INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **22**: 1687–1708 (2002)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/joc.811

A GRAPHICAL SENSITIVITY ANALYSIS FOR STATISTICAL CLIMATE MODELS: APPLICATION TO INDIAN MONSOON RAINFALL PREDICTION BY ARTIFICIAL NEURAL NETWORKS AND MULTIPLE LINEAR REGRESSION MODELS

ALEX J. CANNON[a,*] and IAN G. McKENDRY[b]
[a] *Meteorological Service of Canada–Pacific and Yukon Region, 700-1200 West 73rd Avenue, Vancouver, B.C. V6P 6H9, Canada*
[b] *Department of Geography, University of British Columbia, 217-1984 West Mall, Vancouver, B.C. V6T 1Z2, Canada*

Received 3 October 2001
Revised 20 May 2002
Accepted 23 May 2002

ABSTRACT

A form of sensitivity analysis is described that illustrates the effects that inputs have on outputs of statistical models. The strength and sign of relationships, the types of nonlinearity, and the presence of interactions between inputs can be diagnosed using this technique. Intended for interpreting flexible nonlinear models, the graphical sensitivity analysis is applied to artificial neural networks (ANNs) in this study. As ANNs are increasingly being used for climate prediction, the discussion focuses on specific problems associated with their use in this context. The technique is illustrated using a real-world, long-range climate prediction example. Principal components (PCs) of circulation fields prior to the Indian summer monsoon are related to rainfall during monsoon months for the 1958–98 period. The skill of multiple linear regression and ensemble ANNs are compared using a resampling procedure. Interpretation of the models is then conducted using traditional diagnostic tools and graphical sensitivity analysis. This provides an improved investigation of precursor circulation field–summer monsoon rainfall relationships identified in a previous modelling study. The relatively stable, linear relationship identified between the May 200 hPa geopotential height field and summer monsoon rainfall is confirmed. Correlations previously identified between 850 hPa geopotential heights during January and rainfall by ANNs are shown to be the result of a weakly nonlinear, interactive relationship involving the first and second PCs of this field. An analysis of out-of-sample model predictions suggests that this relationship does not persist over the entire study period. This may result from a modulation of the strength of the circulation–rainfall relationship by El Niño–southern oscillation. Stratification of the results also reveals a relatively strong, nearly linear relationship with monsoon strength during years exhibiting positive scores of the second PC. On extending the analysis to longer lead-times, the surface pressure and 850 hPa geopotential height fields during November show relatively strong, persistent precursor relationships with summer monsoon rainfall. Sensitivity analyses suggest a mildly nonlinear relationship that is common to both fields.

KEY WORDS: sensitivity analysis; artificial neural network; nonlinear model; seasonal forecasting; India

1. INTRODUCTION

The climate system is extremely complex. Until recently, statistical relationships between climate variables were modelled using relatively simple linear methods (e.g. multiple linear regression (MLR) models or canonical correlation analysis) (Hsieh and Tang, 1998). Though linear models are fully specified and their parameters have simple meanings (von Storch and Zwiers, 1999), they are poorly suited to modelling the complex, often nonlinear, relationships that exist between climate variables.

* Correspondence to: Alex J. Cannon, Meteorological Service of Canada–Pacific and Yukon Region, 700-1200 West 73rd Avenue, Vancouver, BC V6P 6H9, Canada; e-mail: alex.cannon@ec.gc.ca

Recently, flexible models capable of representing nonlinear relationships have started being used in climatological research; however, unlike linear methods, many of these models can be difficult to interpret. Connectionist models, such as artificial neural networks (ANNs), lack widely accepted, straightforward methods for describing the functions they represent. Certain techniques, like generalized additive models (Hastie and Tibshirani, 1990), multivariate adaptive regression splines (Friedman, 1991), and recursive partitioning trees (Breiman *et al*., 1984), do contain built-in methods for interpreting relationships between inputs and outputs. Comparing and contrasting different models can, however, be a challenging task without a common method of interpretation. Also, with the adoption of ensemble averaging as a means of improving the stability of flexible nonlinear models (Hsieh and Tang, 1998), even methods that are usually simple to interpret become difficult to decipher.

This perceived lack of interpretability has slowed the adoption of nonlinear methods in certain areas of climate analysis. In many cases, ANNs and other flexible statistical models applied to climate prediction tasks have been thought of as little more than 'black boxes'. Models are frequently deemed acceptable for forecast use but are rarely used to analyse climate relationships (Hewitson and Crane, 1994). For example, Cannon and McKendry (1999) used ANNs to relate circulation conditions in southeast Asia to Indian summer monsoon rainfall, but then used linear methods to investigate the circulation–rainfall relationships. Kumar *et al*. (1995) pointed out that most nonlinear models applied to monsoon forecasting rely on linear methods to first identify predictors. As a result, nonlinear predictive relationships may be missed and the full capabilities of flexible nonlinear models may not be realized.

The current study describes a method for visualizing the input–output mappings of statistical models. Unlike other techniques, graphical sensitivity analysis is not specific to any one statistical prediction model. The method is conceptually simple, can be used to diagnose complex model relationships, and provides results in both graphical and numerical formats. The usefulness of the technique is demonstrated by analysing results from a real-world statistical climate forecasting problem, specifically the seasonal prediction of Indian monsoon rainfall using ANNs and MLR models. This provides an improved analysis of circulation–rainfall relationships identified by Cannon and McKendry (1999). To highlight the suitability of the method for use in statistical climate modelling, the study addresses complicating factors such as pre-processing using principal component analysis (PCA), model validation through resampling, and the ensemble averaging of models.

2. METHOD

2.1. Multi-layer perceptron ANN

ANNs form a class of models loosely based on the biological nervous system. Clustering, classification, and regression tasks can be performed using different varieties of ANN (Sarle, 1994). Of these varieties, the multi-layer perceptron (MLP) ANN is commonly used for nonlinear classification and regression. Practical guidelines for their use are given by Smith (1993) and Reed and Marks (1999). Reviews of MLP ANN applications in meteorology and climatology have been conducted by Gardner and Dorling (1998) and Hsieh and Tang (1998).

MLP ANN models in the current study related X_i input variables to a single output Y using a single hidden layer of j hyperbolic tangent processing nodes. The model architecture is described by the equation

$$Y = \sum_j \left[\tanh\left(\sum_i X_i \, {}^1w_{ij} + {}^1b_j \right) \right] {}^2w_j + {}^2b \tag{1}$$

Model parameters in the MLP ANN are the hidden-layer weights ${}^1w_{ij}$, the hidden-layer biases ${}^1b_j$, the output-layer weights ${}^2w_j$, and the output layer bias ${}^2b$. Given this architecture, the MLP ANN can account for nonlinear relationships and interactions between input variables. The exact form of these relationships need not be specified prior to modelling. With a sufficient number of hidden-layer nodes, the MLP ANN model is capable of approximating any continuous function to an arbitrary level of accuracy (Hornik *et al*., 1989).

2.2. Sensitivity analysis

Sensitivity analyses assess the effects that changes in model inputs have on the model output. For example, in the equation defining a simple MLR model

$$Y = \sum_i X_i a_i + b \tag{2}$$

the slope parameters a_i can be directly interpreted as sensitivities. A single unit change in input X_i contributes an a_i unit change in the output Y, assuming that all other inputs are fixed. However, in most flexible nonlinear models the parameters cannot be interpreted in this way. Instead, perturbation-based sensitivity analyses have been proposed as an indirect means of analysing relationships between model inputs and outputs.

Hewitson and Crane (1994) used a perturbation-based sensitivity analysis to interpret an ANN MLP model relating principal components (PCs) of synoptic-scale circulation conditions to local precipitation in southern Mexico. To identify the relative importance of circulation variables at the surface and at 500 hPa to variability in precipitation, they measured the sensitivity of the model output to small perturbations in PC scores on each day in the record. Vector lengths of sensitivities for the PCs were used as measures of the importance of each atmospheric level on a given day. To avoid small or zero lengths resulting from sensitivities of opposite sign, vector lengths were computed using absolute values of the individual sensitivities. As this approach only considers the sensitivity of the network to small changes in input, irrespective of the magnitude of the output variable, the same sensitivity value can have very different meanings depending on the value of the output. Hewitson and Crane (1994) recognized this problem and scaled the sensitivities by a logarithmic function of precipitation, thereby giving more weight to sensitivities associated with low precipitation amounts. Time series plots of sensitivities were then constructed to gain insight into the relative importance of SLP and 500 hPa conditions during different seasons. Spatial sensitivity maps were also plotted, showing the influence of different regions during different periods of time. Novel in this approach was the mapping of the PC sensitivities back to the original grid-point locations.

Tangang *et al*. (1998) applied a similar form of sensitivity analysis to ANN MLP models used to predict sea-surface temperature (SST) anomalies in the equatorial Pacific Ocean. They assessed the sensitivity of the model output to removal or pruning of input variables. Unlike the analysis performed by Hewitson and Crane (1994), this method considers the net importance of a given input, not sensitivities of individual cases. Root-mean-squared errors and correlation coefficients between outputs from the original network and the pruned networks were computed and used as indicators of input variable importance. Based on these measures of importance, an improved model was constructed containing four of the eight original input variables.

While capable of estimating the influence of different input variables on the output, these methods do not provide any insight into the types of relationship present between inputs and outputs. Trends, nonlinearities, and interactions between variables cannot easily be diagnosed. Recently, Plate *et al*. (2000) proposed a modified form of sensitivity analysis for visualizing and interpreting input–output mappings from ANN models. Combining aspects of the sensitivity analyses used by Hewitson and Crane (1994) and Tangang *et al*. (1998), the importance of inputs, nonlinearity of modelled relationships, interactions between inputs, and trends in the effects of inputs can be readily identified using this method. This form of sensitivity analysis was used in the current study to help interpret predictive relationships between atmospheric circulation fields and Indian monsoon rainfall. Further details on the modelling study are given in subsequent sections. Because this form of analysis has not been widely used in the atmospheric sciences, a relatively detailed description is given here. Following the description offered by Plate *et al*. (2000), the method is also compared with those introduced by Hewitson and Crane (1994) and Tangang *et al*. (1998). In addition, details relevant for use in climatological applications are considered.

In graphical sensitivity analysis, the effects of changing model inputs X_i from some arbitrary baseline values b_i to their original values are calculated and plotted for a set of cases. Output effects Δ_i are defined as

$$\Delta_i = Y(X) - Y(X_1, \ldots, X_{i-1}, b_i, X_{i+1}, \ldots, X_k) \tag{3}$$

where k is the index of the last input variable. For standardized X_i, values of b_i are typically set to zero, the mean of each input variable. By definition Δ_i equals zero when X_i equals the baseline value. As in the study conducted by Tangang *et al.* (1998), the effect of removing input variables from the network forms the basis for the sensitivity analysis described here. However, unlike their method, values of the effects are calculated and retained for individual cases. This allows the sensitivities to be plotted graphically (i.e. a scatter plot of Δ_i versus X_i) thereby giving more insight into the nature of the modelled input–output relationships.

For complicated relationships involving both nonlinearities and interactions between inputs, simple scatter plots of the effects can be hard to interpret. Effects are instead plotted as short line segments, with slopes of the segments given by partial derivatives of the output with respect to X_i. By plotting the effects as line segments instead of points, the trends and the types of nonlinear relationships that are present are easier to identify. Partial derivatives used to define the line segment slopes can be calculated analytically (Egmont-Petersen *et al.*, 1994) or they can be approximated using finite differences (Bishop, 1995). In the sensitivity analysis conducted by Hewitson and Crane (1994), partial derivatives of ANN outputs were approximated using finite differences. The same approach was taken in the current study. For climatological applications, partial derivatives can be used to construct time-series plots and maps of sensitivities or sensitivity vector lengths in the manner described by Hewitson and Crane (1994). However, more detailed investigations of modelled relationships can be performed by analysing the partial derivatives in conjunction with values of Δ_i.

By combining the sensitivities, slopes, and input values on one graph, the effects plot, insight into the nonlinearity of modelled relationships, trends in the effects of input variables, the importance of variables, and the presence of interactions between variables can be gained. Trends and nonlinearity in effects plots relate directly to trends and nonlinearity in effects of a particular variable on model output. The overall vertical range indicates the variable's importance and the vertical spread at points along the abscissa indicates interactions between the plotted input and at least one other input variable. Variables with no effect on the output appear as horizontal lines on the sensitivity plots. Additive variables plot as single lines or curves. To illustrate these characteristics, effects plots for a simple, synthetic function of six variables

$$Y = 5\sin(10X_1X_2) + 20(X_3 - 0.5)^2 - 10X_4 + 20X_5X_6 \tag{4}$$

are given in Figure 1. As described above, graphs are constructed by plotting values of variable X_i along the abscissa and variable effects Δ_i along the ordinate. A synthetic example was chosen to allow clear links to be made between the relationships described by the mathematical function and the features of the sensitivity analysis plots. The method is applied to a real-world example in subsequent sections.

For this particular function, variables X_1 and X_2 interact with one another and have a nonlinear effect on Y. X_3 is not involved in any interactions, instead having an additive nonlinear effect on Y. X_4 is also additive, and has a negative linear effect on Y. Variables X_5 and X_6 exhibit a simple multiplicative relationship. These relationships can be seen in the corresponding effects plots for each variable. Vertical spread in the effects for X_1 and X_2 indicates the presence of an interaction. The sinusoidal form of the nonlinearity for these variables can be seen by following the trace of the line segment slopes in their effects plots. Similarly, the parabolic relationship is present in the effects plot for X_3; absence of vertical spread suggests an additive contribution for this variable. The slopes of effects for X_4 are negative. Increases in this variable are therefore associated with decreases in Y. As the effects lie along a straight line, the relationship between X_4 and Y is linear and additive. Vertical spread in effects for X_5 and X_6 suggests the presence of interactions. The linear trace of the slopes indicates a simple multiplicative interaction. In terms of variable importance, roughly measured by the vertical range in effects, variables X_1 and X_2 appear most important, followed by X_5 and X_6, then X_4, and finally X_3. Interaction strength is more difficult to diagnose graphically, but, judging from the degree of vertical spread, X_1 and X_2 appear less additive than X_5 and X_6.

Although variable importance and interaction strength can be estimated from the effects plot, Plate *et al.* (2000) also introduced objective numerical measures for these quantities. Variable importance is measured by computing the variance of the effects for X_i. This statistic is equivalent to the root-mean-squared error measure adopted by Tangang *et al.* (1998). Variable interaction strength, or degree of non-additivity, is

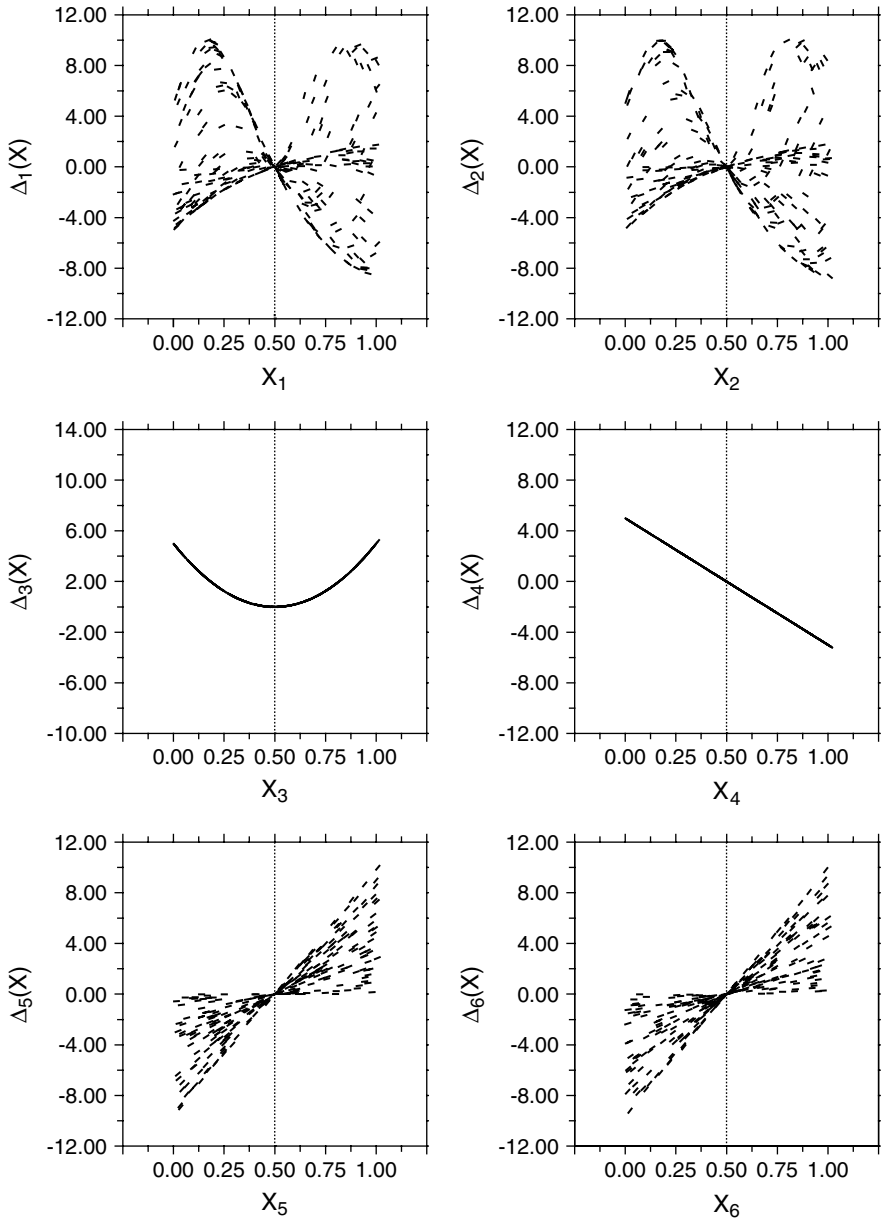

Figure 1. Example effects plots for inputs to the function in Equation (4)

measured by fitting a smooth curve to the scatter plot of the partial derivatives and then calculating the deviation of points from this curve. Partial derivative scatter plots for Equation (4) are given in Figure 2. Partial derivatives for additive variables, such as X_3 and X_4, fall along single lines or curves. As a result, deviations of points from the fitted curve are generally small. Partial derivatives for variables involved in interactions, such as X_1, X_2, X_5, and X_6, plot as clouds of points. In these cases, a fitted curve provides a poor approximation and deviations of points from the curve are generally large. Averaged over all cases, a small deviation, therefore, indicates a near additive relationship and a large deviation indicates a non-additive relationship. Curves can be drawn using any of a number of smoothing techniques (Cleveland and Loader, 1996). Locally weighted linear regression was chosen for this particular example. Figure 3 shows

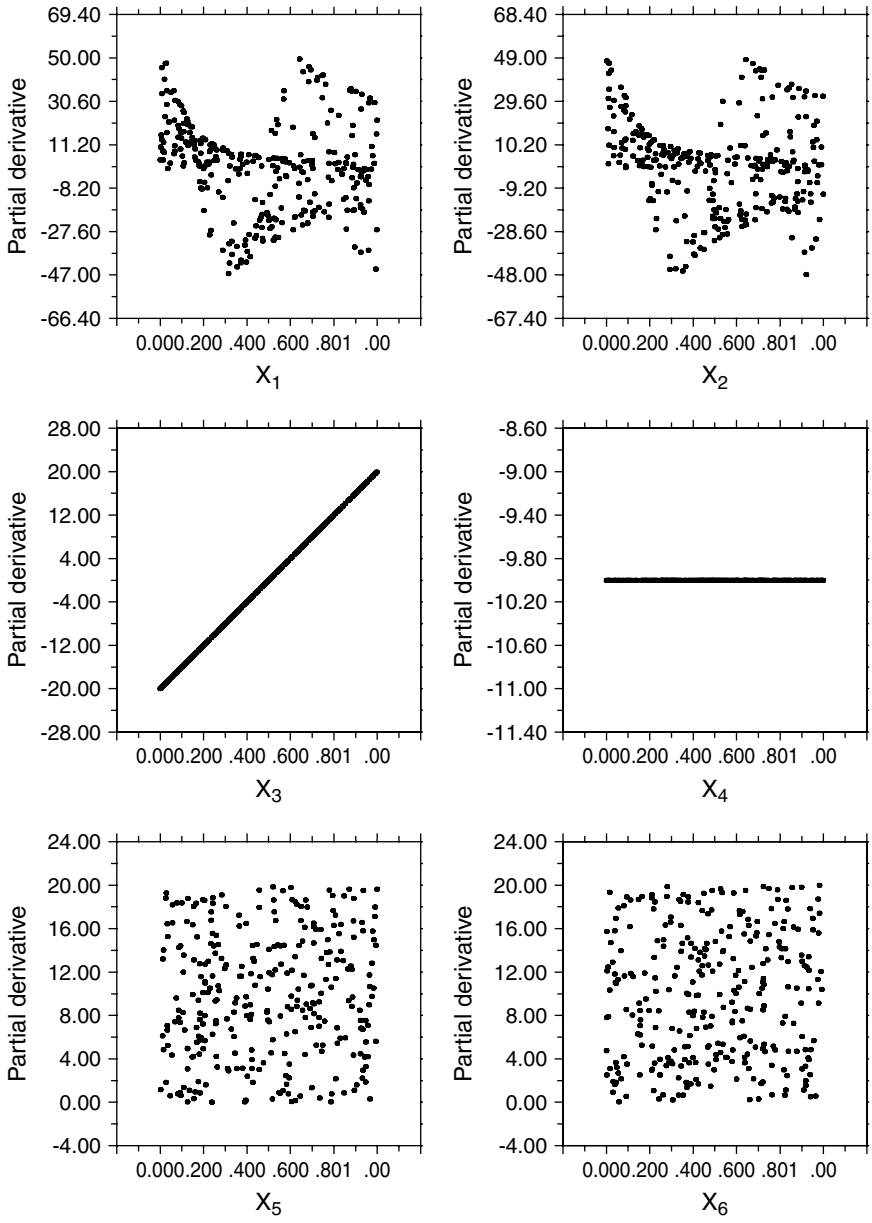

Figure 2. Example partial derivative plots for inputs to Equation (4)

plots of variable importance versus interaction strength for variables X_1 to X_6. The objective measures are consistent with impressions gained through visual inspection of the effects plots and partial derivative plots.

For the example given in Figure 1, determining which variables were involved in specific interactions was relatively straightforward. In practice, relationships may be more complex and may involve multiple variables. In these situations, simple effects plots are not sufficient for identifying which variables are interacting with one another. Instead, stratified effects plots may be used to help identify specific interactions. In stratified effects plots, line segments for a given input variable are coloured according to values of another input variable. If the two inputs are involved in an interaction, colours should appear in distinct bands in the effects

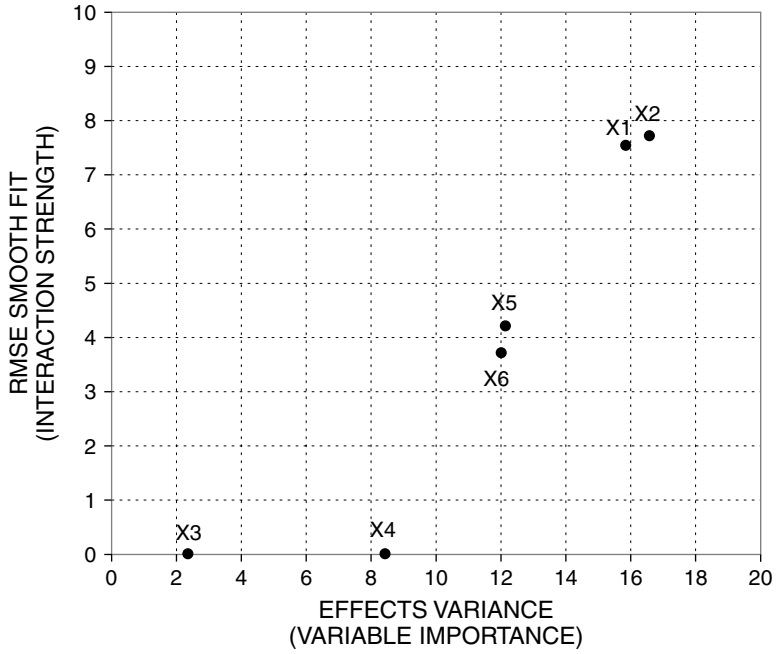

Figure 3. Scatter plot showing quantitative measures of variable importance (abscissa) and variable interaction strength (ordinate) for inputs to Equation (4)

plot. If the two inputs are not involved in an interaction, colours should be distributed randomly across the plot. Examples of stratified effects plots for variables X_1 and X_2 and X_5 and X_6 are given in Figure 4. For X_1 and X_2 the colours show the ranges of variables X_2 and X_3; for X_5 and X_6 the colours show the ranges of X_4 and X_5. The plot indicates the presence of an interaction between X_1 and X_2 and between X_5 and X_6, but not between X_2 and X_3 or between X_4 and X_5.

In situations where split-sample or resampling methods, such as cross-validation or bootstrapping, are used to estimate model performance, modifications to the basic sensitivity analysis procedure described by Plate *et al*. (2000) are required. Owing to the random selection of training and testing cases, variations in sensitivity plots will be present between samples. Rather than choose a single model from one sample to represent the input–output relationships, effects plots are instead modified to include information from all available samples. Conceptually, this is similar to the bootstrap procedure developed by Baxt and White (1995) to assess ANN sensitivities. Median values of effects and partial derivatives are plotted to represent average conditions over the range of models. Lower and upper quartile values are also shown on the same graph to indicate the variability in these quantities across the samples. This modification was used successfully by Cannon and Whitfield (2001) to visualize relationships in models relating precipitation and antecedent climate conditions to water quality. Examples of this approach are presented in Section 4.

Modifications to the sensitivity analysis procedure are also required when PCA is used to reduce the dimension of gridded datasets. In this case, interpreting sensitivities in terms of the original grid-point locations can be difficult. Though PC loadings can help in the spatial interpretation of the effects plots, the presence of nonlinear and interactive effects make this difficult. As an alternative, PC sensitivities can be mapped back to the original grid points (Hewitson and Crane, 1994). Using this approach, the model is expanded to include the projection of the grid-point data onto the PCs. Instead of perturbing the PCs to calculate the effects, the grid-point data are perturbed, the perturbed data are then projected onto the PCs, and the resulting scores are then entered into the model. Examples of spatial effects plots are presented in later sections.

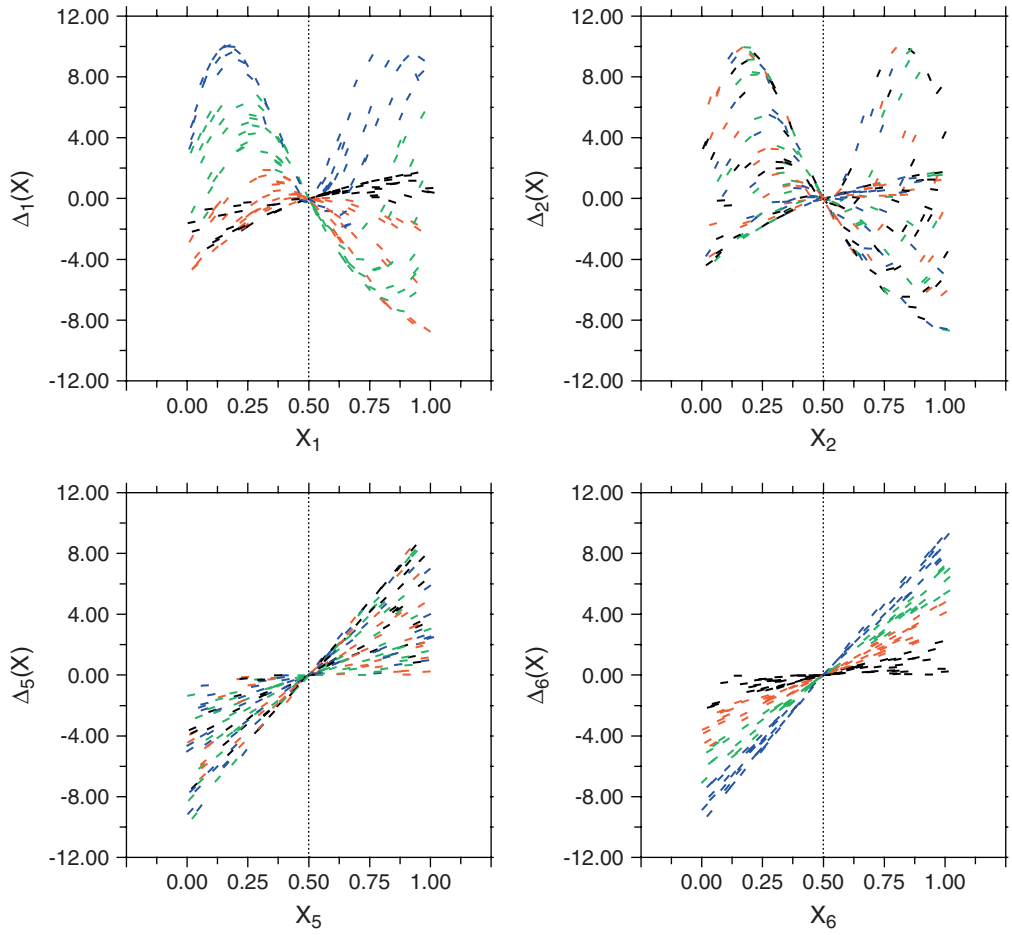

Figure 4. Stratified effects plots for inputs X_1 (top left), X_2 (top right), X_5 (bottom left), X_6 (bottom right) to Equation (4). Colours show the ranges of the variables X_2 (top left), X_3 (top right), X_4 (bottom left), and X_5 (bottom right): black, ≤ 0.25; red, >0.25 and ≤ 0.5; green, >0.5 and ≤ 0.75; blue, >0.75

3. DATA

3.1. Sea-level pressure and geopotential height

To illustrate the sensitivity analysis procedure, relationships between Indian monsoon rainfall and gridded circulation data from four atmospheric levels over the South Asian subcontinent were investigated using MLR and MLP ANN models. Data and models are similar to those considered by Cannon and McKendry (1999).

For use as inputs to models, daily averaged SLP, 850 hPa, 500 hPa, and 200 hPa geopotential height data (2.5 by 2.5 resolution) were obtained from National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis model output (Kalnay *et al.*, 1996). A subset covering the region from 62.5 to 95 E and 7.5 to 35 N was extracted and averages for months preceding the summer monsoon (October–May) were computed for the 41 year period from 1958 to the end of 1998. This extends the analysis conducted by Cannon and McKendry (1999), which only considered the period 1958–94 and months from December–May. Gridded data were averaged in space to a resolution of 5 by 5 . Standardized SLP and geopotential height anomalies were then computed by subtracting the climatological monthly mean from each of the monthly average values and dividing by the monthly standard deviation.

The eight monthly series of standardized anomaly fields (36 grid points) were used as indicators of pre-monsoon circulation at each atmospheric level. Because grid-point values during the study period were

Table I. Cumulative percent explained variance for PCs of gridded circulation data. *N* indicates the number of PCs retained for each atmospheric level

Level	*N*	October	November	December	January	February	March	April	May
SLP	3	93.0	88.5	91.9	92.4	93.0	90.3	89.0	88.7
850 hPa	3	95.5	93.6	95.6	94.9	93.4	93.7	93.4	93.5
500 hPa	4	96.5	95.6	97.5	97.2	97.2	97.3	96.5	96.4
200 hPa	4	97.7	97.2	98.1	97.5	97.7	97.1	96.9	97.4

highly correlated in space, PCA was used to reduce the dimensionality of the original dataset and provide uncorrelated representations of the circulation data. For each series and level, an *S*-mode PCA (Green, 1978) was performed on the correlation matrix of the pre-monsoon circulation anomalies. The number of PCs to retain as predictors was determined by taking the average of three truncation criteria: the rule-*N* test (Overland and Preisendorfer, 1982), the eigenvalue-one rule (Kaiser, 1959), and PCs accounting for 90% of the total variance in the original dataset. Table I shows the cumulative percent variance accounted for by the retained components at each level and month. In all cases the cumulative explained variance exceeded 88%.

Unlike Cannon and McKendry (1999), where PCs were not rotated, retained PCs in the current study were rotated orthogonally using Kaiser's varimax criterion (Kaiser, 1958). Enhanced meteorological and climatic interpretability of rotated PC solutions has been documented by Richman (1986) and Barnston and Livezey (1987). However, in the current study PCA was used mainly for data reduction and filtering purposes; interpretation of PCs as dominant modes of atmospheric variability was of less importance.

3.2. All-India summer monsoon rainfall

The all-India summer monsoon rainfall (AISMR) dataset consists of area-weighted averages of summer (June–September) rainfall totals from 306 district rain gauge stations. Approximately 90% of India is accounted for, with a number of mountainous regions removed from the analysis. Originally developed and described by Parthasarathy *et al.* (1987) for the period 1871–1984, the dataset was recently extended by Parthasarathy *et al.* (1994) to include years through 1993. Data from 1994–98 were obtained from the Indian Institute of Tropical Meteorology. To match the available circulation data record, AISMR totals for years from 1958 to 1998 were extracted from the series. For use in the forecast models, the mean value over the 1958–98 subset was subtracted from each summer's reported total. Anomalies were then standardized by dividing by the standard deviation for the subset data period.

4. RESULTS

4.1. Model predictions

The MLP ANN model architecture was used to relate regional circulation PCs to AISMR. All networks in the current study were trained using the resilient backpropagation algorithm (Riedmiller, 1994). Control parameters were set to suggested default values. Initial values for weights and biases were set to random values ranging from −0.5 to 0.5. To avoid nonlinear instability and overfitting of the training data, ensemble averaging and stopped training procedures were used to build the models (Hsieh and Tang, 1998). Ensemble models in the current study each contained 20 networks.

The bootstrap aggregation (bagging) ensemble averaging procedure (Breiman, 1996) was used to reduce overfitting and improve performance of the ANNs developed in the current study. In bagging, a number of training datasets are generated by sampling with replacement from the available pool of data. Models are then trained on the resampled datasets and results from individual ensemble members are averaged to yield final output values. Ensemble averaging reduces the variance of the outputs due to the random initialization

of weights and biases and the random selection of training set cases. Each resampled training set is selected to be equal in size as the original dataset.

As resampling is done with replacement, approximately 37% of training cases are not included in each of the bagged sets. These 'out-of-bag' cases can be used to estimate the generalization error of the model, thereby allowing training to be stopped at a point prior to overfitting. This procedure is a simple method for ensuring good model performance on data not used in the training process (Finnoff *et al.*, 1993). Rather than choosing the weights and biases that maximize performance on the bagged training set, final weights and biases are chosen to maximize performance on the out-of-bag cases. During the optimization process, network weights and biases are stored after each presentation of the bagged training cases and an error value is calculated using these data. The out-of-bag cases are then presented to the model and another error value is calculated using the stored weights and biases. This provides an estimate of the model's performance on data not included in the training set. Following convergence on the training dataset, weights and biases corresponding to the minimum error value on the out-of-bag cases are retrieved and are used as final parameters in the trained model.

Given the small sample size available (41 years), a resampling procedure based on the repeated application of split-sample validation (Weiss and Kulikowski, 1991) was used to estimate the predictive skill of the models. Parameters were set using three-quarters of the available cases (30 years) and the resulting model was used to predict output values for the remaining cases (11 years). Predictive skill on test set cases was then evaluated using the Pearson product-moment correlation coefficient r. The split and test procedure was repeated 200 times, with each trial using different randomly selected training and testing sets. Empirical distributions of r on the 200 randomly selected test sets were used to estimate the significance of the circulation–AISMR relationships.

For comparison with the ensemble ANN MLP models, MLR models were also constructed using the same input and output data. Because models in each split-sample validation trial used the same training and testing sets, the significance of differences in model performance was also estimated. Split-sample validation results for the MLR and ANN MLP model forecasts are given in Table II. For each atmospheric level, median values of r and differences in r between models are listed, as are empirical estimates of probability values for the significance of r and differences in r calculated based on the split-sample validation trials. The level of significance was set at 0.10. As the primary goal of this study was the demonstration of the graphical sensitivity analysis, statistics are only shown for predictors with statistically significant circulation–AISMR relationships. The full set of results is available from the authors.

Consistent with the findings of Cannon and McKendry (1999), PCs of the 200 hPa geopotential height field in May exhibited the strongest correlations with summer monsoon rainfall. Median values of r over the trials equalled or exceeded 0.65 for MLR and MLP ANN models. For PCs of 850 hPa geopotential height, a maximum in skill was evident in January, although significant correlations were only noted for the MLP ANN models and not the MLR models. For the longer lead-times, only SLP and 850 hPa geopotential height PCs during November showed significant relationships with AISMR. Both MLR and MLP ANN models identified these relationships.

Split-sample validation results were evaluated using training and test sets selected from the entire 41 year period of record. As previous studies have suggested that the strength of predictor–AISMR relationships may

Table II. Split-sample validation results for MLP ANN and MLR model predictions of AISMR. Values of r are medians from the 200 split-sample validation trials. Probability values are empirical estimates from the split-sample validation trials

Level	Month	r MLP	r MLR	r MLP − r MLR
200 hPa	May	0.65 ($p < 0.005$)	0.67 ($p < 0.005$)	−0.01 ($p = 0.68$)
850 hPa	January	0.40 ($p = 0.10$)	0.12 ($p = 0.32$)	0.25 ($p = 0.09$)
850 hPa	November	0.37 ($p = 0.06$)	0.27 ($p = 0.13$)	0.11 ($p = 0.20$)
SLP	November	0.50 ($p = 0.05$)	0.39 ($p = 0.09$)	0.11 ($p = 0.23$)

change on decadal time scales (Parthasarathy *et al*., 1991; Hastenrath and Greischar, 1993), some means of evaluating stationarity of the circulation–AISMR relationships was also needed. In Cannon and McKendry (1999), multiple correlation coefficients between PCs and AISMR were calculated for 11 year and 15 year sliding windows. This method did not, however, accurately describe out-of-sample forecast performance and only measured the strength of linear relationships between circulation PCs and AISMR.

In the current study, split-sample validation was used to evaluate stationarity of both MLR and MLP ANN forecast skill between 1958 and 1998. Thirty years of data were used for model training and 11 years of data were used to test out-of-sample model performance. Training and test sets were not, however, selected randomly from the 41 year period of record. Test sets were instead taken from an 11 year sliding window and training data from the remaining 30 years. To gauge variability in MLP ANN results, ten ensembles were trained and evaluated for each 11 year window. Results are shown in Figure 5. Bars indicate correlation coefficients for MLR models and lines indicate the range in correlation coefficients for the ten ensemble MLP ANNs. The black line indicates the median correlation coefficient for the ten ensemble MLP ANN models. Results are plotted for the centre year in each 11 year window.

Of the relationships identified in Table II, only PCs of the May 200 hPa geopotential height field exhibited long-term periods of statistically significant, positive skill for both linear and nonlinear forecast models. Skill for both model types exceeded $r = 0.8$ through the middle of the 1970s. In the 1980s, however, the skill dropped to the 0.05 significance level. Values below the 0.10 significance level were noted for windows centred between 1989 and 1991. For PCs of the 850 hPa circulation field during January, the skill of the MLP ANN models was consistently higher than that of the linear models. Centred on the mid to late 1970s, ensemble MLP ANNs showed positive skill above the 0.10 significance level for five consecutive sliding window periods. Skill during the 1980s and 1990s declined to near zero. Positive skill was, however, evident

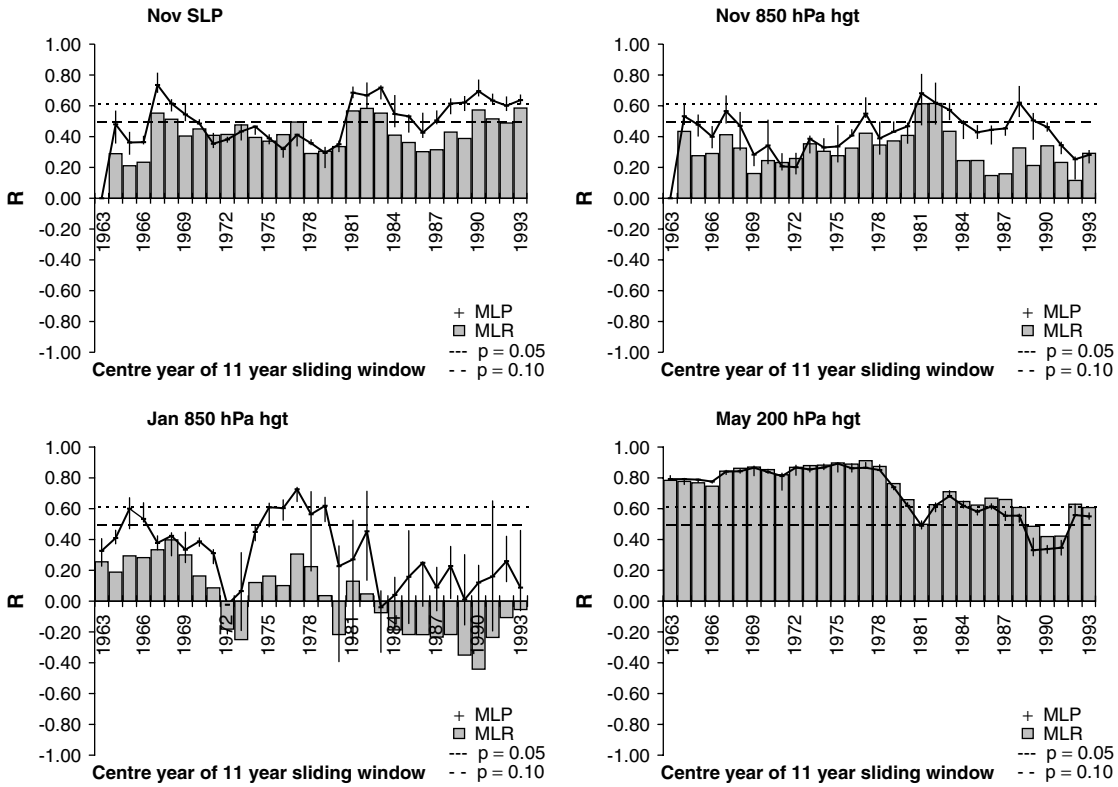

Figure 5. Sliding split-sample validation results for November SLP PCs, November 850 hPa geopotential height PCs, January 850 hPa geopotential height PCs, and May 200 hPa geopotential height PCs

in results for SLP PCs over this latter period of record, nearing the 0.10 significance level from 1982 onward. Ensemble MLP ANNs using November PCs of SLP and 850 hPa geopotential height fields as predictors exhibited positive skill levels over the entire period of record. For SLP predictors, models showed positive skill at the 0.05 significance level since the early 1980s, with a dip below the 0.10 significance level during the mid 1990s. Prior to 1981, skill was positive, but was not statistically significant. For 850 hPa geopotential height predictors, statistically significant skill levels were noted during the early 1980s. For both atmospheric levels, MLR models tended to perform less well than the corresponding MLP ANN models.

4.2. Statistical circulation–AISMR relationships

Statistically significant relationships between pre-monsoon circulation PCs and AISMR were present during the 1958–98 study period. In this section, the nature of these relationships are examined using traditional diagnostic plots and the graphical sensitivity plots described above. Comparisons are made for circulation fields listed in Table II, starting with the 200 hPa level in May. For completeness, relationships between each of the four retained PCs and AISMR are illustrated for this level and lead-time. However, for January and November the discussion focuses on PCs exhibiting the strongest relationships with AISMR. Unless noted, relationships between other PCs and AISMR were not significant during these months. Likewise, large differences between the linear and nonlinear models were not evident for other PCs at these two lead-times.

4.2.1. May. Rotated PC loadings for the 200 hPa geopotential height field in May are given in Figure 6. Contour values show the magnitude of linear correlations between PC scores and the standardized anomaly series at each grid point. Scatter plots between PC scores and AISMR, as well as correlation coefficients between the circulation PCs and AISMR, are also given, as are best-fit linear regression and locally weighted linear regression lines (Cleveland and Loader, 1996). Scores of PC3 and PC4 in May were strongly and significantly correlated ($r = 0.45$, $p = 0.003$ and $r = 0.47$, $p = 0.002$ respectively) with summer monsoon rainfall. Correlations between AISMR and PC1 and PC2 were not statistically significant. Best-fit linear and locally weighted regression lines for May 200 hPa PCs did not suggest the presence of nonlinear relationships between AISMR and these components, a finding consistent with model prediction results presented in Table II.

The rotated PC loadings and scatter plots shown in Figure 6 suggest that the two most important pre-monsoon circulation controls during the study period were (1) the magnitude and sign of 200 hPa geopotential height anomalies over Pakistan (PC3), and (2) the magnitude and sign of the 200 hPa geopotential height anomalies across west-central India (PC4). May months with positive (negative) height anomalies centred over Pakistan were generally associated with positive (negative) AISMR anomalies. May months with positive (negative) height anomalies over west-central India were associated with positive (negative) AISMR anomalies.

For comparison with loading plots and PC–AISMR scatter plots, the sensitivity analysis was applied to MLR and MLP ANN models for AISMR. Effects plots for PC predictors are given in Figure 7, with results from the MLP ANNs plotted next to those from the corresponding MLR models. In these plots, the central dark-coloured lines represent median values of the effects and partial derivatives over the 200 resampling trials; the peripheral light-coloured lines represent the upper and lower quartile values. Because the total number of PC predictors was relatively small, quantitative measures of variable importance and interaction strength were not calculated. To facilitate visual comparison, the ordinate of each plot was scaled to the same range for all levels and lead-times. For the May 200 hPa predictors, sensitivity effects plots for the linear and nonlinear models each showed the same basic trends and magnitudes of variable effects. Consistent with the PC loadings and scatter plots, PC1, PC3, and PC4 of the May 200 hPa geopotential height field showed positive linear trends, with PC2 exhibiting a negative linear trend. There was little evidence for interactions between input variables. Effects were strongest for PC3 and PC4, stronger than for any other PC predictor, agreeing with results of the linear correlation analysis. Median and upper and lower quartiles of effects and partial derivatives were very consistent, showing little variability across the 200 split-sample validation trials.

To illustrate the geographic distribution of model sensitivity better, effects plots were constructed using the grid-point perturbation technique described above. Spatial effects plots for the MLP ANN model are given in

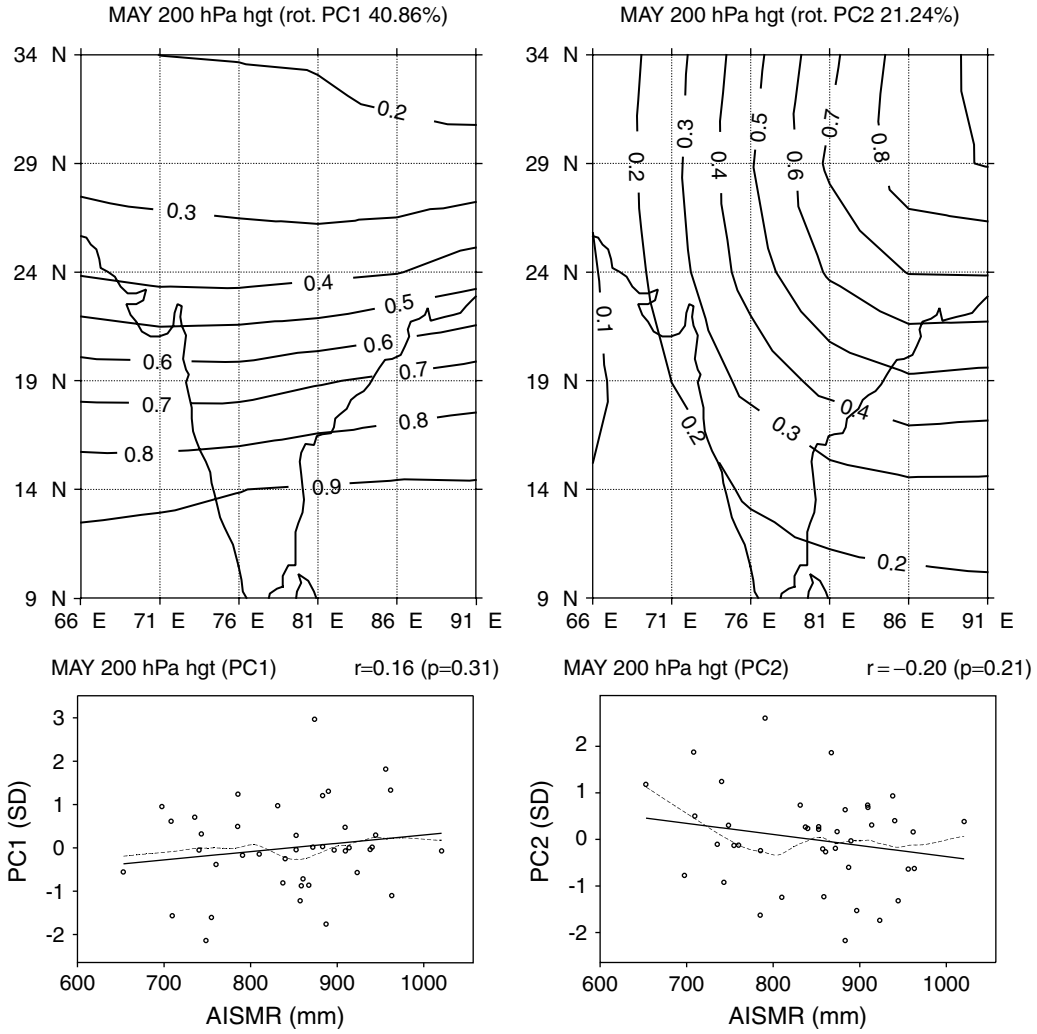

Figure 6. PC loadings for May 200 hPa geopotential height data (top). Scatter plot showing the relationship between May 200 hPa geopotential height PC scores and AISMR (bottom). The solid line shows the best-fit linear regression line. The dashed line shows the locally weighted linear regression curve

Figure 8. Strongest effects were concentrated in the northern and west-central portions of the study domain. The weak negative correlation between PC2 and AISMR, evident in Figure 6, is seen in the area of negative effects centred over the northeastern corner of the region. As expected from Figure 7, relationships were additive and linear at this level and lead-time.

4.2.2. January and November. Loadings for PC2 of the 850 hPa geopotential height anomaly field in January are given in Figure 9, along with scatter plots between scores of PC2 and AISMR. The correlation between PC2 and AISMR was relatively low ($r = -0.28$, $p = 0.08$). The highest PC loadings were present over the northern portion of the domain, decreasing southward toward peninsular India. As the results indicated significant differences between the skill of MLR and MLP ANN models, the true strength of the circulation–rainfall relationship at this level and lead-time may have been underestimated by the linear correlation. However, the scatter plot and best-fit lines between PC scores and AISMR did not seem to indicate the presence of a simple nonlinear relationship between PC2 and AISMR.

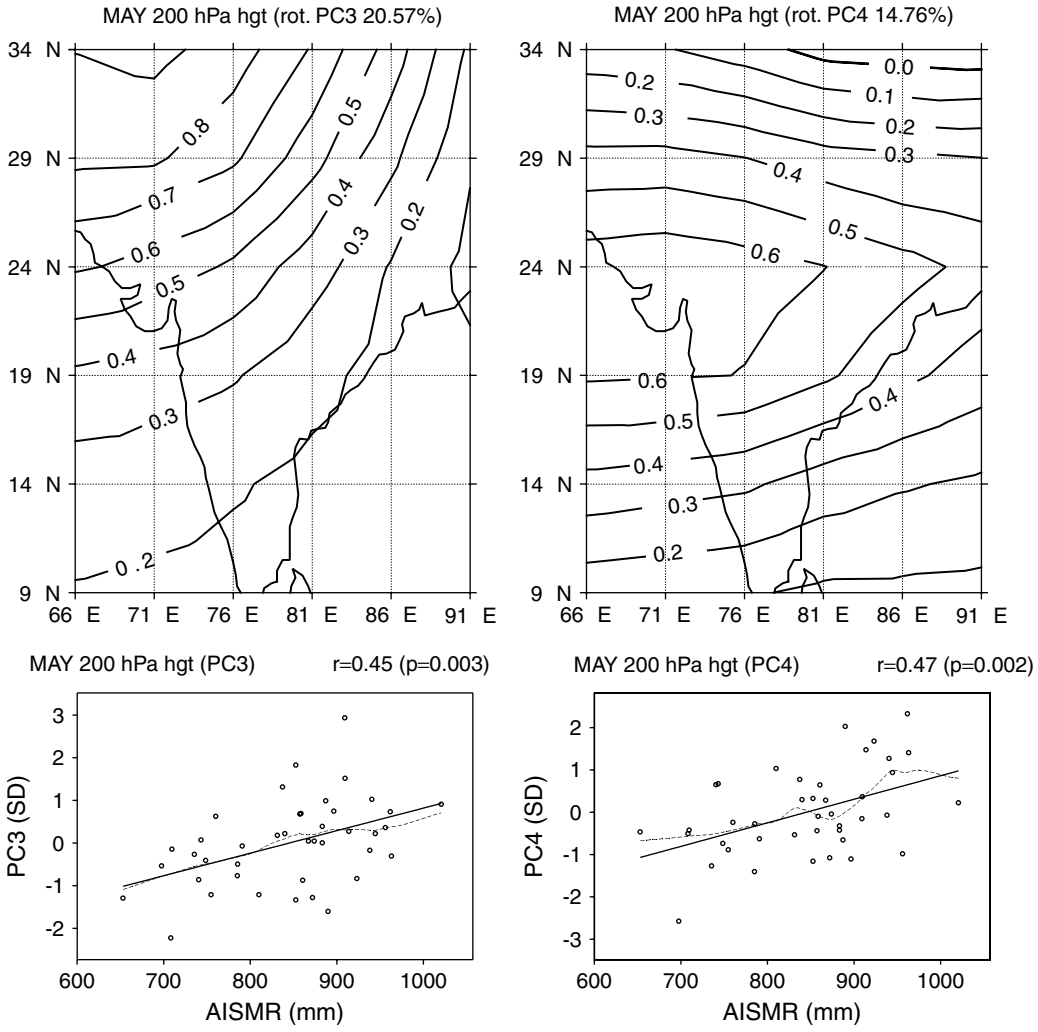

Figure 6. (*Continued*)

Effects plots for PC2 of the January 850 hPa field are given in the top panel of Figure 10. Unlike the results for the May predictors, distinct differences between the MLR and MLP ANN models were present at this level and lead-time. In the MLR model, a large negative effect was noted for PC2. In the MLP ANN model, the relationship between PC2 and AISMR was nonlinear. Perturbing negative scores of PC2 had almost no effect on the model, whereas perturbing positive scores resulted in a significant negative trend in model output. Line-segment slopes suggest that the local sensitivity of the model was greatest when positive values of PC2 were low to moderate (0.5–1 standard deviation unit) and that local sensitivity decreased with higher values of PC2 (>1 standard deviation unit). The nonlinearity in Figure 10 was verified following stratification of the PC scores and recalculation of the correlation coefficients. When scores were greater than or equal to zero, $r = -0.63$ ($p = 0.001$, $N = 24$); when scores were less than zero, $r = -0.19$ ($p = 0.47$, $N = 17$).

The presence of an interaction between PC1 and PC2 was also suggested in the effects plots for the MLP ANN model. Stratified effects plots for these two variables are shown in Figure 11. Unlike previous plots, which showed the effects for observed data points, line segments in this case were plotted at 200 random points spanning the input space. This allowed sparsely populated areas of the input space to be visualized. A multiplicative relationship was present between PC1 and PC2. Negative trends in effects of PC2 were strongest when values of PC1 were high, and weaker when values of PC1 were low. Conversely, effects of

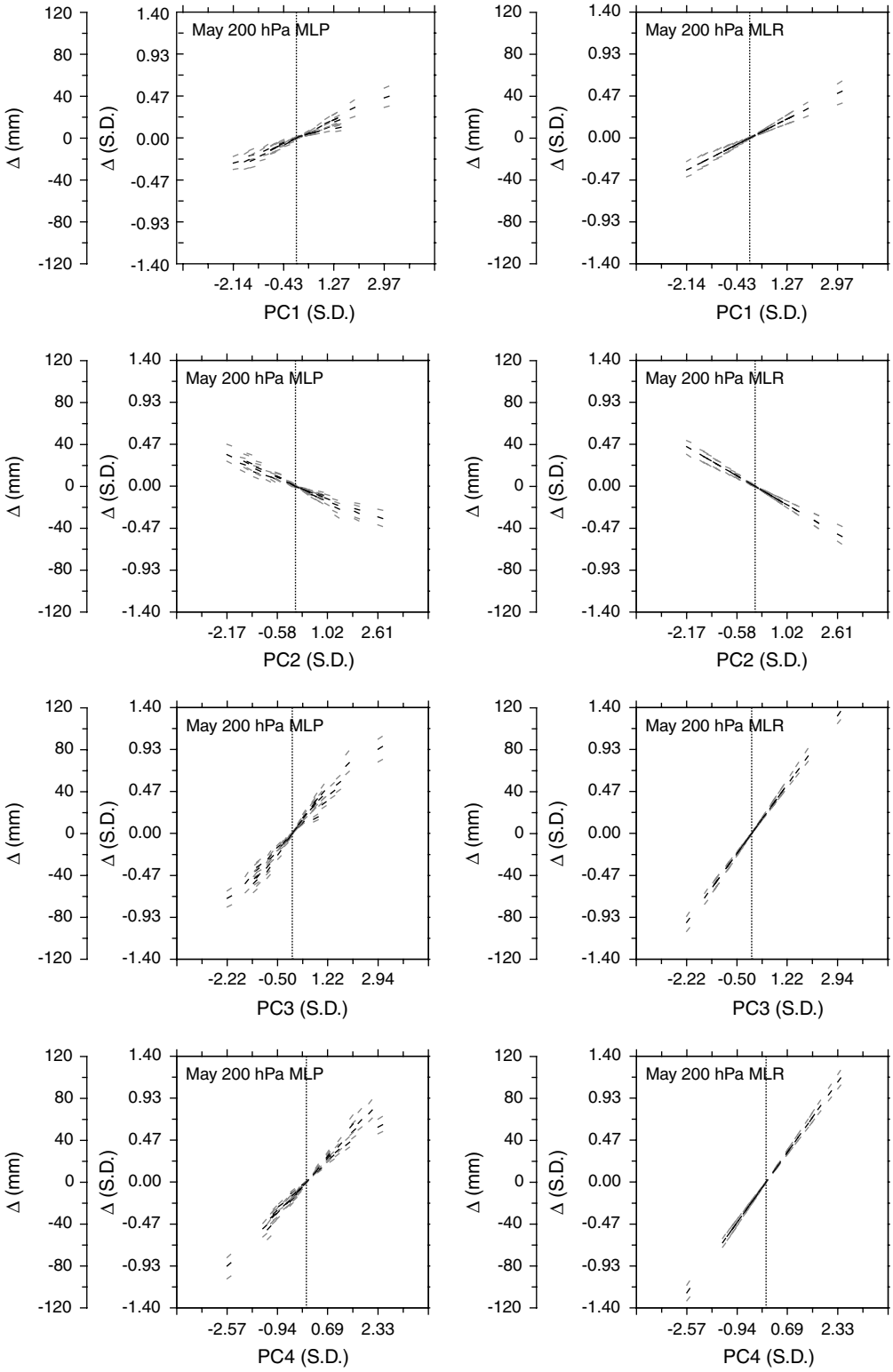

Figure 7. Effects plots for May 200 hPa geopotential height PC predictors to MLP ANN and MLR models

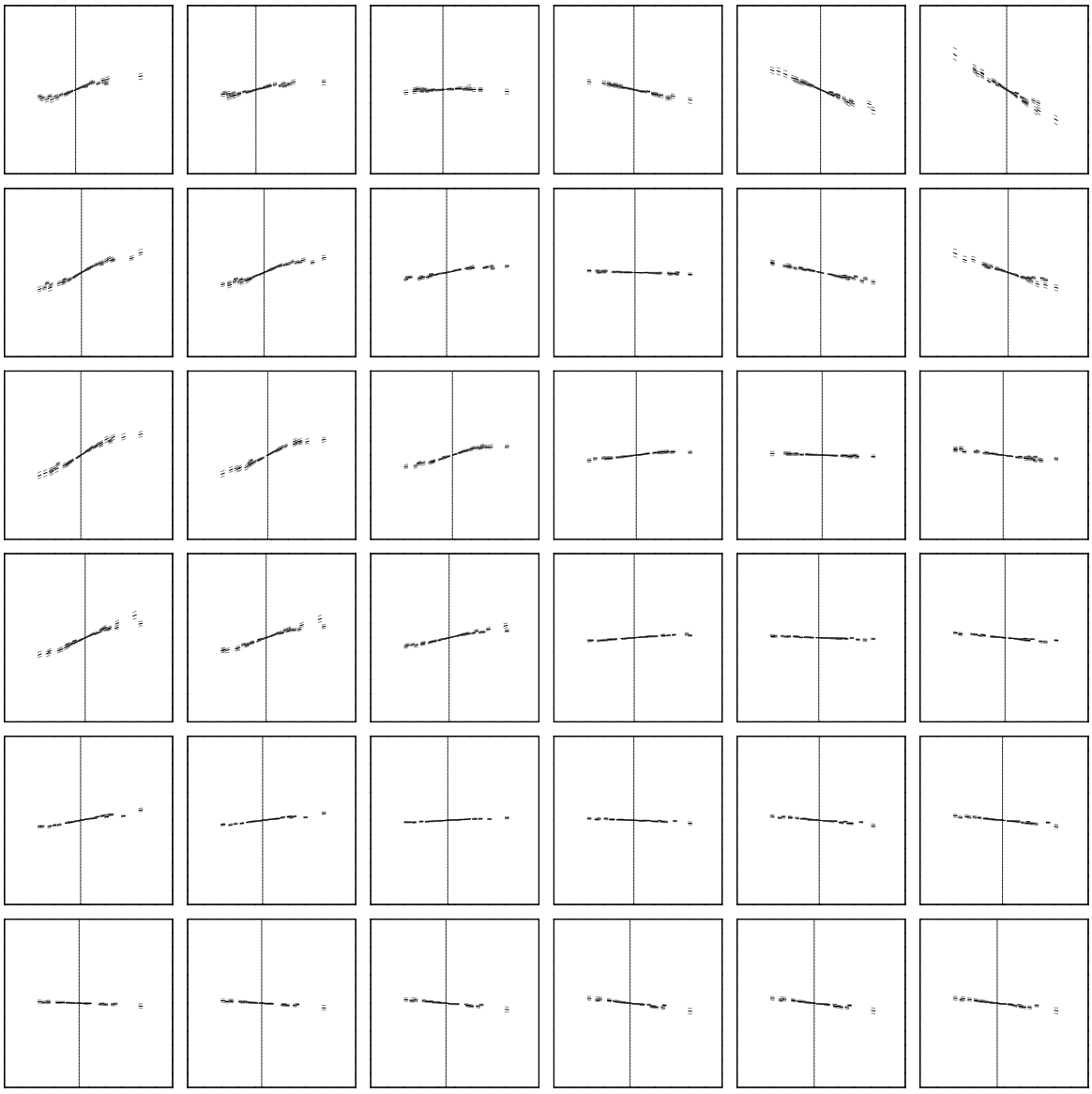

Figure 8. Effects plots for May 200 hPa grid-point predictors in the MLP ANN models. The top left plot corresponds to the grid point centred on 34 N and 66 E. The bottom right plot corresponds to the grid point centred on 9 N and 91 E. The vertical scale on each plot ranges from −60 to +60 mm or −0.7 to +0.7 standard deviation units

PC1 were positive when values of PC2 were lowest, and near zero or negative when values of PC2 were moderate to high.

Relationships between AISMR and 850 hPa geopotential height and SLP anomalies during November are shown in the bottom two panels of Figure 9. For both fields, significant relationships were observed between PC3 and AISMR ($r = -0.40$, $p = 0.01$ for 850 hPa geopotential height and $r = -0.46$, $p = 0.003$ for SLP). Loadings were similar for each field, exhibiting maximum correlations over the Arabian Sea in the southwest portion of the domain. For 850 hPa geopotential heights in January, the skill of MLP ANNs using November circulation PCs as predictors exceeded the skill of the corresponding linear models. Though not statistically significant, differences were notable. Again, inspection of the scatter plots and best-fit lines did not suggest the presence of simple forms of nonlinearity.

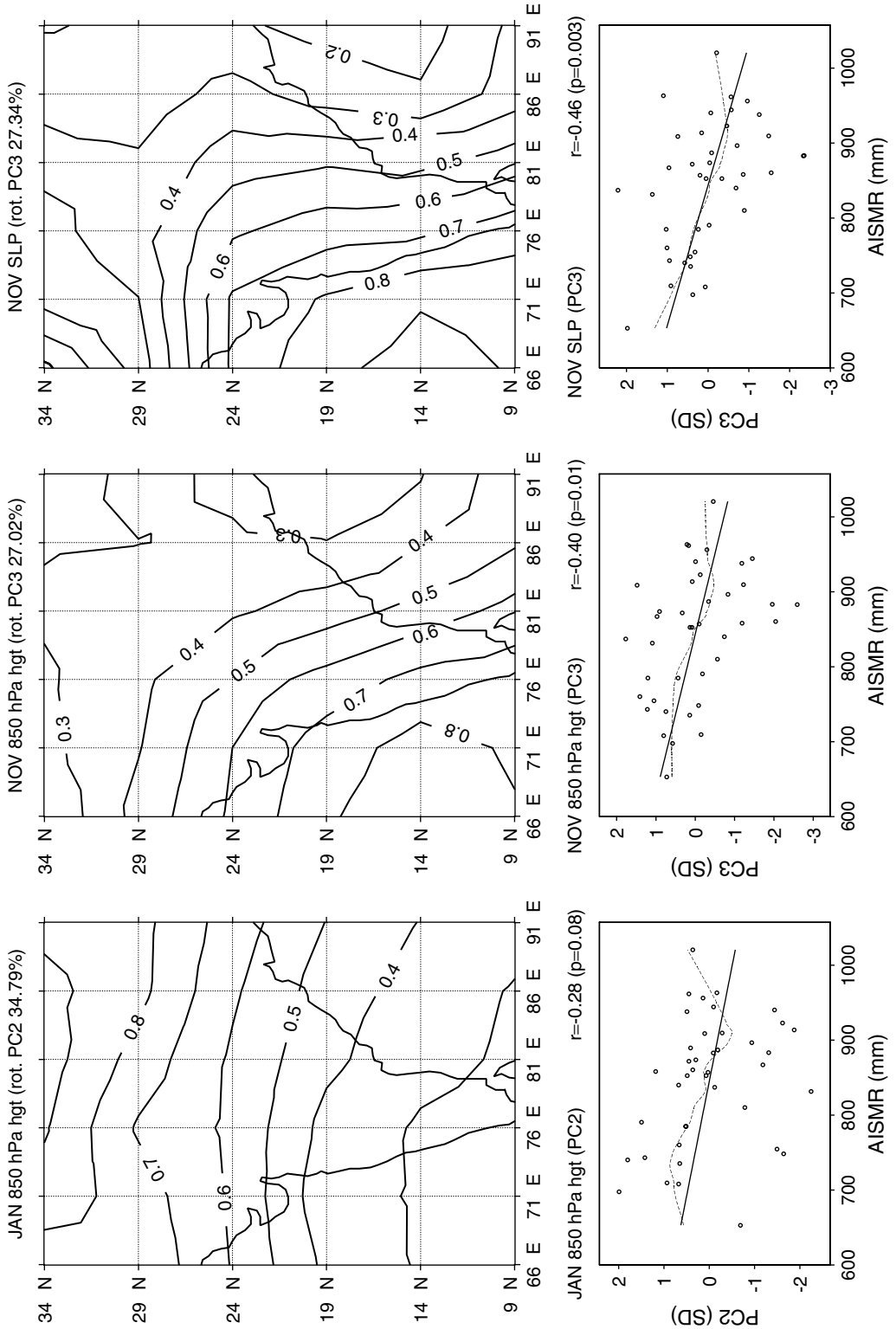

Figure 9. PC loadings for January 850 hPa geopotential height, November 850 hPa geopotential height, and November SLP data (top). Scatter plots showing the relationships between PC scores and AISMR (bottom). The solid lines show the best-fit linear regression lines. The dashed lines show the locally weighted linear regression curve

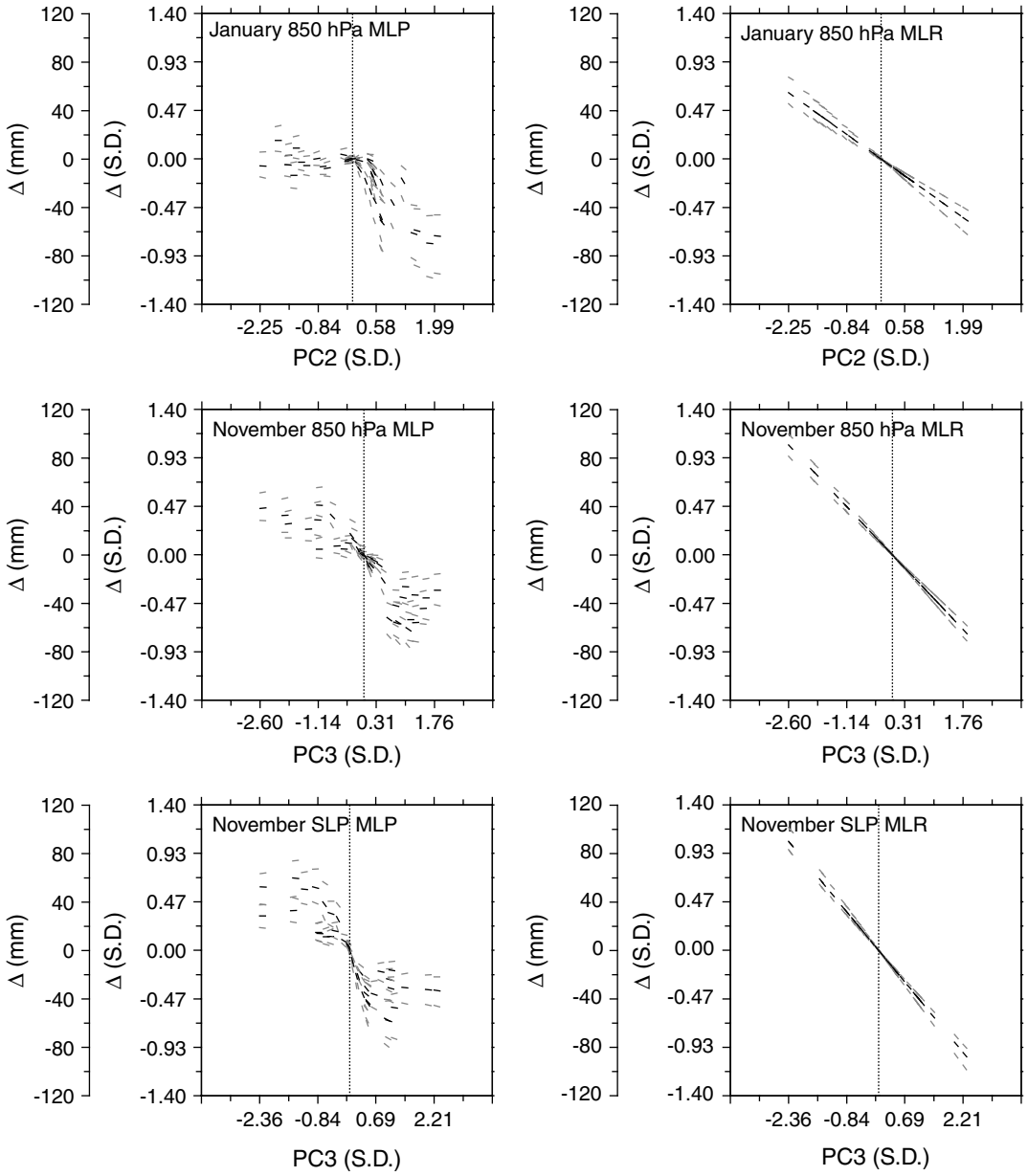

Figure 10. Effects plots for January 850 hPa geopotential height, November 850 hPa geopotential height, and November SLP PC predictors to MLP ANN and MLR models

Effects plots for PCs of 850 hPa geopotential height and SLP fields in November shared the same basic features (Figure 10). This was expected, as the PCs were very similar for these fields (Figure 9). For both SLP and 850 hPa geopotential heights, PC3 had a strong negative effect on model output, almost equal in range to effects observed for PC3 and PC4 of May 200 hPa geopotential heights. Differences in effects of the linear and MLP ANN models were evident for PC3. Trends in effects for both MLR and MLP ANN models were negative, but slopes of effects for the MLP ANN model decreased with absolute magnitude of PC3. This sigmoidal nonlinearity was present in results for both circulation fields. Performance statistics for MLP ANN models using November 850 hPa geopotential height and SLP PCs as predictors were better than corresponding linear models (Table II), although significant differences were not identified. The slight

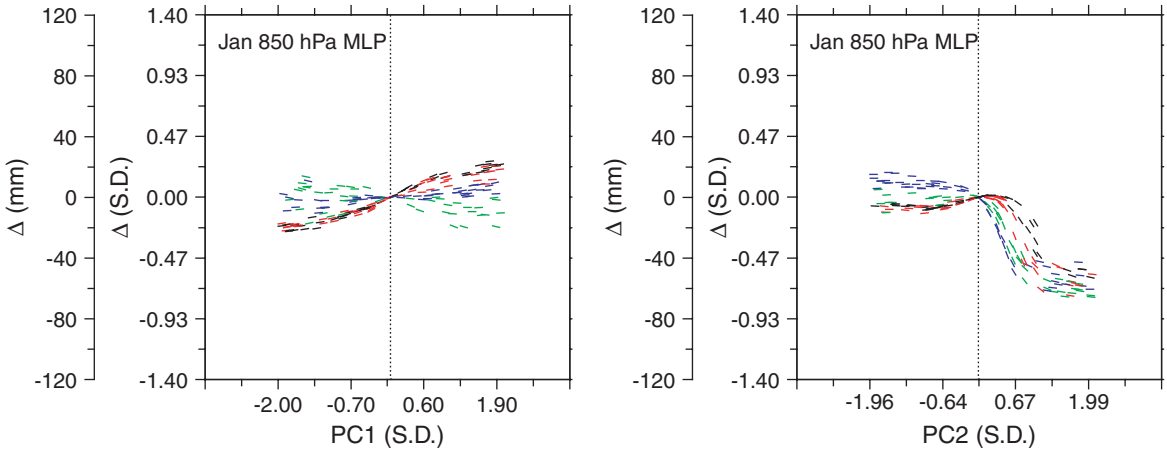

Figure 11. Stratified ANN MLP effects plots for PC1 and PC2 of the January 850 hPa geopotential height field. Colours in each plot correspond to ranges in scores of the other PC: black, ≤ -1; red, > -1 and ≤ 0; green, > 0 and $\leq +1$; blue, $> +1$

nonlinearity in the relationship between the PC3 and AISMR might explain the small difference in skill present between the model types.

5. DISCUSSION AND CONCLUSIONS

The graphical sensitivity analysis described in the current study provides a simple and effective method for investigating input–output mappings of statistical models. Although applied to ensemble MLP ANNs in this study, the method can be used with any statistical modelling technique. Unlike traditional analyses, which concentrate either on average sensitivities (Tangang *et al*., 1998) or local sensitivities (Hewitson and Crane, 1994), the graphical method is based on both types of information. By taking this approach, quantitative and qualitative indicators of variable importance, interaction strength, and nonlinearity can be used to diagnose input–output relationships. In addition, because the method acts on inputs and outputs, without reference to individual model parameters, it can easily be applied to combinations of models, such as ensemble MLP ANNs. To illustrate, the procedure was applied to a seasonal climate forecasting problem. Pre-monsoon PCs of geopotential height and SLP fields over the South Asian subcontinent were used as predictors for AISMR during the 1958–98 period. Predictive skill and stationarity of ensemble MLP ANNs and MLR models were assessed using a split-sample resampling procedure. The graphical sensitivity analysis was then used to gain insight into the form of the modelled relationships.

As found in the study by Cannon and McKendry (1999), pre-monsoon PCs of the 200 hPa geopotential height field in May formed a compact, interpretable, and significant set of predictors for AISMR (median $r = 0.67$, $p < 0.005$). Differences in performance between the MLP ANN and MLR models were small and not statistically significant. Results from the graphical sensitivity analysis suggest linear circulation–rainfall relationships at this lead-time. This is consistent with prior research. In an observational and GCM study, Yang *et al*. (1996) formed composites of spring geopotential height, wind, and temperature anomalies at the 200 hPa level during strong and weak Indian monsoon years. They found evidence of precursor relationships similar to those shown here and consistent with the upper-level circulation features identified in other studies (Verma and Kamte, 1980; Parthasarathy *et al*., 1991; Cannon and McKendry, 1999). Strong monsoon years were associated with positive 200 hPa height anomalies over the entire Indian region, whereas weak monsoon years were associated with negative anomalies. Easterly wind anomalies in spring were associated with strong monsoon years, whereas enhanced upper-level westerlies were associated with weak monsoon years. Prior to strong monsoon years an anomalous upper-level anticyclonic circulation was present over the Tibetan plateau, whereas an anomalous cyclonic circulation existed before weak monsoon years. These monsoon precursors

appear to be related to the nature of the winter–summer transition in upper-level circulation regimes from one dominated by the subtropical westerly jet stream to one dominated by the tropical easterly jet stream.

At longer lead-times, MLP ANNs tended to perform better than MLR models, although differences were only significant for January 850 hPa PCs (median $r = 0.4$, $p = 0.10$; median difference $r = 0.25$, $p = 0.09$). A sliding split-sample model validation for this predictor field showed little indication of a persistent relationship over the 1958–98 study period. However, the sensitivity analysis did show that the relationship between January 850 hPa PCs and AISMR, was strong when values of the second PC were positive, but was near zero when values of the second PC were negative. Interactions between PC1 and PC2 were also present, contributing to the increased skill of the MLP ANN model. A physical mechanism linking 850 hPa geopotential height anomalies in January with AISMR has not been suggested in the monsoon forecasting literature. With the strongest correlations found over the northern portion of the domain, it is possible that the 850 hPa response is related to snow cover, a boundary condition thought to be relevant to subsequent development of the summer monsoon circulation (Dey and Kumar, 1983). Through recent observational and modelling studies have linked regional snow cover and snow depths at various points in Eurasia with AISMR (Parthasarathy and Yang, 1995; Bamzai and Shukla, 1999; Kripalani and Kulkarni, 1999), Bamzai and Shukla (1999) found no significant relationship between Himalayan snow cover and summer monsoon rainfall. Interestingly, Sankar-Rao *et al.* (1996) found that the relationships between winter snow in Eurasia and AISMR were strongest when excluding El Niño years from the analysis. Stratifying the current sample based on El Niño–southern oscillation (ENSO) conditions yields a correlation with AISMR of $r = 0.58$ ($p = 0.04$, $N = 13$) during El Niño years and $r = -0.14$ ($p = 0.48$, $N = 28$) during La Niña and neutral years. This suggests that the relationship between PC2 of the January 850 hPa geopotential height field and AISMR may have been modulated by ENSO, with the strongest coupling tending to occur during El Niño years. Reasons for the weak relationship seen in the current study when anomalously low heights were present over the northern portion of the domain, and for the strong relationship when heights were above average, still need to be determined. Efforts to investigate the correspondence and linkages between ENSO conditions, January 850 hPa circulation conditions, and AISMR are ongoing.

Examination of other months prior to monsoon onset shows that November SLP (median $r = 0.5$, $p = 0.05$) and 850 hPa PCs (median $r = 0.37$, $p = 0.06$) were significantly correlated with AISMR. A slight sigmoidal nonlinearity in relationships between AISMR and the third PC of the two fields was evident, possibly accounting for the small differences in skill between the linear and nonlinear models. Given that rotated PC loadings for SLP and 850 hPa geopotential heights were similar, the nonlinear circulation–AISMR relationship was probably a coherent near-surface precursor signal of monsoon strength. Few significant regional circulation–AISMR relationships have previously been identified for lead-times 5 months prior to the onset of the monsoon. It is possible that correlations between AISMR and November surface circulation over the Arabian Sea are related to anomalous SSTs in this region. In a recent study by Clark *et al.* (2000), significant positive correlations between AISMR and SSTs in the Arabian Sea and central Indian Ocean were noted in the winter and autumn preceding the summer monsoon. Correlations of the opposite sign would be expected for SLP and geopotential height fields. Preliminary analyses suggest a weak, but statistically significant, negative relationship between PC3 of the November SLP field and November SST anomalies at the location defined by Clark *et al.* (2000) for their Arabian Sea index ($r = -0.38$, $p = 0.016$, $N = 40$).

The combined use of PCA, ensemble MLP ANN models, and the graphical sensitivity analysis procedure allowed statistical relationships between regional circulation conditions over South Asia and Indian summer monsoon rainfall to be investigated in detail. In general, wider application of graphical sensitivity analysis is recommended, especially given the increasing use of different nonlinear models in climatology and the simplicity of the procedure. Following the results from this study, further work is needed to identify physical mechanisms responsible for the statistical relationships identified by the sensitivity analysis. Comparisons between PCs exhibiting significant correlations with AISMR and indices identified in other studies might be a simple means of starting this process. A dynamical analysis, similar to the one performed by Yang *et al.* (1996), is another possible approach. As previous work has shown that predictors for AISMR can be highly correlated, an integrated assessment of circulation field predictors in an operational forecast setting is also necessary. Combining circulation predictors from different lead-times with ENSO, cross-equatorial,

global/hemispheric, and other regional predictors would indicate whether or not the PCs identified in the current study are capable of generating real increases in predictive skill.

REFERENCES

Bamzai AS, Shukla J. 1999. Relation between Eurasian snow cover, snow depth, and the Indian summer monsoon: an observational study. *Journal of Climate* **12**: 3117–3132.

Barnston AG, Livezey RE. 1987. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Monthly Weather Review* **115**: 1083–1126.

Baxt WG, White H. 1995. Bootstrapping confidence intervals for clinical input variable effects in a network trained to identify the presence of acute myocardial infarction. *Neural Computation* **7**: 624–638.

Bishop CM. 1995. *Neural Networks for Pattern Recognition*. Oxford University Press: Oxford.

Breiman L. 1996. Out-of-bag estimation. Technical report, Statistics Department, University of California.

Breiman L, Friedman J, Olshen R, Stone C. 1984. *Classification and Regression Trees*. Wadsworth: Belmont, CA.

Cannon AJ, McKendry IG. 1999. Forecasting all-India summer monsoon rainfall using regional circulation principal components: a comparison between neural network and multiple linear regression models. *International Journal of Climatology* **19**: 1561–1578.

Cannon AJ, Whitfield PH. 2001. Modeling transient pH depressions in coastal streams of British Columbia using neural networks. *Journal of the American Water Resources Association* **37**(1): 73–90.

Clark CO, Cole JE, Webster PJ. 2000. Indian Ocean SST and Indian summer rainfall: predictive relationships and their decadal variability. *Journal of Climate* **13**(14): 2503–2519.

Cleveland WS, Loader C. 1996. Smoothing by local regression: principles and methods. In *Statistical Theory and Computational Aspects of Smoothing*, Hardle W, Schimek MG (eds). Physica Verlag: Heidelberg.

Dey B, Kumar OSRUB. 1983. Himalayan winter snow cover area and summer monsoon rainfall over India. *Journal of Geophysical Research* **88**((C9)): 5471–5474.

Egmont-Petersen M, Talmon JL, Pelikan E, Vogelsang F. 1994. Contribution analysis of multi-layer perceptrons. Estimation of the input sources' importance for the classification. In *Pattern Recognition in Practice IV*, Gelsema ES, Kanal LN (eds). Elsevier Science: 347–358.

Finnoff W, Hergert F, Zimmermann HG. 1993. Improving model selection by nonconvergent methods. *Neural Networks* **6**: 771–783.

Friedman JH. 1991. Multivariate adaptive regression splines. *The Annals of Statistics* **19**(1): 1–141.

Gardner MW, Dorling SR. 1998. Artificial neural networks (the multilayer perceptron) — a review of applications in the atmospheric sciences. *Atmospheric Environment* **32**(14–15): 2627–2636.

Green PE. 1978. Introductory aspects of factor analysis. In *Analyzing Multivariate Data*. The Dryden Press: 341–390.

Hastenrath S, Greischar L. 1993. Changing predictability of Indian monsoon rainfall anomalies? *Proceedings of the Indian Academy of Sciences (Earth and Planetary Sciences)* **102**(1): 35–47.

Hastie T, Tibshirani R. 1990. *Generalized Additive Models*. Chapman & Hall: London.

Hewitson BC, Crane RG. 1994. Precipitation controls in southern Mexico. In *Neural Nets: Applications in Geography*, Hewitson BC, Crane RG (eds). Kluwer Academic: Dordrecht; 121–143.

Hornik K, Stinchcomb M, White H. 1989. Multilayer feedforward neural networks are universal approximators. *Neural Networks* **2**: 359–366.

Hsieh WW, Tang B. 1998. Applying neural network models to prediction and data analysis in meteorology and oceanography. *Bulletin of the American Meteorological Society* **79**(9): 1855–1870.

Kaiser HF. 1958. The varimax criterion for analytic rotation in factor analysis. *Psychometrika* **23**: 187–200.

Kaiser HF. 1959. Computer program for varimax rotation in factor analysis. *Educational and Psychological Measurement* **19**: 413–420.

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77**(3): 437–471.

Kripalani RH, Kulkarni A. 1999. Climatology and variability in historical Soviet snow depth data: some new perspectives in snow–Indian monsoon teleconnections. *Climate Dynamics* **15**: 475–489.

Kumar KK, Soman MK, Kumar KR. 1995. Seasonal forecasting of Indian summer monsoon rainfall: a review. *Weather* **50**(12): 449–467.

Overland JE, Preisendorfer RW. 1982. A significance test for principal components applied to a cyclone climatology. *Monthly Weather Review* **110**(1): 1–4.

Parthasarathy B, Yang S. 1995. Relationships between regional Indian summer monsoon rainfall and Eurasian snow cover. *Advances in Atmospheric Sciences* **12**(2): 143–150.

Parthasarathy B, Sontakke NA, Munot AA, Kothawale DR. 1987. Droughts/floods in the summer monsoon season over different meteorological subdivisions of India for the period 1871–1984. *Journal of Climatology* **7**: 57–70.

Parthasarathy B, Kumar KR, Deshpande VR. 1991. Indian summer monsoon rainfall and 200-mbar meridional wind index: application for long-range prediction. *International Journal of Climatology* **11**: 165–176.

Parthasarathy B, Munot AA, Kothawale DR. 1994. All-India monthly and seasonal rainfall series: 1871–1993. *Theoretical and Applied Climatology* **49**: 217–224.

Plate T, Bert J, Grace J, Band P. 2000. Visualizing the function computed by a feedforward neural network. *Neural Computation* **12**(6): 1355–1370.

Reed RD, Marks RJI. 1999. *Neural Smithing*. The MIT Press: Cambridge, MA.

Richman MB. 1986. Rotation of principal components. *Journal of Climatology* **6**: 293–335.

Riedmiller M. 1994. Advanced supervised learning in multi-layer perceptrons — from backpropagation to adaptive learning algorithms. *Computer Standards and Interfaces* **16**(3): 265–278.

Sankar-Rao M, Lau KM, Yang S. 1996. On the relationship between Eurasian snow cover and the Asian summer monsoon. *International Journal of Climatology* **16**: 605–616.

Sarle WS. 1994. Neural networks and statistical models. In *Proceedings of the Nineteenth Annual SAS Users Group International Conference*, SAS Institute Inc., Cary, NC; 1538–1550.
Smith M. 1993. *Neural Networks for Statistical Modeling*. Van Nostrand Reyngold: New York, NY.
Tangang FT, Tang B, Monahan AH, Hsieh W. 1998. Forecasting ENSO events: a neural network-extended EOF approach. *Journal of Climate* **11**: 29–41.
Verma RK, Kamte PP. 1980. Statistical technique for long-range forecasting of summer monsoon activity over India. In *Proceedings of Symposium on the Probabilistic and Statistical Methods in Weather Forecasting*, 8–12 September, Nice. WMO: Geneva; 303–307.
Von Storch H, Zwiers FW. 1999. *Statistical Analysis in Climate Research*. Cambridge University Press: Cambridge, UK.
Weiss SM, Kulikowski CA. 1991. *Computer Systems That Learn: Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning, and Expert Systems*. Morgan Kaufmann: San Mateo, CA.
Yang S, Lau KM, Sankar-Rao M. 1996. Precursory signals associated with the interannual variability of the Asian summer monsoon. *Journal of Climate* **9**: 949–964.

