2010 IEEE/ACM International Conference on Green Computing and Communications & 2010 IEEE/ACM International Conference on Cyber, Physical and Social Computing

Statistical Power and Performance Modeling for Optimizing the Energy Efficiency of Scientific Computing

Balaji Subramaniam and Wu-chun Feng
Department of Computer Science
Virginia Tech
{balaji, feng}@cs.vt.edu

***Abstract*—High-performance computing (HPC) has become an indispensable resource in science and engineering, and it has oftentimes been referred to as the "third pillar" of science, along with theory and experimentation. Performance tuning is a key aspect in utilizing HPC resources to the fullest extent. However, recent exascale studies suggest that power and energy consumption will be a major impediment to HPC in this coming decade. Therefore, performance tuning should evolve and take energy consumption into account. Unfortunately, the increase in system complexity and the number of tunable parameters in applications makes the performance tuning of an application cumbersome.**

To address these issues, we propose energy-efficient tuning via statistical regression techniques. Such techniques can be used to model the power and performance of a scientific application, and then the application parameters can be tuned to achieve the best energy efficiency possible, based on metrics such as the performance-to-power ratio.

In this paper, we utilize multi-variable regression to model the power and performance of the high-performance LINPACK (HPL) benchmark. We then tune the HPL parameters for energy efficiency and compare them to the energy efficiency achieved at maximum possible performance (R_{max}). Our results show that statistical regression modeling can be used for predicting the HPL configuration for achieving the maximum energy efficiency with very high accuracy.

I. Introduction

The list of fastest supercomputers is maintained by the Top500 [2], based on the high-performance LINPACK (HPL) benchmark [1]. The list ranks the supercomputers based on performance using the metric of floating-point operations per second (FLOPS). Many analytical and statistical models have been developed to automate the process of tuning the parameter space of HPL for achieving maximum performance (R_{max}). The use of the FLOPS metric has been the primary driver to increase the performance of HPC systems by adding more compute units without taking energy consumption into consideration. However, a recent exascale study projects that power and energy consumption will be major impediments on the road to exascale computing [3]. Therefore, energy efficiency should also be taken into account while tuning application parameters as the maximum energy efficiency of an application may not necessarily be achieved while executing at R_{max}.

Adding more concern is the growing number of tunable parameters in applications. In HPL, there are 18 parameters and eight of these have predefined values and changing even a single parameter can introduce high variations in performance [11]. Therefore, analytical modeling becomes a tedious process. Further, analytical models use assumptions on application input, e.g., range for input size, thus failing to capture all possible scenarios of the application. Accurate prediction of power and performance will lead to a proper understanding of the power-performance trade-offs associated with an application and potentially save expenses in the order of millions of dollars. Work like [11] seek to address this issue by identifying the parameters that affect the overall performance using several feature selection techniques. However, how these parameters affect the accurate modeling of power and performance is still unclear.

To address this issue, we use multi-variable regression techniques [5] to model the power consumed and performance of the HPL benchmark. Regression techniques are commonly used for prediction of power and performance of a system. These models require a set of observations and the model is constructed based on the observed values once the explanatory variables (variables which have high correlation with the response variable) are determined. The relevant explanatory variables are

978-0-7695-4331-4/10 $26.00 © 2010 IEEE
DOI 10.1109/GreenCom-CPSCom.2010.138

identified by using set of test for significance of the variable. Further, we predict the HPL parameters for best energy efficiency based on popularly used energy efficiency metrics [6] such as performance to power ratio by using these models.

Our contributions in the paper are three-fold, they are:

- First, we accurately analyze the statistical significance of the HPL parameters for creating the regression model for power and performance.
- We then create a model for performance and power of the system for executing HPL benchmark by using multi-variable regression. Our models achieve high accuracy in predicting both power and performance.
- Finally, we predict the HPL parameters for best energy efficiency using the regressed models. Our results indicate that the best energy efficiency is not always achieved when the system is executed at the maximum possible performance.

The rest of the paper is organized as follows. Section II describes the HPL benchmark and its parameters. Section III throws light on the multi-variable regression theory. Here we discuss about the metrics used for choosing the variables (HPL parameters) that are suitable for creating the regression model. We then describe about the metrics used for assessing the goodness of the fit in the same section. In Section IV, a discussion about popularly used energy efficiency metrics is presented. We use these metrics to perform energy efficiency optimizations. Experimental platforms and setup is described in Section V. The experimental results is discussed in Section VI. Here we provide the results for analyzing the statistical significance of the predictor variables and assessing the "goodness of the fit" of the models. We then present the results for optimizing the HPL benchmark for energy efficiency using the statistical models. Section VII presents the related work. The final section concludes the paper and describes possible future work.

II. High Performance LINPACK (HPL)

The HPL benchmark solves a dense linear system of equations. It performs LU factorization on the coefficient matrix and computes the solution by backward substitution. The data is distributed in a block-cyclic fashion to ensure load balance. The benchmark reports performance in terms of FLOPS. The number of floating-point operations is calculated by using Equation (1), where N is the problem size [1]. Based on Equation (1), the number of floating-point operations is only dependent on the problem size.

$$\text{Floating Point Operations} = \frac{2}{3} * N^3 + \frac{3}{2} * N^2 \quad (1)$$

Table I provides the list of parameter inputs into HPL along with their associated predefined values, if any.

#	Parameter Name	Adjustable Values
1	Problem Size (N)	NA
2	Block Size (NB)	NA
3	Process Mapping (PMAP)	0=Row-major, 1=Column-major
4	Rows of Process Grid (P)	NA
5	Columns of Process Grid (Q)	NA
6	Panel Factorization (PFACT)	0=left, 1=Crout, 2=Right
7	Minimum Columns For Recursion (NBMIN)	NA
8	Subpanel Division (NDIV)	NA
9	Recursive Factorization (RFACT)	0=left, 1=Crout, 2=Right
10	Broadcast Algorithm (BCAST)	0=1rg, 1=1rM, 2=2rg, 3=2rM, 4=Lng, 5=LnM

TABLE I
List of HPL Parameters. NA = Not Applicable

III. Multi-Variable Regression Analysis

We use multi-variable regression modeling to predict the power and performance of HPL. We derive models where the response (HPL power and performance) is modeled as linear combinations of the predictor variables (HPL parameters).

A. Modeling

The regression modeling expresses a response variable in terms of predictor variables as shown in Equation (2) where y is the response variable, x_1,...,x_n are the predictor variables, β_1,...,β_n are the partial regression coefficients, β_0 is the intercept and e is the error.

$$y = \sum_{i=1}^{n} \beta_i x_i + \beta_0 + e \quad (2)$$

The least square method is usually preferred to find n coefficients in β based on the observations. The model allows us to predict the response y given the predictor variables x_i.

B. Predictor Selection and Optimization

Predictor selection is the process of selecting the predictor variables which have high correlation with the response. HPL performance has very high correlation with N, NB, P and Q as shown in [11] and analytical models such as [4] use only these parameters. Therefore we use only N, NB, P and Q in our regression analysis and apply statistical analyses only for these parameter. Further, these parameters might have some non-linear relationships with the response. Therefore, some mathematical transformations can be applied either to response and/or the predictor variables. The rest of this section describes about the methodologies used for testing the significance of each of these parameters and gives an overview of cases where transformation might be required on the variables.

1) Test of Significance: The analysis of variance (ANOVA) forms the basis for test of significance by quantifying the variability accounted by the model. F-test is usually preferred to quantify the significance of the predictor variables in the fit. The test statistic for the F-test is the F-ratio from the ANOVA table. Equation (3) shows how F-ratio is calculated where y is observed value, $\hat{y}$ is the predicted value, $\bar{y}$ is the mean of response, n is the number of observations, p is number of predictor variables, SSM is the sum of squares model and SSE is the sum of squares error. Large F-ratio values provide evidence against the null hypothesis which is defined as β_1,...,β_n = 0 in Equation (2).

$$\begin{aligned} \text{F-ratio} &= \frac{\frac{1}{p}\sum(\hat{y}_i - \bar{y})^2}{\frac{1}{(n-p-1)}\sum(y_i - \hat{y}_i)^2} \\ &= \frac{\text{SSM/p}}{\text{SSE/(n-p-1)}} \end{aligned} \tag{3}$$

However F-test does not indicate that all the coefficients in β_i are not zero, it simply suggests that one or more parameters used is linearly related to the response variable. Therefore p-values along with F-test are used to determine whether each parameter is statistically relevant. p-value tells us how likely the coefficient of a predictor variable does not describe a statistically significant relationship. A p-value of 0.01 or lesser is accepted for the rejection of the null hypothesis.

Stepwise selection is also commonly applied to understand what parameters to retain. The selection is based on the correlation and the F-ratio value. The process starts with only one predictor in the regression model and proceeds to add all the significant parameters into the model. Each predictor is added to the model based on correlation with the response and the decision whether it needs to be retained is based upon the F-ratio. In this paper we apply all these techniques to analyze whether the parameters N, NB, P and Q are statistically relevant to predict the response (both power and performance of the HPL benchmark).

2) Non-Linearity and Collinearity: One of the fundamental assumptions of linear regression is that the predictor variables vary linearly with the response but this is not the case in most problems. The easiest way out of this problem is to apply transformations on the variables. For example if a predictor variable x_i varies exponentially with a response variable y, then applying a logarithmic transformation on x_i will make it linearly vary with the response variable. Transformation like logarithmic and square root can also be applied simply to reduce the magnitude of variance in the data.

Another assumption in regression modeling is that all the predictor variables are not correlated to each other i.e. each of the predictors explain a separate portion of the variability in response. However, in practice we might see some correlation among predictors. The solution to this problem is to decide on what variable to use in the model. F-test can come in handy in such situation. For example, if two variables are severely correlated, then F-ratio of the model including only one of the variables at a time can help in determining which variable to retain.

C. Goodness of Fit

The fit can be assessed statistically by quantifying how well the model captures the trends. The model's fit is assessed by using residual plots. Residual plots show us how the error is distributed. These plots can be helpful in checking whether the residual is correlated with any of the predictors. In such cases, the fit can be improved by applying appropriate transformation to that variable. Usually a random distribution of residuals is preferred.

Another metric used for assessing the fit is R^2. R^2, defined in Equation (4), shows the variability in observed values accounted by the model. R^2 varies from 0.0-1.0 where a value of 1.0 refers to perfect fit. However, R^2 increases with increase in number of predictor variables. So in practice adjusted R^2 (defined in (5)) is used to measure the goodness of the fit. Adjusted R^2 will only increase if the variable added to the model has statistical significance.

$$R^2 = 1 - \frac{\text{SSE}}{\text{SST}} \tag{4}$$

$$\text{Adjusted } R^2 = 1 - \frac{\text{SSE/(n-p-1)}}{\text{SST/(n-1)}} \quad (5)$$

Goodness of fit can also be assessed by the average error percentage. The metric shows the accuracy of our predictions. Overall p-value is also used to assess the fit of the model as the metric provides a view on the overall statistical significance of the model. While each of these metrics individually assesses the fit of the model, generally it is accepted that the model is a good fit if most of these metrics are in a acceptable range.

IV. Metrics for Energy Efficiency

The most popular metric for energy efficiency is the performance-per-watt metric (i.e., FLOPS/watt). It is used by the Green500 to rank the most energy-efficient supercomputers in the world. Another commonly used metric is the energy-delay product (EDP), which is calculated as amount of energy consumed multipled by the execution time. However, both of these metrics have their own disadvantages [6]. A closer look at the FLOPS/watt metric shows that the metric has a indirect energy component to it, as shown in Equation (6). In this paper, we use the performance-per-watt metric to evaluate the energy efficiency of a system and apply our tuning for energy efficiency on the HPL benchmark.

$$\begin{aligned} \text{FLOPS/Watt} &= \frac{\text{Floating Point Operations/Second}}{\text{Joules/Second}} \\ &= \frac{\text{Floating Point Operations}}{\text{Joule}} \end{aligned} \quad (6)$$

V. Experimental System and Setup

To evaluate the use of statistical techniques on performance and power and to perform energy efficiency tuning on the HPL benchmark, we use a cluster platforms called SystemG. Each node in the cluster consists of two quad-core 2.8 GHz Intel Xeon 5462 processors and 8GB of RAM. The nodes are connected over a QDR InfiniBand interconnect technology. We use 64 nodes from SystemG. We use OpenMPI 1.4.1 as the communication library.

Figure 1 shows the experimental setup. The power meter acts as an intermediate device between the power supply and the system under test. The power and energy values necessary can be extracted from the power meter using the measuring machine. In our experiments, we use a "Watts UP? Pro ES" power meter which has a maximum resolution of one second. The power values for all the 64 nodes are extrapolated from a single node in all our experiments. Note that dynamic voltage-frequency techniques were not applied for any of the experiments. The performance CPU frequency governor was used if applicable.

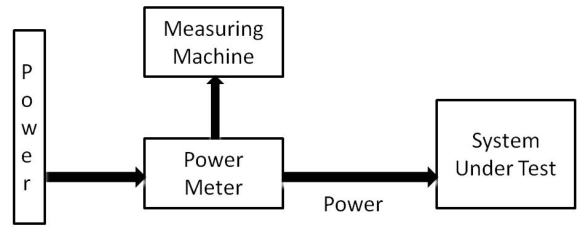

Fig. 1. Power Meter Setup

#	Parameter Name	Observations used
1	N	10 to 50% of memory
2	NB	64, 96, 128, 160, 176, 192, 224, 240, 248, 256, 264, 272, 280, 296
4	PxQ	All P and Q combinations $\forall P * Q =$ Number of cores

TABLE II
List of Observations Used for Creating the Models

#	Model	F-ratio
1	Time-Baseline Plot	555.9187
2	Time-Transformed Plot	2900.5090
4	Energy-Baseline Plot	612.7036
4	Energy-Transformed Plot	3203.3270

TABLE III
F-ratio For Various Models

VI. Experimental Results

In this section, we discuss about the results for predictor selection and optimizations and evaluate the goodness of the fit as described in Section III. We then use these models to perform energy efficiency tuning on the HPL benchmark.

We leverage the knowledge about the HPL benchmark to model the power and performance indirectly (Refer Section II). We create a model for energy and execution time of the benchmark and calculate the power and performance (FLOPS) from these models. The FLOPS is calculated indirectly from the execution time of the benchmark by using Equation (1). This is done to achieve higher accuracy of prediction as energy and execution time vary relatively more linearly with respect to the predictors as compared to power and performance.

A. Predictor Selection and Optimization

In this section, we evaluate the statistical significance of the predictor variables used for creating the model for performance and power. Table II shows the observations used to create the model. The same observations are also used for studying the statistical significance of the predictor variables.

#	Model	R^2	Adjusted R^2
1	Time-Baseline Plot	.764	.763
2	Time-Transformed Plot	.944	.944
4	Energy-Baseline Plot	.781	.780
4	Energy-Transformed Plot	.949	.949

TABLE IV
R^2 AND ADJUSTED R^2 VALUES FOR VARIOUS MODELS

#	Performance Achieved	MFLOPS/Watt
1	R_{max}	187.70
2	R_{eff}	188.02

TABLE V
OBSERVED ENERGY EFFICIENCY AT R_{max} AND R_{eff}

As mentioned in Section III-B, there are several techniques to evaluate the statistical significance of a predictor variable. Table III shows the F-ratio for various models used in this paper. There are two sets of observed values used for creating the required models in this paper, they are:

- Baseline: Observed values with no transformation on the predictor or the response variables
- Transformed: Observed values with logarithmic transformation on the response variable and no transformation on the predictor variables.

Logarithmic transformations are commonly applied to stabilize the variance in a variable. As seen, the F-ratio for all the models are good and lie in an acceptable range. The transformed plots for both time and energy modeling have a higher F-ratio suggesting that the predictor variables gain more significance in the transformed models. Another metric used for testing the significance of the predictor variables is p-value. In all the models, both baseline and transformed, the p-value for individual variables are less than 0.0001. This suggests that there is 0.01% chance that these variables do not account for a statistically significant part of the variability in the response. Correlation between the predictor must also be analyzed to understand whether the variable accounts for a separate portion of variability in the response. In our case, all the variables have a correlation of the magnitude less than one except P and Q. This is because of the way the observations are chosen as all possible combination of P and Q are chosen. However, deeper analysis reveals that $P * Q$ and $Q * P$ for the same P and Q can have different impact on performance and power. Thus we retain both the variables and construct our models using all the four variables.

#	Value	MFLOPS/Watt
1	Predicted R_{eff}	187.07
2	Observed R_{eff}	179.27

TABLE VI
PREDICTED ENERGY EFFICIENCY AT R_{eff}

B. Modeling

As mentioned earlier, we use two plots for both time and energy modeling. We apply log transform as a mechanism to stabilize the variance in the plots. In this section, we analyze the statistical significance of the entire model using metrics discussed in Section III-C. The model can be assessed by plotting the observed values vs. the predicted value and using the regression line as the reference for understanding the goodness of the fit. Note that time and energy are measured in seconds and joules respectively in all the plots. Figure 2 shows the plot for modeling both the baseline and transformed response (time). As observed, more predicted values on the transformed model lie closer to the regression line which suggests that the variance stabilizing transform has increased the accuracy of the model. Similar results can be seen for modeling energy in Figure 3. Transformed models for both energy and time have more accuracy than baseline model.

Residual plots can be useful in determining whether the residuals are correlated to any of the variables. A random distribution of the residual over all the predicted values is considered ideal. Figures 4 and 5 show the residual distribution of time and energy for both the baseline and the transformed plot. There is no specific pattern that can be seen from both the plots. The maximum outliers of transformed plots are far less than the maximum outliers in the baseline plots for both the

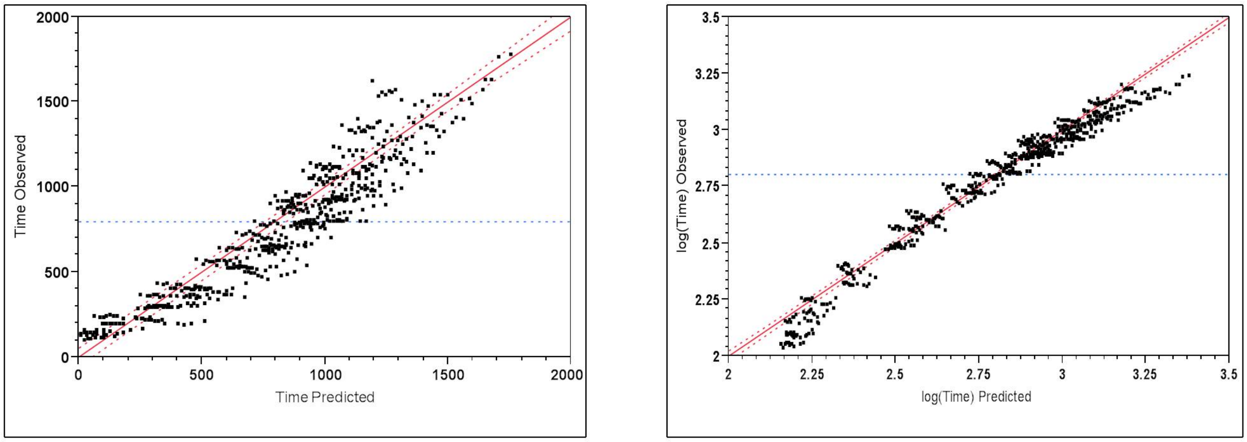

Fig. 2. Description of Fit for Modeling Time A. Baseline Plot B. Transformed Plot

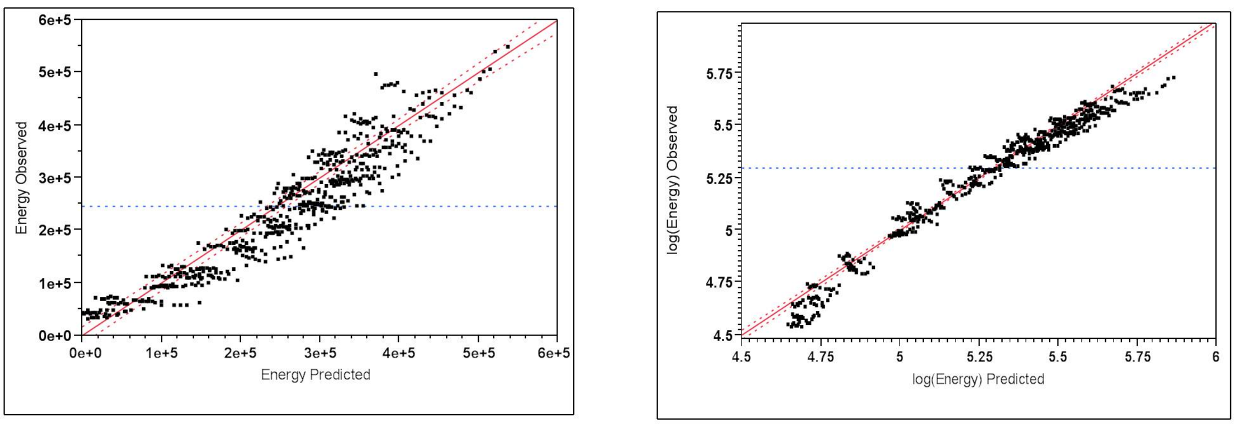

Fig. 3. Description of Fit for Modeling Energy A. Baseline Plot B. Transformed Plot

models suggesting that the transformed plots are more accurate. R^2 and adjusted R^2 are used to understand the variability of the response that is explained by the model. Table IV shows R^2 and adjusted R^2 values for all the models in this paper. Clearly, the R^2 and adjusted R^2 values for the transformed models are better than the baseline plots. Adjusted R^2 values show that the fit in the model is achieved by using proper predictor variables and number of observations and not by introducing any variables artificially just to increase the accuracy of the model. Another metric that needs to be looked at is the overall p-value. The overall p-value for all the models is less than .001 reflecting the accuracy of these models.

C. *Energy Efficiency Tuning*

In this section, we use the models created for time and energy to tune the HPL benchmark to achieve the maximum energy efficiency possible. Table V shows the observed values for energy efficiency at R_{max} and R_{eff}. We define R_{eff} as the performance achieved while executing the HPL benchmark at the maximum energy efficiency possible based on particular energy efficiency metric. As seen from Table V the best energy efficiency is not always achieved while executing the benchmark at R_{max}. The R_{max} for this system stands at 3.65 TFLOPS. However, the R_{eff} while using MFLOPS/Watt is 3.62 TFLOPS. The R_{eff} while using MFLOPS/Watt is very close to R_{max} and has energy efficiency slightly higher than R_{max}. Investigations into such trade-offs are not involved in performance tuning.

As shown, in some cases applications execute at the best energy efficiency while operating at performance less than R_{max}. This is due to the fact that components other than the CPU are contributing to energy consumed

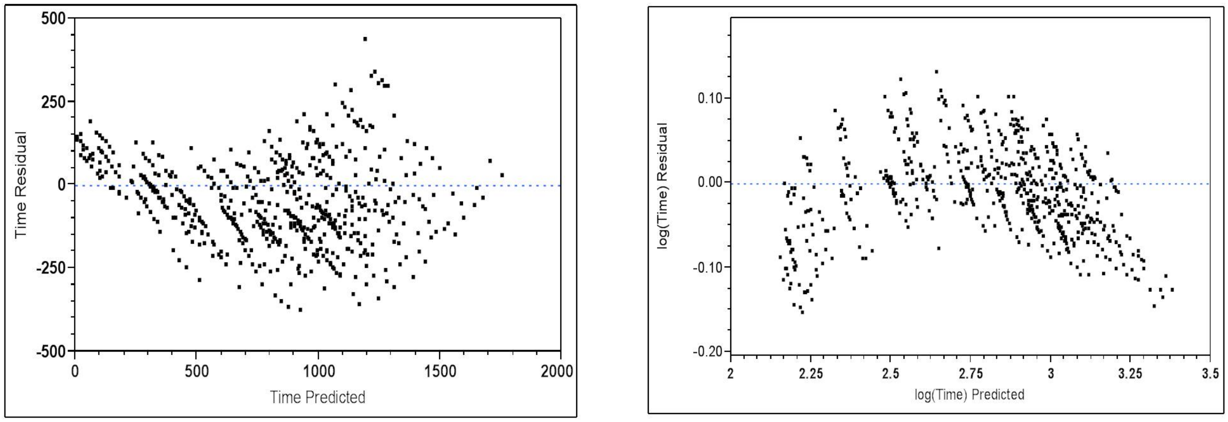

Fig. 4. Residual Distribution for Time Model A. Baseline Plot B. Transformed Plot

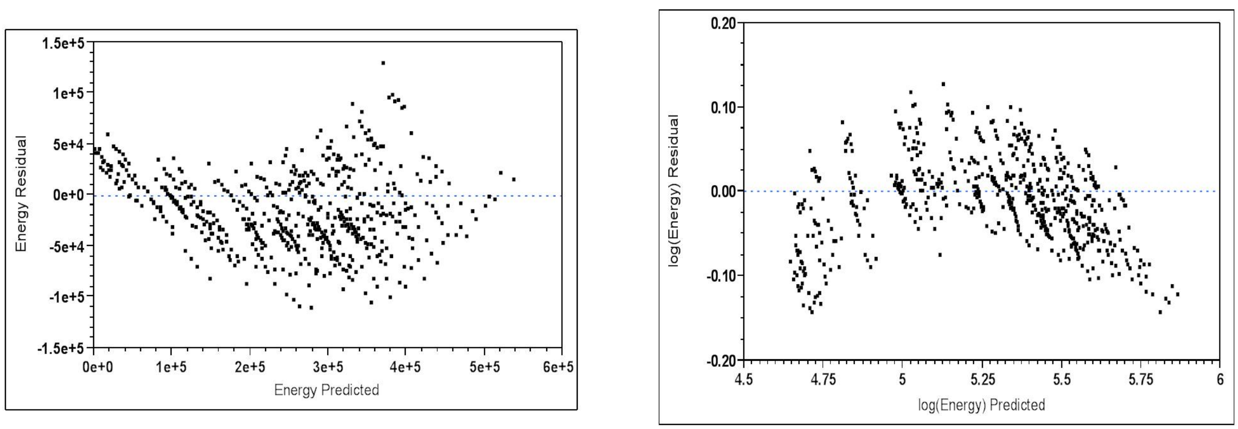

Fig. 5. Residual Distribution for Energy Model A. Baseline Plot B. Transformed Plot

while executing the application. Given the fact that energy efficiency is a major obstacle in achieving large scale performances, there is a need to understand the contributions to energy consumption from components other than the CPU. From another perspective, there is an opportunity for optimizing the applications based on the energy efficiency while remaining in an acceptable performance range. We optimize the energy efficiency of HPL benchmark by using the models to automate the process of finding the parameters of HPL benchmark to achieve maximum possible energy efficiency. Table VI shows the R_{eff} values predicted by the model created in the previous section. The model predicts that the performance for R_{eff} using MFLOPS/Watt as 3.58 TFLOPS. However, the actually observed performance while executing the benchmark with the predicted parameters is 3.42 TFLOPS. The error percentage for predicting the performance at R_{eff} is 4.6%. These models predict the R_{eff} at a reasonably high accuracy. The error percentage in predicting the MFLOPS/Watt for R_{eff} is 4.8%. This process simplifies the search space for finding the parameters for executing at the best energy efficiency possible.

VII. Related Work

Several analytical and statistical models have been developed for predicting the performance of HPL and to automate the process of finding the HPL parameters to achieve R_{max}. In [4], Chau-Yi et al. illustrate the use of analytical modeling to predict the performance of HPL benchmark. They propose a semi-empirical model to predict R_{max}. An improvement over existing model is proposed by dividing the performance modeling into computational cost and message passing overhead and using the communication model proposed in [12]. Their model achieves a prediction error of less than 5% on different cluster platforms. However, analytical models are accurate only when the underlying assumption while modeling is met.

Due to the disadvantages of using analytical models, several researchers have predicted the performance of HPL by using statistical approaches. In [11], Tuan et al. find the HPL parameters that has highest relevance to performance and validate the existing analytical and popularly used rules of thumb to tune the performance of HPL by using data mining techniques such as linear regression, M5P, Multilayer perceptron (MLP) and Support Vector Machine (SVM). The authors also try to tune HPL at smaller problem sizes and validate if the same configuration can be applied for larger problem sizes. Authors of [10] use artificial neural networks and linear regression techniques to predict the performance of Semicoarsening Multigrid and HPL. They use hierarchical clustering, association analysis and correlation analysis to characterize the parameter space of the applications. While both these work highlight the importance of accurately predicting the performance of applications, the authors do not address the problem of energy efficiency.

Several performance modeling and predictions for scientific applications have been implemented in works such as [8], [7] and [13]. In [9], the authors propose the use of regression techniques to predict the micro-architectural power and performance. The paper addresses the problem of the micro-architectural simulation costs by reducing the number of observation required and eases the process of power and performance prediction. The techniques used achieve a mean error of 4.9% for performance and 5.6% for power. In this paper, we use a similar approach but use the techniques to predict the performance and power for scientific application.

VIII. Conclusion and Future Work

In this paper, we emphasize the need for energy efficiency tuning of scientific application. We find the parameters of the HPL benchmark that can be used for creating a model to predict the power and performance of the benchmark. We show that maximum energy efficiency is not always achieved at highest possible performance and construct models to predict power and performance of the HPL benchmark. Finally, we use the model to automate the process of finding the parameters for achieving maximum possible energy efficiency. As a future work, we would like to investigate the feasibility of energy efficiency tuning in other scientific applications.

Acknowlegements

This research project was supported by National Science Foundation under grant CCF-0848670.

References

[1] High Performance LINPACK (HPL). Available at http://www.netlib.org/benchmark/hpl.

[2] The Top500 list. Available at http://top500.org.

[3] K. Bergman, S. Borkar, D. Campbell, W. Carlson, W. Dally, M. Denneau, P. Franzon, W. Harrod, K. Hill, J. Hiller, S. Karp, S. Keckler, D. Klein, R. Lucas, M. Richards, A. Scarpelli, S. Scott, A. Snavely, T. Sterling, R. S. Williams, K. Yelick, and P. Kogge. Exascale Computing Study: Technology Challenges in Acheiving Exascale Systems.

[4] C. Chou, H. Chang, S. Wang, K. Huang, and C. Shen. An Improved Model For Predicting HPL Performance. In *Proceedings of the 2nd international conference on Advances in grid and pervasive computing*, pages 158–168, Paris, France, 2007. Springer-Verlag.

[5] F. Harrell. *Regression Modeling Strategies*. Springer, 2001.

[6] C. Hsu, W. Feng, and J. S. Archuleta. Towards Efficient Supercomputing: A Quest for the Right Metric. In *Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS'05) - Workshop 11 - Volume 12*, page 230.1. IEEE Computer Society, 2005.

[7] E. Ipek, B. R. D. Supinski, M. Schulz, and S. A. Mckee. An Approach to Performance Prediction for Parallel Applications. *EURO-PAR, SPRINGER LNCS*, 3648:2005, 2005.

[8] D. J. Kerbyson, H. J. Alme, A. Hoisie, F. Petrini, H. J. Wasserman, and M. Gittings. Predictive Performance and Scalability Modeling of a Large-Scale Application. In *Proceedings of the 2001 ACM/IEEE conference on Supercomputing*, pages 37–37, Denver, Colorado, 2001. ACM.

[9] B. C. Lee and D. M. Brooks. Accurate and Efficient Regression Modeling for Microarchitectural Performance and Power Prediction. In *Proceedings of the 12th international conference on Architectural support for programming languages and operating systems*, pages 185–194, San Jose, California, USA, 2006. ACM.

[10] B. C. Lee, D. M. Brooks, B. R. de Supinski, M. Schulz, K. Singh, and S. A. McKee. Methods of Inference and Learning for Performance Modeling of Parallel Applications. In *Proceedings of the 12th ACM SIGPLAN symposium on Principles and practice of parallel programming*, pages 249–258, San Jose, California, USA, 2007. ACM.

[11] T. Z. Tan, R. S. M. Goh, V. March, and S. See. Data Mining Analysis to Validate Performance Tuning Practices for HPL. In *2009 IEEE International Conference on Cluster Computing and Workshops*, pages 1–8, New Orleans, LA, USA, 2009.

[12] Z. Xu and K. Hwang. Modeling Communication Overhead: MPI and MPL Performance on the IBM SP2. *Parallel & Distributed Technology: Systems & Applications, IEEE*, 4(1):9–24, 1996.

[13] L. T. Yang, X. Ma, and F. Mueller. Cross-Platform Performance Prediction of Parallel Applications Using Partial Execution. In *Proceedings of the 2005 ACM/IEEE conference on Supercomputing*, page 40. IEEE Computer Society, 2005.