between the two sferics pulses is -0.65 μs.

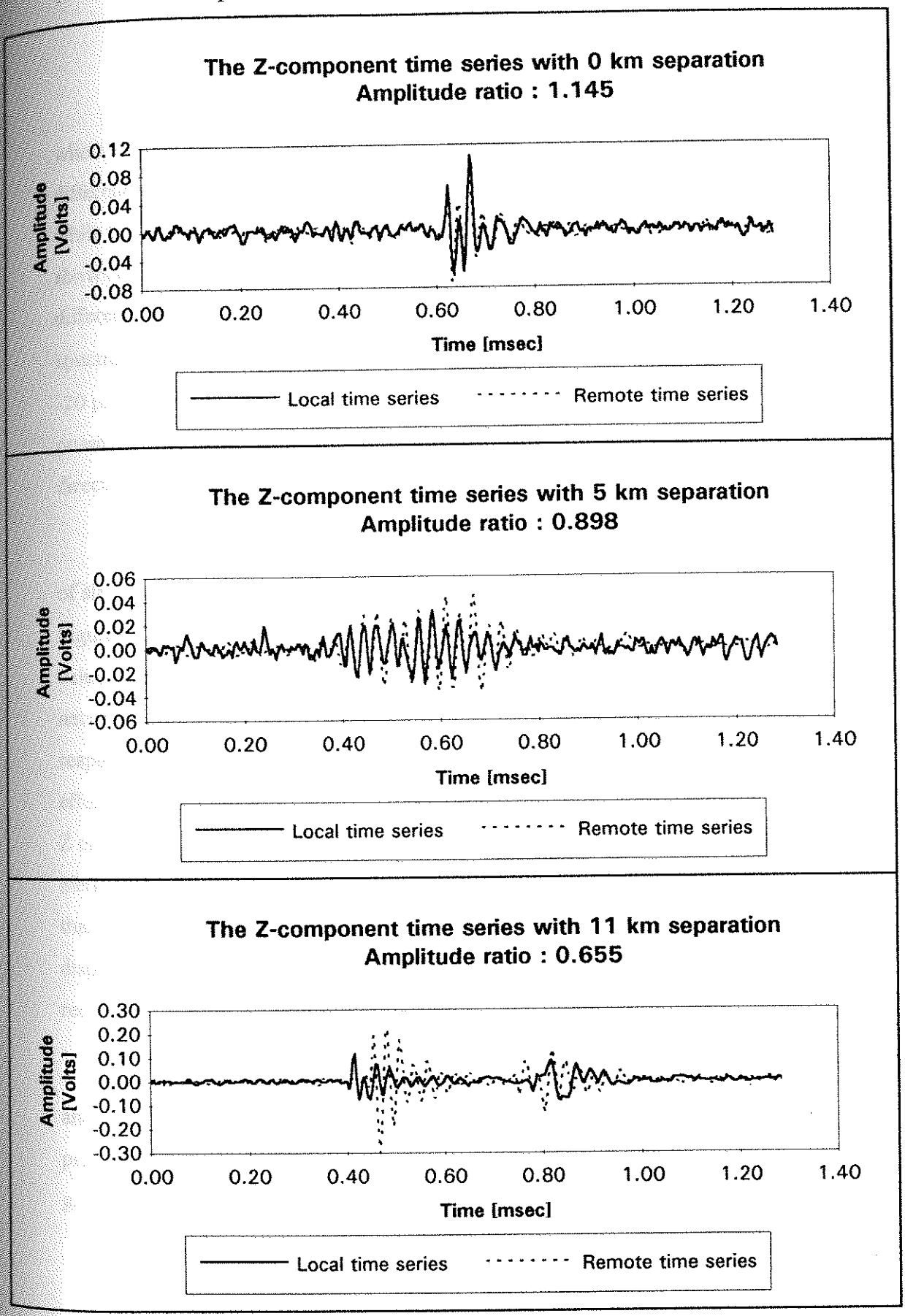

Figure 5.11 Effect of ground conductivity change on the Z component of sferics noise measured simultaneously at two stations with station separations of 5 and 11 km in addition to two sferics pulses measured simultaneously at the same station. These measurements were carried out at Koolpinyah, near Darwin in December 1994.

where the amplitude ratio of 0.63 was obtained after compensating for the sensor difference (Figure 5.12). In this measurement, one receiver was placed near the sea while the second receiver was placed inland. The cross-correlation of two time series shows anti-correlation, and the time delay calculated from the phase spectrum has different values at different frequencies as shown by the varying slope of this spectrum. For example, the time delay in the frequency range from ~20 to 30 kHz is -20 μs. In the frequency range from ~30 to 42 kHz the time delay is +12 μs. It is possible that this response consists of two sferics pulses originating in quite different directions, and hence its structure is more complicated.

Figures 5.13 and 5.14 show examples of an RNPF applied to the Z component of sferics. The residual time shift between the two measured time series shown in Figure 5.13 is -1.76 μs, while the residual time shift in Figure 5.14 is 2.39 μs. An *NRF* value of 0.4 is obtained from the simple subtraction in both examples. The neural network filter achieves *NRF* values of 0.6 and 1.2 in Figures 5.13 and 5.14, respectively. In both cases, the RNPF and simple subtraction methods do not produce effective noise reduction, unlike their application to the X or Y component. Since the Z component of sferics is generally ten times smaller than the X or Y component of sferics, uncorrelated random electronics noise is relatively greater on the Z component than on the horizontal component. This uncorrelated random electronic noise and the dispersion effect explained in Section 2.4.5 make it difficult for the neural network to recognise the amplitudes and time shifts between the local and remote time series.

A description of the analysis of cross-correlation between the local vertical and remote horizontal (X and Y) components, and of a neural network configured to predict the local vertical component using the remote X, Y, and Z components is given in Section 7.2.2.2 of Chapter 7.

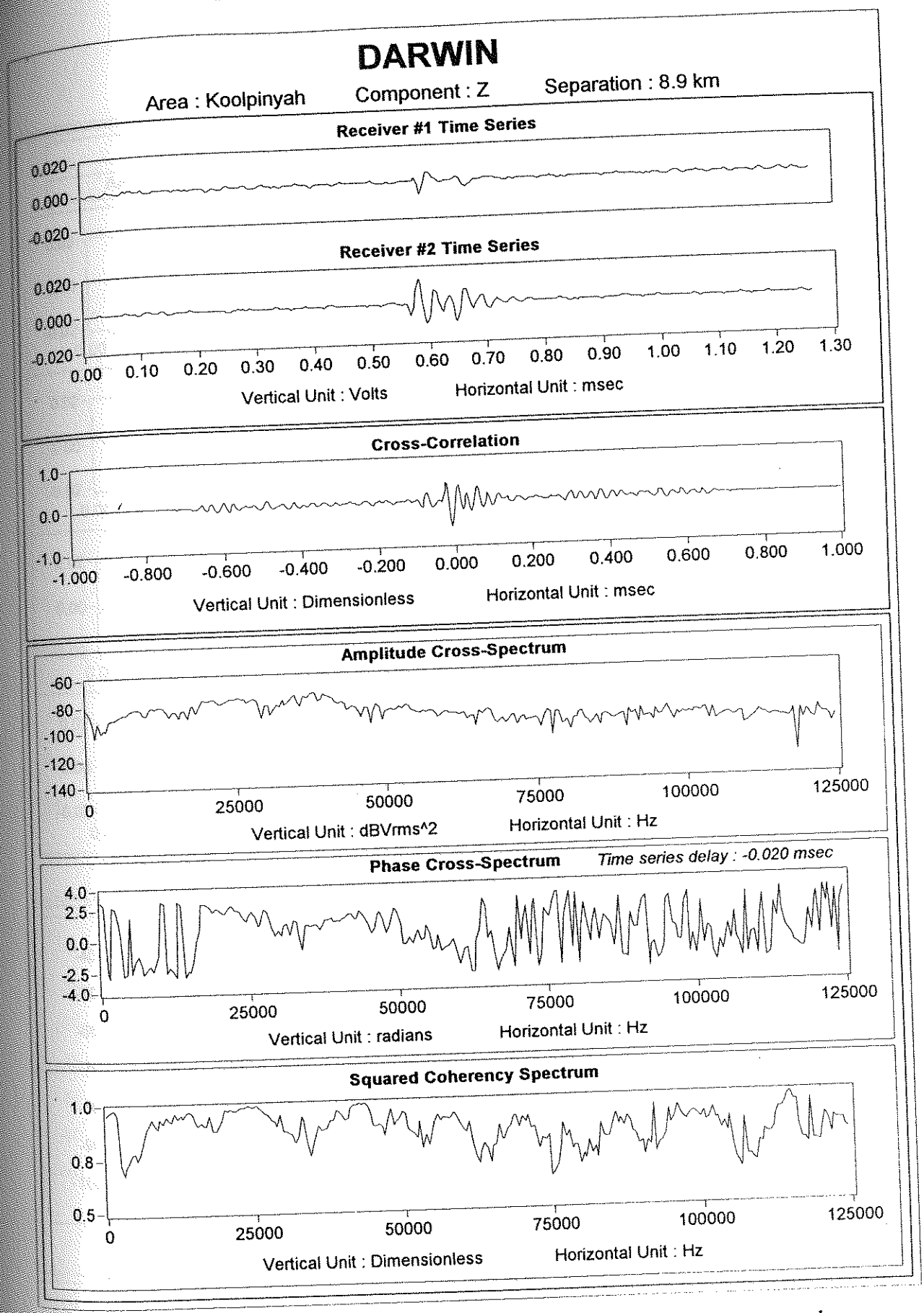

Figure 5.12 Example of time series, correlation function, and cross-power and cross-phase spectra for the Z component of high-frequency sferics data recorded in December 1994 at a station separation of 11 km at Koolpinyah, near Darwin.

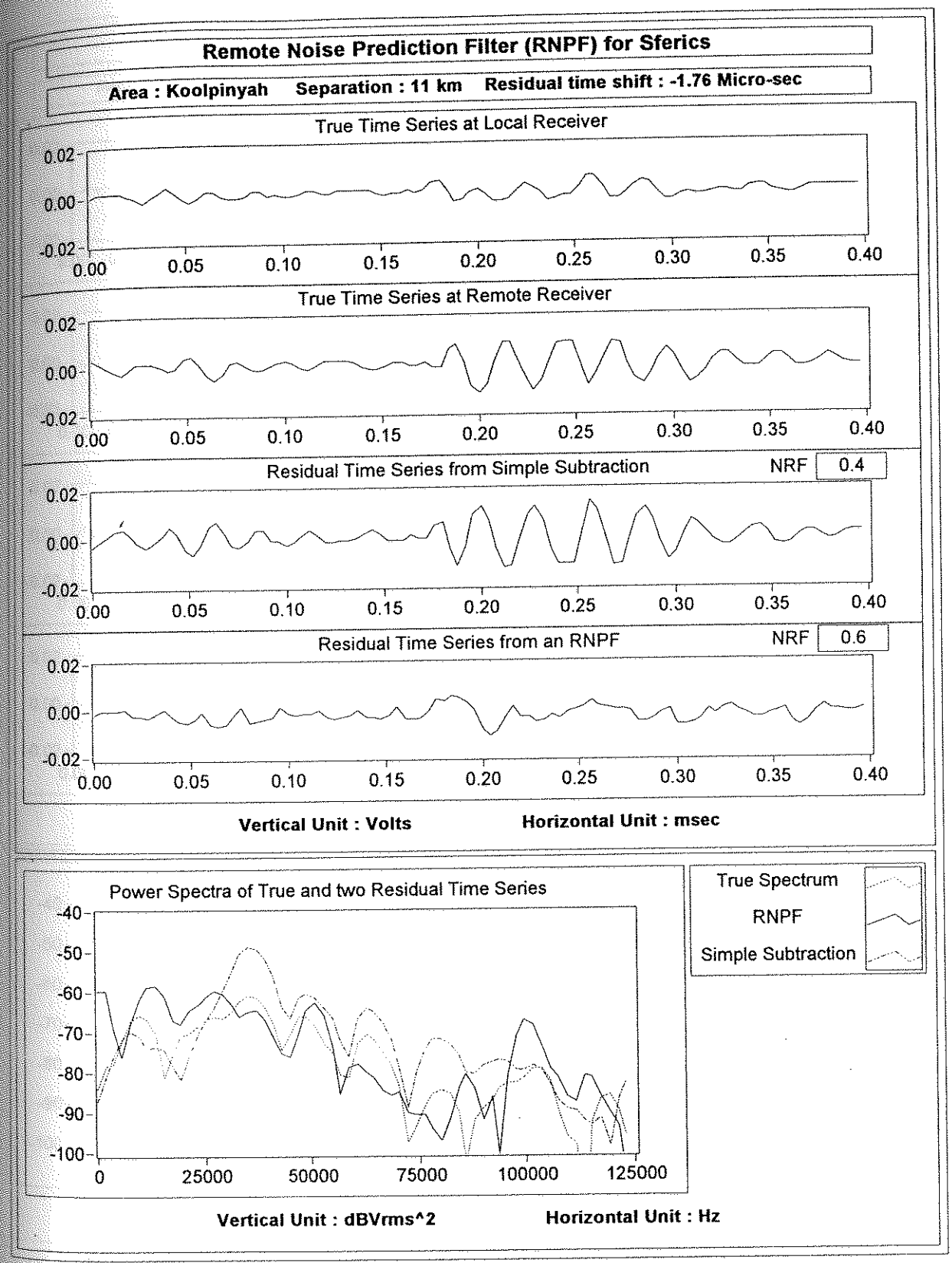

Figure 5.13 Example of performance of an RNPF and simple subtraction used to reduce the Z component of high-frequency sferics noise. Two sferics pulses were measured simultaneously at the local and remote receivers with a separation of 11 km and the residual time shift between the two sferics pulses is -1.76 μs.

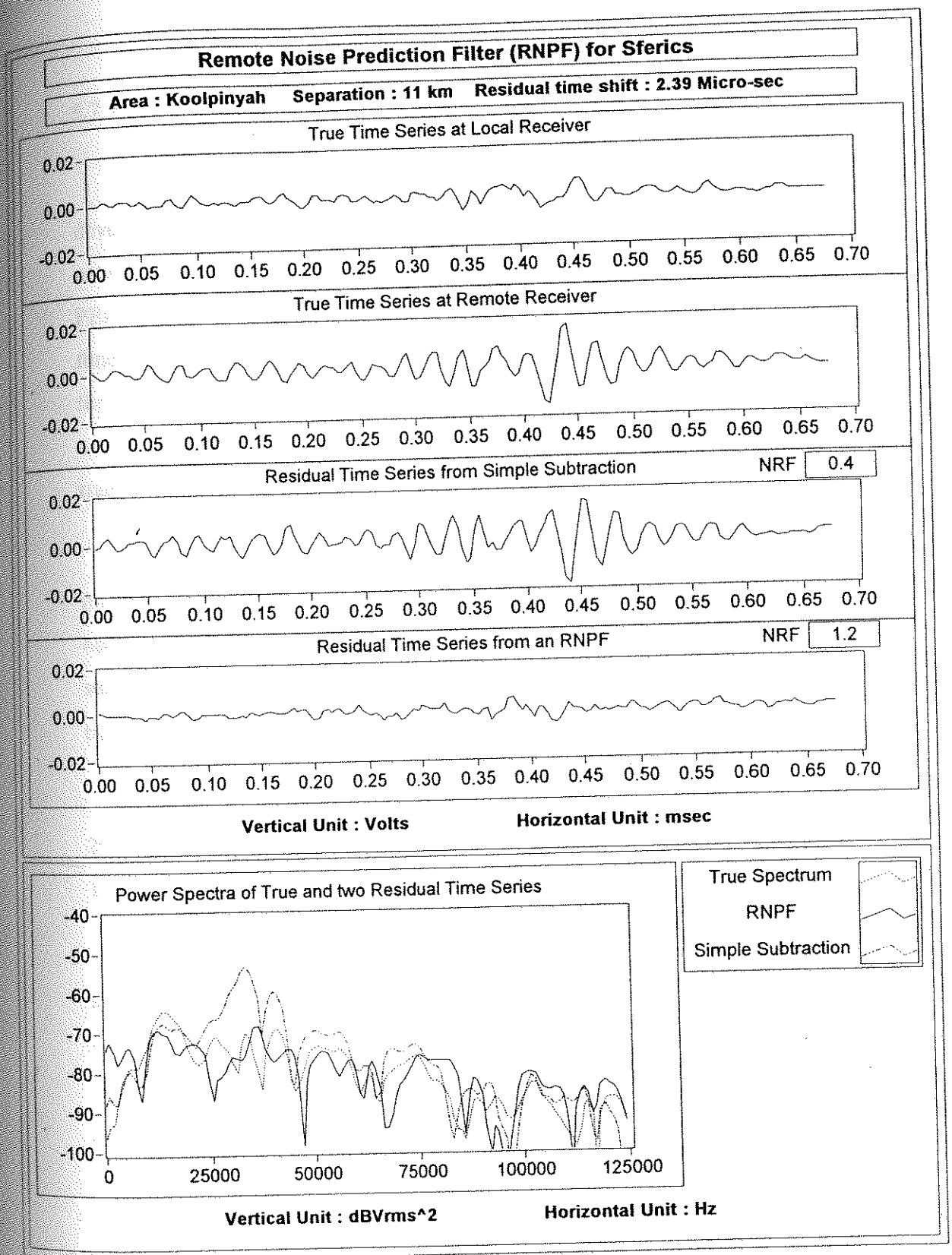

Figure 5.14 Example of performance of an RNPF and simple subtraction used to reduce the Z component of high-frequency sferics noise. Two sferics pulses were measured simultaneously at the local and remote receivers with a separation of 11 km and the residual time shift between the two sferics pulses is 2.39 μs.

5.7 Conclusion

A neural network-based remote noise prediction filter (RNPF) has been developed for both background EM noise and high-frequency sferics. The optimal network for both background EM noise and high-frequency sferics consists of 6 input PEs, 25 hidden PEs, and 5 output PEs. The nonlinearities are located in the activation function of hidden PEs for sferics pulse prediction and in both hidden and output PEs for background EM noise prediction. A hyperbolic tangent is used as the network activation function. The networks for background EM noise and sferics noise are trained for 170 and 250 cycles, respectively.

For two time series of given X and Y components of high-frequency sferics measured at two separated stations, the noise attenuation produced with a neural network RNPF is much greater than the noise attenuation obtained from simple subtraction of one time series from the other. The neural network RNPF can be trained to accommodate any residual time shift smaller than the sample interval (4 μs).

An RNPF achieves an attenuation in the reduction of background EM noise by 20 dB (i.e., by a factor of 10 in amplitude) at the VLF frequency of ~20 kHz (NWC). The performance of an RNPF is invariant over time.

For the reduction of the horizontal (X and Y) components of high-frequency sferics, an RNPF attenuates the power of sferics noise by a factor of more than 25 dB (i.e., by a factor of 18 in sferics amplitude) at frequencies in the range of 5 to 50 kHz. When there is a time shift between corresponding sferics pulses measured at two stations, the attenuation of any accompanying VLF noise may be less than that of sferics because correction for the time shift of a given sferics pulse introduces an inappropriate time shift in the VLF signals accompanying the sferics pulse.

The vertical (Z) component of sferics pulse was the component most affected by a ground conductivity change between two separated stations in the case studied. For the reduction of this component of high-frequency sferics, both an RNPF and simple subtraction method do not show the same performance obtained when the same filter is applied to the X and Y components.