JET PROPULSION LABORATORY INTEROFFICE MEMORANDUM

331.1-95-048
November 20, 1995

To: Distribution

From: Dave Watola

Subject: Differential Discriminative Diagnosis for Downlink Analyzer

REFERENCES:

1. J. B. Hampshire II, *A Differential Theory of Learning for Efficient Statistical Pattern Recognition*. PhD. thesis, Carnegie Mellon University, Department of Electrical & Computer Engineering, September 1993.

2. Y. Le Cun, J. S. Denker, and S. A. Solla, "Optimal Brain Damage," in *Advances in Neural Information Processing Systems 2*, D. S. Touretzky, ed., pp. 598-605, Morgan Kaufmann, San Mateo, CA, 1990.

3. B. A. Pearlmutter, "Fast Exact Multiplication by the Hessian," *Neural Computation* 6(1), January 1994.

4. D. Watola, "Overview of Downlink Analyzer Neural Network Technology," JPL IOM 331.1-95-047, November 20, 1995.

1. Purpose

This document is the penultimate in a series of five memoranda detailing the background theory, implementation, and issues surrounding the Downlink Analyzer task. Focusing exclusively on the Differential Discriminative Diagnosis procedure used by Downlink Analyzer and its ***DDD*** software to produce a ranked list of causes after the discovery of an anomaly in the system under observation, this memo represents the first available coherent documentation of the Discriminative Diagnosis algorithm; all previous written work comprises only raw notes and proposals.

2. Discriminative Diagnosis

Discriminative Diagnosis (DDIAG) is the brainchild of John Hampshire, and is based loosely on the Optimal Brain Damage (OBD) algorithm [2] for reducing functional complexity in neural networks by excising unnecessary weights from the network with the intent of reducing empirical discriminant variance, thereby improving generalization. OBD computes and ranks the importance (saliency) of weights to the average classification performance of a neural network over an ensemble of training vectors by modelling parameter removal as a perturbation of the weight vector that induces a degradation in the ensemble-averaged objective function score. DDIAG resembles OBD in that it quantifies the penalty (or benefit) associated with each component of an input perturbation by decomposing the change in objective function score accompanying a transition in output classification. However, DDIAG possesses many subtle but critical differences from OBD. These subtle variations are explored below. The notation and background of [4], where OBD is fully developed in the context of the differential learning environment of Downlink Analyzer and ***dlaNet***, are assumed.

For the purposes of diagnosis, define the external state of a neural network model at any given instant as the 3-tuple ($\mathbf{X}$, $\mathbf{Y}$, Ω), where $\mathbf{X}$ is the input feature vector driving the network, $\mathbf{Y}$ is the concomitant output vector, and Ω is the true health state of the system being observed, i.e. the classification indicated by $\mathbf{Y}$ if the network model is correct. Given an initial state ($\mathbf{X}_0$, $\mathbf{Y}_0$, Ω_0) followed by a later state ($\mathbf{X}_1$, $\mathbf{Y}_1$, Ω_1), a measure of how desirable or undesirable $\mathbf{X}_1$ is relative to the reference $\mathbf{X}_0$ is obtainable via a Taylor Series expansion of the objective function, $\eta_0(\mathbf{X})$, about $\mathbf{X}_0$:

$$\begin{aligned}\eta_0(\mathbf{X}_1) &= \eta_0(\mathbf{X}_0) + (\mathbf{X}_1 - \mathbf{X}_0)^T \cdot \nabla_X \eta_0(\mathbf{X}_0) \\ &\quad + \frac{1}{2}(\mathbf{X}_1 - \mathbf{X}_0)^T \cdot \nabla_X^2 \eta_0(\mathbf{X}_0) \cdot (\mathbf{X}_1 - \mathbf{X}_0) + \mathrm{O}\left(\|\mathbf{X}_1 - \mathbf{X}_0\|^3\right), \qquad (1)\end{aligned}$$

where the gradient and Hessian are computed with respect to the inputs rather than the weights. All derivatives are evaluated at $\mathbf{X}_0$, assuming a correct mapping of $\mathbf{X}_0$ into class Ω_0 by the model. Ω_1 and intermediate state involved in computing $\mathbf{Y}_1$ are never used, nor are derivatives resulting from input $\mathbf{X}_1$. Accordingly, (1) is *not* an attempt to approximate $\eta_1(\mathbf{X}_1)$; doing so would inappropriately combine unrelated factors since that quantity is associated with a classification of Ω_1. This distinction is particularly profound for the case of CFM [1], where $\eta_0(\cdot)$ and $\eta_1(\cdot)$ may be different functions of $\mathbf{Y}$. The subscript in $\eta_0(\cdot)$ reinforces the fact that (1) quantifies a change in the objective function associated with a classification of Ω_0. It also underscores another important difference between OBD and DDIAG: whereas the former is driven by an objective function score averaged over the ensemble of training vectors, the latter is based solely on the value of η for a single input.

As in OBD, some assumptions are necessary in order to render (1) tractable. DDIAG shares the quadratic assumption adopted by OBD, so third- and higher-order terms are considered to be negligible. In contrast to OBD, however, there is no reason to assume that the gradient component of (1) also vanishes. Consequently, the value of $\Delta\eta_0=\eta_0(\mathbf{X}_1)-\eta_0(\mathbf{X}_0)$ is not constrained to either exclusive non-negativity or non-positivity.

Defining $\Delta\mathbf{X}=\mathbf{X}_1-\mathbf{X}_0$ then yields the more compact expression

$$\Delta\eta_0(\Delta\mathbf{X}) = \Delta\mathbf{X}^T\cdot\nabla_x\eta_0(\mathbf{X}_0) + \frac{1}{2}\Delta\mathbf{X}^T\cdot\nabla_x^2\eta_0(\mathbf{X}_0)\cdot\Delta\mathbf{X}. \tag{2}$$

Rewriting using only scalar quantities allows the contribution from each input perturbation component to be isolated:

$$\Delta\eta_0(\Delta\mathbf{X}) = \sum_{i=1}^{N_I}\left(\Delta X_i\cdot\frac{\partial\eta_0(\mathbf{X}_0)}{\partial X_i} + \frac{1}{2}\Delta X_i\sum_{j=1}^{N_I}\Delta X_j\cdot\frac{\partial^2\eta_0(\mathbf{X}_0)}{\partial X_i\partial X_j}\right). \tag{3}$$

Individual saliencies for each network input can be extracted from (3) as

$$\Psi_{X_i} = \Delta X_i\cdot\left(\frac{\partial\eta_0(\mathbf{X}_0)}{\partial X_i} + \frac{1}{2}\sum_{j=1}^{N_I}\Delta X_j\cdot\frac{\partial^2\eta_0(\mathbf{X}_0)}{\partial X_i\partial X_j}\right), \tag{4}$$

where the components due to the gradient and Hessian are clearly distinguishable.

In the Downlink Analyzer task, the primary interest is diagnosis of the causal factors prompting passage from a classification of Ω_0=GOOD to a state with Ω_1=BAD. When this transition occurs, (4) is used to find which components of the input cause the largest degradation in η given that the desired class should still be $\Omega_1=\Omega_0$=GOOD. The only additional constraint is that the classifier be robust, i.e. the neural network must provide a reliable estimate of the true health state of the system for $\mathbf{X}=\mathbf{X}_1$ and $\mathbf{X}=\mathbf{X}_0$, otherwise there is no justification for applying (2). Thus, DDIAG cannot shed insight into a situation where Ω_0=GOOD and Ω_1=BAD if $\mathbf{Y}_1$ erroneously testifies that the new health state is still GOOD. The appearance of these scenarios are minimized by providing adequate network training. Their inevitable manifestation, if sufficiently infrequent, is not a problem because DDIAG always has access to the true Ω and can therefore avoid using information from misclassified inputs.

3. Diagonal-Hessian Discriminative Diagnosis

As with Optimal Brain Damage, a simplified version of DDIAG can be implemented by assuming that the Hessian is diagonal and ignoring all cross-derivatives in (4). Diagonal-Hessian DDIAG is slightly simpler than diagonal-Hessian OBD since it only requires a single additional equation. This extra equation, applied during an augmented backward pass of the backpropagation algorithm, is identical to one used for Diagonal-Hessian OBD:

$$\frac{\partial^2\eta}{\partial {y_i}^2} = \begin{cases} \dfrac{\partial}{\partial y_i}\left(\dfrac{\partial\eta(\mathbf{Y})}{\partial y_i}\right) = \dfrac{\partial^2\eta(y_{N-N_O+1},\ \ldots,\ y_N)}{\partial {y_i}^2} & \text{for } i > N-N_O \\ \dfrac{\partial}{\partial y_i}\left(\displaystyle\sum_{j:c_{j,i}=1}\dfrac{\partial\eta}{\partial x_j}\dfrac{\partial x_j}{\partial y_i}\right) = \displaystyle\sum_{j:c_{j,i}=1} w_{j,i}^2\left(\dfrac{\partial\eta}{\partial y_j}F_j''(x_j) + \dfrac{\partial^2\eta}{\partial {y_j}^2}F_j'(x_j)^2\right) & \text{for } i \le N-N_O \end{cases} \tag{5}$$

where the diagonal Hessian values corresponding to inputs are

$$\frac{\partial^2 \eta}{\partial X_i^2} = \frac{\partial^2 \eta}{\partial y_i^2}, \quad i = 1, \ldots, N_I. \tag{6}$$

The desired saliencies in (4) are then approximated by

$$\hat{\Psi}_{X_i} = \Delta X_i \cdot \frac{\partial \eta_0(\mathbf{X}_0)}{\partial X_i} + \frac{1}{2} \Delta X_i^2 \cdot \frac{\partial^2 \eta_0(\mathbf{X}_0)}{\partial X_i^2}. \tag{7}$$

4. Full-Hessian Discriminative Diagnosis

The algorithm for implementing the full-Hessian variant of DDIAG is closely related to the scheme presented in [4] for effecting full-Hessian OBD; it will be helpful to review that reference before continuing with this section since the derivations are parallel and the presentation below is not as explicit. Once again, a linear differential operator is introduced, *viz.*,

$$R_{\Delta X;X}\{f(\mathbf{X})\} = \frac{\partial}{\partial r} f(\mathbf{X} + r \cdot \Delta \mathbf{X})\Big|_{r=0}. \tag{8}$$

This time, the operator is a functional affecting functions of inputs rather than weights. Consult [4] for a comprehensive summary of the useful properties of this operator *vis a vis* neural networks and their state perturbations.

Equation (8) can be used to implement DDIAG based on the complete Hessian by supplementing the backpropagation algorithm with four new equations. The forward equations are

$$\begin{aligned} R_{\Delta X;X}\{x_i\} &= R_{\Delta X;X}\left\{ \sum_{j:c_{i,j}=1} w_{i,j} y_j \right\} \\ &= \sum_{j:c_{i,j}=1} \left(w_{i,j} \cdot R_{\Delta X;X}\{y_j\} + y_j \cdot R_{\Delta X;X}\{w_{i,j}\} \right) = \sum_{j:c_{i,j}=1} w_{i,j} \cdot R_{\Delta X;X}\{y_j\} \end{aligned} \tag{9}$$

and

$$R_{\Delta X;X}\{y_i\} = \begin{cases} R_{\Delta X;X}\{F_i(x_i)\} = F_i'(x_i) \cdot R_{\Delta X;X}\{x_i\}, & N_I < i \le N \\ \Delta X_i, & 1 < i \le N_I \end{cases} \tag{10}$$

where (9) is simpler than its OBD counterpart because $R_{\Delta X;X}\{w_{i,j}\}=0$ for all (i, j), a consequence of the fact that adaptation has been terminated, severing any dependence that the weights had on the network inputs.

The backward equations are:

$$R_{\Delta X;X}\left\{\frac{\partial\eta}{\partial y_i}\right\} = \begin{cases} \displaystyle\sum_{j=N-N_O+1}^{N} \frac{\partial}{\partial y_j}\left(\frac{\partial\eta}{\partial y_i}\right)\cdot R_{\Delta X;X}\{y_j\}, & N-N_O < i \le N \\ \displaystyle\sum_{j:c_{j,i}=1} w_{j,i}\cdot R_{\Delta X;X}\left\{\frac{\partial\eta}{\partial x_j}\right\}, & 1 \le i \le N-N_O \end{cases} \tag{11}$$

and

$$R_{\Delta X;X}\left\{\frac{\partial\eta}{\partial x_i}\right\} = R_{\Delta X;X}\left\{\frac{\partial\eta}{\partial y_i}\frac{\partial y_i}{\partial x_i}\right\} = F_i'(x_i)\cdot R_{\Delta X;X}\left\{\frac{\partial\eta}{\partial y_i}\right\} + \frac{\partial\eta}{\partial y_i}\cdot F_i''(x_i)\cdot R_{\Delta X;X}\{x_i\}. \tag{12}$$

Like (9), (11) is also simpler than its OBD counterpart because the connection strengths no longer vary with inputs or input perturbations.

The property that

$$R_{\Delta X;X}\left\{\nabla_x\eta(\mathbf{X})\right\} = \nabla_x^2\eta(\mathbf{X})\cdot\Delta\mathbf{X}, \tag{13}$$

arising from the definition in (8), permits the desired saliencies to be computed as

$$\Psi_{X_i} = \frac{1}{2}\Delta X_i\cdot R_{\Delta X;X}\left\{\frac{\partial\eta}{\partial X_i}\right\} = \frac{1}{2}\Delta X_i\cdot R_{\Delta X;X}\left\{\frac{\partial\eta}{\partial y_i}\right\},\quad 1 \le i \le N_I. \tag{14}$$

5. Differential Discriminative Diagnosis

Differential DDIAG is a variant of DDIAG employing an objective function that engenders differential learning rather than probabilistic learning. Only the CFM objective function is considered here because it is employed by Downlink Analyzer software.

The definition of CFM ensures that all derivatives of η_{CFM} are zero except those associated with two of the network outputs. Derivatives also disappear for discriminant differentials that equal or exceed the confidence parameter, δ, associated with the CFM function $\sigma_\delta(\cdot)$. Thus, it is mandatory that the value of δ used with DDIAG exceed the discriminant differential accompanying ($\mathbf{X}_0$, $\mathbf{Y}_0$, Ω_0); it is not sufficient to fix δ at its final value from the network training phase. Selection of the optimum value is a problem that has not yet been fully explored.

Note that it is not desirable to set δ=1 (the maximum value) since this extreme value makes the CFM function nearly linear, causing all Hessian elements to vanish in the region of interest.

A further consequence of the use of CFM is that derivatives of η_{CFM} with respect to a network output y_i may depend on another output y_j in addition to y_i. This is in sharp contrast to most commonly used objective functions, where each output displays an error derived from strictly local information.

A more subtle aspect of this phenomenon is the fact that which two specific outputs are involved varies depending on the true class and the result vector **Y**. Thus, $\eta_{CFM}(\cdot)$ will be a different function between the classification trials ($\mathbf{X}_0$, $\mathbf{Y}_0$, Ω_0) and ($\mathbf{X}_1$, $\mathbf{Y}_1$, Ω_1) when $\Omega_0 \neq \Omega_1$. Even when the two classes are equal, the forms of their trial objective functions may differ if there are more than two classes.

7. Differential DDIAG and Downlink Analyzer

A prototype software package that performs Differential Discriminative Diagnosis for the two-class, CFM-based problem has been developed. This tool, ***DDD***, currently reads feature vectors from files written by the ***DLA*** tool and performs user-directed diagnoses with a neural network trained by the ***dlaNet*** program. In the future, ***DDD*** will be able to accept real-time data directly from ***DLA*** and execute diagnoses when necessary without human intervention.

A sample diagnosis is illustrated in the remainder of this section by Figures 1-10, which are graphics extracted from the standard ***DLA*** and ***DDD*** tool displays. This sample diagnosis is based on data recorded at DSS-13 on January 16, 1993 during the Ka-band Link Experiment (KaBLE) conducted with the Mars Observer spacecraft. The principal health metric is the X-band P_c/N_0 statistic. Available monitor data for causal factors are limited to pointing metrics and weather station data. Specifically, the time series used were azimuth, elevation, azimuth error, elevation error, wind speed, wind direction, temperature, and water vapor density.

Figure 1 depicts the transition of the link from GOOD to BAD. Because no attempt was made to optimize the neural network or the choice of feature vectors, the model is not robust during the transition. Therefore, the two vectors selected as a reference state and anomalous state are not adjacent. Instead, one sample exhibiting high discriminant differentials was selected from each side of the transition.

Figure 2 shows the raw time series from which features were extracted using ***DLA***. The principal health metric, P_c/N_0, wavers about the threshold of the GOOD/BAD SNR boundary several times in this graph; the vertical cursor indicates the location of the selected anomalous state, and a short interval around the region of interest is highlighted in gray. In the vicinity of the anomaly, SNT rises slightly. However, the most dramatic change is a sudden shift in wind direction when the wind is at low velocity. This phenomenon is known to excite underdamped modes in the antenna pointing system, leading to oscillations in pointing error that can be a significant source of loss at X-band frequencies. Indeed, oscillations in azimuth error are visible in the time series immediately after the wind event. The other time series plots are omitted because they display no interesting characteristics.

Figure 3 shows all of the feature vector component values associated with the reference state. Note that the network uses some derived features in addition to (or in lieu of) the raw time series data. For example, it was anticipated that the magnitude of the pointing error metrics would be more useful than the bipolar-valued pointing errors themselves. Also, foreknowledge of the existence of underdamped pointing system modes encouraged the selection of wind speed variance and direction variance as features.

Figure 4 is a graphical depiction of the entire network internal state for the reference input. Nodes are shaded according to the output of the represented neuron, with maximum shading (black) indicating a value of -1.0 and minimum shading (white) signalling +1.0. Input nodes are shaded according to the normalized value of the input since all features are mapped to the interval [-1.0, +1.0] before being presented to the network. The unscaled, unshifted feature values are shown at left, beneath the feature name. On the right is a summary of classifier performance, including the value of discriminant differential. This display allows at-a-glance determination of classification performance as well as indicating which hidden nodes respond most strongly to the current input pattern.

Figures 5 and 6 show the same information as Figures 3 and 4, except that their displays are based on the anomalous feature vector. Side-by-side comparison of input node shading in Figures 4 and 6 reveals that wind-related features have changed dramatically, while others have remained quiescent.

Saliencies computed by applying both the approximate diagonal-Hessian and full-Hessian DDIAG procedures are shown in Figure 7. Both methods select wind direction variance as the primary contributor to the degradation in performance, demonstrating a top-ranked saliency that is at least an order of magnitude larger than the next most salient perturbation component. In a robust application, more difference between salient parameters and those to which the procedure should be insensitive (e.g. relative wind direction) would be expected. Unfortunately, severely limited data for training precludes this from happening here.

Finally, Figure 8 is a graphical summary of the saliencies computed using the full-Hessian DDIAG technique. Here, the most salient parameter is shaded with intensity 1.0 (white). Other features that caused a degradation in performance are shaded with intensities proportional to the magnitude of their saliencies. The remaining features, which have saliencies sporting the opposite sign, are shaded flat black. This display is mainly useful as an immediate indicator of the top-ranked perturbation component.

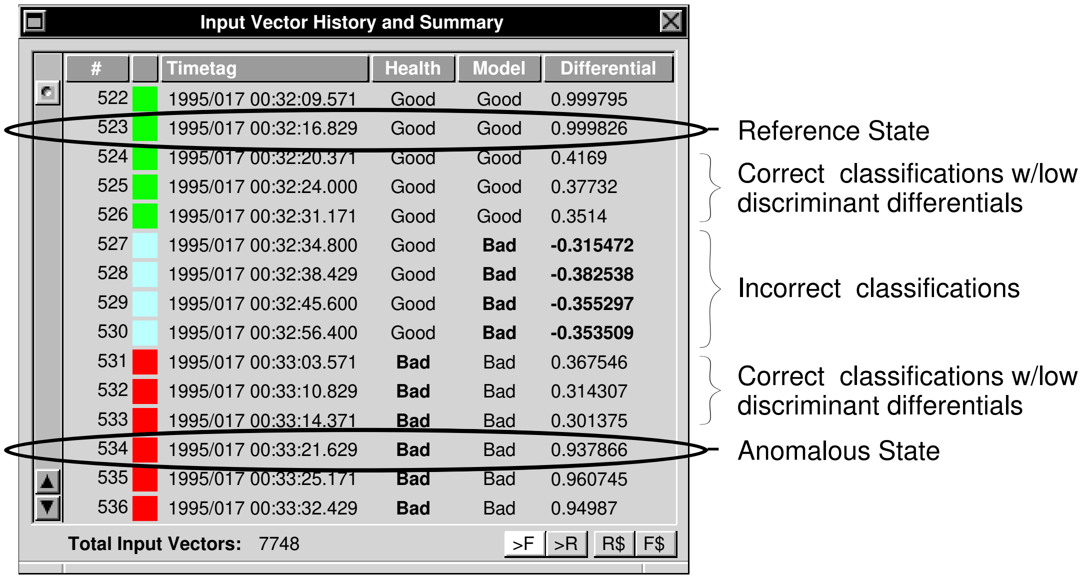


Figure 1. ***Reference and anomalous vectors selected for DDIAG.***

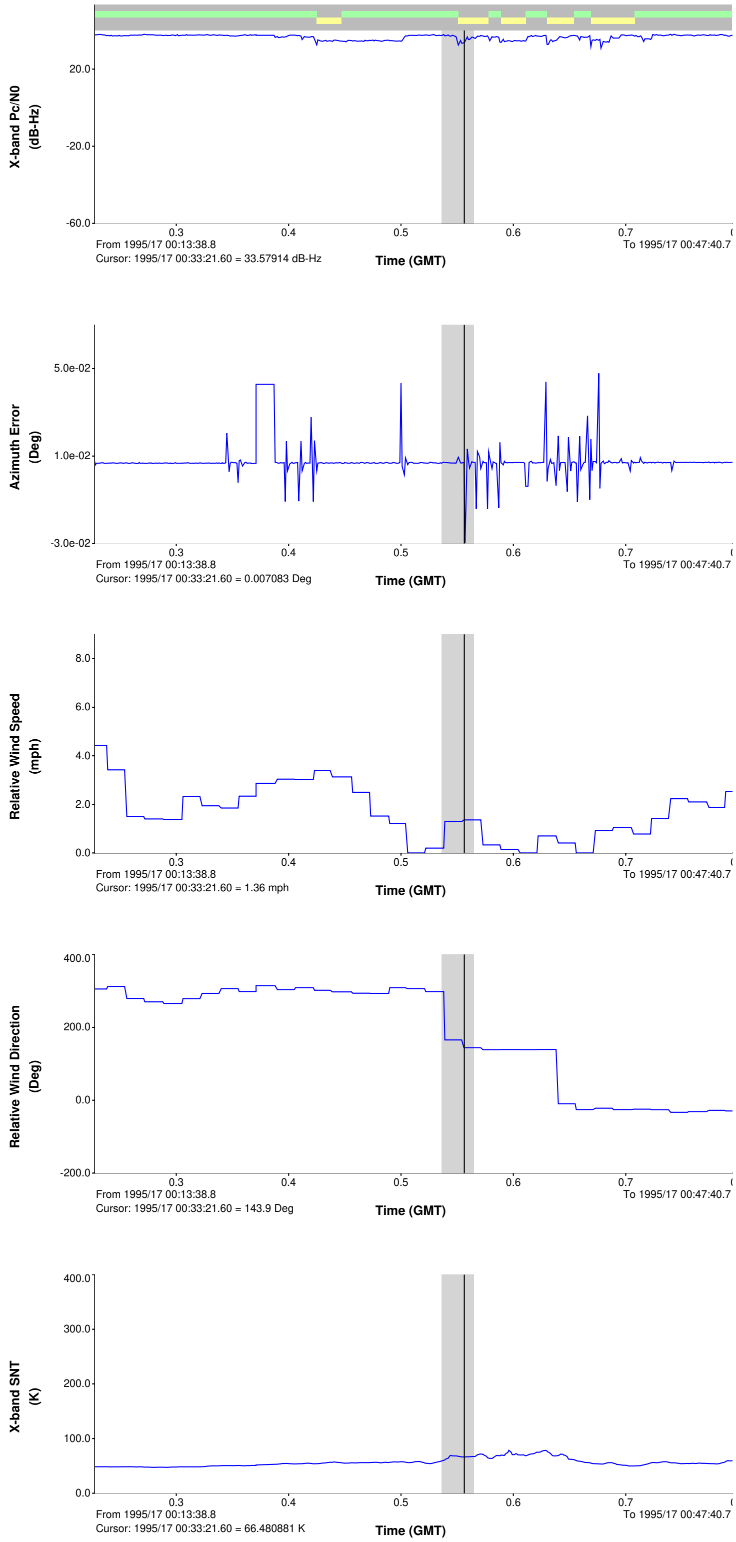


Figure 2. ***Raw time series data depicting conditions at anomaly onset.***

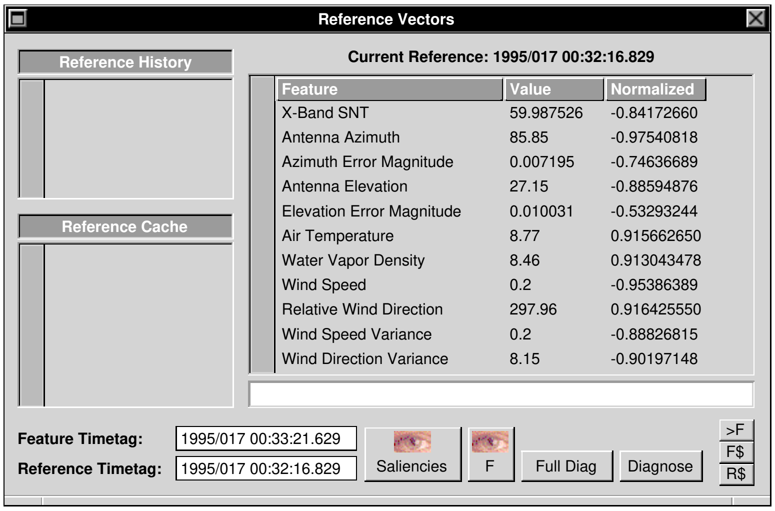


Figure 3. ***Reference vector component display*.**

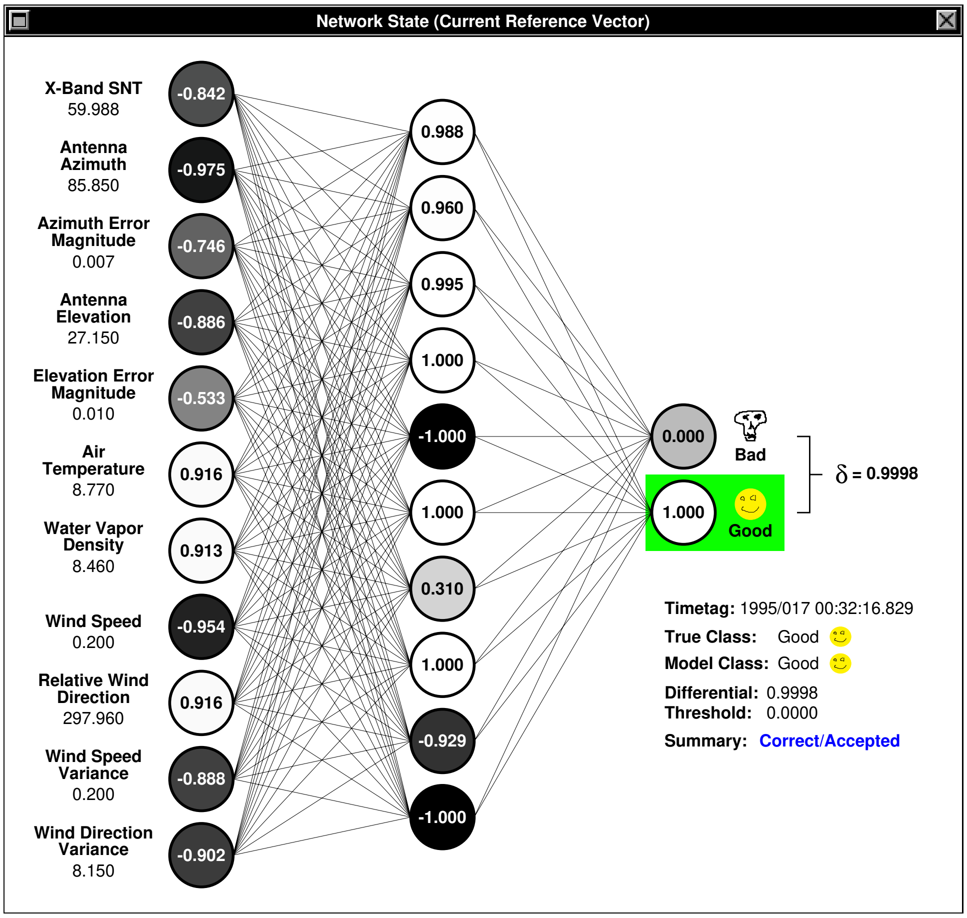


Figure 4. ***Neural network model state display for reference vector*.**

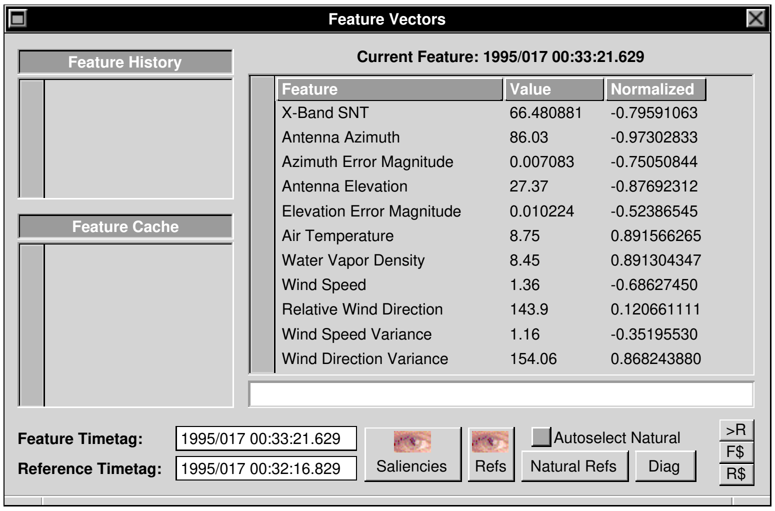


Figure 5. *Anomalous feature vector component display*.

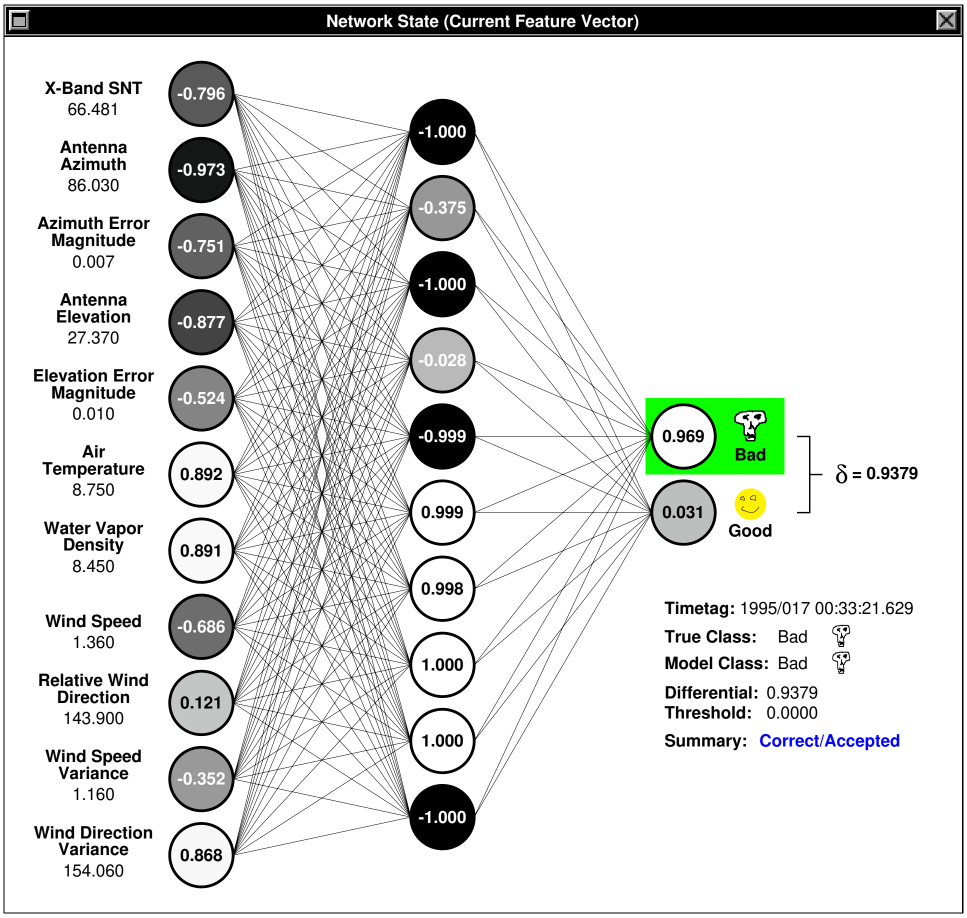
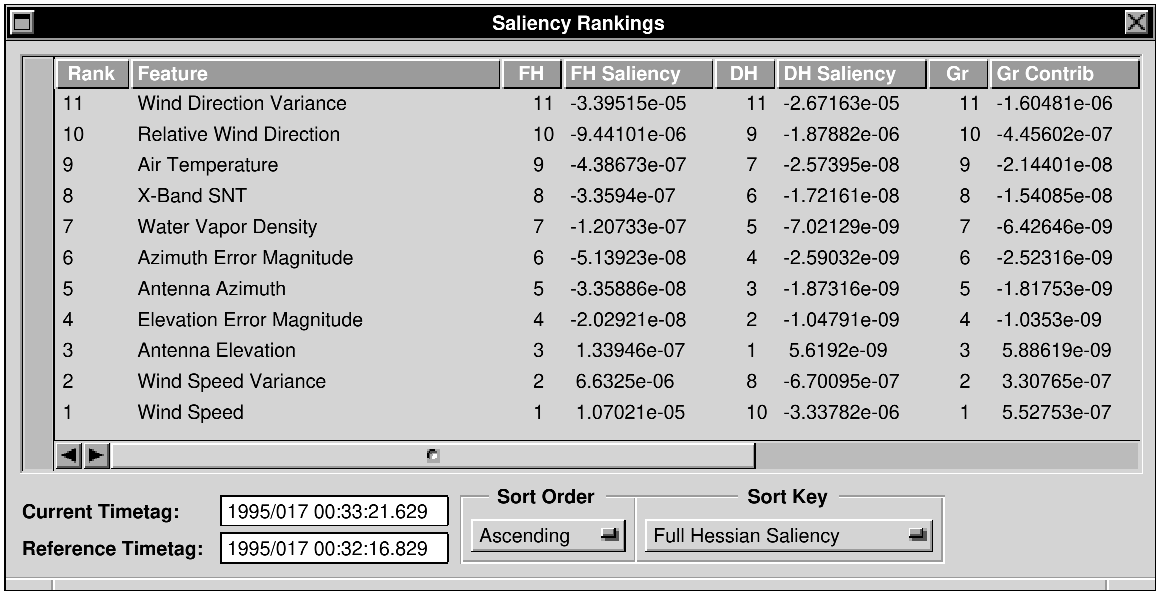


Figure 6. *Neural network model state display for anomalous feature vector*.



Rank	Feature	FH	FH Saliency	DH	DH Saliency	Gr	Gr Contrib
11	Wind Direction Variance	11	-3.39515e-05	11	-2.67163e-05	11	-1.60481e-06
10	Relative Wind Direction	10	-9.44101e-06	9	-1.87882e-06	10	-4.45602e-07
9	Air Temperature	9	-4.38673e-07	7	-2.57395e-08	9	-2.14401e-08
8	X-Band SNT	8	-3.3594e-07	6	-1.72161e-08	8	-1.54085e-08
7	Water Vapor Density	7	-1.20733e-07	5	-7.02129e-09	7	-6.42646e-09
6	Azimuth Error Magnitude	6	-5.13923e-08	4	-2.59032e-09	6	-2.52316e-09
5	Antenna Azimuth	5	-3.35886e-08	3	-1.87316e-09	5	-1.81753e-09
4	Elevation Error Magnitude	4	-2.02921e-08	2	-1.04791e-09	4	-1.0353e-09
3	Antenna Elevation	3	1.33946e-07	1	5.6192e-09	3	5.88619e-09
2	Wind Speed Variance	2	6.6325e-06	8	-6.70095e-07	2	3.30765e-07
1	Wind Speed	1	1.07021e-05	10	-3.33782e-06	1	5.52753e-07

Figure 7. ***Ranked saliency list*.**

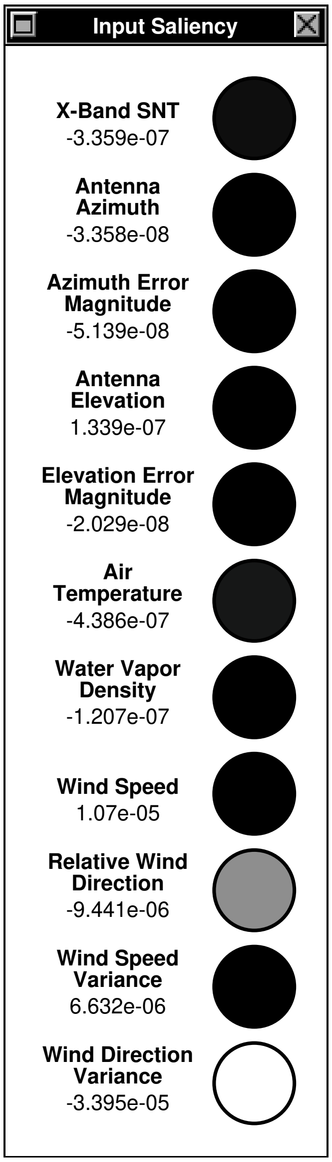


Figure 8. ***Graphical ranked saliency display*.**

8. Summary of Dual-class CFM-based DDIAG Backpropagation Equations

Standard forward pass:

$$x_i = \begin{cases} \sum_{j:\, c_{i,j}=1} w_{i,j} y_j & \text{for } i > N_I \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

$$y_i = \begin{cases} F_i(x_i) & \text{for } i > N_I \\ X_i & \text{otherwise} \end{cases} \tag{16}$$

$$\frac{\partial y_i}{\partial x_i} = \begin{cases} F_i'(x_i) & \text{for } i > N_I \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

$$\frac{\partial^2 y_i}{\partial x_i^2} = \begin{cases} F_i''(x_i) & \text{for } i > N_I \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

Augmented forward pass for full-Hessian DDIAG:

$$R_{\Delta X;X}\{x_i\} = \begin{cases} \sum_{j:c_{i,j}=1} w_{i,j} \cdot R_{\Delta X;X}\{y_j\} & \text{for } i > N_I \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

$$R_{\Delta X;X}\{y_i\} = \begin{cases} F_i'(x_i) \cdot R_{\Delta X;X}\{x_i\} & \text{for } i > N_I \\ \Delta X_i & \text{otherwise} \end{cases} \tag{20}$$

Standard backward pass:

$$\frac{\partial \eta}{\partial y_i} = \begin{cases} \sum_{j:c_{j,i}=1} w_{j,i} \frac{\partial \eta}{\partial x_j} & \text{for } i \le N - N_O \\ +\sigma_\delta'(Y_{(1)} - Y_{(2)}) & \text{for } i = (1) \\ -\sigma_\delta'(Y_{(1)} - Y_{(2)}) & \text{for } i = (2) \end{cases} \tag{21}$$

$$\frac{\partial\eta}{\partial x_i} = \frac{\partial\eta}{\partial y_i}\cdot F_i'(x_i) \tag{22}$$

Augmented backward pass for diagonal-Hessian OBD:

$$\frac{\partial^2\eta}{\partial y_i^2} = \begin{cases} \displaystyle\sum_{j:c_{j,i}=1} w_{j,i}^2\left(\frac{\partial\eta}{\partial y_j}\cdot F_j''(x_j) + \frac{\partial^2\eta}{\partial y_j^2}\cdot F_j'(x_j)^2\right), & 1 < i \le N - N_O \\ \sigma_\delta''(Y_{(1)} - Y_{(2)}), & N - N_O < i \le N \end{cases} \tag{23}$$

Augmented backward pass for full-Hessian OBD:

$$R_{\Delta X;X}\left\{\frac{\partial\eta}{\partial y_i}\right\} = \begin{cases} \displaystyle\sum_{j=N-N_O+1}^{N} \frac{\partial}{\partial y_j}\left(\frac{\partial\eta}{\partial y_i}\right)\cdot R_{\Delta X;X}\{y_j\}, & N - N_O < i \le N \\ \displaystyle\sum_{j:c_{j,i}=1} w_{j,i}\cdot R_{\Delta X;X}\left\{\frac{\partial\eta}{\partial x_j}\right\}, & 1 \le i \le N - N_O \end{cases} \tag{24}$$

$$R_{\Delta X;X}\left\{\frac{\partial\eta}{\partial x_i}\right\} = F_i'(x_i)\cdot R_{\Delta X;X}\left\{\frac{\partial\eta}{\partial y_i}\right\} + \frac{\partial\eta}{\partial y_i}\cdot F_i''(x_i)\cdot R_{\Delta X;X}\{x_i\} \tag{25}$$

Distribution:

M. Belongie	238-420
F. Pollara	238-420
S. Townes	238-420
N. Xiao	238-420
D. Watola	82-104