# A dynamic statistical experiment for atmospheric interactions

Devdutta S. Niyogi, Sethu Raman *, Kiran Alapaty and Jongil Han

*Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, NC 27695, USA*

Received July 1996; revised May 1997

Interactions among atmospheric parameters exist at different scales. The pristine approach for observational or model data analysis involves changing the input parameters one at a time (OAT) and studying the effect on the system. Limitations of this approach for atmospheric applications are discussed. A fractional factorial (FF) based study is evolved and a methodology is outlined involving dynamic graphical analysis. Observational data from the FIFE and HAPEX-MOBILHY experiments are utilized with a vegetation and soil moisture scheme dynamically coupled in a planetary boundary layer model to demonstrate the robustness of this approach. Both low-resolution and high-resolution designs are considered. Various aspects of the vegetation-atmosphere interactions are delineated. Results obtained from the interaction-based FF approach differ considerably from the earlier OAT-type studies.

**Keywords:** FIFE, HAPEX-MOBILHY, planetary boundary layer, biospheric analysis, fractional factorial design

## 1. Introduction

The atmosphere is a dynamic system where various energy transfer mechanisms act simultaneously at different scales. Over the years, our knowledge of this system has evolved from various field experiments and rigorous modeling exercises. Efforts to understand the atmospheric processes started with the assumption of a homogeneous and uniform bare surface. Presently, one of the biggest challenges in atmospheric and climate modeling is to efficiently represent surface features such as vegetation, and soil moisture and associated surface temperature variation [4,20]. This knowledge has helped in understanding and realistically simulating planetary boundary layer (PBL) processes. To study the effect an input parameter has on an entire modeling system, that parameter is generally varied while all the others are held constant. However, we feel that to gain more knowledge of the system, better methods of analysis than the "one-at-a-time" approach must be applied.

Our present study proposes the use of a dynamic graphical statistical method such as main-effect and Pareto plots, which can be efficiently employed for extracting information on various interactions within the atmospheric processes. The effectiveness of the proposed method is demonstrated through a simulation study using the land-surface scheme of Noilhan and Planton [19] (hereafter referred to as NP89) in a columnar version of the North Carolina State University (NCSU) planetary boundary layer (PBL) model [1].

Our overall aim is to show that interaction explicit analysis is useful, if not essential, for atmospheric studies.

* Corresponding author. Address: Department of MEAS, North Carolina State University, Raleigh, NC 27695-8208, USA. E-mail: sethu_raman@ncsu.edu.



## 2. Experimental design

Proper experimental design is crucial for any analysis. Presently there are three ways of designing this kind of an experiment, based on the following approaches: (1) one at a time (OAT), (2) factor separation [21], and (3) fractional factorial [6]. To elucidate the bases of these approaches, consider a system with input parameters $P_1$, $P_2$, and $P_3$. In an OAT approach, the effect a parameter $P_1$ has on the system is treated as

$$E(P_1) = K(P_1),$$

i.e., the changes in the system effect are attributed to the changes in $P_1$ alone through a function $K$. $P_1$ is altered over a possible range in steps and the resulting $E$ (OAT) for each $P_1$ value is obtained. The same is done for $P_2$ and $P_3$, and by comparing the values of $E(P_i)$ the role or significance of each parameter on the entire system is pictured.

The factor separation (FacSep) approach attempts to resolve the effects of $P_1$ on the system into those that are directly and interactively dependent on $P_1$. Thus, the setup involves simulating a system with $2^n$ combinations (see [21]).

In comparing the OAT and FacSep approaches, we find that the OAT approach has the limitations that (1) it is conceptually incorrect, as it assumes an independence of the events, and (2) the outcome exaggerates the significance of the parameter. For example, if vegetal cover is altered in a model, corresponding variations are expected in soil moisture and soil temperature. The net effect due to the vegetal cover change is thus a combined one and not just that due to change in vegetal cover alone (see [3] for a discussion). This could lead to an erroneous hypothesis for the development of parameterizations and for an understanding of the process. The FacSep method is informative, but it may not be directly applicable for a system with a large number of parameters.

Compared to the OAT and FacSep approaches, the fractional factorial (FF) approach [6] is a practical and useful compromise, particularly for systems with numerous parameters. In the FF approach, the effect due to $P_1$ is analyzed as

$$E(P_1) = K_1(P_1) + K_2(P_1 : P_2, P_1 : P_3, P_1 : P_2, P_1 : P_2 : P_3).$$

The $K_1(\cdot)$ part is the "main-effect" term and is similar to the RHS term in the OAT analysis. The $K_2(\cdot)$ part is the interaction term. For a full fractional experiment, this method is similar to the FacSep approach, where $2^n$ simulations would be performed. However, as its name suggests, the FF approach treats a fraction of the combinations. This makes the design of the form $2^{n-p}$, where $p$ is taken to yield the most information in the fewest simulations [8]. The trade-off is between sample size and interactions; most three-way and higher interactions are considered negligible (which is valid for most atmospheric cases). Depending on the information sought in terms of parameter interactions (one, two, or three-way), the FF design has "resolutions." A resolution V design has all the main effects and two-factor interactions plus principal three-way interactions. In a resolution III design, all the main effects and principal two-factor interactions are resolved, which makes it a desirable design for screening when the sample size is large and interactions alone do not dominate the scenario. A resolution IV design is intermediate between V and III. Obviously, with increasing resolution the number of simulations required increases.

Below, we describe the application of the FF approach in context of our problem.

## 3. Methodology

This study has the following objectives: (1) to illustrate the applicability of the dynamic statistical FF-related approaches for land-surface schemes such as that in NP89, (2) to test the sensitivity of the vegetation and soil moisture scheme, and (3) to demonstrate the use of graphical methods to identify significant input parameters for various effects. Thus, the overall objective is to understand the role of interactions for the purpose of improving parameterizations and analysis of atmospheric-biospheric-related experiments. Though the emphasis is on NP89, this should not limit the application of the method for other atmospheric processes or systems.

The theory behind the FF design is well documented [6], and for brevity, therefore, only the key steps about the design approach applied in this study are mentioned here:

1. Decide on the parameters to be included in the study.
2. Decide on the resolution to be applied (III, IV, or V).
3. Conduct simulations/experiments with the combinations provided by the design and generate output.

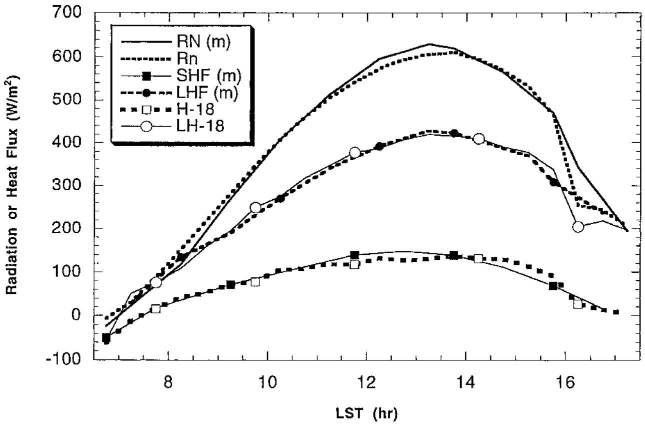


Figure 1. Comparison between FIFE-observed and model-predicted surface parameters. Rn and RN (m) are observed and modeled net radiation, H-18 and SHF (m) are the observed and modeled sensible heat fluxes, LH-18 and LHF (m) are the observed and modeled latent heat fluxes for the FIFE site 18.

For atmospheric data, the next aspect we propose is to perform a graphical analysis, such as the main-effect and Pareto plots, or the active-contrast plots (see [9] for details), in order to (1) rank the parameters, and (2) identify the significant ones; the latter step is crucial in developing computationally efficient models and for measurement protocol [18]. We illustrate this broad methodology using the land-surface scheme of NP89 and data from the First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment (FIFE) Golden Day 2 (11 July 1987) as the reference data set [20].

The model used is a 1-D, 30-layer modular version of the NCSU PBL model [1]. It utilizes a 1.5-order TKE closure scheme [10,13] with the vegetation and land-surface parameterization of NP89. Both the turbulence closure scheme and the vegetation parameterizations are employed in a number of atmospheric models globally [15]. A single-layer vegetation is prescribed and prognostic equations for ground temperature and soil moisture are solved. The equations are given in NP89 and Jacquemin and Noilhan [11] (hereafter JN90) and are not repeated here. The important inputs to the model include initial profiles of atmospheric parameters such as wind, temperature, and humidity; initial surface conditions for pressure, moisture availability, and ground temperature; and physiological details such as vegetation cover, leaf area index, and minimum stomatal resistance. Both the model and the reference data set have been well tested using OAT-like analyses (see [20] for the FIFE results and analysis; and [1,11] for the model scheme under consideration).

The initial conditions are obtained from the FIFE Golden Day 2. This is to ensure that the results obtained from the statistical analysis are perturbations of a real situation and therefore physically possible, as opposed to a synthetic case study. The model was run for a scenario of 10 hours from 0700 to 1700 LST with emphasis on the overall daytime interactions. Figure 1 shows the observed and predicted radiation and surface fluxes. These outcomes

Table 1
Input parameters and higher (+) and lower (−) values selected for use in the factorial designs.

| Symbol* | Input parameter | − | + |
|---|---|---|---|
| A | RGL: threshold radiation value ($W\,m^{-2}$) used in stomatal/surface resistance (Rs) calculation | 60 | 100 |
| B | Rsmin: minimum stomatal resistance ($s\,m^{-1}$) | 60 | 120 |
| C | LAI: leaf area index of vegetation present | 0.9 | 2.8 |
| D | VEG: amount of vegetal cover over the domain | 0.4 | 0.9 |
| E | SWG: surface soil moisture ($m^3\,m^{-3}$) | 0.16 | 0.4 |
| G | SSWG: subsurface soil moisture ($m^3\,m^{-3}$) | 0.14 | 0.25 |
| H | $z_0$: surface roughness for momentum (m) | 0.065 | 0.1 |
| I | Tg: surface soil temperature (K) | 290 | 300 |
| J | ALB: net albedo (foliage and soil) of the domain | 0.15 | 0.3 |

*These are the symbols used in the main-effect and Pareto plots resulting from the resolution III study.

confirm the ability of the model to simulate observations closely.

Following the FF approach outlined earlier, we first chose the input parameters to be included. We then identified the resolution, choosing a screening-type resolution III design. Third, we chose the design FF0916 (FF design for 9 factors with 16 runs, see [8]). Table 1 lists the parameters and their values corresponding to "+" (higher) or "−" (lower) settings used in the FF design. These values were based on the FIFE Golden Day 2 (11 July 1987) observations; additional values for physiological representations (Rsmin, Rgl, emissivity) are from the model initial conditions used for FIFE simulations in the literature [1,7]. Finally, we ran the model using the specified inputs and analyzed the resulting outcomes.

Following discussion of the screening-type experiment, where nine surface input parameters are considered simultaneously (section 4), a high-resolution analysis of three output parameters (PBL height, ground temperature, and soil moisture) is addressed in section 5. Section 6 presents our conclusions.

## 4. Screening experiment (resolution III design)

We analyzed the FF0916 screening experiment results primarily using graphical outcomes such as main-effect and the Pareto plots [8]. In our study, the output parameters considered are the model-predicted values of surface resistance (Rs), net radiation (RN), PBL height (PBL ht), surface latent heat flux (LHF), surface sensible heat flux (SHF), ground temperature (Tg), air temperature in the surface layer (Ta), specific humidity of air close to the soil surface (Qs), surface layer specific humidity (Qa), soil moisture content at the surface soil layer (SWG), and soil moisture content in the deep reservoir or the subsurface (SSWG). Note that the model has prognostic equations for all the above mentioned parameters and hence table 1 shows only the initial conditions and the data for the interaction study is the time-varying model predicted outcome (see figure 1 as an example for energy fluxes).

First, some background on our main-effect and Pareto plots. The main-effect plots (figure 2) provide a "mean view" of the changes in the response or effect due to changes in the input parameters. Pareto plots are based on the principle that 90% of the effects in a system can be attributable to 10% of the parameters [8]. The terms A–J (barring F, which is reserved for the statistical programming logic) represent a model input parameter (table 1) and each takes a higher (+) or a lower (−) value as per the design FF0916. In main-effect analysis, the interaction effect is implicit; in Pareto plots (figure 3), however, it is explicit.

We carried out these analyses for the 11 model predicted output parameters (Rs, RN, PBL ht, LHF, SHF, Tg, Ta, Qs, Qa, SWG and SSWG) for 1300 LT. Many interesting features are apparent from the main-effect plots (figure 2) and Pareto plots (figure 3). The subsections below highlight some of the key findings.

### *4.1. Radiation limit for photosynthesis (Rgl)*

We found that Rgl is quite significant as an interaction term (figure 3) for almost every output parameter considered. However, even a detailed OAT validation on this model (JN90) could not identify this term as crucial. The reason our study found it to be important for the NP89 scheme is primarily because of the interactive nature of this term. In other words, a high radiation limit by itself does not produce any significant changes in the PBL structure; rather, it is the interaction with other terms, like LAI and surface soil moisture (SWG), that is crucial.

### *4.2. Soil moisture (SSWG and SWG)*

The deep soil moisture (SSWG) plots in figures 2 and 3 show that any error in assigning initial subsurface soil moisture apparently does not get smoothed out and remains during the entire run. Looking simply at the sign associated with each term (figure 3), (and confirming they are not just a statistical artifact), it appears that, for longer time periods, the surface change, if triggered through vegetative processes, could be more significant than any other. With ground temperature (Tg) as a main effect and the two Rgl interactions, we see a tendency toward depletion of deep soil moisture (SSWG). On the other hand, higher vegetal cover (VEG) would tend to conserve the deep soil moisture. In effect, the radiative process seems to have the greatest effect on the complex deep soil moisture variation, and a strong interaction with vegetation is also important. There appears to be a compensating process among the surface parameters in the net outcome. Physiological *interaction* depletes the subsurface water content, while the vegetal cover as a *main effect* conserves it. Thus, as vegetal cover increases, generally the physiological interaction will also increase, and as the two act in an opposite manner, the

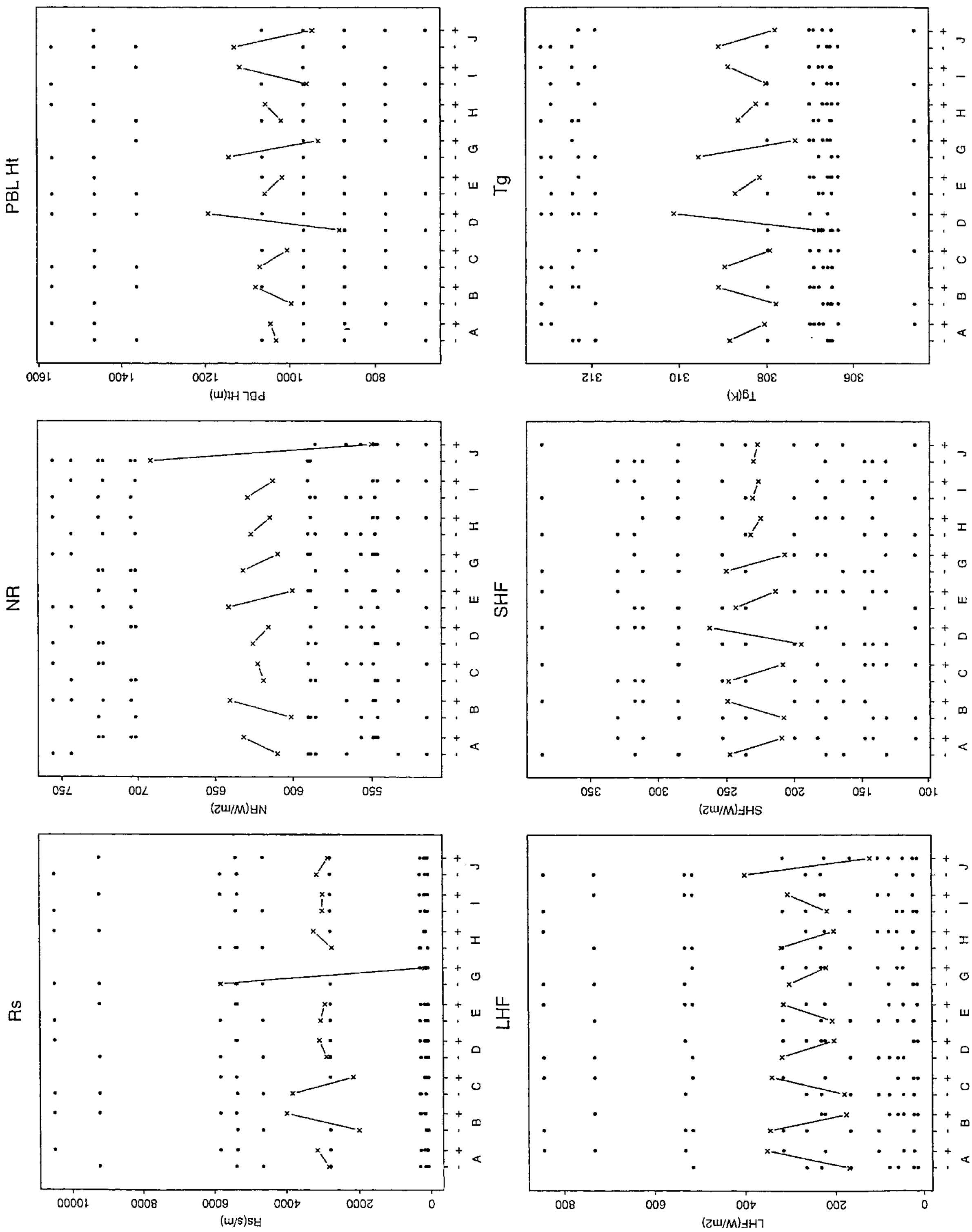


Figure 2. Main-effect plots for the impact of the 9 input parameters on the 11 output parameters (resolution III experiment). The symbols A–J are defined in table 1.

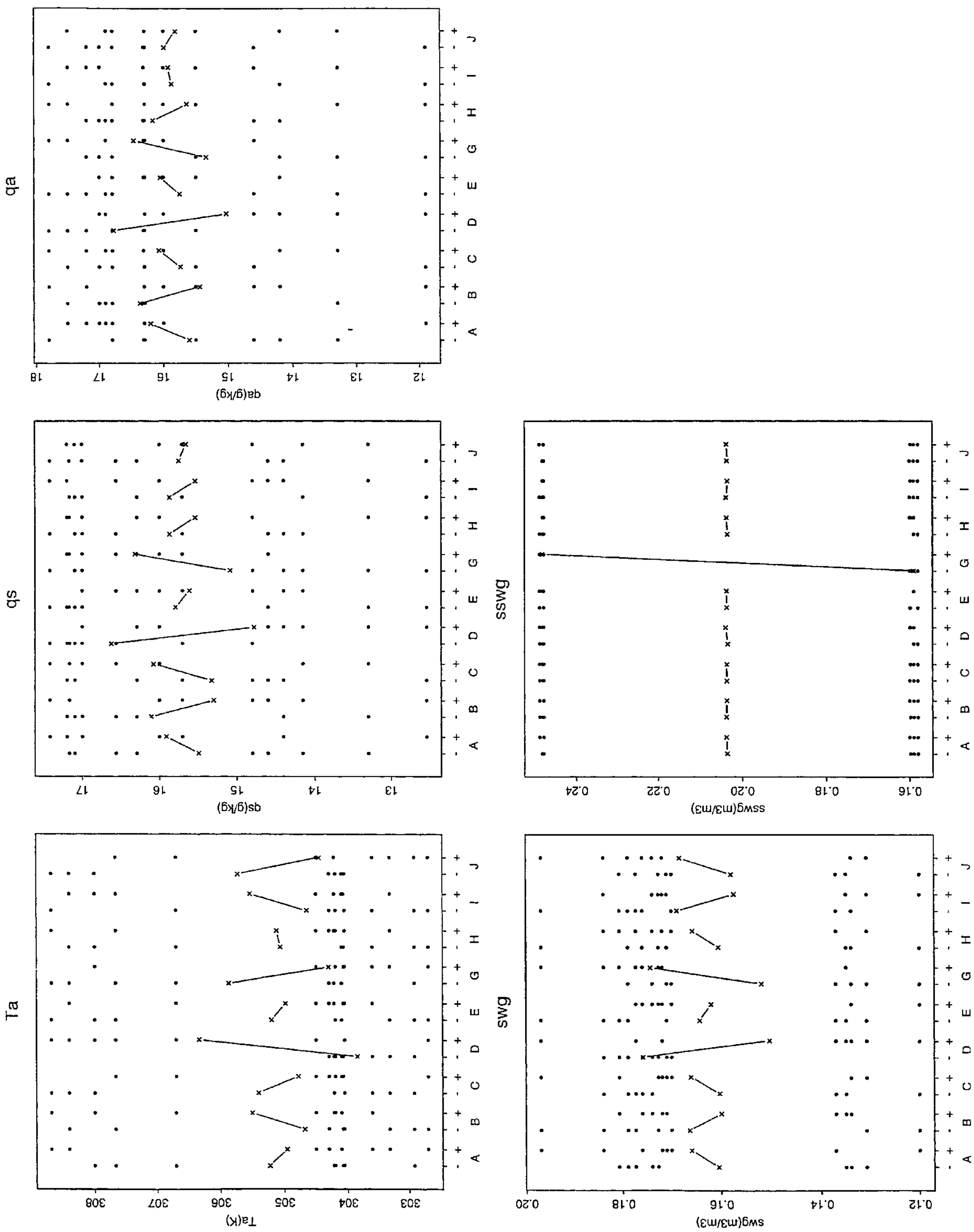


Figure 2. (Continued).

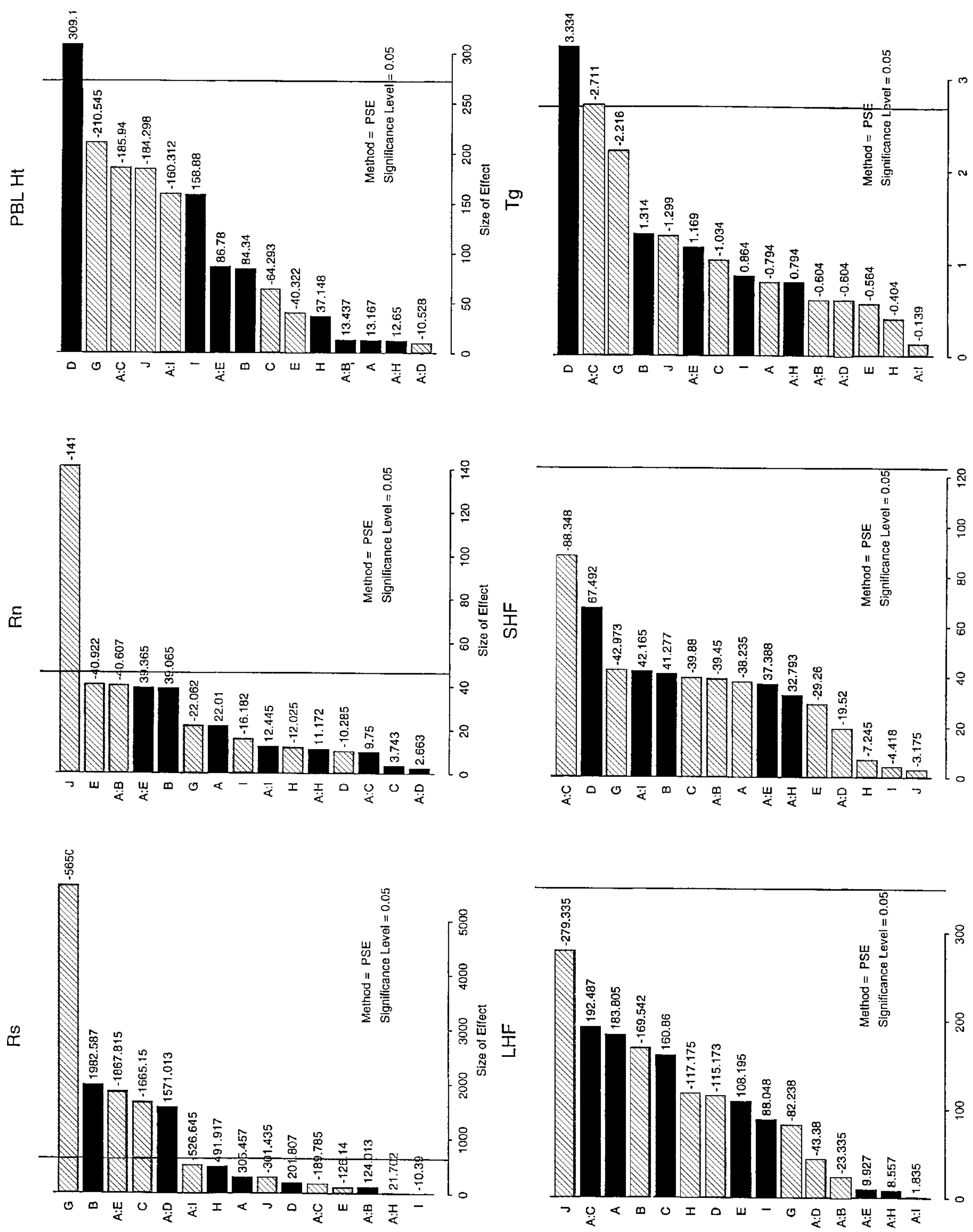


Figure 3. Pareto plots corresponding to the main-effect plots in figure 2 (resolution III experiment). PSE stands for pseudo standard error.

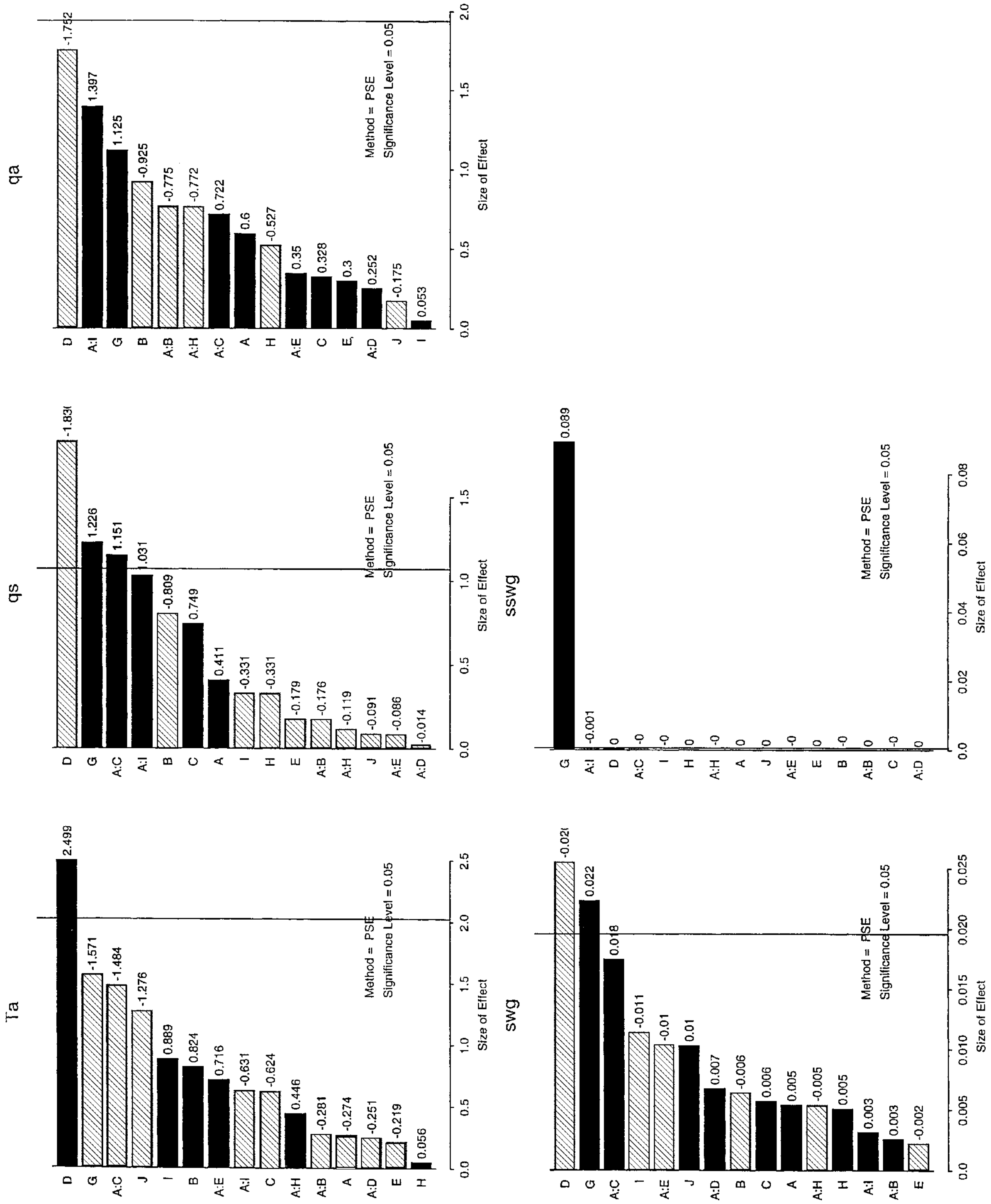


Figure 3. (Continued).

net effect may not be as sensational as the OAT variations generally predict.

Interestingly, although closely related to deep soil moisture, SWG has a markedly different response to changes in surface features. The major difference occurs for changes in the vegetal cover, which is found to be the most important parameter for both surface and deep soil moisture. For deep soil moisture there is a positive correlation, but for the surface soil moisture vegetal cover is negatively linked. However, the Rgl : LAI interaction term showed a positive correlation with surface soil moisture and a negative one with subsurface moisture. This again implies that the omission of this interaction would exaggerate the effect of vegetal cover in OAT predictions. Also, a direct relation between albedo and surface soil moisture is seen, and a higher LAI conserves surface moisture. However, surface roughness ($z_0$) depletes and Rsmin restores the surface moisture in the scheme. Generally, for a growing crop, both Rsmin and $z_0$ increase [15], so the net effect would be less dramatic than for either parameter considered individually. However, if the fractional vegetal cover also increases, the tendency would be to reduce the surface moisture more rapidly than it could be made up by the restoring capacity of other parameters.

### 4.3. Surface fluxes (SHF and LHF)

To get a base comparison with other observational and simulation studies, we compare our results with JN90. In a detailed sensitivity (OAT-type) study, JN90 obtained relative variations of surface latent heat flux (LHF) for different input parameters ranging over a $-90\%$ to $+90\%$ deviation of the base value considered. Their analysis resulted in the following order of significance of the parameters: subsurface soil moisture, vegetal cover, LAI, Rsmin, and $z_0$ for the 0 to $-90\%$ range and subsurface soil moisture, vegetal cover, Rsmin, LAI, and $z_0$ for the 0 to $+90\%$ deviation from base value. In our study, considering main effects alone (interactions implicitly considered) (figure 2), the order of importance is surface albedo, Rgl, Rsmin, LAI, vegetal cover, $z_0$, surface soil moisture, and subsurface soil moisture, and ground temperature. The Pareto analysis outcome (figure 3), which considers interactions explicitly, brings out the following order: surface albedo, interaction between Rgl and LAI, Rgl, Rsmin, LAI, $z_0$, vegetal cover, surface soil moisture, surface temperature, and subsurface soil moisture. There is a remarkable difference in the order of importance between a study that considers interactions and a simple OAT experiment that does not.

The difference between this study and the OAT-type JN90 can be attributed to (1) the sensitivity approach used, and/or (2) the driving PBL model. To examine this, we performed an OAT-type experiment (as employed by JN90) for LHF. Table 2 lists the OAT quartile ranges for the inputs. The 1300 LST predictions of LHF from different runs were compiled and compared with the reference simulation (FIFE observations as input). The order of significance (figure 4) is subsurface soil moisture, vegetal cover, Rsmin, LAI, and surface albedo. This is quite consistent with JN90 results. It is therefore clear that the interaction-explicit FF approach alone is responsible for the differences in the order of significance.

Table 2
Parameter ranges for the OAT experiment for LHF predictions.

| Input parameter | Levels I | II | III | IV |
|---|---|---|---|---|
| RGL ($W\,m^{-2}$) | 60 | 70 | 80 | 90 |
| Rsmin ($s\,m^{-1}$) | 90 | 120 | 150 | 180 |
| LAI | 0.90 | 1.50 | 2.00 | 2.50 |
| VEG | 0.10 | 0.30 | 0.60 | 0.90 |
| SWG ($m^3\,m^{-3}$) | 0.14 | 0.22 | 0.32 | 0.40 |
| SSWG ($m^3\,m^{-3}$) | 0.16 | 0.22 | 0.30 | 0.35 |
| $z_0$ (m) | 0.13 | 0.19 | 0.25 | 0.30 |
| Tg (K) | 290 | 293 | 297 | 300 |
| ALB | 0.15 | 0.22 | 0.27 | 0.30 |

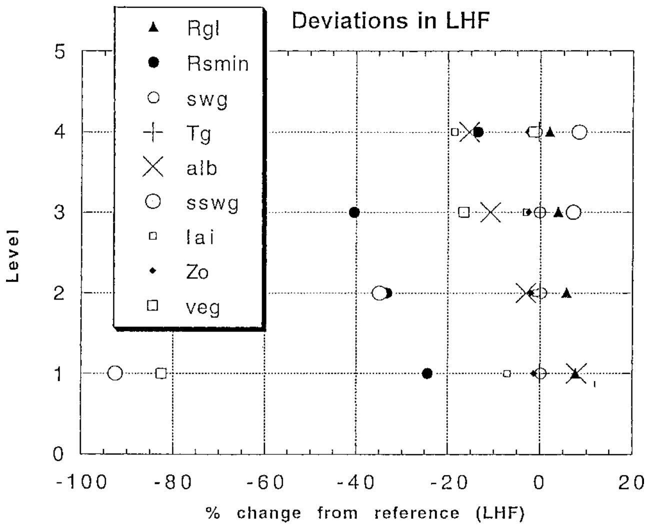


Figure 4. An OAT outcome for LHF ($W/m^2$) predictions from the model. "Level" indicates the parameter quartile settings shown in table 2.

Arguments similar to that presented earlier in this paper and by Alpert et al. [3] highlight the exaggerated impact caused by OAT variation. Thus, the factorial-based Pareto plot results from the present study should be more acceptable than the JN90 conclusions. Additionally, the "magnitude" of the effect could be another reason for the altered order of the significant variables. Consider the effect of surface albedo and surface roughness ($z_0$) (two physical parameters that are assumed as fairly constant for a given surface condition) on LHF (figure 3). The value of the effect of ALB ($-279$) is approximately twice as much as that of $z_0$ ($-117$). Translated for the OAT framework, this implies that for a consistent effect the range over which albedo could be varied should be half of that for $z_0$; conversely, if the range for albedo is set at 40%, then for consistency in perturbation $z_0$ should vary by about 80%. However, this is not practical because it could involve more speculation regarding the inputs for the sensitivity study than realization in terms of the interpretation of the output. Hence the present methodology and the outcome are probably more universal in nature than the OAT-type studies like JN90. Any study of this nature with an objective of viewing relative variations could benefit more from using the FF approach paired with Pareto plots than from an OAT analysis.

In principle, one would expect the latent heat flux (LHF) to depend primarily on the vegetation. It was therefore interesting to see that for LHF, the two vegetation-related parameters (LAI and vegetal cover) ranked 5th and 7th respectively in their relative significance for LHF. The radiation physical parameters such as albedo were the most significant ones. Similarly for sensible heat flux (SHF), the Rgl : LAI interaction and vegetation cover were the most significant parameters, with vegetal cover showing a positive effect while the subsurface soil moisture and Rgl : LAI interaction indicated negative effects. Also, it is the LAI rather than the vegetal cover that is directly related to the evaporative flux, that is, for an increasing LAI, the LHF increases while SHF decreases (opposite to the effect of increasing vegetal cover) for the model. Increasing surface roughness decreases both SHF and LHF; the effect is more dominant for LHF than SHF. Stomatal resistance acts in an opposite sense to LAI, but similar to surface soil moisture in altering the surface heat fluxes. The ground temperature (Tg) and LHF effects are directly related, while for lower Tg higher SHF values are obtained. The Rgl : LAI interaction is one of the most important aspects deciding the Bowen ratio for the system.

We reviewed these and associated features using field observations from HAPEX-MOBILHY [4] referred to in JN90. Figures 15–17 of JN90 (cases 1–3) show a direct relation between vegetal cover and SHF. For a vegetal cover around 0.2, the maximum SHF is 100–150 $\mathrm{W\,m^{-2}}$, while for a vegetal cover value around 0.5, it is $\sim 200$ $\mathrm{W\,m^{-2}}$. This should help clear up a possible misconception that higher vegetal cover alone yields lower sensible heat flux and enhanced LHF, within the energy transfer process. The interactions (particularly for soil moisture and possibly leaf temperature) and their nonlinear propagation in the system restrict us from having a straightforward “rule of thumb” of this nature. Though not presented numerically, it can be inferred that for the same fractional vegetal cover in cases 2 and 3 ($= 0.2$), the maximum SHF for a lower subsurface moisture is about 25 $\mathrm{W\,m^{-2}}$ more than that for a higher subsurface moisture. In another case of JN90, subsurface moisture is constant, while the fractional vegetal cover changes from 0.5 to 0.7. Our earlier analysis would lead us to expect higher SHF values for the VEG $= 0.7$ case. However, the observations indicate a *decrease* by about 100 $\mathrm{W\,m^{-2}}$. This is a clear indication of the dominance of the LAI : Rgl interaction over the fractional vegetal cover as a main effect. The value of the effects from the Pareto plot (figure 3) obtained in this study indicate that the Rgl : LAI interaction ($-88.3$) is about 1.4 times more pronounced than the vegetal cover main effect (67.4). Hence, even though the fractional vegetal cover value has been raised from 0.5 to 0.7, an increase in LAI from 1.0 to 2.0 decreases the maximum SHF.

### *4.4. Temperature and humidity (Ta, Tg, Qs, and Qa)*

Predicting and understanding the link between temperature, humidity, and soil moisture variation is another challenging aspect for weather forecasting (cf. [5,12] with NP89 scheme).

Figures 2 and 3 show that vegetation is the key feature for all these output parameters (Tg, Ta, Qs, Qa, SWG, and SSWG). Higher vegetal cover, as a main effect alone, yields higher ground temperature (Tg) while the interaction between Rgl : LAI and the main effect SSWG both lower Tg. This is consistent with the SHF results discussed earlier. In addition, we found that Rsmin is positively related while albedo is negatively related to Tg.

For Tg, the order obtained was fractional vegetal cover, Rgl : LAI interaction, subsurface soil moisture, Rsmin, and surface albedo. While an increase in VEG causes an increase in Tg, increased LAI causes decreased Tg; this opposing behavior between LAI and vegetal cover was also seen for both the surface energy fluxes. The Rgl : surface soil moisture interaction term is also important, implying that the interaction between transpiration and surface soil moisture increases Tg. For air temperature (Ta), the order was fractional vegetal cover, deep soil moisture, interaction between Rgl and LAI, deep soil moisture, surface albedo, and Tg.

For the humidities (both near surface and air), the effect of vegetal cover is strong and negative. Deep soil moisture availability is positively but less strongly linked. The vegetation present seems to have an effect similar to that of Rsmin, which is also negatively related to the humidity. Low Rsmin would result in higher diffusion through vegetation, yielding higher humidities [15]. For higher vegetal cover, the effect is equivalent to retaining moisture and keeping it away from the atmosphere. Thus, the moisture-retaining capacity of vegetal cover appears to have an overriding effect compared to soil moisture for predicting humidity.

## 5. Higher resolution experiment (resolution V design)

Through the screening experiment, we were able to assess the performance of the land-surface scheme (NP89) within the NCSU PBL model (columnar version) by simulating several features observed in the field. Resolution III (that includes all main effects and some principal two-way interactions) outcomes are useful and informative. However, it is also of interest to use a resolution V design, in which all interactions and all main effects are explicit. Additionally, this would help us gather information that may not have been resolved explicitly in the resolution III design.

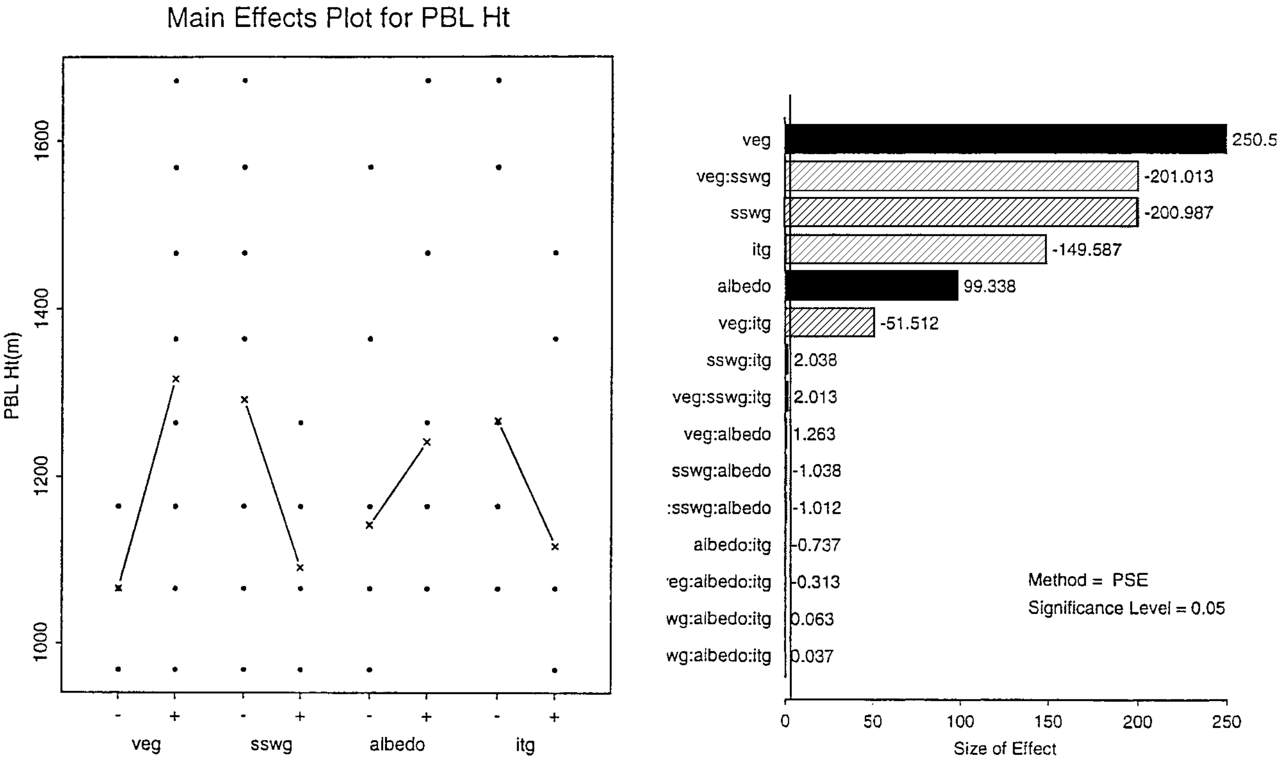


Figure 5. Main-effect and Pareto plots for PBL height (resolution V experiment).

Table 3
Parameters and design chosen for the higher resolution (resolution V) experiment.

| Predicted output parameter | Initial input parameter | Design |
|---|---|---|
| PBL height | VEG, SSWG, ALB, itg | FF0416 |
| Tg | VEG, SSWG, Rsmin, ALB | FF0416 |
| SWG | VEG, SSWG | FF0204 |

For our resolution V design, we considered three sample output parameters: PBL height, ground temperature (Tg), and surface soil moisture (SWG). We used our resolution III results to choose a subset of principal input parameters for this experiment (table 3). The specific designs utilized for this high resolution analysis are FF0416 (for PBL height and Tg) and FF0204 (for SWG) (see [14] for design details). The higher (+) and lower (−) values assigned to the parameters are still the same as in table 1. We present and analyze main-effect and Pareto plots, and draw attention to some key interactions and model features.

### *5.1. PBL height*

Figure 5 shows the main-effect and Pareto plots for the output parameter PBL height. The main-effect outcome shows a significant variation in their mean values and confirms that all four of the input parameters are important, supporting the hypothesis that the four parameters chosen are appropriate for the analysis. From the plots, larger vegetal cover resulted in a higher PBL height, while higher deep soil moisture (SSWG) availability leads to a lower PBL height. The resolution III experiment indicated the moisture retentive tendency of vegetal cover (figures 2 and 3). Also, lower albedo causes greater heating which in turn increases PBL heights, while warmer initial ground temperature (itg) in the model tends to predict relatively lower PBL heights. This ground temperature related result is somewhat surprising, but could be due to the increase in latent heat flux for warmer initial temperatures and consequent decrease in sensible heat flux.

The Pareto plot (figure 5) confirms the outcomes from the main-effect plot, giving the following order of priority to the input parameters as main effects: vegetal cover (VEG), subsurface soil moisture (SSWG), initial ground temperature (itg), and surface albedo (ALB). The VEG main effect is about 2.5 times higher than that for ALB. Also, the VEG : SSWG and VEG : itg interaction terms are quite prominent. Note that the importance of these interactions can also be deduced from the resolution III experiment. Further, it can be seen that none of the higher interaction terms, such as the VEG : SSWG : itg (vegetal cover : subsurface soil moisture : initial ground temperature) triple interaction, are significant, which confirms an implicit assumption we had to make when accepting the resolution III design. Thus, the resolution III and resolution V experiments are statistically appropriate in the case of PBL height, and the conclusions from both supplement the information obtained about this output parameter.

Both the experiments show the importance of vegetation in determining the PBL structure. However, using a comprehensive vegetation scheme in analysis is computa-

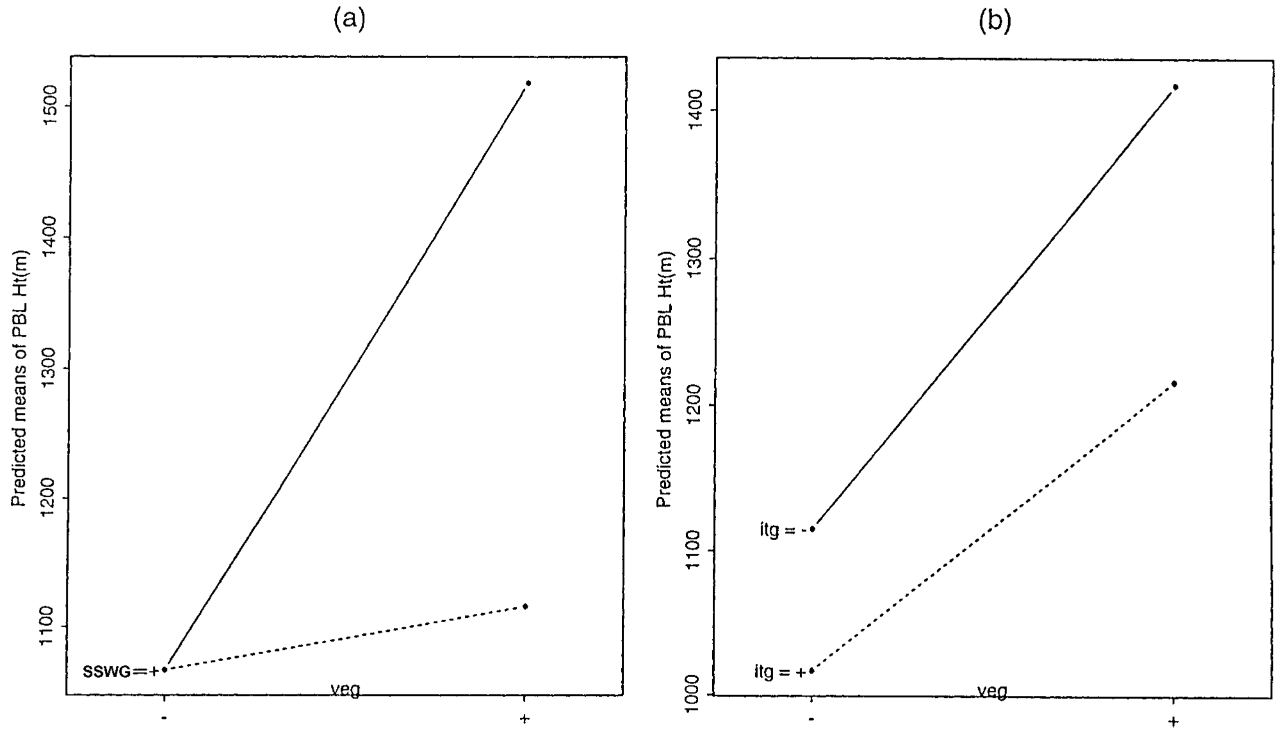


Figure 6. Two-factor interaction plots for (a) deep soil moisture ($m^3/m^3$) and vegetal cover interaction, and (b) initial ground temperature (K) and vegetal cover interaction, for PBL height prediction. (Confidence level is $\sim$90% based on pseudo standard error.)

tionally expensive. Therefore, we need an answer to the following question: When can we expect the omission of vegetation and soil moisture processes to be very significant? Obviously, we can say that it is not significant when the vegetation itself is insignificant. Figure 6 deals with this aspect, showing how the two-factor interaction between the PBL height changes with (a) variation in deep soil moisture (SSWG) and vegetal cover (VEG), and (b) variation in initial ground temperature (itg) and VEG. The PBL height outcome increases with vegetal cover. This increase is intense for drier soil moisture cases as against for cases with higher soil moisture availability. Also it can be inferred, when the vegetal cover is lower, the changes in soil moisture alone do not affect the PBL height much. It is also interesting to note that when SSWG is high, even a significant change in vegetal cover does not really seem to affect the simulation of the PBL depth. The two-factor interaction plot for vegetal cover and initial ground temperature interaction (figure 6b) confirms the earlier observation that for lower initial ground temperatures, the PBL heights generated are consistently higher (for the scheme and the data considered).

Figure 7 shows the half-normal, normal, and active-contrast plots obtained for the output parameter PBL height. These plots are used to confirm (i) the analysis outcomes from earlier methods (such as main-effect or Pareto analysis), (ii) there are no "outliers" in the dataset, and (iii) all four of the input parameters considered are statistically significant. To test representatives of the four parameters of the entire system (for the output parameter PBL height), we utilized the analysis of variance (ANOVA) approach [6] (figure 8). The good fit shown in figure 8 confirms the representativeness. However, there is a "fanning" in the scatter plots for residuals (observed minus fitted) indicating some nonlinear interactions ($P_1^n$ type) that are not resolved in the present analysis. Figure 8 confirms our hypothesis that the data set generated from the limited parameters is representative of the entire model (for the output parameter PBL height), so conclusions based on this limited data should be statistically and physically relevant for the NP89 scheme (see [14] for details).

### 5.2. Ground temperature

For Tg, figure 9 shows the main-effect, Pareto, two-factor interaction, half-normal, normal, and active-contrast, and two-factor interaction plots for the high-resolution analysis. The main-effect plot shows that deep soil moisture (SSWG) is the principal main effect, but minimum stomatal resistance (Rsmin) and surface albedo (ALB) also have significant slopes. The Pareto plot reveals that the SSWG : Rsmin interaction is also important. The active-contrast plot supports this conclusion, and confirms that the VEG : Rsmin interaction may also be significant. The normal and half-normal plots are continuous, without any break or abnormality in the center, thus satisfying the data representativeness condition. (The two-factor interaction plot is discussed in depth in the paragraph below.) The diagnostic plots (not shown) for a reduced model from the cleaned dataset verify the representativeness of both the high resolution parameters chosen and the data considered.

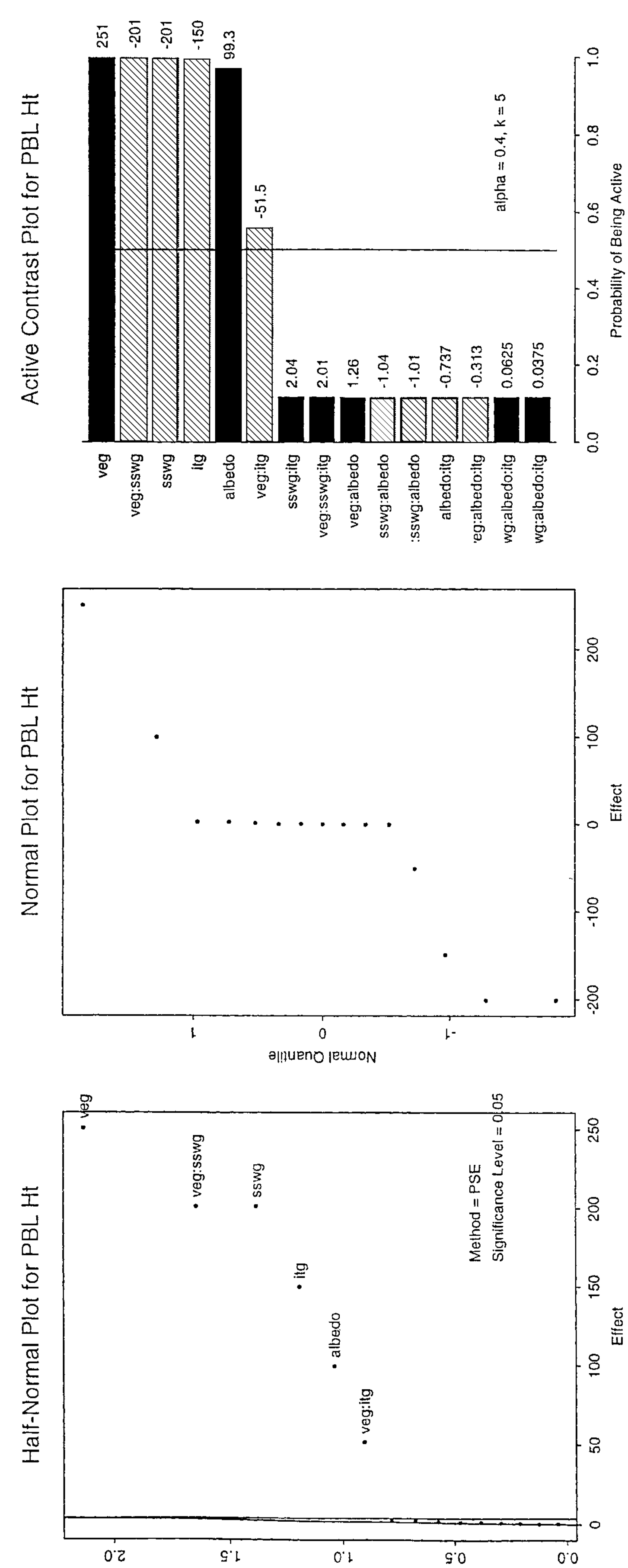


Figure 7. Half-normal, normal, and active contrast plots for PBL height (resolution V experiment).

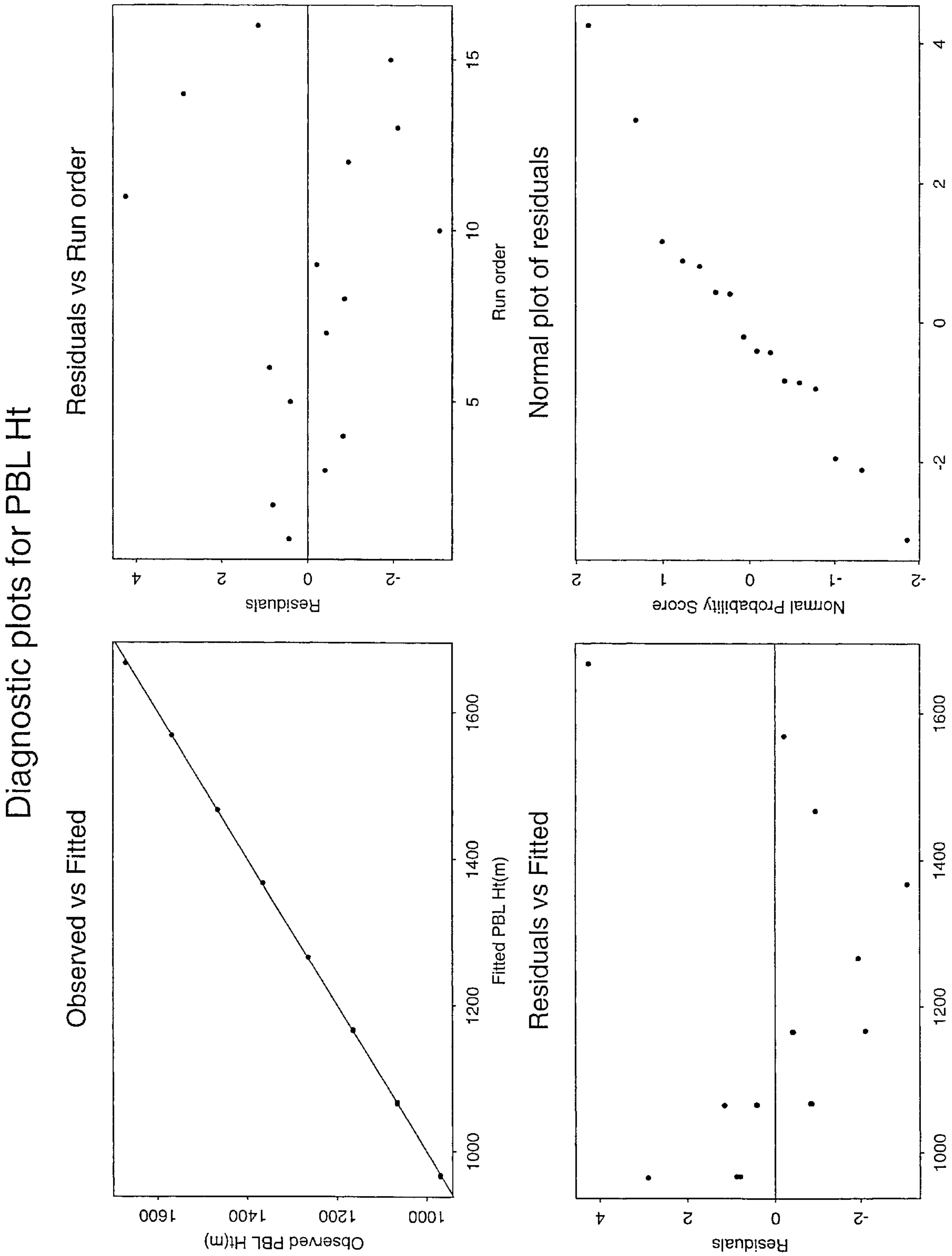


Figure 8. Diagnostic plots for checking the representativeness of the set of four input parameters chosen from the entire model to represent PBL height outcome.

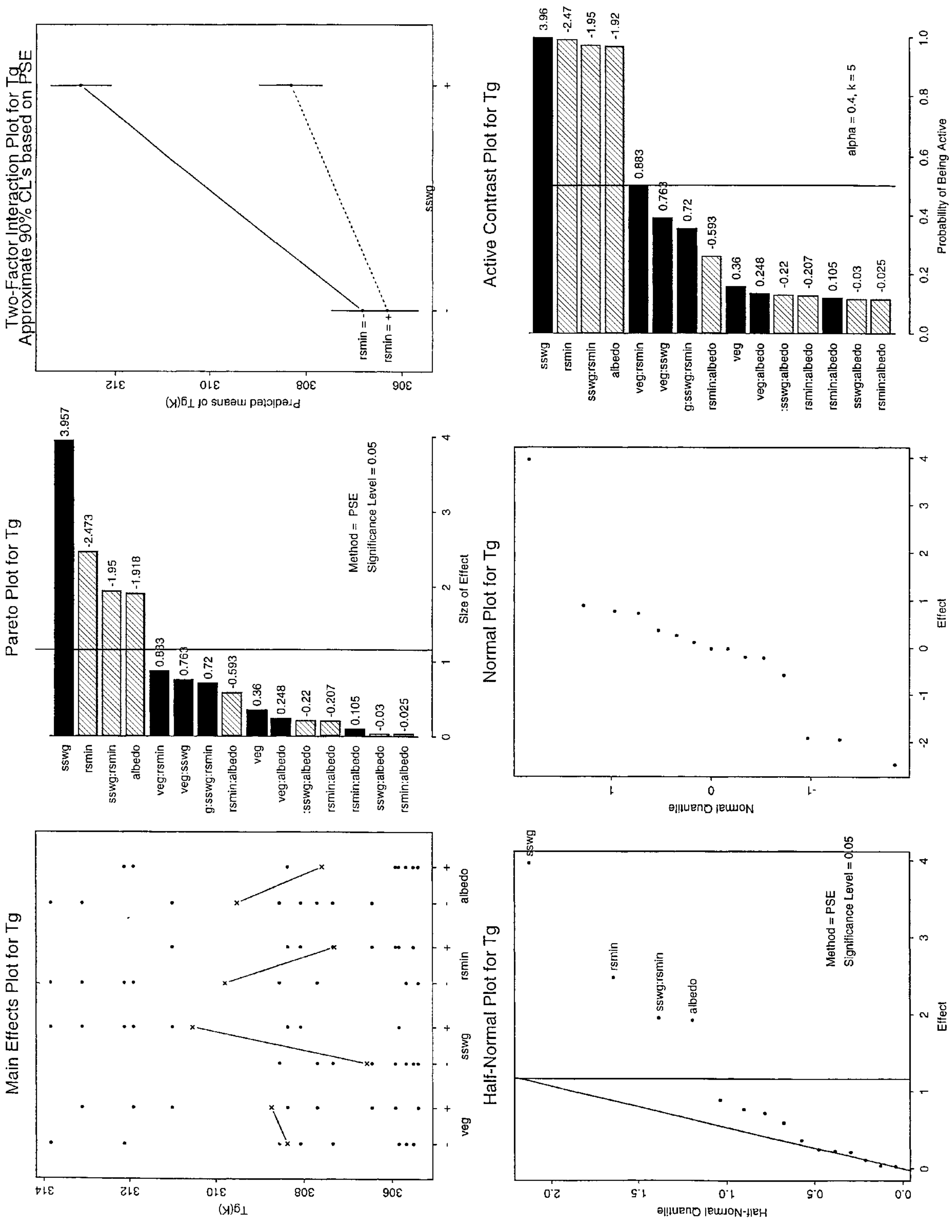


Figure 9. Graphical analyses from ground temperature dataset (resolution V experiment).

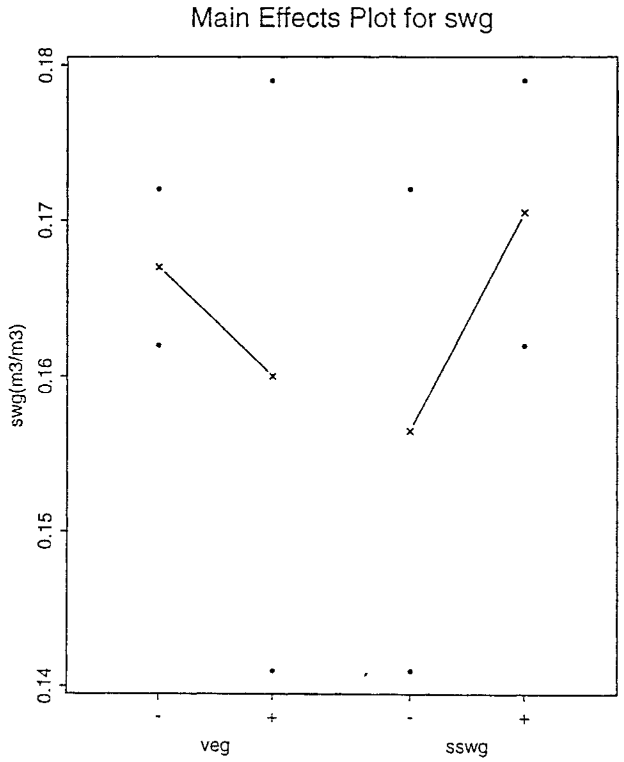


Figure 10. Main-effect plot for surface soil moisture (resolution V experiment).

The two-factor interaction plot for SSWG and Rsmin (figure 9) provides additional information. Previous studies (e.g., [2,15,16,18]) have demonstrated that vegetation modules are sensitive to Rsmin specification for PBL and hydrological processes. The figure 9 two-factor interaction plot shows that the sensitivity of Rsmin specification increases with the deep soil moisture availability. Also, Rsmin specification seems to be affecting the bare soil temperature through an interaction. The resulting change in temperature could be due to the net surface (soil and vegetation combined) temperature representation in NP89. Although a lowered SSWG reduces the predicted Tg, the transpiration rate could be more significant than the SSWG change for a short period. The association of the inputs Tg and Rsmin has attached the interaction term within the stomatal resistance estimation to the output parameter Tg. Some triple-order interactions are also seen (figure 9 active contrast plot), such as VEG : SSWG : Rsmin, but they essentially highlight the importance of the Rsmin–Tg association.

### *5.3. Soil moisture*

Figure 10 shows the main-effect plot for surface soil moisture. The two factors considered result in only a main-effect plot and no Pareto plot. It suggests that higher deep soil moisture (SSWG) tends to give higher surface moisture (SWG) values, and SWG could decrease due to seepage with the growth in vegetal cover. The slopes in the main-effect plot suggest that SSWG is more significant than VEG for determining surface soil moisture in the model. Vegetation could result in using up the water in the top layer, which is then replenished by the deep soil moisture. The resolution V result is in congruence with that from the resolution III design.

## 6. Conclusions

We have demonstrated the application of a robust methodology for the analysis of atmospheric data. The FF-based approach is recommended for model sensitivity and validation studies. With the Noilhan and Planton scheme [19] and the NCSU 1-D PBL model [1], we have demonstrated that the approach is particularly useful for dealing with a dynamically interactive system such as the atmosphere.

For the model and the NP89 scheme employed, many interesting features are apparent. We found that the term Rgl (radiation limit for initiating photosynthesis) is an important interaction term. It was also seen that subsurface soil moisture needs to be specified quite accurately, as errors in its initial value do not smooth out. Temperature and vegetation generally have opposing effects on the surface and subsurface moisture. With regard to the heat flux output parameters, the "ranking" (in order of significance) of the various input parameters in this study, which considered interactions, was different from that for JN90 [11], which assumed independence of events or an OAT approach. Radiation-related physical parameters such as albedo were found to be the most important input parameters for latent heat flux, while for sensible heat flux the vegetation-related parameters were important. These features are supported by observations from HAPEX-MOBILHY.

For the vegetation-PBL interaction, subsurface soil moisture and vegetation cover are identified as crucial input parameters, with albedo and Rgl being important radiation parameters. For temperature and humidity predictions, vegetation is the key feature. However, studies that use surface parameters in a sequential assimilation method [12] need to consider interaction as well, as highlighted by this study. For most of the output parameters, LAI and vegetal cover showed opposing tendencies. LAI seems to have a more direct impact on the water vapor flux than does vegetal cover alone.

We conducted higher resolution designs to complement the information from the resolution III design. From these we concluded that for high subsurface soil moisture availability, changes in vegetal cover may not significantly affect the simulations. Two related hypotheses are suggested: (1) results from earlier sensitivity studies may be biased, and (2) it is possible that the uncertainties in the input parameters might "correct" each other, with the net effect being more acceptable than the individual outcomes (see [18]). Overall, we found that the main effect and the interaction term tend to produce opposite effects on the outcome for the vegetative scheme considered.

We found the use of graphical analyses, such as Pareto, main-effect, two-factor interaction, active contrast, and diagnostic plots, to be beneficial for interpreting the atmospheric processes and recommend their use. We also showed that detailed statistical tests are required (even if they would provide seemingly redundant information) to check data and validate conclusions.

In summary, this study emphasized the role of interactions in atmospheric processes and suggested that both the main effects and the interactions should be considered when evaluating atmospheric data. The data could be from either simulations or actual observations. However, this brings us to some fundamental questions: Have we been able to successfully tackle these interactions in our present parameterizations? Further, how are interactions different from the "feedback" mechanisms? This study suggests that a prognostic feedback process alone does not depict interactions. Iterative solutions, empirical equations, and budget approach help reduce the errors in an OAT analysis. But when we are now developing models or particularly non-budget equations for physiological-hydrological and climate change interactions, we need to explicitly consider an interaction term in the model to realistically simulate the phenomenon.

## Acknowledgements

The authors would like to gratefully acknowledge the financial support of this study in part by the Division of International Programs (INT-93-00751) and the Atmospheric Sciences Division (ATM-92-12636) of the National Science Foundation. Many interesting interactions with Drs. Perry Halaand and Mike O'Connel of the Becton Dickinson Research Center, and Dr. J.C. Lu of the Statistics Department of the North Carolina State University, helped in formulating the experimental design. We also thank Prof. A. Prabhu, Center for Atmospheric Sciences (CAS), Indian Institute of Science, Bangalore, and Prof. U.C. Mohanty, CAS, Indian Institute of Technology, Delhi, for their review and comments and Ms. Jeanne Eichenger for her editorial assitance.

The statistical analysis was performed using S-PLUS and numerical simulations were performed on a Cray T-90 through the North Carolina Supercomputing Center.

## References

[1] K. Alapaty, J. Pleim, S. Raman, D. Niyogi and D. Byun, J. Appl. Meteor. 36 (1997) 214.
[2] K. Alapaty, S. Raman and D. Niyogi, Uncertainty in the specification of surface characteristics: A study of prediction errors in the boundary layer, Bound. Lay. Meteor. 82 (1997) 473.
[3] P. Alpert, M. Tsidulko and U. Stein, J. Atmosph. Sci. 52 (1995) 597.
[4] J. André, B. Goutorbe and A. Perrier, Bull. Amer. Meteor. Soc. 67 (1986) 138.
[5] F. Boultier, J. Mahfouf and J. Noilhan, J. Appl. Meteor. 32 (1993) 1335.
[6] G. Box, W. Hunter and J. Hunter, *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building* (Wiley, New York, 1978).
[7] J. Dorman and P. Sellers, J. Appl. Meteor. 28 (1989) 833.
[8] P.D. Haaland, *Experimental Design in Biotechnology* (Marcel Dekker, New York, 1989).
[9] P. Haaland and M. O'Connel, Technometrics 37 (1995) 82.
[10] T. Holt and S. Raman, Rev. Geophys. 26 (1988) 761.
[11] B. Jacquemin and J. Noilhan, Bound. Lay. Meteor. 52 (1990) 93.
[12] J. Mahfouf, J. Appl. Meteor. 30 (1991) 1534.
[13] G.L. Mellor and T. Yamada, J. Atmosph. Sci. 31 (1974) 1791.
[14] D.S. Niyogi, Dynamic interactions in soil–vegetation–atmosphere-transfer processes, M.S. thesis, North Carolina State University, Raleigh (1997).
[15] D.S. Niyogi and S. Raman, J. Appl. Meteor. 36 (1997) 903.
[16] D. Niyogi, S. Raman and K. Alapaty, Interaction between planetary boundary layer and vegetation: philosophy to a paradigm, in: *22nd Conf. on Fluid Mech. and Fluid Power* (Indian Inst. of Tech., Madras, 1975).
[17] D.S. Niyogi, S. Raman and K. Alapaty, Towards a dynamic parameterization of vegetation in PBL models, in: *12th Conference on Biometeorology and Aerobiology*, 27 January–2 February, Atlanta (Amer. Meteor. Soc., Boston, 1996).
[18] D. Niyogi, S. Raman and K. Alapaty, Uncertainty in specification of surface characteristics, Part 2: Hierarchy of interaction – explicit statistical analysis, Bound. Lay. Meteor. (1997) (submitted).
[19] J. Noilhan and S. Planton, Mon. Wea. Rev. 117 (1989) 536.
[20] P. Sellers and F. Hall, J. Geophys. Res. 97(D17) (1989) 19091.
[21] U. Stein and P. Alpert, J. Atmosph. Sci. 50 (1993) 2107.