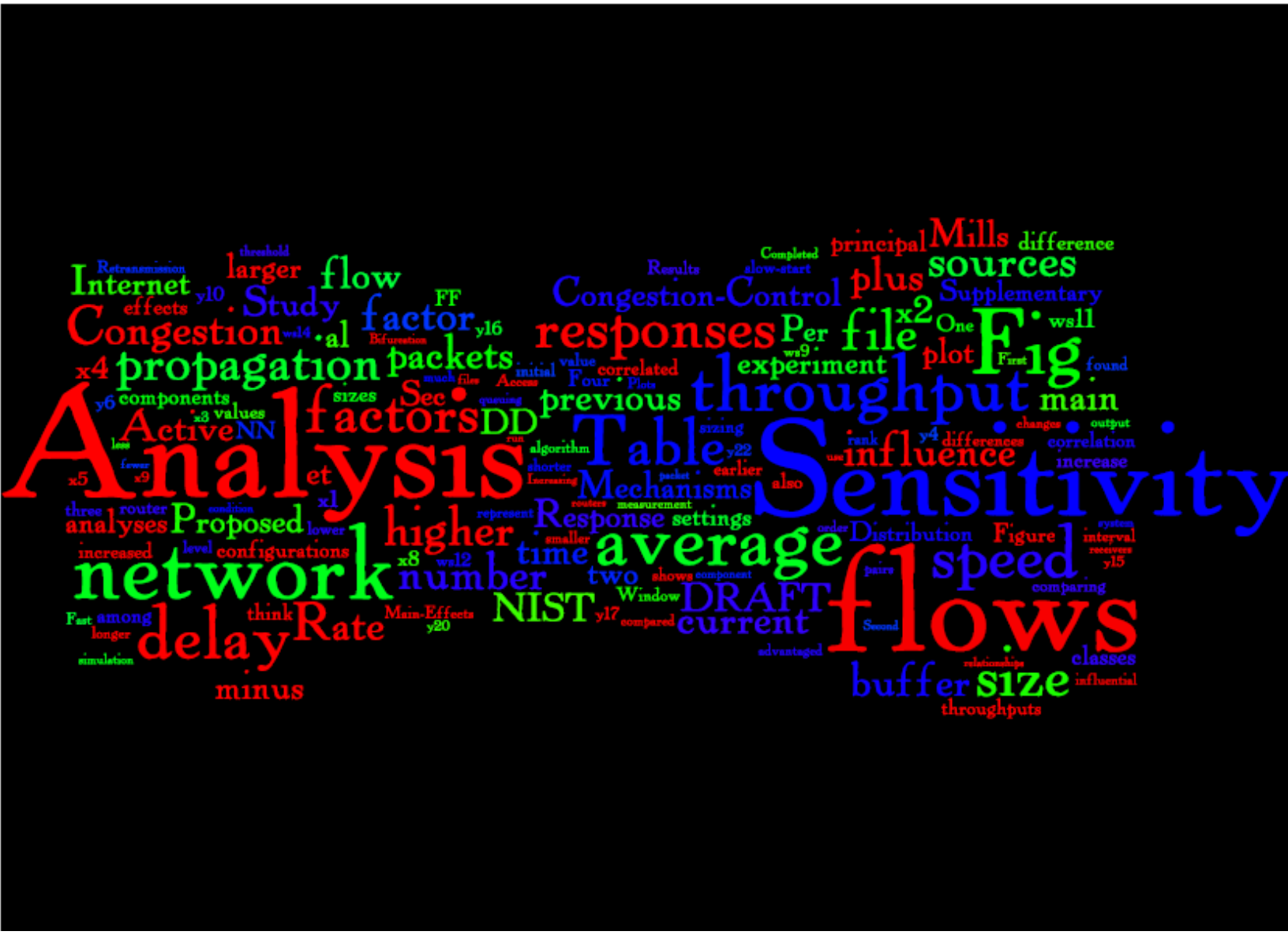
Appendix C – Supplementary Sensitivity Analysis Results



Appendix C Supplementary Sensitivity Analysis Results

As pointed out in Chapter 2 and Chapter 4, two-level experiment designs exhibit some limitations, arising from the small number of values assigned to each parameter. First, conclusions drawn from such experiments are valid only for the combinations of levels investigated. Second, the system under investigation is assumed to exhibit monotonic behavior between the lower and higher values assigned to each factor. We addressed these shortcomings in part by running several experiments to explore system behavior with different pairs of values for selected factors and by subjecting the system to various scenarios. In this section, we provide a supplementary sensitivity analysis, repeating the experiment described in Chapter 4, but selecting alternate values for the two levels used for each of the 11 input factors. We expect this second sensitivity analysis to increase our confidence in MesoNet by confirming relationships consistent with earlier analyses, by identifying any new relationships not revealed previously and by allowing us to explain variances with previously established relationships. Increasing confidence in the validity of MesoNet should also increase confidence in our other experiments comparing behavior among various congestion control algorithms proposed for use on the Internet.

We begin in Sec. C.1 by describing our experiment design, which follows the same general approach explained in Chapter 4. In Sec. C.2, we outline how we executed the simulations to collect the required data. Next, in Sec. C.3, we present results regarding MesoNet sensitivity with respect to the two selected levels for each of 11 factors. In Sec. C.4, we compare and contrast our findings with those reported for the earlier sensitivity analysis (described in Chapter 4). We conclude in Sec. C.5.

C.1 Experiment Design

We adopt a 2^{11-5} orthogonal fractional factorial design, encoded with the same template shown previously in Fig. 4-1, and we use the 11 factors identified in Table 4-10, but here we select different values for the two levels assigned to each factor. Table C-1 identifies values we chose for the two levels of each factor. The reader may compare this with Table 4-11 to identify similarities and differences in factor settings between the current and earlier sensitivity analyses. Table C-2 defines values for selected fixed parameters. The three parameters highlighted in red have changed[1] from the previous sensitivity analysis: the network now contains more sources (**baseSources** is 10^3 instead of 100) and faster sources (**Hbase** is 8 p/ms instead of 1 and **Hfast** is 80 p/ms instead of 8). We deployed our sources over the same topology (recall Fig. 3-1), possessing defined link propagation delays (see Table 3-1) and leading to specific minimum round-trip times on designated routes (recall Table 3-2).

Three factors parameterize network properties, including propagation delay (x1), speed (x2) and buffer sizing (x3). Here, most markedly, we increase the network speed by an order of magnitude over the previous sensitivity analysis. The current experiment simulates network backbone speeds approximating up to 192 Gbps, while the previous experiment topped out at 9.6 Gpbs. Further, we increase the difference in speed to eightfold between the minus and plus levels, whereas the previous experiment evinced

[1] Unless otherwise noted, we assigned fixed parameters the same values used in Chapter 4.

only a doubling. The speed increase occurs for all router classes, as shown in Table C-3, because the speed of the backbone routers determines the relative speed of other routers. The increase in network speed justifies our choice to increase the number and speed of simulated sources and receivers in order to match the additional network capacity. We also extend, over the previous sensitivity analysis, the difference in propagation delays considered, increasing by a factor of two (plus level) or reducing in half (minus level) the base delays encoded within the topology. While we use the same two buffer-sizing algorithms adopted previously, we reduce buffer size computed from the RTTxC algorithm (plus level) by ½ and we increase buffer size computed from the RTTxC/sqrt(*n*) (minus level) by a factor of 2. The net effect of this modest change in buffer-sizing algorithm is overwhelmed by the increases in network speed and changes in propagation delay, both of which influence buffer size. In the previous sensitivity analysis, backbone buffer size averaged below 10^3 packets in 32 configurations and averaged above 16×10^3 packets in the other 32 configurations, reaching a maximum of just over 65×10^3 packets. In the current experiment (see Table C-4) backbone buffer size averaged below 10^3 packets in only 8 configurations, while exceeding 65×10^3 packets in 24 configurations. In general, the current experiment provides increased buffers under most configurations because the network speed has increased substantially.

Table C-1. Two-Level Settings for Each of 11 Factors in Sensitivity Analysis

	Factor	Factor Name	Plus Level	Minus Level
Network Factors	x1	Propagation Delay	2.5 times base delay	0.5 times base delay
	x2	Network Speed	16×10^3 p/ms	2×10^3 p/ms
	x3	Buffer Sizing	RTTxCx(Qfactor = 0.5)	RTTxC/sqrt(*n*)x(Qfactor = 2)
User Factors	x4	File Size	200 packets	25 packets
	x5	Think Time	10 seconds	1.25 seconds
	x6	Large File Probability	0.04	0.005
Source & Receiver Factors	x7	Fast Host Probability	0.80	0.10
	x8	Number of Sources	3 times base sources	1 times base sources
	x9	Source Distribution	0.1/0.3/0.6	0.3/0.1/0.6
	x10	Receiver Distribution	0.1/0.3/0.6	0.3/0.1/0.6
Protocol Factors	x11	Initial Slow-Start Threshold	1.07×10^9 packets	20 packets

User behavior is defined by three factors: average file size (x4), average think time (x5) and probability of transferring a (**Fx** = 10 times) larger file (x6). We increased the spread among the plus and minus values for each factor, when compared with the values chosen in Chapter 4. These choices implement a general strategy to increase the distance between the plus and minus settings for each factor in order to determine if such increases reveal strengthened relationships between factors and responses. For example, for the sole protocol factor, initial slow-start threshold (x11), we lower the minus level from 43 packets in the previous sensitivity analysis to 20, while keeping the plus value at the same arbitrarily high value used in Chapter 4.

Table C-2. Selected Fixed Parameters

Parameter	Name	Fixed Value
BBspeedup	Backbone Router Speed Multiplier	1
R2	POP Router Speed Divisor	4
R3	Access Router Speed Divisor	10
Bfast	Fast Access Router Speed Multiplier	2
Bdirect	Directly Connected Access Router Speed Multiplier	10
Hbase	Speed of Normal Network Interfaces	8 p/ms
Hfast	Speed of Fast Network Interfaces	80 p/ms
baseSources	Base Number of Sources Per Access Router	10^3
Fx	Large File Size Multiplier	10
alpha	File Size Shape Parameter	1.5

Table C-3. Router Speeds (p/ms) by Router Class for Each Level of Network Speed (x2)

Router Type	Plus	Minus
Backbone	16×10^3	2×10^3
POP	4×10^3	500
Typical Access	400	50
Fast Access	800	100
Directly Connected Access	4×10^3	500

Table C-4. Average Buffer Size (in packets) by Router Class for Specific Combinations of Propagation Delay (x1), Network Speed (x2) and Buffer-Sizing Algorithm (x3)

x1	x2	x3	Backbone Router Buffers (avg.)	POP Router Buffers (avg.)	Access Router Buffers (avg.)
-	-	-	368	140	48
+	-	-	3.36×10^3	1.306×10^3	435
-	+	-	4.453×10^3	1.789×10^3	600
+	+	-	12.335×10^3	5.308×10^3	1.751×10^3
-	-	+	20.800×10^3	5.200×10^3	827
+	-	+	102.182×10^3	25.545×10^3	4.062×10^3
-	+	+	166.400×10^3	41.600×10^3	6.614×10^3
+	+	+	817.455×10^3	204.363×10^3	32.492×10^3

The remaining factors determine the speed, number and distribution of the sources and receivers deployed in the topology. In this experiment, fast network interfaces for sources and receivers operate at a maximum of 80 p/ms (960 Mbps), while we parameterized slower network interfaces to operate at a maximum of 8 p/ms (96 Mbps). The probability that a given source or receiver has a fast network interface is determined by the fast host probability (x7), which we set to either .8 (plus) or .1 (minus), a .7 difference compared with the .2 difference used in the earlier sensitivity analysis. The previous experiment determined that this factor had little influence on system responses so we decided to increase the difference in probabilities in order to probe the invariance of this finding.

A combination of three factors, number of sources (x8) and distribution of sources (x9) and receivers (x10), determine the probability that flows go between specific combinations of access router classes: directly connected to directly connected (DD), directly connected to fast (DF), directly connected to normal (DN), fast to fast (FF), fast to normal (FN) and normal to normal (NN). We call these flow classes. Table C-5 shows the influence of these factors on the number and distribution of sources in the topology, while Table C-6 gives similar information regarding the number and distribution of receivers.

Table C-5. Relation between Factors and Number and Distribution of Sources

x8	x9	x10	Total Sources	% under D Routers	% under F Routers	% under N Routers
1	+	+	67.500×10^3	16	37.33	46.67
3	+	+	202.395×10^3	16	37.35	46.64
1	-	-	113.700×10^3	9.5	7.4	83.11
3	-	-	341.072×10^3	9.5	7.4	83.11
1	+	-	67.500×10^3	16	37.33	46.67
3	+	-	202.396×10^3	16	37.33	46.67
1	-	+	113.700×10^3	9.5	7.4	83.11
3	-	+	341.072×10^3	9.5	7.4	83.11

Table C-7 reports the influence of x8, x9 and x10 on the probability of various flow classes. Four combinations of parameters, the rows highlighted in purple, represent traffic patterns consistent with Web browsing augmented with some peer-to-peer (P2P) exchanges. Two parameter combinations, the rows highlighted in rose, represent traffic patterns with a slightly increased proportion of Web browsing compared to P2P exchanges. The remaining (white) rows show traffic patterns shifted substantially toward P2P traffic. The probabilities in Table C-7 represent a shift toward more Web-based traffic patterns when compared with the previous sensitivity analysis (see Table. 4-15), which had an even balance of configurations with Web and P2P traffic patterns. In addition, the P2P configurations in the current experiment represent a somewhat more

pronounced probability of DN, FN and NN flows. Further, the current experiment increases the proportion of DD flows for all configurations. Finally, while increasing the number of sources and receivers by about an order of magnitude over the previous sensitivity analysis, the current experiment expands the difference in number of sources and receivers among the configurations.

Table C-6. Relation between Factors and Number and Distribution of Receivers

x8	x9	x10	Total Receivers	% under D Routers	% under F Routers	% under N Routers
1	+	+	270.000×10^3	16	37.33	46.67
3	+	+	809.856×10^3	16	37.34	46.66
1	-	-	454.800×10^3	9.5	7.4	83.11
3	-	-	136.437×10^4	9.5	7.4	83.11
1	+	-	454.800×10^3	9.5	7.4	83.11
3	+	-	136.437×10^4	9.5	7.4	83.11
1	-	+	270.000×10^3	16	37.34	46.67
3	-	+	809.856×10^3	16	37.34	46.67

Table C-7. Relation between Factors and Distribution of Flow Classes

x8	x9	x10	% DD Flows	% DF Flows	% DN Flows	% FF Flows	% FN Flows	% NN Flows
1	+	+	2.56	11.95	14.93	13.94	34.84	21.78
3	+	+	2.56	11.95	14.93	13.94	34.84	21.76
1	-	-	0.9	1.4	15.79	0.55	12.28	69.08
3	-	-	0.9	1.4	15.79	0.55	12.28	69.08
1	+	-	1.52	4.73	17.73	2.76	34.48	38.79
3	+	-	1.52	4.73	17.73	2.76	34.48	38.79
1	-	+	1.52	4.73	17.73	2.76	34.48	38.79
3	-	+	1.52	4.73	17.73	2.76	34.48	38.79

To summarize the differences from the earlier sensitivity analysis: we increased network speed and size by an order of magnitude, we stretched the range of parameter values covered by the plus and minus settings of each factor, and we shifted the traffic patterns slightly to generate more DD flows and to give a higher prominence to Web browsing activity over P2P exchanges. These changes provided a very different set of configurations under which we could evaluate the relationship between model input

parameters and responses. We made no changes in fixed parameters controlling the simulation duration or the length of measurement intervals.

To permit ready comparison between results from both the earlier and current sensitivity analyses, we elected to measure the same responses in both experiments. One set of responses (repeated here as Table C-8) measured macroscopic behavior of the entire network and a second set of responses (see Table C-9) measured average instantaneous throughput on each of the six flow classes.

Table C-8. Responses Characterizing Macroscopic Network Behavior

Response	Definition
y1	Active Flows – flows attempting to transfer data
y2	Proportion of potential flows that were active: Active Flows/All Sources
y3	Data packets entering the network per measurement interval
y4	Data packets leaving the network per measurement interval
y5	Loss Rate: y4/(y3+y4)
y6	Flows Completed per measurement interval
y7	Flow-Completion Rate: y6/(y6+y1)
y8	Connection Failures per measurement interval
y9	Connection-Failure Rate: y8/(y8+y1)
y10	Retransmission Rate
y11	Congestion Window per Flow
y12	Window Increases per Flow per measurement interval
y13	Negative Acknowledgments per Flow per measurement interval
y14	Timeouts per Flow per measurement interval
y15	Smoothed Round-Trip Time
y16	Relative queuing delay: y15/(x1x41)

Table C-9. Responses Characterizing Average Instantaneous Throughput by Flow Class

Response	Definition
y17	Average Throughput for Active DD Flows
y18	Average Throughput for Active DF Flows
y19	Average Throughput for Active DN Flows
y20	Average Throughput for Active FF Flows
y21	Average Throughput for Active FN Flows
y22	Average Throughput for Active NN Flows

C.2 Experiment Execution and Data Collection

The experiment plan required 64 simulation runs, each simulating a different combination of factor settings, as constructed by mapping values from Table C-1 into the template shown in Fig. 4-1. We had 48 physical processors on which we could run our experiments, so we conducted simulations in parallel. We were sharing these processors with other projects, so we could not always use all of the available processors. Below, we give a brief discussion of the resource requirements for the simulations.

Table C-10 reports the characteristics of the 48 processors[2] available for our sensitivity analysis. Since MesoNet is implemented in SLX, each of the processors had access to an SLX simulation environment. SLX comes in two varieties: one configured to

[2] These 48 processors amounted to 6 servers that each had 8 processor cores.

run in a 32-bit address space and one configured to run in a 64-bit address space. We chose to run all our simulations using the 32-bit version of SLX because our simulations could fit easily within a 32-bit address space and 32-bit simulation runs faster than 64-bit simulation.

Table C-10. Configuration of Compute Servers for Simulations

Node	Processor Count	Speed (GHz)	Processor Type	Memory (GB)	Operating System
ws9	8	2.6	Dual-Core AMD Opteron 8218	32	Windows Server 2003 R2 x64 Edition SP2
ws10	8	2.6	Dual-Core AMD Opteron 8218	32	Windows Server 2003 R2 x64 Edition SP2
ws11	8	3.0	Dual-Core AMD Opteron 8222SE	32	Windows Server 2003 R2 x64 Edition SP2
ws12	8	3.0	Dual-Core AMD Opteron 8222SE	32	Windows Server 2003 R2 x64 Edition SP2
ws13	8	3.0	Dual-Core AMD Opteron 8222SE	32	Windows Server 2003 R2 x64 Edition SP2
ws14	8	3.0	Dual-Core AMD Opteron 8222SE	32	Windows Server 2003 R2 x64 Edition SP2

We executed the simulations continuously over about three days, starting as many simulations as available processors and then initiating a new simulation when one finished. Table C-11, organized into four three-column, color-coded groups of 16 simulations, reports the number of processor hours required by each simulation on a specified compute-server node. The average processor time for a simulation was 46.4 hours, about 5.5 times more than the average processor time required for the earlier sensitivity analysis, which simulated a slower and smaller network. The average memory used for a simulation was 1.1 Gbytes, nearly a tenfold increase over the earlier sensitivity analysis. We collected and summarized data using the same techniques adopted for the earlier sensitivity analysis. Refer to Sec. 4.3.2 for the details.

Table C-11. Execution Time (Processor Hours) Required for Each Simulation Run

Run	Node	Time	Run	Node	Time	Run	Node	Time	Run	Node	Time
1	ws9	36.6	17	ws11	11.4	33	ws14	36.4	49	ws11	7.8
2	ws9	14.8	18	ws11	1.3	34	ws14	23.4	50	ws11	3.7
3	ws9	14.7	19	ws11	1.3	35	ws14	29.8	51	ws11	3.2
4	ws9	70.3	20	ws11	8.0	36	ws14	46.8	52	ws12	7.4
5	ws9	27.7	21	ws11	2.6	37	ws14	17.8	53	ws12	1.9
6	ws9	32.2	22	ws11	6.8	38	ws14	34.6	54	ws9	14.9
7	ws9	56.8	23	ws11	4.9	39	ws14	96.0	55	ws9	15.8
8	ws9	19.8	24	ws11	2.4	40	ws11	14.2	56	ws9	1.8
9	ws10	31.0	25	ws12	22.9	41	ws11	30.7	57	ws14	19.5
10	ws10	47.1	26	ws12	30.3	42	ws11	42.9	58	ws14	36.3
11	ws10	233.7	27	ws12	48.7	43	ws11	218.9	59	ws11	100.0
12	ws10	156.8	28	ws12	20.7	44	ws13	78.4	60	ws11	15.6
13	ws10	42.5	29	ws12	35.8	45	ws14	36.5	61	ws9	43.3
14	ws10	32.7	30	ws12	11.0	46	ws13	31.4	62	ws12	27.7
15	ws10	97.0	31	ws12	10.7	47	ws11	172.8	63	ws12	23.1
16	ws10	238.6	32	ws12	70.2	48	ws11	239.0	64	ws12	55.1

C.3 Results

Below, we report results from subjecting (22 x 64 =) 1408 responses to three treatments: correlation analysis, principal components analysis and main effects analysis. We also employed some exploratory analyses, as used in Chapter 4. We discuss each analysis in turn.

C.3.1 Correlation Analysis

Given 64 average values (one per run) for 22 responses, we conducted a correlation analysis to investigate the degree to which pairs of responses are linearly correlated. We used the same techniques applied in the earlier sensitivity analysis (Sec. 4.4). We began by generating a scatter plot and computing the correlation for each pair of responses, as plotted together in Fig. C-1, which should be interpreted as explained earlier in Sec. 4.1.3.3. Of particular interest, correlations with magnitudes of .8 and above are colored red, magnitudes between .3 and .79 are blue and magnitudes below .3 are green.

Figure C-1. Combined Matrix of Scatter Plots and Correlation Values for 22 Responses

Scanning Fig. C-1 reveals some mutual correlations, for example among responses y17 through y22, which represent throughput for various flow classes. The figure also reveals some strongly correlated pairs: y1 and y2 (active flows and proportion of sources that are active), y3 and y4 (packets input and output), y5 and y10 (loss rate and retransmission rate), y8 and y9 (connection failures and connection-failure rate) and y13

and y14 (negative acknowledgments and timeouts). A few responses (e.g., y6 and y16) appear largely uncorrelated with other responses. Comparing Fig. C-1 with Fig. 4-13 from the earlier sensitivity analysis, one finds fewer strong correlations overall in the current experiment. Also of note, unlike the previous experiment, which correlated throughput for flow classes into three groupings, the current experiment shows strong positive correlation in throughput among all flow classes.

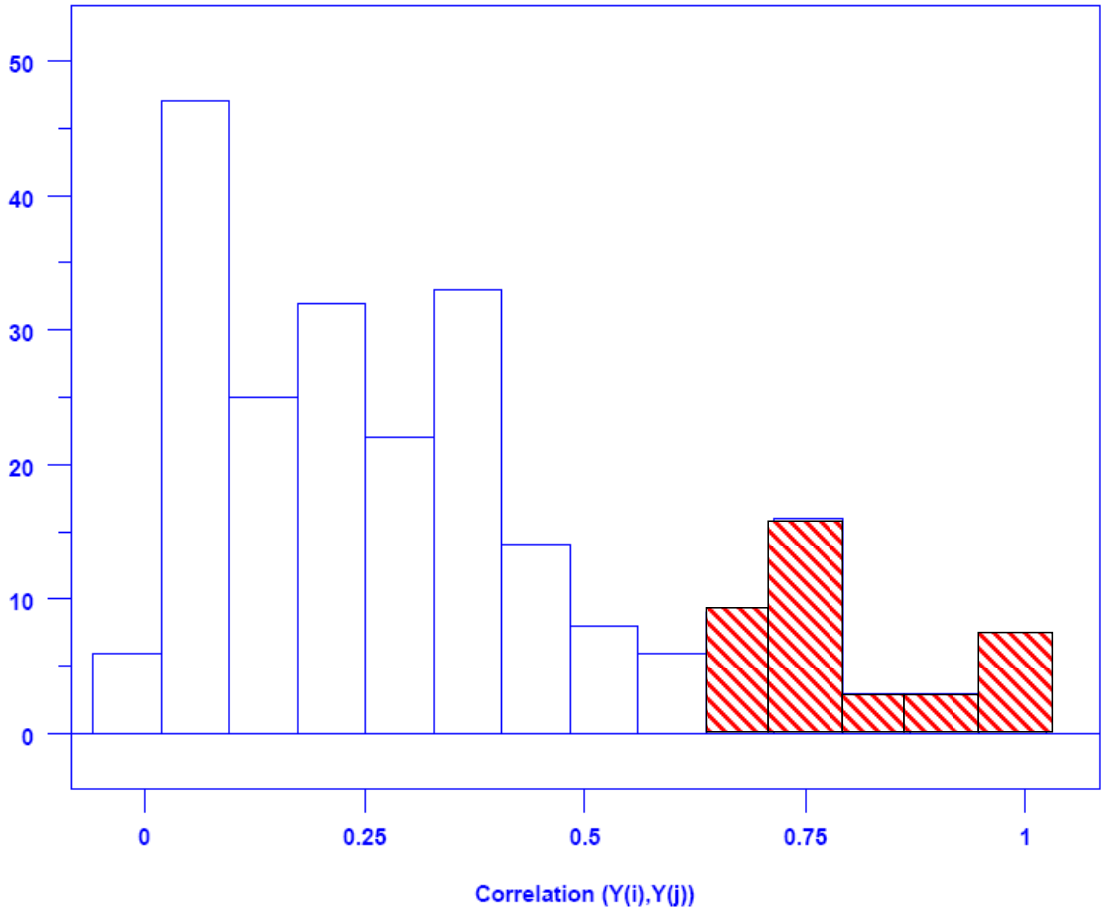


Figure C-2. Frequency Distribution of the Absolute Value of Correlations for All Pairs of Responses

Fig. C-2 shows the frequency distribution of the absolute value of correlations for all pairs of responses. The five, highlighted bins represent correlations with magnitude greater than 0.65, which we chose (consistent with our choice in Chapter 4) as the cutoff for correlations to be considered significant. Comparing Fig. C-2 with Fig. 4-14 from the earlier analysis indicates that the current analysis has 36 pairs that are significantly correlated, while the earlier analysis found 42 such pairs. We show (Fig. C-3) the 36 significantly correlated pairs as an index-index plot, ordering the indices on both axes as ordered in Fig. 4-14 in order to facilitate comparison. Comparing Fig. C-3 with Fig. 4-14 confirms that throughput among flow classes, previously organized into three groups ([y17], [y18, y20] and [y19, y21, y22]) are now mutually correlated. Other changes can also be noted. For example, y15 (round-trip time) and y16 (relative queuing delay) are no longer correlated. Further, y8 (connection failures), y9 (connection-failure rate) and y14 (timeouts) remain correlated but are no longer correlated with y5 (loss rate) and y10

(retransmission rate). In the current experiment, y5 (loss rate) and y10 (retransmission rate) are correlated with the y1 and y2 (active flows and proportion of sources active); the number of active flows was not a factor in the previous sensitivity analysis. This difference likely arises because the current experiment uses a faster network, requiring more active flows to generate load. In the current sensitivity analysis, congestion seems driven by the number of active flows. Throughput among all flow groups now seems correlated and thus likely driven by some common factors, but note that congestion window size (y11) is not correlated with throughput on DD flows (y17). In fact, the y11-y17 correlation is 0.60, which falls just below our cutoff (0.65). Finally, round-trip time (y15) and queuing delay (y16) are now uncorrelated. These correlation changes are considered in the discussion (Sec. C.4) after assessing the main factors that influence model responses.

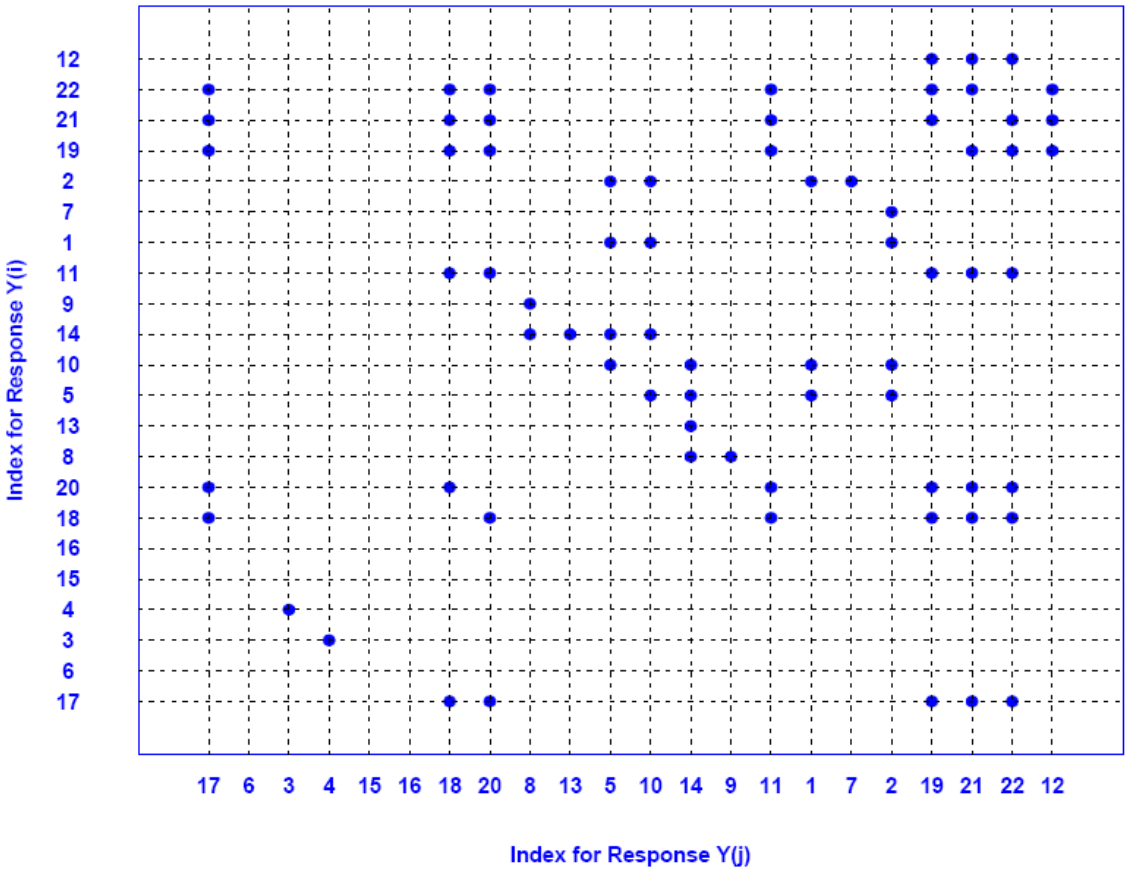


Figure C-3. Index-Index Plot for Correlation Pairs where |Correlation (Y_i, Y_j)| > 0.65

C.3.2 Principal Components Analysis

Given the changes noted in the correlation analysis, we should expect to see some changes in the principal components analysis (PCA) as well. Fig. C-4 shows the PCA for all 22 responses generated in the current experiment. The first four principal components account for 95 % of the response variance and thus we select these components for

further examination, as given in Fig. C-5, where we plot the relevant weight vectors. Comparing Fig. C-5 with Fig. 4-17 from the previous sensitivity analysis illustrates that the top four principal components have changed configurations. In fact, the principal components from the current experiment appear more difficult to interpret than those from the previous experiment. To provide additional information, we introduce Fig. C-6, containing four main effects plots, one per principal component, analogous to Figs. 4-26 through 4-29. Fig. C-6 shows that many of the same factors influence the top four principal components, suggesting we will be unable to find a clear and satisfying interpretation. We will do the best we can.

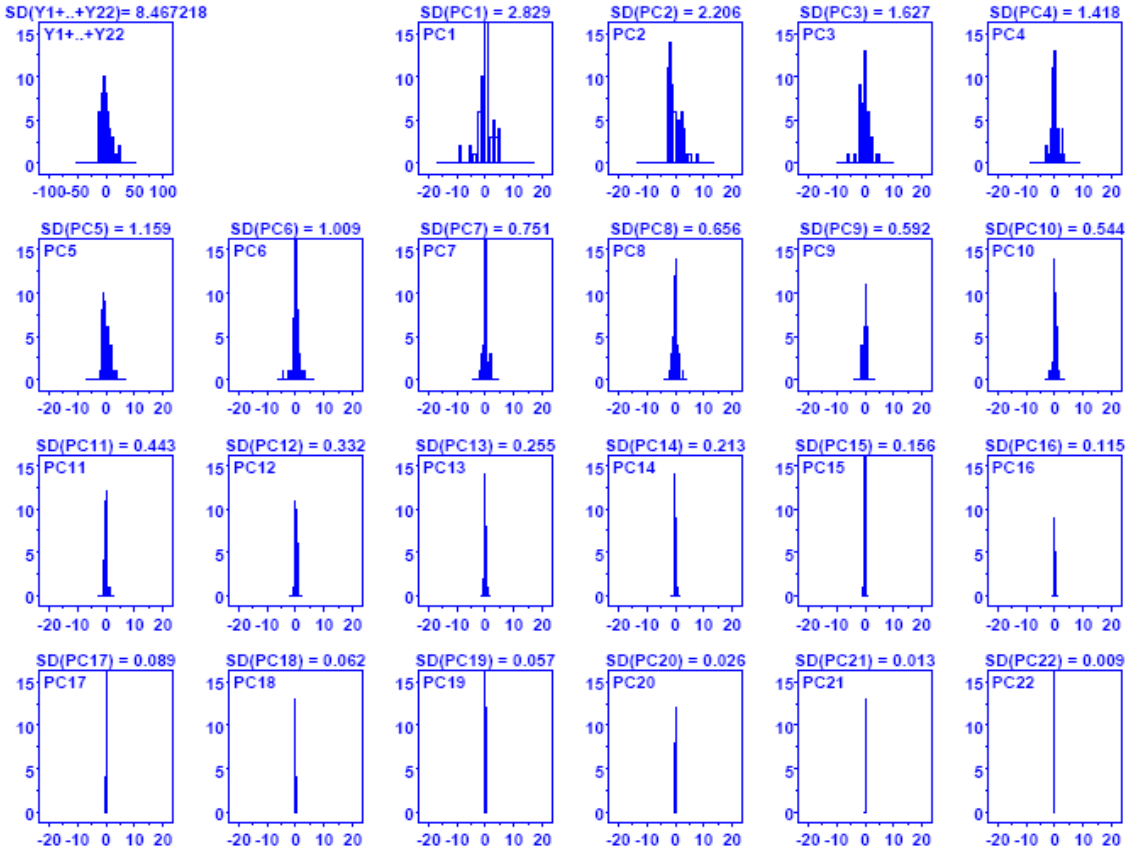


Figure C-4. Histograms for 22 Principal Components (x axis of each sub-plot identifies bins of normalized component values ranging from -20 to +20 and y axis the count of values within each bin). Above each sub-plot is the standard deviation in the data accounted for by the Principal Component. The first sub-plot gives the distribution of the normalized responses.

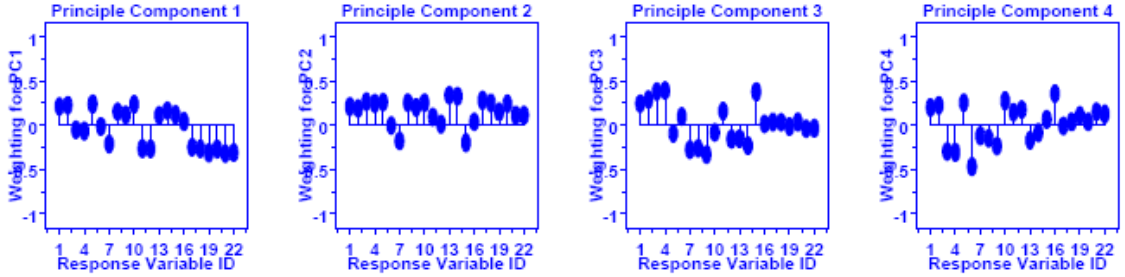


Figure C-5. Weight Vectors for the First Four Principal Components

The first principal component groups 12 responses in four main categories: active flows (y1 and y2), loss and retransmission rates (y5 and y10), congestion window size and increase rate (y11 and y12) and throughput among all flow classes (y17-y22). The main effects plot for PC1 indicates that slower network speed, longer propagation delay, shorter think time and more sources lead to a positive component value – a negative component value is produced by the opposite setting for these factors. From this, we may infer that higher network congestion yields a positive value for PC1 and lower network congestion yields a negative value. This inference suggests that PC1 represents the influence of network congestion (y1, y2, y5 and y10) on congestion window size (y11) and increase rate (y12), which determines throughput on flows of all classes (y17-y22).

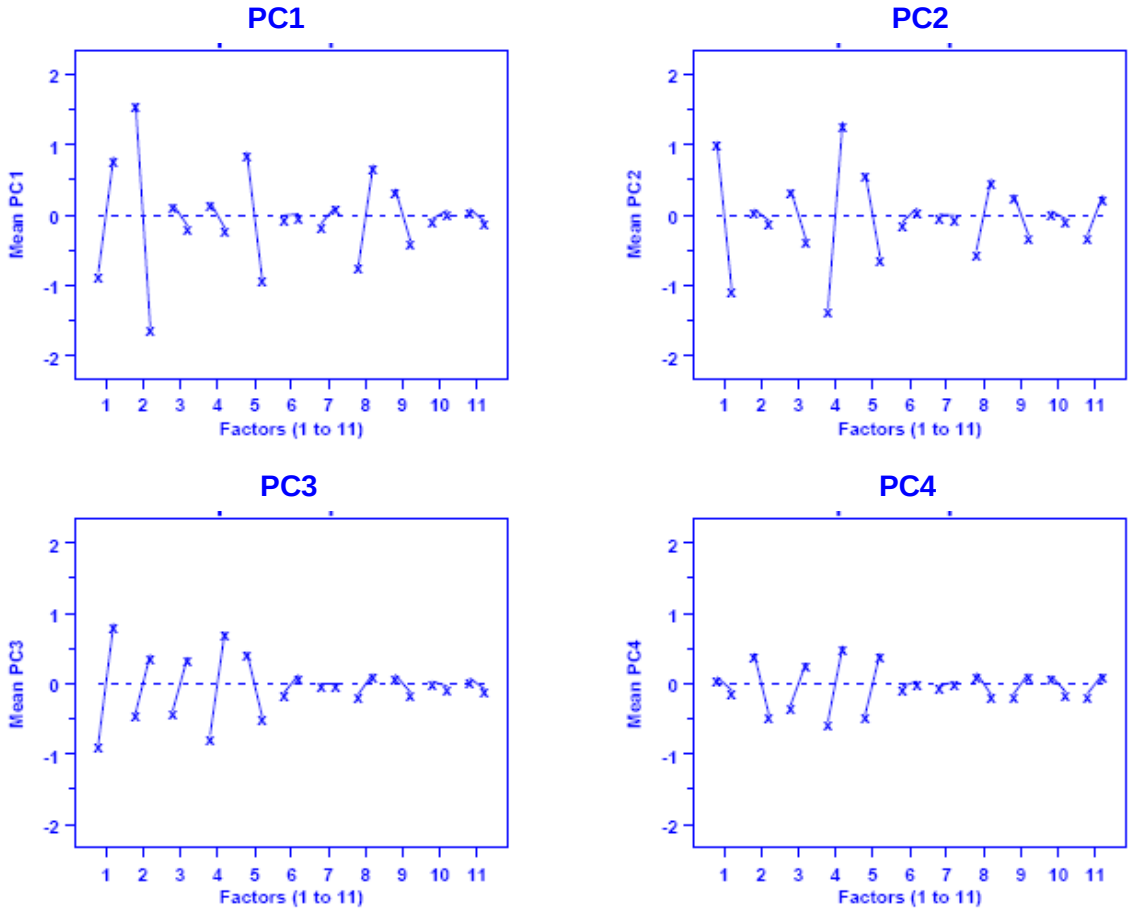


Figure C-6. Main Effects Plots for Top Four Principal Components

The second principal component appears to characterize throughput on more advantaged DD (y17), DF (y18) and FF (y20) flows. Provided congestion is not too heavy, larger file sizes (x4) and shorter propagation delays (x1) should lead to higher throughputs, especially for advantaged flows. The main effects plot for PC2 suggests that higher throughputs are coincident with a positive value of the component. At the same time, the network would transport more packets (y3 and y4) for more flows (y1 and y2), which would lead to more losses (y5) and retransmissions (y10), especially for less advantaged flows (y19, y21 and y22), which would experience more negative acknowledgments (y13) and timeouts (y14).

The third principal component appears to characterize queuing delay (y16), which is grouped together with the number of packets flowing in (y3) and out (y4) of the network. Referring to the main effects plot for PC3, longer propagation delay (x1), faster network speed (x2), larger buffer sizes and larger file sizes (x4) generate a positive value for the component. We can infer that such conditions permit the larger network buffers to hold more packets, leading to longer queuing delays.

The fourth principal component appears to characterize network throughput measured in terms of flows completed (y6) and packets transferred (y3 and y4). Shorter files sizes (x4) combine with higher network speed (x2) and shorter think times (x5) to permit more flows to be completed per unit time (y6). Comparing the main effects plot for PC4 from Fig. C-6 with the main effects plot for PC4 from Fig. 4-29 reveals similarity.

The forgoing discussion illustrates that the PCA conducted for the second sensitivity analysis produced results more difficult to interpret than was the case for the original sensitivity analysis reported in Chapter 4. For this reason, we postpone until the discussion (Sec. C.4) further consideration of the principal components.

C.3.3 Exploratory Analysis of y7-y22 Scatter Plot Bifurcation

In Fig. 4-10 of Sec. 4.1.6, we reported a scatter plot of y7 (flow completion rate) vs. y22 (throughput on NN flows) that showed a bifurcation. We used an exploratory technique, altering the plot symbols to represent minus and plus settings for each factor, to discover that the bifurcation arose due to factor x4 (average file size). Shorter file sizes resulted in higher completion rates (y7) and yet led to lower average throughputs for NN flows (y22). This made sense because shorter files spend a higher percentage of their transfer in TCP slow start, during which throughputs are lower. On the other hand, shorter files are generally transferred more quickly because they involve fewer packets. Since shorter files take less time, more flows complete per unit of time and the flow completion rate is higher. Longer files spend a higher percentage of their transfer beyond TCP slow start, during which throughputs are higher. On the other hand, longer files require transferring more packets, taking more time and completing fewer flows per unit of time. Since we increased the distance between the minus and plus file sizes (from 50/100 to 25/200), we expected the bifurcation to appear in enhanced form in this sensitivity analysis.

Fig. C-7 shows twelve y7-y22 scatter plots generated from the 64 simulations in the current experiment. The first scatter plot contains the bifurcation data. Each of the next 11 plots distinguishes the minus and plus level settings for a given factor. The plots clearly identify average file size (x4) as the factor responsible for the bifurcation. Comparing Fig. 4-10 with Fig. C-7 shows that, as expected, the bifurcation is enhanced in the current sensitivity analysis.

C.3.4 Main Effects Analysis

In this section, we provide main effects plots generated from the data captured in the current experiment. To facilitate comparison we plot main effects for the same responses used in the original sensitivity analysis, including those listed in Table 4-19, and adding the average congestion window size, used previously as an example in Fig. 4-9. Analyzing the same responses should allow us to identify similarities and differences in factor-response relationships between the two sensitivity analyses. We expect that most

of the factor-response relationships will remain unchanged, but differences in the correlation and principal components analyses suggest that some relationships might be different. We organize the exposition into four categories, responses related to congestion, responses related to delay, responses related to macroscopic throughput and responses related to throughput for advantaged flows, as shown in Table C-12.

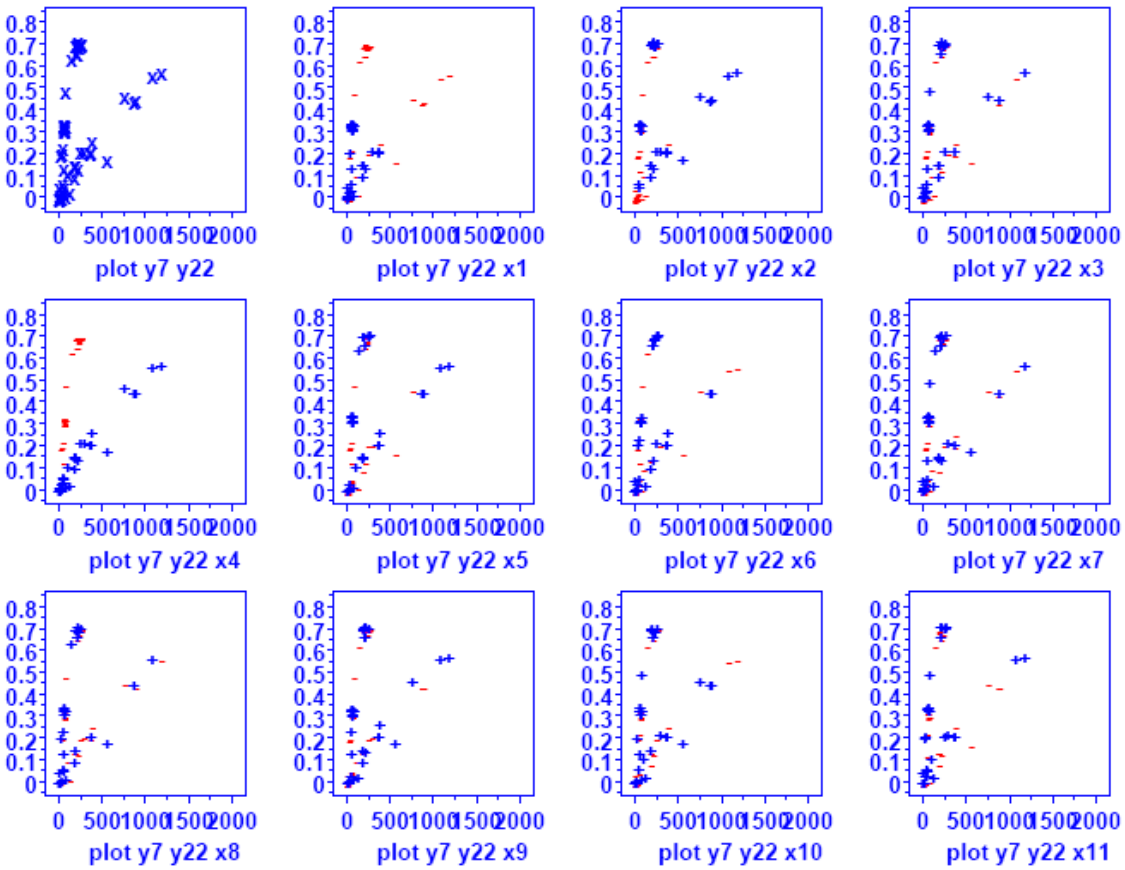


Figure C-7. Y-Y-X plot for Responses y7 and y22

Table C-12. Responses Selected for Investigation in Sensitivity Analysis

	Response	Definition
Congestion	y1	Average number of active flows
	y10	Average retransmission rate
	y11	Average congestion window size
	y22	Average instantaneous throughput for NN flows
Delay	y15	Average smoothed round-trip time
Macroscopic Throughput	y4	Average number of packet output per measurement interval
	y6	Average number of flows completed per measurement interval
Advantaged Flows	y17	Average instantaneous throughput for DD flows
	y20	Average instantaneous throughput for FF flows

C.3.4.1 Congestion-Related Responses. We begin by examining the main factors influencing the average number of active flows (y1). Fig. C-8 shows the relevant main effects plot, which can be compared with Fig. 4-18. We find the same main factors influencing the number of active flows in both experiments, though the order of the factors shifts slightly. The main factors appear to fall into three categories: (a) number of sources underneath N-class access routers, (b) duration for which flows remain active and (c) idle interval for those sources. The number of sources (x8) claims the main influence, followed by the average file size (x4). The longer it takes to transfer files, the more likely flows are to be active. The duration of transfer time is influenced not only by file size, but also by network speed[3] (x2) and congestion, which is influenced by number (x8) and distribution (x9) of sources and by average think time (x5). These relationships are evident in Figs. C-8 and 4-18.

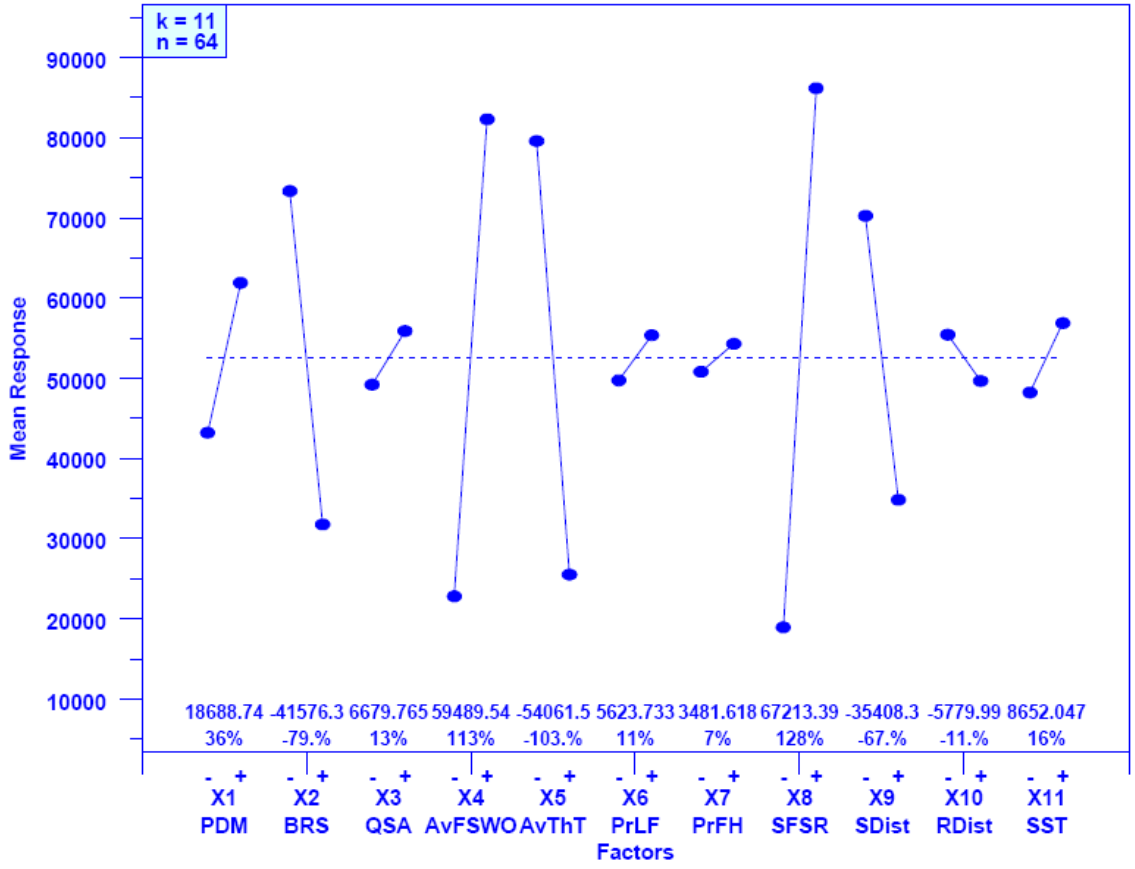


Figure C-8. Main-Effects Plot for Response y1 (Average Number of Active Flows) – x axis lists 11 model parameters with a – and + value for each parameter, y axis gives average number of active flows, two averages are given for each parameter, one average when the parameter is set to its – level and one average when the parameter is set to its + level, and a line connects the pair of average sizes for each parameter. Dashed line is the overall average number of active flows (about 52.250 x 10^3 here)

[3] Recall that in the previous sensitivity analysis we miscoded network speed (minus was higher network speed) and distribution of sources (plus was a more P2P-like traffic pattern). In this sensitivity analysis the levels were properly coded, so care should be taken in comparing the slope for these factors on the main effects plots in Chapter 4 against the main effects plots in Appendix C.

Fig. C-9 gives the main-effects plot for retransmission rate (y10). Comparing this with Fig. C-8 shows that the same factors influence both the number of active flows and retransmission rate. This mirrors the same relationship found in Chapter 4, where Fig. 4-18 and Fig. 4-19 also show the same influential factors. Figs. C-9 and 4-19 exhibit one main difference: buffer sizing algorithm does not play as significant a role in Fig. C-9. This makes sense because buffer sizing for the former sensitivity analysis led to very small buffer sizes under the minus setting. Buffer sizes were not as constrained under the minus setting in the current sensitivity analysis. One other difference can also be discerned: the number of sources plays a larger role and the distribution of sources a smaller role in Fig. C-9 than in Fig. 4-19. This makes sense because in the current sensitivity analysis the variation in number of sources was larger and the variation in the distribution of sources was smaller.

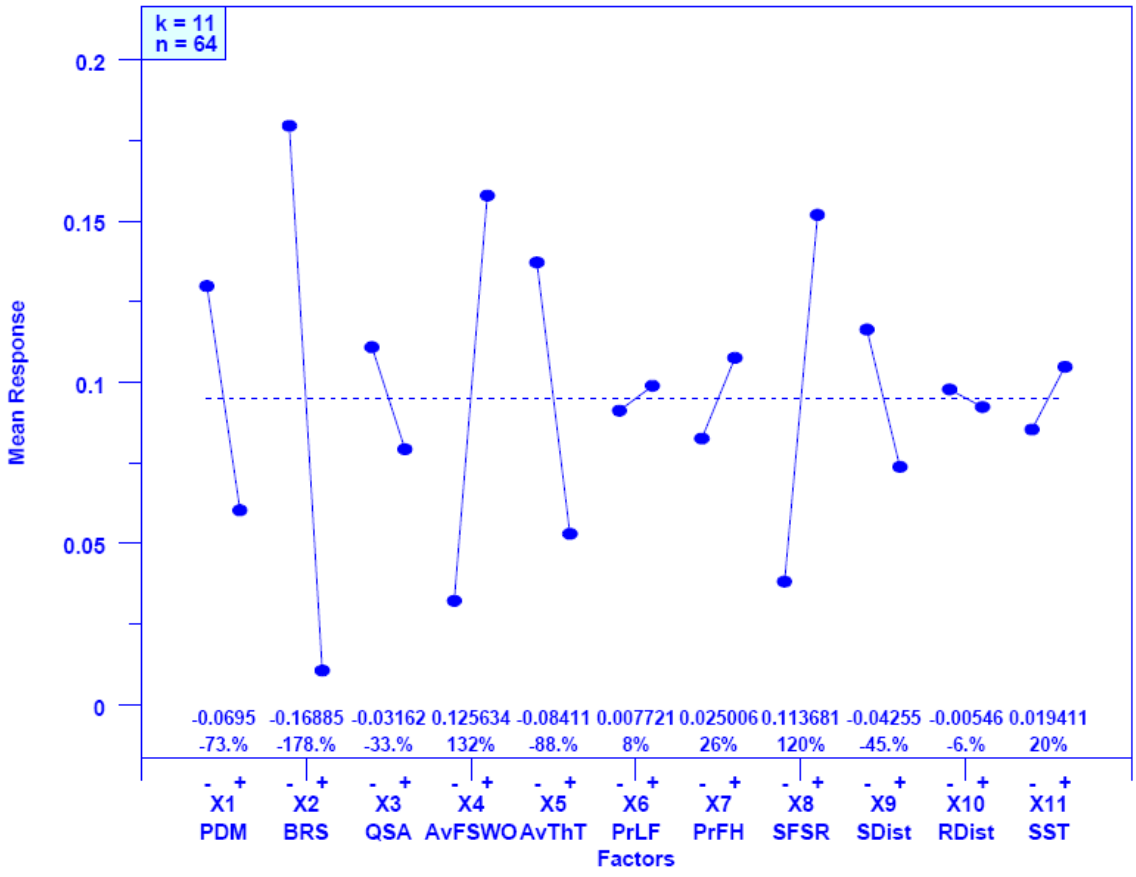


Figure C-9. Main-Effects Plot for Response y10 (Average Retransmission Rate) – y axis gives the proportion of packets resent

Fig. C-10 gives the main-effects plot for average congestion window (CWND) size. This figure can be compared with Fig. 4-9 when assessing similarities and differences among the two sensitivity analyses. Fig. C-10 reveals the CWND size is influenced mainly by two factors: average file size (x4) and network speed (x2). This differs somewhat from Fig. 4-9, which identified network speed (x2) as the main factor followed by four closely grouped factors: buffer sizing algorithm (x3), initial slow-start

threshold (x11), think time (x5) and distribution of sources (x9). We previously explained why buffer sizing algorithm and distribution of sources should be less influential with respect to congestion in the current sensitivity analysis. What about other differences between Figs. C-10 and 4-9?

Fig. C-10 identifies average file size (x4) as a significant factor, while Fig. 4-9 does not. In the current sensitivity analysis the difference between the plus and minus settings for average file size was increased substantially, which accounts for the increase in influence of factor x4. Average think time (x5) is the third most significant factor in both sensitivity analyses. What accounts for the diminished influence in the initial slow-start threshold? This appears related to increased congestion. The factor settings for the current sensitivity analysis allowed higher levels of congestion, as expressed for example by retransmission rates, which averaged around 9.5 % compared with only 8.8 % in the previous sensitivity analysis. Higher levels of congestion reduce the influence of the initial slow-start threshold because flows can incur lost packets sooner, and thus transition to congestion avoidance sooner.

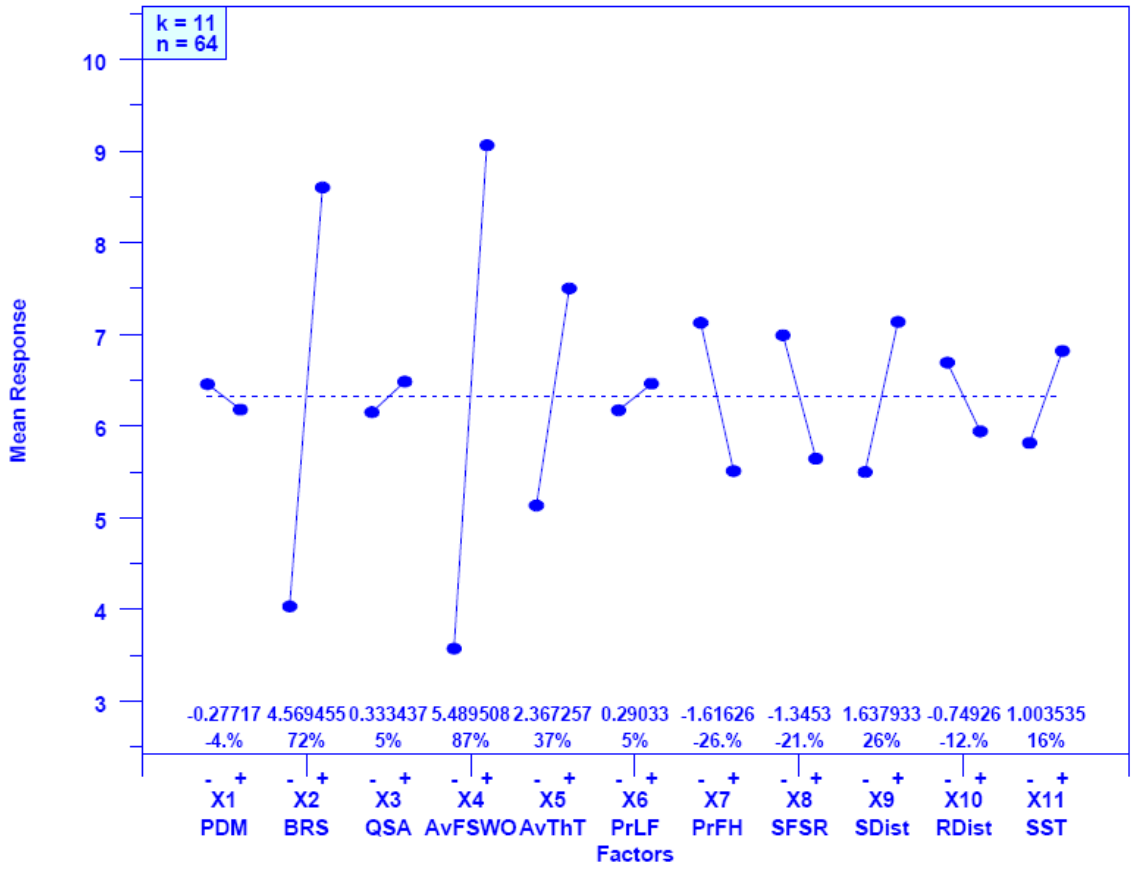


Figure C-10. Main-Effects Plot for Response y11 (Average Congestion Window Size) – y axis gives average congestion window size in packets

Fig. C-11, which can be compared with Fig. 4-20, displays main effects driving throughput on NN flows. The previous sensitivity analysis showed throughput on NN flows to be driven primarily by a relationship between available bandwidth (network speed) and number of active flows. Fig. C-11 also reflects this relationship: throughput is

higher under increased network speed (x2) when fewer flows are active. Factors leading to fewer active flows include a lower number of sources (x8 minus) and a source distribution (x9 plus) that leads to fewer NN flows, as well as longer think times (x5 plus). The main differences between Fig. C-11 and Fig. 4-20 relate to buffer sizing algorithm (previously explained) and average file size (x4). In the previous sensitivity analysis we found, surprisingly, that smaller file sizes led to higher throughputs on NN flows. This finding was surprising because larger file sizes can generally achieve higher throughputs. We attributed this to the fact that smaller files finished more quickly, which helped to reduce the number of active flows. This attribution made sense because the difference in average file size between the plus and minus settings was relatively small (50 packets), so file size could not have much influence on throughput. In the current sensitivity analysis, the difference in average file size between the plus and minus settings was significantly larger (175 packets). These larger files can achieve much higher instantaneous throughput than the much smaller files.

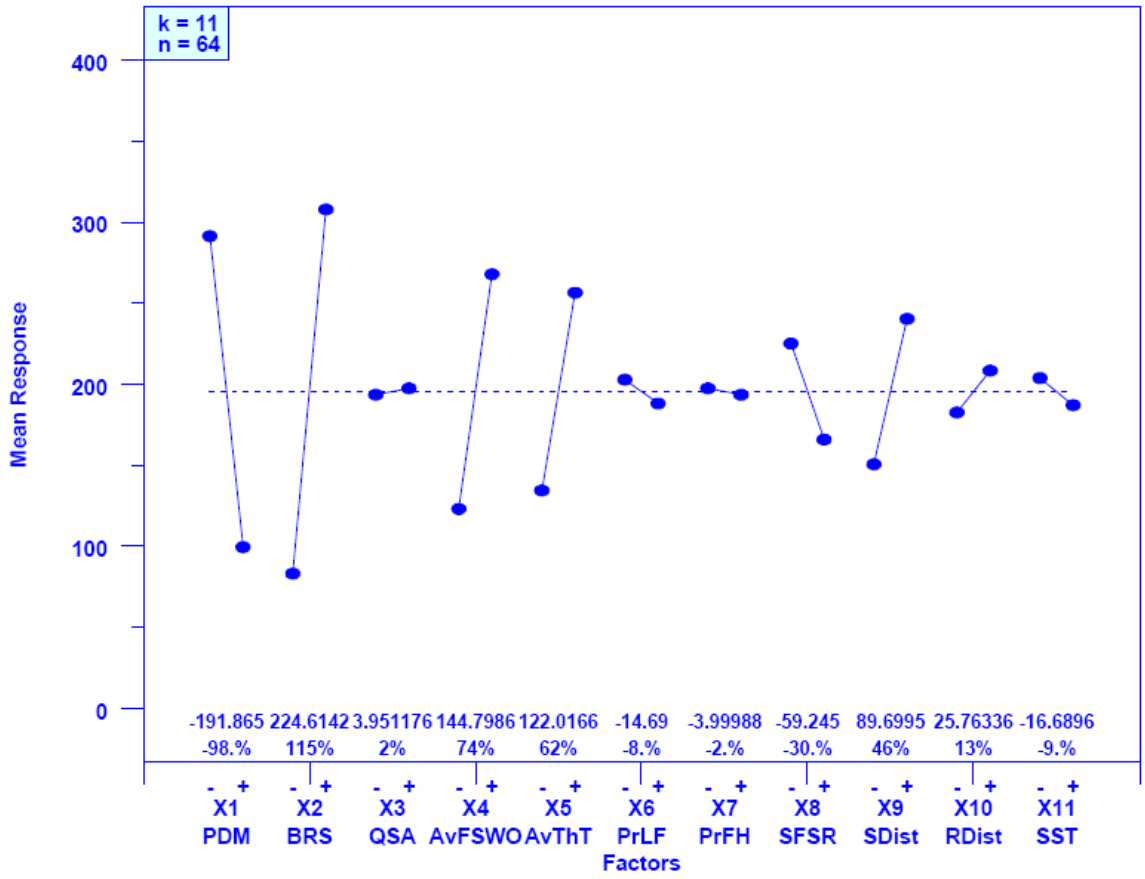


Figure C-11. Main-Effects Plot for Response y22 (Average Throughput on NN Flows) – y axis gives average goodput in pps

C.3.4.2 Delay-Related Responses. Fig. C-12 reports the influence of each input factor on response y15: average smoothed, round-trip time (SRTT). Both Fig. C-12 and the comparable Fig. 4-21 identify propagation delay (x1) and buffer sizing algorithm (x3) as the two main factors influencing SRTT. The influence of buffer sizing algorithm is less

significant in Fig. C-12 because fewer configurations exhibit small buffer sizes in the current sensitivity analysis.

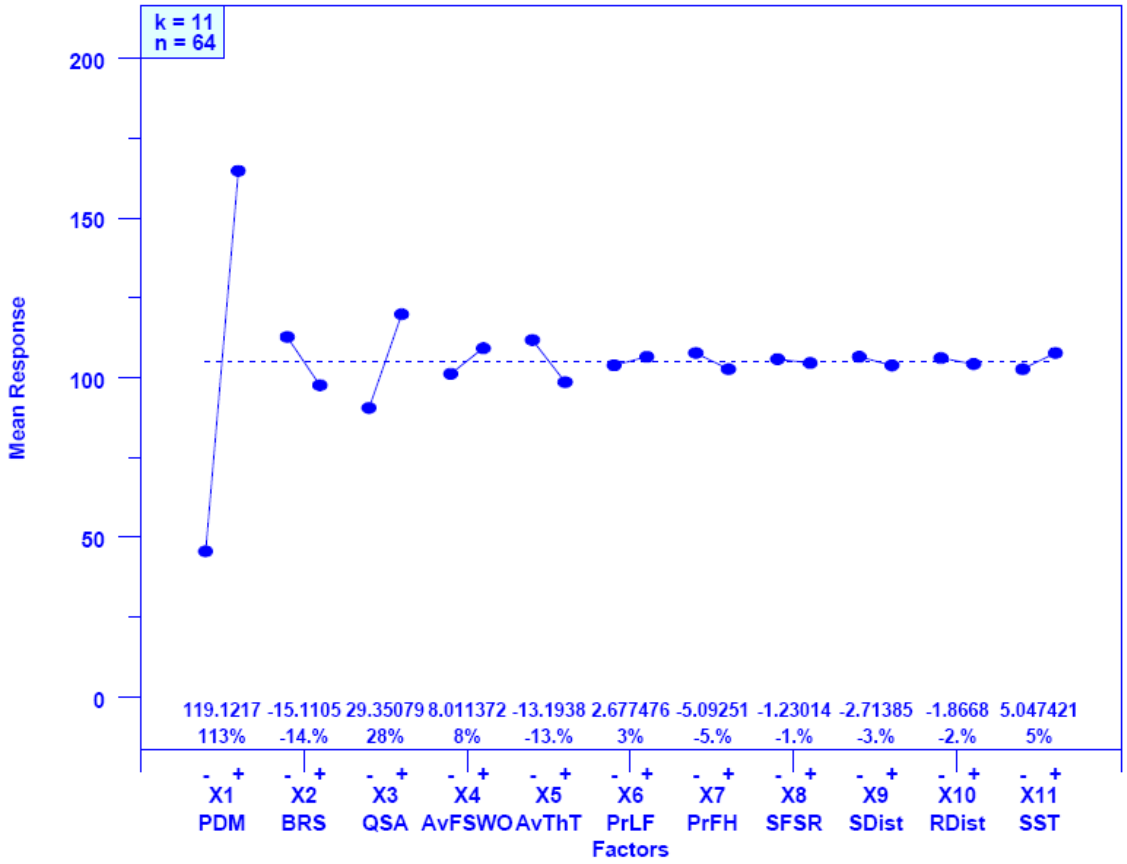


Figure C-12. Main-Effects Plot for Response y15 (Average Smoothed Round-Trip Time) – y axis gives average round-trip time in ms

C.3.4.3 Responses Related to Macroscopic Throughput. To represent macroscopic network throughput, we selected two responses: data packets output per interval (y4) and flows completed per interval (y6). The first response represents the rate at which packets are flowing through the network, while the second response represents the rate at which flows are being completed by the network. We begin by considering the rate of packet output.

Fig. C-13 identifies the same main factors influencing rate of packet output as revealed in Fig. 4-22. The main influence on the rate of packet output is network speed: higher network speed (x2 plus) means a greater rate of packet output. This stands to reason in a network with a sufficient number of active flows. The combination of shorter think times (x5 minus) and more sources (x8 plus) leads to an increase in the number of active flows and the higher network speed implies that each flow can transmit faster. Thus, the aggregate rate of packet output should be greater under these circumstances. File size is another factor significantly affecting the rate of packet output. Larger file sizes (x4 plus) lead to greater throughputs because a smaller portion of the transfer occurs during slow-start, the transfer phase during which a flow's congestion window is lowest.

Flows transferring with a larger congestion window achieve higher throughput, which helps to increase the aggregate network throughput.

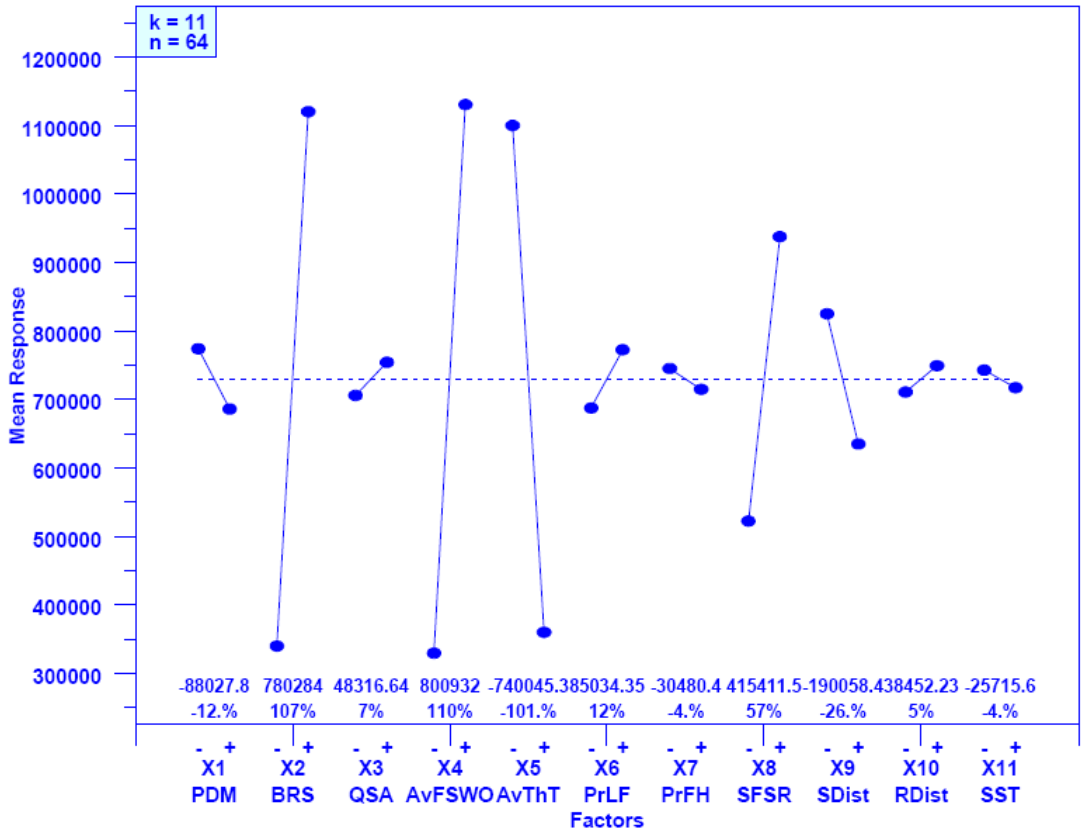


Figure C-13. Main-Effects Plot for Response y4 (Average Packets Output per Measurement Interval) – y axis gives average number of packets output per 200 ms

As shown in Fig. C-14 (and also Fig. 4-23), with one major exception, the story regarding the rate of flow completions is quite similar to the story regarding the rate of packet outputs. A sufficient number of connections (x5 minus and x8 plus) combined with higher network speed (x2 plus) contributes to a higher rate of flow completion. The exception involves file size (x4). In the case of packets output, larger file sizes (x4 plus) led to higher throughputs and thus to more packets output. On the contrary, for flows completed, a smaller average file size led to a higher completion rate. This stands to reason; smaller flows will be completed sooner. The sooner flows can be completed, the more flows can be completed per unit of time.

C.3.4.4. Responses Related to Advantaged Flow Classes. The final two responses we investigate represent throughputs achieved over advantaged flow classes, which are flows that transit between sources and receivers located under directly-connected and fast access routers. We examine the average instantaneous throughput of DD (y17) and FF (y20) flows. We begin by considering DD flows.

In the previous sensitivity analysis (see Fig. 4-24) we found that throughput on DD flows was influenced by only two factors: propagation delay (x1) and file size (x4). Shorter propagation delay (x1 minus) permitted faster feedback on DD flows, which

allowed the congestion window to increase more quickly. The rate of feedback was most important during the initial slow-start phase, where the congestion window started at a small size but doubled with each acknowledgment received. The influence of file size was also clear. Larger file sizes (x4 plus) allowed more of the packets in a file to be transferred after the flow reached its peak sending rate. Smaller file sizes (x4 minus) implied that more of the packets in a file were sent early in the slow-start phase, when a flow is building up toward its peak sending rate. Throughput early in slow-start will be much smaller than throughput after a flow reaches its peak rate.

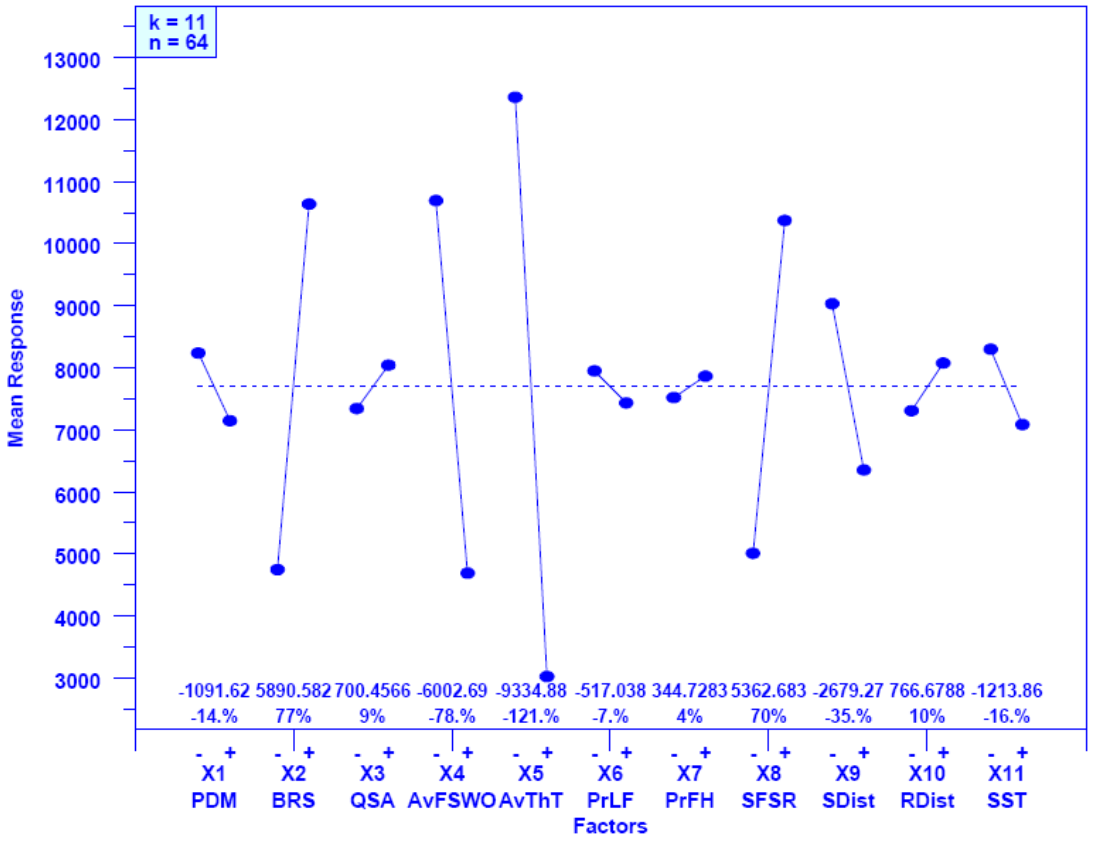


Figure C-14. Main-Effects Plot for Response y6 (Flows Completed per Measurement Interval) – y axis gives average number of flows completed per 200 ms

The current sensitivity analysis (Fig. C-15) also identifies propagation delay (x1) and file size (x4) as the two main factors influencing throughput on DD flows. Fig. C-15 also identifies network speed (x2) as a significant factor. This makes sense because the speed difference between the plus and minus settings was much larger (14×10^3 p/ms) here than in the previous sensitivity analysis (400 p/ms). The higher difference in network speed would certainly contribute to a larger throughput on DD flows. Fig. C-15 identifies a few other factors having some influence. For example, lower congestion arising from fewer sources (x8 minus) and longer think times (x5 plus) allow for higher throughputs on DD flows. Finally, the larger difference in file sizes (175 packets instead of 50 packets) enables a higher initial slow-start threshold to contribute more to higher throughputs on DD flows.

In the previous sensitivity analysis (see Fig. 4-25) we found throughput on FF flows to be influenced by a more complex mix of factors than DD flows. The significance of propagation delay (x1) and file size (x4) were two clear common factors between all advantaged flow classes. Shorter propagation delay meant quicker feedback, which led to faster increase in the congestion window for flows that were not impeded by congestion. Larger file sizes allowed more of a flow's packets to be transferred at a higher sending rate. Less advantaged (DN, FN and NN) flows were influenced mainly by congestion, so propagation delay had less affect on those flows.

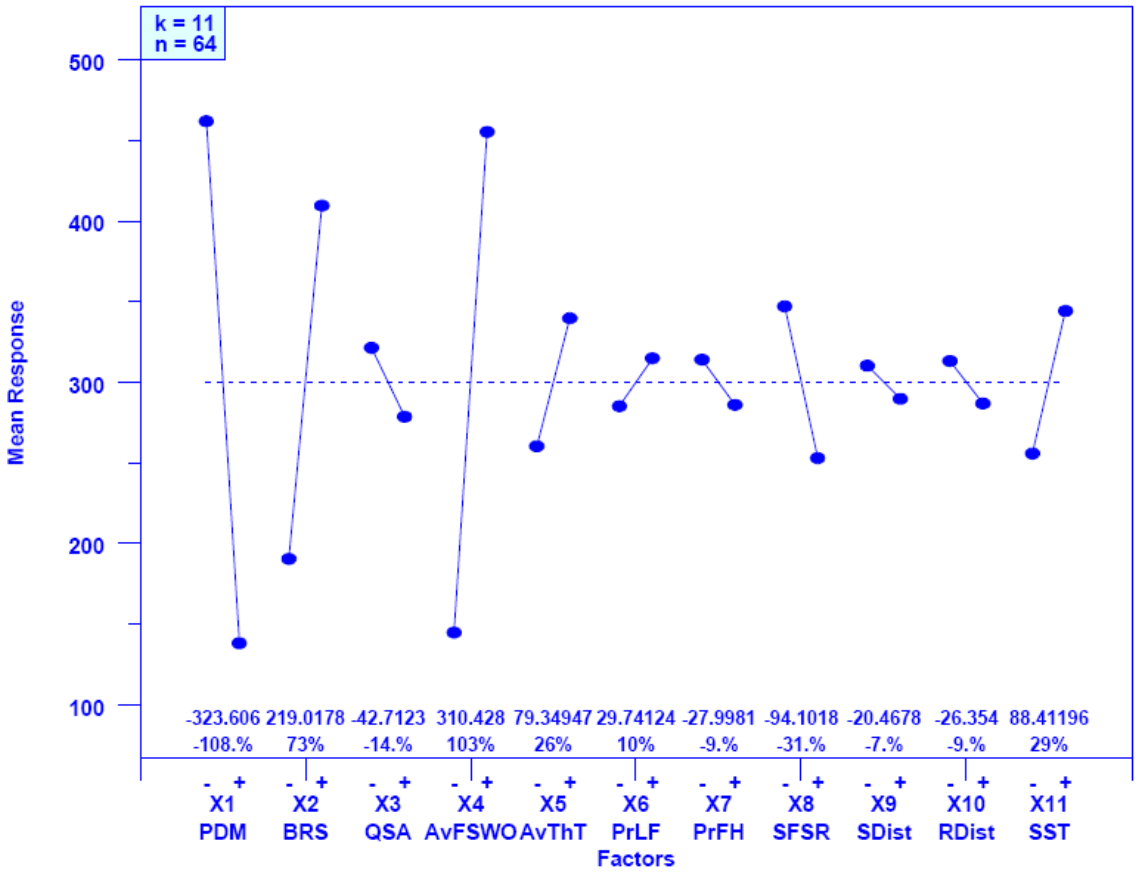


Figure C-15. Main-Effects Plot for Response y17 (Average Instantaneous Throughput on DD Flows) – y axis gives average goodput in pps

As shown in both Figs. C-16 and 4-25, FF flows, unlike DD flows, can face some congestion because selected source distributions lead to higher numbers of FN flows. Specifically, a source distribution (x9 plus) that gives the network a Web-centric characteristic leads to more FN flows, which compete for throughput with FF and DF flows. In addition, more sources (x8 plus) and lower average think time (x5 plus) lead to more active flows that can compete for throughput. Under these circumstances, higher network speed (x2 plus) allows competing flows to achieve higher throughputs. In the current sensitivity analysis buffer sizing algorithm has less influence on throughput because buffer sizes tend to be larger and initial slow-start threshold has more influence on throughput because the spread in file sizes and the higher network speed enabled increased use of initial slow-start.

C.3.5 Summary of Findings from Sensitivity Analysis

We use a rank analysis to summarize the findings from the current sensitivity analysis. We use one response to represent each characteristic: packet throughput (y4), flow completion throughput (y6), congestion (y10), delay (y15) and throughput of DD (y17) and FF (y20) flows. Table C-13 shows the results of our rank analysis, where the relative influence of each factor on each of the six responses is assigned a rank from one (most influential) to 11 (least influential) based upon the degree to which the factor altered the response when moving from a plus to a minus setting. The average rank is computed for each factor, and then the average rank is converted into an ordinal ranking based on ordering the factors from most (one) to least (11) influential. The table shows that network speed (x2) is the most influential factor, followed by file size (x4) and propagation delay (x1). Next is think time (x5), followed by number of sources (x8). These five factors (x2, x4, x1, x5 and x8) are the same five factors identified as most influential in the earlier sensitivity analysis (see Table 4-25). The only difference is one of ordering: propagation delay has jumped from fifth to third most influential.

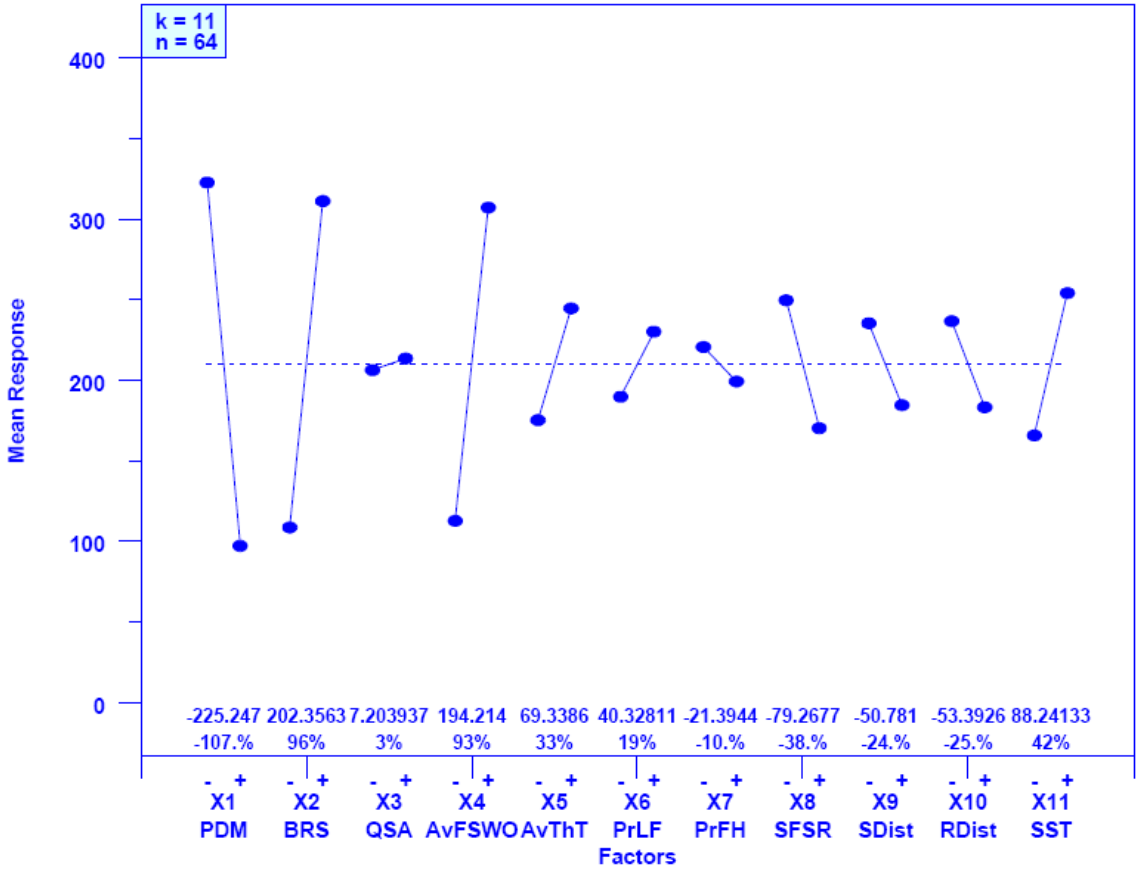


Figure C-16. Main-Effects Plot for Response y20 (Average Instantaneous Throughput on FF Flows) – y axis gives average goodput in pps

Table C-13. Rank Analysis of Sensitivity Analysis Responses

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11
y4	6.5	2	8	1	3	6.5	10.5	4	5	9	10.5
y6	7	3	9	2	1	10	11	4	5	8	6
y10	5	1	7	2	4	10	8	3	6	11	9
y15	1	3	2	5	4	8.5	6.5	11	8.5	10	6.5
y17	1	3	7	2	6	8	9.5	4	11	9.5	5
y20	1	2	11	3	6	9	10	5	8	7	4
Average Rank	3.58	2.33	7.33	2.50	4.00	8.67	9.25	5.17	7.25	9.08	6.83
Ordinal Rank	3	1	8	2	4	9	11	5	7	10	6

This ordering shift can be justified. First, propagation delay affects many aspects of flow operation and the difference between the plus and minus settings in propagation delay has increased from 2 fold in the previous sensitivity analysis to 5 fold in the current sensitivity analysis. Second, the number of active sources (y1) is influenced by a relationship between think time and number of sources. The current sensitivity analysis has an order of magnitude more sources but these sources can see more than an order of magnitude greater network capacity and can have much longer think time (up to 10 instead of 5 seconds). The combination of sources with potentially longer think times and operating in a network with up to 20 times more capacity leads to a slightly higher proportion (74 % vs. 72 %) of sources in the thinking state at any given time. Since we scaled number of sources, think time and network speed to match, we would expect propagation delay to exert increased influence in the current sensitivity analysis.

In examining the less influential factors, we find that the sixth most influential parameter has become the initial slow-start threshold (x11). This makes sense because file sizes can be much larger and network speeds can be much higher in the current experiment, so more flows have opportunity to exploit the potential of a higher initial slow-start threshold.

C.3.6 Exploring Effects of Buffer Sizing

We decided to repeat the exploration of the effects of buffer sizing, which we applied to the earlier sensitivity analysis, as described in Sec. 4.7.2. Specifically, we consider the relative influence of propagation delay (x1), network speed (x2) and buffer-sizing algorithm (x3) on selected responses, chosen to represent macroscopic network behavior and user experience. To represent macroscopic behavior, we use packet throughput (y4), flow completion throughput (y6), retransmission rate (y10) and relative queuing delay (y16). To represent user experience, we use average throughput from three different flow classes: DD flows (y17), FF flows (y20) and NN flows (Y22). We aim to determine which of the three factors (x1, x2 or x3) has largest influence on the combined responses.

We use a rank analysis to study the effects of our chosen factors on our selected responses. In this particular analysis, we elected to use a larger number to indicate higher rank and a smaller number to indicate lower rank. We began by combining our three factors into a condition that can be assigned one of eight settings, as illustrated in Table C-14 (which can be compared with Table 4-27). Next, we computed the average value for each of our responses under each condition. Table C-15 (comparable to Table 4-28) displays the results of this averaging.

Table C-14. Mapping of Factor Settings to Eight Conditions (M = minus; P = plus)

	Values			
Condition	**Factor Settings x1:x2:x3**	**Propagation Delay Multiplier**	**Backbone Router Speed**	**Buffer Sizing Algorithm**
C1	**M:M:M**	**1**	**2×10^3**	***RTT*x*C*/SQRT(*n*)x2**
C2	**P:M:M**	**2**	**2×10^3**	***RTT*x*C*/SQRT(*n*)x2**
C3	**M:P:M**	**1**	**16×10^3**	***RTT*x*C*/SQRT(*n*)x2**
C4	**P:P:M**	**2**	**16×10^3**	***RTT*x*C*/SQRT(*n*) x2**
C5	**M:M:P**	**1**	**2×10^3**	***RTT*x*C*/2**
C6	**P:M:P**	**2**	**2×10^3**	***RTT*x*C*/2**
C7	**M:P:P**	**1**	**16×10^3**	***RTT*x*C*/2**
C8	**P:P:P**	**2**	**16×10^3**	***RTT*x*C*/2**

Using the average responses from Table C-15, we next rank each condition from high (8) to low (1) for each response, based on the appropriate ordering criteria. For retransmission rate (y6) and relative queuing delay (y16) a lower value would be ranked higher. For the other responses in Table C-15, a higher value would be ranked higher. After ranking the conditions with respect to each response, we compute an average ranking. The results of our ranking are shown in Table C-16 (comparable to Table 4-29).

Table C-15. Average Response Values for Each Condition

	Response						
Condition	**y4**	**y6**	**y10**	**y16**	**y17**	**y20**	**y22**
C1	**358 946.069**	**5561.8**	**0.211**	**1.6199**	**176.28**	**196.53**	**132**
C2	**316 517.564**	**3579.9**	**0.1946**	**1.3866**	**399.78**	**43.253**	**41.1**
C3	**1 308 862.84**	**8632.8**	**0.0379**	**1.9674**	**637.31**	**450.52**	**435**
C4	**839 113.623**	**11593**	**6E-05**	**1.4241**	**176.28**	**135**	**166**
C5	**347 508.785**	**4416.3**	**0.2679**	**2.9595**	**158.16**	**107.59**	**100**
C6	**336 533.732**	**5428.9**	**0.0445**	**2.0917**	**131.96**	**87.617**	**59.3**
C7	**1 080 811.29**	**14340**	**0.0024**	**2.143**	**652.1**	**535.56**	**498**
C8	**1 251 852.86**	**7983.3**	**0.0021**	**1.5581**	**172.49**	**123.35**	**132**

We can assign the average rank for each condition to the vertex of a cube, where each vertex represents a specific combination of settings for propagation delay, network speed and buffer size. Fig. C-17 (comparable to Fig. 4-32) shows the cube corresponding to Table C-16. Moving along the edges among the vertices on the cube allows us to determine changes in ranking attributable to each factor. The change in each factor (x1, x2 and x3) across all conditions is represented by a set of four different edges from among the 12 edges contained in the cube. We extract the relevant changes in ranking and display them in Table C-17 (comparable to Table 4-30).

Interpreting Table C-17 we see that changing network speed has the largest effect on the responses we selected. This agrees with the previous sensitivity analysis. Changing buffer sizing has the second largest effect, nearly the same as changing propagation delay, which has the smallest effect. Further, Fig. C-17 shows that changing from fewer to more buffers has a larger effect when network speed is low and propagation delay is short. This finding is counter to the earlier sensitivity analysis that found that changing from fewer to more buffers had a larger effect when network speed is high and

propagation delay is long. While counter to the previous findings, the effect can still be explained. In the previous sensitivity analysis, the network speeds were quite close and the small buffer size was very small for half the configurations. In the current sensitivity analysis, the network speeds are quite disparate. At the faster network speed (16×10^3 p/ms) packets rarely need to be buffered and so increasing to a larger buffer size does not matter much. At the slower network speed the potential for buffering packets increases and so the increase in buffer size has a larger influence on responses. In addition, packets can be "buffered in flight" on network transmission links in proportion to the bandwidth-delay product. For this reason, higher network speeds coupled with longer propagation delays mean more potential for in-flight buffering. Conversely, lower network speeds coupled with shorter propagation delay means less potential for in-flight buffering. For this reason, adding buffers when network speeds are lower and propagation delays are shorter should have more influence on responses. If nothing else, the two sensitivity analyses reveal a complicated interrelationship among network speed, propagation delay and buffer sizing. This interrelationship merits additional study.

Table C-16. Ranking for Each Condition vs. Each Response

	Condition							
Response	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**
y4	4	1	8	5	3	2	6	7
y6	4	1	6	7	2	3	8	5
y10	2	3	5	8	1	4	6	7
y16	5	8	4	7	1	3	2	6
y17	4	6	7	5	2	1	8	3
y20	6	1	7	5	3	2	8	4
y22	5	1	7	6	3	2	8	4
Average Rank	4.3	3.0	6.3	6.1	2.1	2.4	6.6	5.1

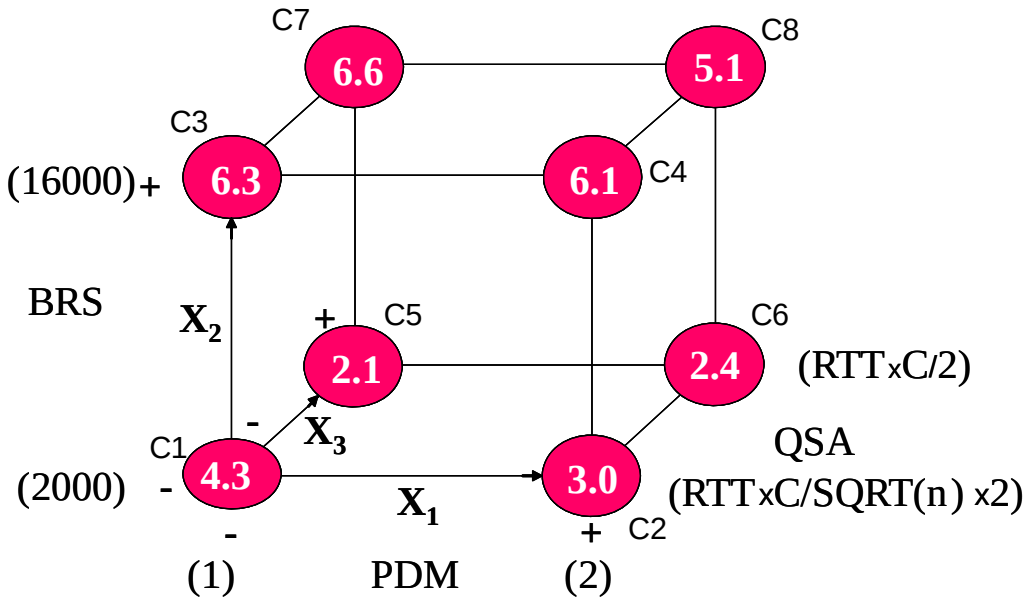


Figure C-17. Average Condition Ranking Displayed on Vertices of a Cube

Table C-17. Changes in Ranking Attributable to Each Factor

	Propagation Delay (x1)	Network Speed (x2)	Buffer Sizing (x3)
Edge 1	1.3	2.0	2.2
Edge 2	0.3	3.1	0.6
Edge 3	0.2	4.5	0.3
Edge 4	1.5	2.7	1.0
Average	0.8	3.1	1.0

C.4 Discussion

The supplementary sensitivity analysis confirmed the primary findings about main effects that were revealed in the previous sensitivity analysis, covered in Chapter 4. Both analyses found that system response was driven by the same primary input factors: network speed, file size, propagation delay, think time and number of sources. Propagation delay was found to be more influential in the current sensitivity analysis because the plus-minus level difference in propagation delay was substantially increased over the previous sensitivity analysis. Initial slow-start threshold moved to sixth (from eighth) because more flows had the opportunity to exploit a high initial slow-start threshold under the current sensitivity analysis, which allowed larger file sizes and much higher network speeds. The influence of buffer-sizing algorithm switched to eighth (from seventh) because in the current sensitivity analysis only 12.5 % of the configurations created average buffer sizes below 10^3 packets, compared with 50 % of the configurations in the previous sensitivity analysis. In comparing the main-effects plots, both sensitivity analyses identified the same input factors driving each system response. For particular responses, as explained above, small differences in factor influences could be discerned – all such differences were justified.

With respect to correlation among responses, both sensitivity analyses identified the same six response pairs to be correlated with magnitude greater than 0.95. The supplementary sensitivity analysis identified a seventh response pair (y17-y18) correlated above 0.95. In the range of 0.90 to 0.95, the supplementary sensitivity analysis found only three correlated pairs, while the previous sensitivity analysis indentified 10 such pairs. Among the seven missing pairs, four appeared at lower strength in the supplementary sensitivity analysis.

Two main differences were noted among correlations in the sensitivity analyses. First, the supplementary analysis found throughputs in all flow classes were correlated, whereas the previous analysis found throughput on DD flows and to be uncorrelated with throughput on other flow classes, and also found throughputs on DF and FF flows to be correlated with each other but uncorrelated with the throughputs of other flow classes. Comparing the relevant main-effects plots for these responses found that the supplementary analysis showed all flow throughputs to be driven by the same top three factors: propagation delay, network speed and file size. The order shifted among the three flow classes: propagation delay the main influence on DD and FF flows and network speed the main influence on NN flows. While the main influence (propagation delay or network speed) also appeared in the previous sensitivity analysis, the second most influential factors differed for DD flows (average file size), FF flows (source distribution) and NN flows (average think time). This shows that throughputs for FF and NN flows

were driven more by congestion in the first sensitivity analysis and less so in the supplementary analysis. Further, in the supplementary analysis, increased differences in network speed and propagation delay had a greater influence on flow throughputs than congestion. Guided by this finding, we defined our experiments comparing congestion control algorithms to include situations where congestion played a significant role, as well as situations where congestion was less significant.

A second main difference in correlation appeared in the supplementary sensitivity analysis. In the previous analysis, round-trip time (y15) and queuing delay (y16) were moderately correlated (0.70), while the supplementary analysis found no correlation (-0.08). This correlation change is largely due to changes in the relationship between propagation delay and buffer sizing between the two analyses. In the supplementary sensitivity analysis, propagation delay had a much bigger influence on SRTT, which increases 113 % when moving from a minus to plus setting vs. only 52 % in the first sensitivity analysis. As a result, higher propagation delays in the supplementary analysis increased the numerator in the computation for relative queuing delay (y16), which reduces y16 more than is the case in the previous sensitivity analysis. As a result, in the supplementary analysis with generally increasing SRTT (driven by both propagation delay and buffer size), y16 increases for 32 conditions (smaller propagation delay) then drops drastically upon reaching the larger propagation delay, increasing again along with increasing buffer size. These two functions, round-trip time (y15) and queuing delay (y16), do not exhibit the same relationship under the settings associated with the previous sensitivity analysis, where a generally increasing SRTT mirrors a generally increasing buffer size, except for occasional dips associated with specific parameter combinations. From this, we concluded that we should not estimate queuing delay by dividing SRTT by average propagation delay. For the experiments comparing congestion control algorithms we decided to estimate queuing delay by subtracting the average round-trip propagation delay from the average SRTT.

Similar to results from the correlation analysis, results from the principal components analysis (PCA) exhibited several differences between the two sensitivity analyses we conducted. Principal components analyses are notoriously difficult to interpret in many domains, especially when investigating a complex system with many factors. In the initial sensitivity analysis, we identified four principal components: congestion, delay, throughput on advantaged flows and network-wide throughput (flows and packets). For the supplementary PCA, identifying the top four principal components proved more difficult. As we explained above, the supplementary sensitivity analysis led us to group responses comprising four principal components that influence: throughput on all flow classes, throughput on advantaged flow classes, delay and network-wide throughput (flows and packets). The top four components bear general similarity among the two analyses, though they differ in the grouping of specific responses. The first component in both analyses could be said to characterize congestion and throughput for most typical users of the network. Delay is also a common component, playing more prominence in the first sensitivity analysis because half the configurations had very small buffers (compared to only 1/8 of the configurations in the supplementary analysis). The remaining two principal components were similar between the two analyses: throughput on advantaged flows and network-wide throughput.

We also compared two exploratory analyses: (1) seeking the source of a bifurcation in a (y7-y22) scatter plot and (2) investigating the relative influence of buffer sizing on system responses. As shown above, both sensitivity analyses revealed average file size (x4) as the cause of the y7-y22 bifurcation. Further, as expected, the supplementary sensitivity analysis led to an increased angle of bifurcation. With respect to the relative influence of buffer size, both sensitivity analyses found network speed to have much greater influence on system response than either propagation delay or buffer size. The supplementary and initial sensitivity analyses did vary with respect to the conditions under which increasing buffer size had larger influence on system responses, but these differences were explainable.

C.5 Conclusions

In this appendix, we conducted a supplementary sensitivity analysis following the general plan presented in Chapter 4, but changing the level settings for the 11 parameters. Specifically, we increased network speed and size by about an order of magnitude, we stretched the range of parameter values covered by the plus and minus settings of each factor, and we shifted the traffic patterns slightly to generate more DD flows and to give a higher prominence to Web browsing activity over P2P exchanges. We subjected the results to the same analyses applied in Chapter 4. Comparing our findings with those from Chapter 4, we identified general agreement in the main factors driving model responses. Where differences arose, we were able to attribute them to changes in level settings or relationships among level settings. On the whole, the results from this supplementary sensitivity analysis increased our confidence in the MesoNet simulation model. In addition, we gained increased confidence in our analysis methods, which were able to identify differences arising from parameter variations, as well as relationships remaining invariant across our two sensitivity analyses. Finally, comparing results from the two sensitivity analyses guided us to change our technique for estimating queuing delay and to select experiment designs to include configurations with little congestion, as well as configurations with significant congestion.