Technical Paper

ISSN 1997-1400 Int. J. Pavement Res. and Technol. 1(3):85-93
Copyright @ Chinese Society of Pavement Engineering

Statistical Models for Determination of the Resilient Modulus of Subgrade Soils

Pranshoo Solanki[1+], Ali Ebrahimi[2], and Musharraf M. Zaman[3]

Abstract: A combined laboratory and modeling study was undertaken to develop a database for subgrade soils in Oklahoma and to develop relationships or models that could be used to estimate resilient modulus (M_R) from commonly used subgrade soil properties in Oklahoma. Sixty-three soil samples from 14 different sites throughout Oklahoma were collected and tested for the development of the database and the statistical models. Additionally, thirty-four soil samples from 3 different sites, located in Rogers and Woodward counties, were collected and tested to evaluate the developed models. The routine material parameters selected in the development of the models included moisture content (w), dry density (γ_d), plasticity index (PI), percent passing No. 200 sieve (P_{200}), and unconfined compressive strength (U_c). Bulk stress (θ) and deviatoric stress (σ_d) were used to identify the state of stress. Several statistical models were developed in this study. These models include: stress-based, multiple regression, polynomial, and factorial. Each model was ranked based on its R^2 (goodness of fit) and F values (significance of the model) for the development dataset. Based on the R^2 and F values, the second order polynomial and factorial models were further considered for the evaluation dataset. An evaluation of the two models indicated that for the combined development and evaluation datasets, a second order polynomial model is a good statistical model for evaluating M_R from the selected routinely determined properties. The models developed in this study are expected to be useful in the Level 2 and Level 3 designs of pavements in Oklahoma.

Key words: *Factorial model; Hierarchical approach; Polynomial model; Resilient modulus; Statistical model; Subgrade.*

Introduction

Empirical design methods for flexible pavement structures are primarily based on the equations that were developed largely from the AASHO Road Tests conducted in 1950's. These methods fail to reflect the dynamic nature of traffic loads. Therefore, the mechanistic design methods referred to as the "AASHTO Guide for Design of Pavement Structure" [1] recommended the use of resilient modulus (M_R), a dynamic-strength parameter, to characterize the flexible pavement materials. The M_R accounts for the cyclic nature of vehicular traffic loading, and is defined as the ratio of deviatoric stress to recoverable elastic strain.

Several laboratory and field procedures are currently either used or evaluated for determining a design M_R value of subgrade soil. Direct laboratory methods used for evaluating M_R during the past two decades includes resonant column, torsional shear, gyratory, and repeated load triaxial testing [1-4]. Among these testing procedures, the M_R from repeated load triaxial test (RLTT) is used most frequently because of the repeatability of test results and its representation of field stress in controlled laboratory environments. RLTT is conducted in the laboratory on remolded or undisturbed samples according to different AASHTO test methods of which AASHTO T-307-99 is used frequently [5]. The AASHTO T-307-99 test method can be a time consuming and expensive test method, particularly for small projects.

In the new 2002 AASHTO guide, which is currently in the evaluation stage, a hierarchical approach is used to determine different design inputs including M_R [5]. It requires the evaluation of the engineering properties of subgrade soils in laboratory or field to pursue a Level-1 (most accurate) design. For a Level-2 (intermediate) design, however, the design inputs are user selected, possibly from agency database or from limited testing program or could be estimated through correlations [5]. A Level-3 design, which is the least accurate and generally not recommended, uses only the default values. For Level-2 designs a regression model for M_R can be very useful as it provides the designer with significant flexibility in obtaining the design inputs for a project.

In the present study, conventional laboratory tests were conducted on some commonly encountered subgrade soils in Oklahoma, and statistical analyses were conducted to develop regression models for M_R for Level-2 pavement design applications. The models developed herein consider both stress (deviatoric stress and bulk stress) and commonly used properties (unconfined compressive strength, dry density (γ_d), moisture content (MC), gradation, and Atterberg limits). The strengths and the weaknesses of the developed statistical models were also examined using additional M_R test results that were not used in the development of these models.

Review of Previous Studies

Several pertinent studies have previously been undertaken to develop empirical correlations to estimate M_R values in terms of other soil properties. One of the commonly used models to represent

[1] School of Civil Engineering and Environmental Science, University of Oklahoma, Norman, Oklahoma, USA.
[2] Burgess Engineering and Testing, Inc., Moore, Oklahoma, USA.
[3] College of Engineering, University of Oklahoma, Norman, Oklahoma, USA
[+] Corresponding Author: E-mail pranshoo@ou.edu
Note: Submitted April 22, 2008; Revised May 28, 2008; Accepted May 29, 2008.

M_R is the power model (see e.g., [6-14]). Dunlap [6] proposed the following correlation for M_R:

$$M_R = k_1(\sigma_3/P_a)^{k_2} \tag{1}$$

where σ_3 is a confining pressure, P_a is a reference pressure (e.g., atmospheric pressure) and k_1 and k_2 are the regression coefficients.

A number of researchers (see e.g., [15-20]) have utilized other soil property indices to estimate M_R. For example, Drumm et al. [21] developed two regression models for M_R of fine-grained soils as a function of deviator stress and soil-index properties, namely, percentage passing No. 200 sieve, plasticity index (PI), γ_d, and unconfined compressive strength. A relatively small (twenty-two) number of these samples were used in developing these models.

In a similar study, Lee et al. [22] investigated the M_R of cohesive soils, mainly clayey subgrade soils, with RLTT. Specimens were compacted using standard and modified proctor methods at near optimum moisture content (OMC) in a mold with a diameter of 38mm (1.5inches) and a height of 100mm (4 inches). It was seen that the custom-compaction results were in close agreement with the maximum dry density (MDD) and the OMC from the standard and modified Proctor tests. Regression analyses were conducted to obtain a relationship between M_R and the stress in unconfined compressive strength (U_C) test causing 1% strain ($S_{U1.0\%}$) in laboratory compacted specimens. The relationship between M_R and $S_{U1.0\%}$ for a given soil was found to be unique regardless of MC and compaction effort. The results showed that the M_R and $S_{U1.0\%}$ vary with the MC in a similar manner. Furthermore, four different compactive efforts were used in that study, but a single relationship between M_R and $S_{U1.0\%}$ was obtained, as presented in Eq. (2):

$$M_R = 695.4\,(S_{U1.0\%}) - 5.93\,(S_{U1.0\%})^2 \tag{2}$$

where M_R = resilient modulus at maximum axial stress of 41.4kPa and confining pressure of 20.7kPa; and $S_{U1.0\%}$ = stress (in kPa) causing 1% strain in conventional U_C test.

Moreover, the relationship was similar for different cohesive soils, indicating that it may be applicable for different types of clayey soils. The limited data suggested that the same correlation might be used to estimate the M_R for both laboratory and field compacted conditions.

In a field study, Yau and Von Quintus [10] proposed the following correlation using the M_R data obtained from the LTPP test sections:

$$M_R = k_1 P_a (\theta/P_a)^{k_2}[(\tau_{OCT}/P_a)+1]^{k_3} \tag{3}$$

where τ_{oct} is the octahedral shear stress, and k_1, k_2, and k_3, are the regression constants. Yau and Von Quintus expressed these regression constants as a function of MC, γ_d, optimum dry density, liquid limit, percent silt, percent clay, and percent passing different sieve sizes. The soils were classified into three different groups (coarse grained sandy soils, fine grained silty soils, and fine grained clayey soils), and the regression constants were developed for each soil type.

Most recently, Khazanovich et al. [14] used M_R results for 23 samples from several locations in Minnesota and evaluated the regression constants for use in the mechanistic-empirical-based pavement designs. However, because the mineralogical and textural characteristics of soils in Oklahoma are different from those in

Table 1. Summary of USCS Soil Classification Results for the Development and Evaluation Dataset.

Soil Classification		Number of Soils	
		Development Dataset	Evaluation Dataset
Unified Soil Classification System (USCS)			
Fat clay	CH	2	1
Fat clay with sand		0	1
Sandy fat clay	CH	1	0
Lean clay	CL	23	8
Lean clay with sand	CL	22	8
Gravelly lean clay	CL	2	0
Sandy lean clay	CL	8	10
Sandy lean clay with gravel	CL	0	2
Silty clay with sand	CL-ML	1	0
Sandy silty clay	CL-ML	1	3
Clayey Sand	SC	1	1
Clayey Sand with gravel	SC	2	0
	Total :	63	34
AASHTO Classification System			
A-4		12	4
A-6		34	18
A-7-6		17	12
	Total :	63	34

Table 2. Basic Statistical Parameters for Specimen PI, P200, MC, DD and UCS.

Dataset	*No of Soil*	PI			P200			MC			DD			UCS		
		S_D	SK	KU	S_D %	SK	KU	S_D %	SK	KU	S_D kg/m^3	SK	KU	S_D kPa	SK	KU
Dev	126	8.4	0.97	0.63	15.0	-1.05	0.37	3.0	0.39	-0.14	106.7	-0.26	-0.44	70.8	0.68	0.50
Eva: RC+WC	68	8.9	-0.03	-1.22	13.1	-0.49	0.27	2.9	-0.24	-0.94	74.8	0.40	-0.60	64.4	0.78	0.81
Eva: RC	58	8.4	-0.32	-0.89	13.9	-0.59	0.16	2.6	-0.49	-0.37	66.0	0.58	-0.02	56.6	0.20	-0.57
Eva: WC	10	3.0	-1.43	1.58	8.0	1.88	3.89	1.2	-0.25	-0.88	46.0	0.13	-0.18	95.9	0.99	-0.61

S_D: Standard Deviation; SK: Skewness; KU: Kurtosis;
PI: Plasticity Index; P200: Percentage passing #200 sieve; MC: Specimen moisture content; DD: Specimen dry density;
Dev: Development dataset; Eva: Evaluation dataset; RC: Rogers County; WC: Woodward County; UCS: Unconfined compressive strength

Minnesota, those results may not be directly used for pavements in Oklahoma for a Level 2 design.

Sources and Characteristics of Subgrade Soils

In the present study, a total of 97 bulk soils samples were collected from 16 different counties in Oklahoma. Of these, 63 samples from 14 different counties were used in the development of the statistical models and are collectively referred to as the *development dataset*. These sites were located in Adair, Alfalfa, Choctaw, Delaware, Greer, Jefferson, Kingfisher, Lincoln, Major, McClain, Noble, Okfuskee, Osage, and Rogers counties in Oklahoma. The remaining 34 soils from two different counties namely, Rogers and Woodward counties, were used for the evaluation of the regression models. Data for these soils are collectively referred to as the *evaluation dataset*. A majority of soils in the development dataset was lean clay and lean clay with sands (Table 1). A majority of soils in the evaluation dataset, on the other hand, was lean clay, lean clay with sand and sandy lean clay.

Laboratory Testing and Result

The laboratory-testing program included routine laboratory tests, namely grain size distribution (AASHTO T11 and AASHTO T27), Atterberg limits (AASHTO T89 and AASHTO T90) and standard proctor (ASTM D698), as well as resilient modulus (AASHTO T307) and unconfined compression (AASHTO T208). Using the proctor test results, two samples were prepared for each soil with different compaction conditions. One of these samples was compacted at the OMC and 95% of the MDD. For the other sample, the MC and γ_d were set at 2% wet of OMC, representing the Oklahoma Department of Transportation (DOT) in-construction stage requirements [23]. Specimens having a moisture variation of more than ±0.5 percent from the targeted MC and γ_d less than 95% of MDD were discarded and new samples were compacted, evaluated, and tested. Thus, a total of 126 M_R tests were conducted for 63 soils used in the development dataset. Likewise, 68 M_R tests were conducted for 34 soils in the evaluation dataset. A static compaction method (a modified version of the double plunger method) was used in sample preparation [5]. The unconfined compressive strength (UCS) test was conducted on the same sample, following the M_R testing. It is assumed that since the M_R strain is in the range of ten thousands (mm/mm) the influence of M_R test on the UCS test would be negligible [23].

A summary of the basic statistical parameters for PI, percentage passing #200 sieve, specimens MC, γ_d, and UCS is tabulated in Table 2. Further statistical details of different parameters (i.e. liquid limit, plastic limit, PI, percentage passing #4, #10, #40, and #200 sieve, specimens MC, γ_d, and UCS) used in this study are given in Ebrahimi [23]. Montgomery et al. [24] recommended that datasets deviating from normal distribution would not affect the outcome of the analysis and the results would not be critically affected. Hence, in this study Kurtosis and Skewness were determined to select the input parameter for regression modeling. Kurtosis parameter is an indicator of heaviness of the tail. A perfectly normal distribution of data has a Kurtosis of zero. A positive Kurtosis is an indication of more observations on the tail end of the distribution curve, while a negative Kurtosis is an indication of fewer observations on the tail end of the distribution curve. Skewness is a measure of distribution of the data. A skewness of zero indicates perfectly normal distribution of data. Negative value of skewness indicates the data skewed left and positive value indicates the data skewed right.

Grain Size Distribution and Plasticity Index

As recommended by the Oklahoma DOT specifications, only selected sieves (#4, #10, #40, and #200) were used in the grain size distribution tests. The Skewness and Kurtosis values for #4, #10, and #40 were fairly high indicating that these results were not normally distributed [23]. For #200 sieve, however, the overall skewness (-0.49 to -1.05) and Kurtosis (0.27 to 0.37) values were much smaller, indicating that these data could be assumed normally distributed (Table 2). Thus, from the grain size distribution tests, only percent passing #200 sieve (P_{200}) was used as an input parameter in the regression modeling. From Table 2, the Skewness (-0.03 to 0.97) and Kurtosis (0.63 to -1.22) for both datasets were negligible, so the distributions of the PI for both datasets may be considered normal. The PI was used as an input parameter in the regression analysis.

Moisture Content (MC) and Dry Density (γ_d)

The Skewness of MC data for the development and the evaluation datasets were 0.39 and -0.24, respectively (Table 2). The corresponding Kurtosis values were -0.14 and -0.94, respectively. These values indicate that the MC data were also approximately normally distributed. For both datasets, the γ_d values were normally distributed, the Rogers County soils being closer to the development dataset than the Woodward County soils (Table 2).

Unconfined Compressive Strength

The Skewness and Kurtosis for the development dataset were 0.68 and 0.50, respectively (Table 2). The corresponding Skewness and Kurtosis for the evaluation dataset were 0.78 and 0.81, respectively. Overall, the U_c values were normally distributed and used in the regression analysis.

Resilient Modulus

The M_R test results for the development and the evaluation datasets are presented in Table 3. For the development dataset, a very high standard deviation, more than 340MPa (49.3ksi), is seen for the loading sequences 1, 6, and 11. A high standard deviation (more than 89MPa or 12.9ksi) is also observed for the evaluation dataset for the same loading sequences. For each of these loading sequences, the applied axial stress is the lowest (13.8kPa or 2psi), resulting in very small deformation of the sample that are difficult to measure due to electrical noise. As a result, the M_R values for these loading sequences were not used in developing the regression model.

Model Development

Table 3. Basic Statistical Parameters for Resilient Modulus at Each Sequence for Development and Evaluation Dataset.

Sequence No.	Confining Pressure (kPa)	Axial Stress (kPa)	Resilient Modulus (MPa)							
			Mean		Minimum		Maximum		Standard Deviation	
			Dev	Eva	Dev	Eva	Dev	Eva	Dev	Eva
1	41.4	13.8	298.6	172.8	54.6	58.1	2042.3	409.1	347.6	89.7
2	41.4	27.6	86.8	80.9	34.5	28.2	229.0	131.3	31.7	21.9
3	41.4	41.4	72.3	66.4	24.6	19.6	163.8	122.1	28.2	24.6
4	41.4	55.2	63.4	54.5	24.3	17.2	155.7	117.8	28.0	21.2
5	41.4	68.9	57.9	49.0	20.7	16.1	152.6	104.4	26.8	19.1
6	27.6	13.8	269.4	161.8	41.9	57.1	2160.8	494.2	341.2	93.1
7	27.6	27.6	84.1	79.0	28.7	33.9	245.9	136.7	33.2	22.9
8	27.6	41.4	70.6	64.2	23.2	22.7	159.5	121.0	28.6	24.4
9	27.6	55.2	63.2	53.5	22.3	18.8	151.7	115.9	28.3	21.2
10	27.6	68.9	58.3	49.4	20.9	16.8	149.2	106.2	27.0	19.6
11	13.8	13.8	312.4	190.8	36.1	54.1	1892.2	826.4	393.3	146.3
12	13.8	27.6	93.3	80.4	23.7	36.2	979.9	135.7	87.9	23.4
13	13.8	41.4	76.9	64.9	19.8	23.9	727.9	126.0	65.9	25.1
14	13.8	55.2	63.9	53.8	19.6	19.4	159.8	116.3	29.6	21.7
15	13.8	68.9	59.4	49.8	18.2	17.3	190.9	107.4	29.3	20.0

Dev: Development Dataset (126 specimens); Eva: Evaluation Dataset (68 specimens)
1kPa = 0.145psi

In the present study, mainly four statistical models were developed, namely stress-based, multiple regression, polynomial, and factorial. The M_R values were predicted using the evaluation dataset and then compared to the experimental M_R values. The R^2 and F values were utilized as the basis of comparing the developed models in regard to the goodness of fit and significance of the model, respectively [25,26].

Stress-Based Model (SBM)

As noted previously, there are several stress-based models (SBM) available for prediction of M_R [6-12]. In this study, to develop stress-based model, bulk stress (θ) and deviatoric stress (σ_d) were used as the model parameters:

$$M_R/P_a = k_1(\theta/P_a)^{k_2}(\sigma_d/P_a)^{k_3} \quad (4)$$

where P_a represents atmospheric pressure, and k_1, k_2, and k_3 are regression constants. These regression constants are determined for each M_R test in the development dataset. The R^2 values for individual tests in the development dataset ranged from 0.512 to 0.996 [23]. Then, using the multiple linear regression option in Statistica 7.1[27], these regression constants are correlated with the specimen and soil parameters. The following expressions are obtained for k_1, k_2, and k_3 from the development dataset:

$$k_1 = 0.08789 + 0.1773\,(U_c/P_a) + 0.005048\,PI - 0.3967\,P_{200} + 1.2652\,w \quad (5)$$

$$k_2 = 0.5074 - 0.01336\,PI + 2.3432\,w - 0.3868\,(\gamma_d/\gamma_\omega) \quad (6)$$

$$k_3 = -0.6612 + 0.1589\,(U_c/P_a) - 0.2254\,P_{200} \quad (7)$$

where w and γ_ω are moisture content and dry density of molded sample.

Fig. 1 shows a plot of the experimental and predicted M_R/P_a values for this model. The overall R^2 value was found to be as low as 0.3226 suggesting that the overall model did not back-predict the M_R/P_a values favorably. Fig. 1 also shows a comparison between experimental and predicted M_R/P_a for three selected specimens, MA-3B, NO-7A, and OS-1B, respectively. The M_R results from these specimens covered the full range of M_R response for the development dataset. Specimen NO-7A shows the best prediction, followed by specimen OS-1B. Specimen MA-3B shows the worst back-prediction. The soil classification results for these specimens indicate lean clay with AASHTO classification of A-6(10), A-6(16), and A-6(21) for OS-1B, NO-7A, and MA-3B, respectively. The U_c results for OS-1B, NO-7A, and MA-3B were 161, 272, and 310kPa (23.3, 39.4, and 45.1 psi), respectively [23]. Thus, even though these soils are all classified.

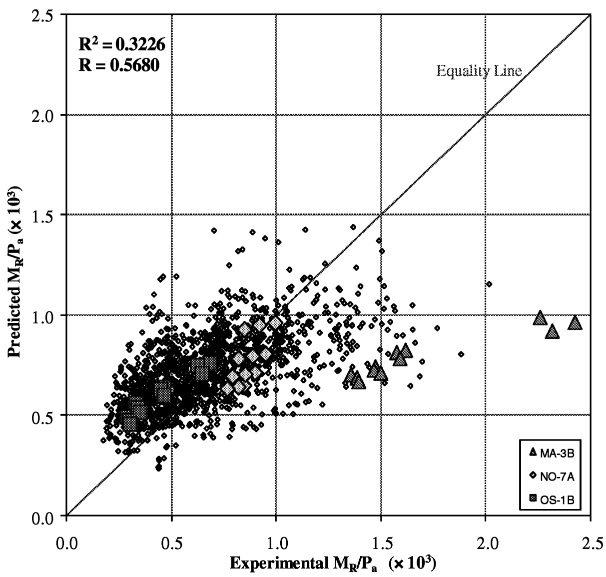

Fig. 1. Comparison of Experimental and Predicted M_R/P_a for Development Dataset: Stress-Based Model.

Table 4. Summary of R^2 and F Values for the Statistical Modeling Using Development Dataset.

Statistical Model	R^2	F
Stress-based	0.3226	253.37
Multiple Regression	0.4357	165.88
Polynomial	0.4858	101.02
Factorial	0.6595	23.74
Separate Slope	0.8722	56.45
Mixture Surface	0.5522	67.79
Homogeneity of Slope	0.9430	83.38

as A-6 soils, their UCSs were quite different. Overall, it was observed that the M_R values increased with increasing UCS. This may have been a contributing factor for the three specimens exhibiting different levels of correlations between the experimental and predicted M_R.

From Table 4, the F value for this model is 253.37, which is an indicator that a more complicated model may be desired for the development dataset used here. Based on the R^2 and F values, as well as from the aforementioned findings, it was concluded that the stress-based model is not appropriate for prediction of M_R.

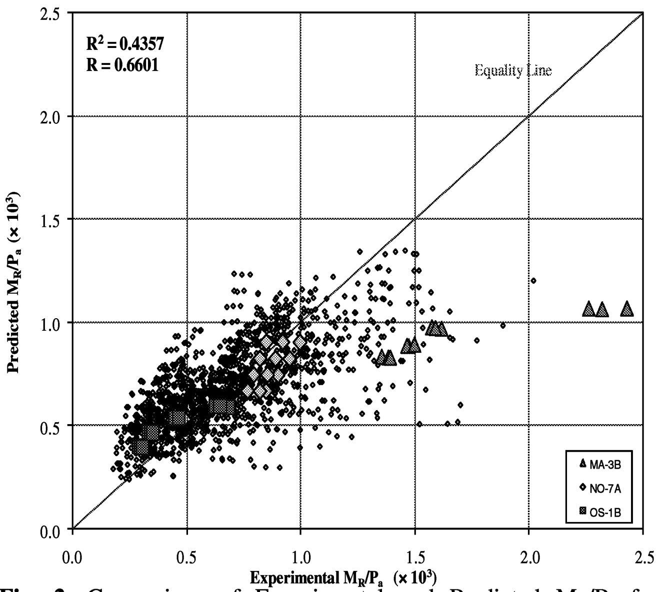

Fig. 2. Comparison of Experimental and Predicted M_R/P_a for Development Dataset: Multiple Regression Model.

Multiple Regression Model (MRM)

Multiple regression model (MRM) represents a class of simple and widely used linear regression models for more than two continuous variables [24, 27]. Using the same development dataset in Statistica 7.1, the following MRM was developed:

$$M_R/P_a = 1.8050 - 0.4904\,w - 0.5747\,(\gamma_d/\gamma_w) + 0.008083\,PI - 0.5123\,P_{200} + 0.2191\,(U_c/P_a) - 0.6401\,(\sigma_d/P_a) - 0.0009399\,(\theta/P_a) \quad (8)$$

The R^2 and F values for the MRM improved to 0.4357 and 165.88, respectively, which is a significant improvement over SBMs. These values are in agreements with Carmichael and Stuart [28] and indicate the importance of the size of the database.

Fig. 2 shows a comparison between experimental and predicted M_R/P_a values for this model. It is evident that the level of scatter in data points reduced significantly for this model as compared to SBM. The trend of the behavior of specimen MA-3B, NO-7A, and OS-1B is the same as that observed for SBM predictions. Hence similar reasons, as mentioned in the preceding section, can be used to justify this prediction. Also, it is evident that the predicted values are closer to the equality line when the M_R/P_a values are less than 1,000. This observation may be due to the distribution of dataset. Only 140 M_R/P_a values out of 1512 M_R/P_a values (approximately 9%) are in the upper range of 1,000. The remaining 91% of the M_R/P_a values for this study are in the lower range of the development dataset. As a result lower order MRMs appear to exhibit difficulty in back-predicting a majority of the resilient modulus values in the dataset that are in the lower range of the M_R/P_a values [24, 27, 29].

Polynomial Model (PM)

A polynomial model (PM) includes the basic components of a MRM with the addition of higher order effects for the independent variables. Although a second order model may be adequate for many problems, a general polynomial model can have higher than second order terms [23, 28]. In polynomial regression, higher order terms are added to the model to determine if they increase the associated R^2 significantly [24, 29, 30]. However, in most cases, orders of PMs greater than three are not practical [30].

Using the polynomial modeling option in Statistica 7.1, the resulting second order PM is given by the following equation:

$$M_R/P_a = 15.8002 + 2.9994\,w - 7.4142\,w^2 - 18.3291\,(\gamma_d/\gamma_w) + 5.4596\,(\gamma_d/\gamma_w)^2 + 0.02191\,PI - 0.0003142\,PI^2 - 0.3705\,P_{200} - 0.009229\,P_{200}^2 + 0.2628\,(U_c/P_a) - 0.01050(U_c/P_a)^2 - 2.0332(\sigma_d/P_a) + 1.62950(\sigma_d/P_a)^2 - 0.01181\,(\theta/P_a) + 0.004735(\theta/P_a)^2 \quad (9)$$

The R^2 and F values for this model were found to be 0.4858 and 101.02, respectively. These values were better than those of the MRM (0.4357 for R^2 and 165.88 for F value). To examine if a higher order model was desired, a third order polynomial regression model was developed for the same development dataset. The R^2 and F values for the third order polynomial model changed to 0.4101 and 254.75, respectively. Specifically, the R^2 value for the third order polynomial regression model was worse than the corresponding values for both the multiple regression and the second order polynomial regression models. Also, the F value increased from the second order to the third order polynomial regression model indicating that the second order PM was a better model [30, 31]. Fig. 3 presents a comparison of experimental and the M_R/P_a values back-predicted by the second order polynomial model. As seen for the other statistical models in the preceding sections, prediction for specimen MA-3B appears to be the worst, while the prediction for specimen NO-7A appears to be the best. Prediction for the third specimen OS-1B appears to be intermediate.

Factorial Model (FM)

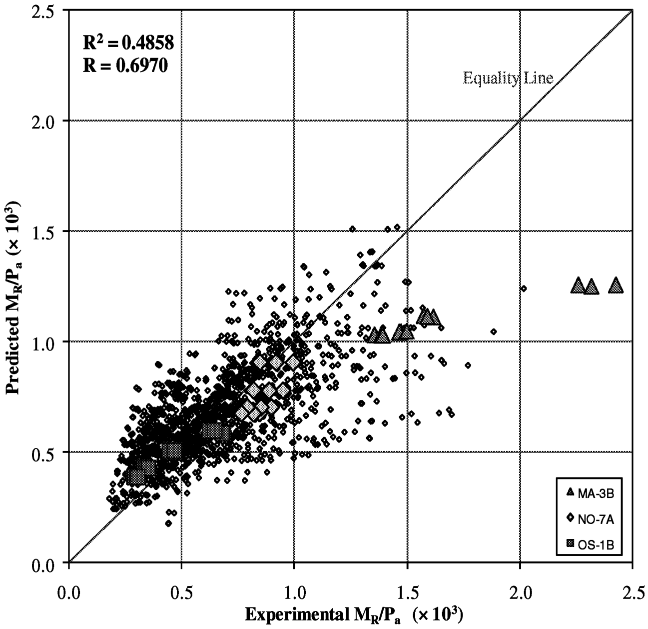

Fig. 3. Comparison of Experimental and Predicted M_R/P_a for Development Dataset: Polynomial Model.

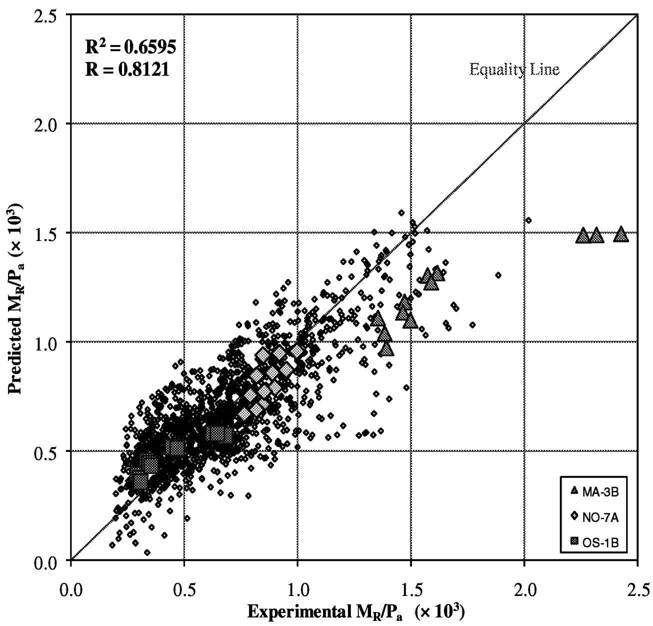

Fig. 4. Comparison of Experimental and Predicted M_R/P_a for Development Dataset: Factorial Model.

Similar to the polynomial model, a factorial model (FM) also includes the components of a MRM. However, instead of considering higher order effects of the independent variables, it accounts for interactions among different variables in the model. Different levels of interactions may be incorporated such as interactions between two variables, among three variables, and so on (i.e. $w \times \gamma_d$, $PI \times U_c \times \sigma_d$, $w \times \gamma_d \times PI \times \sigma_d \times \theta$, etc.). A full-factorial regression model consists of all possible products of the independent variables. Moreover, a factorial regression model can be fractional (i.e., fractional exponent) (see e.g., [24, 29]).

A full-factorial model is used in the present study. With seven independent variables and all possible products of the independent variables, the FM is a long equation with 128 terms. All the regression constants for this model were determined using Statistica 7.1. The resulting equation of the FM is presented in Appendix. The R^2 and F values for the FM were 0.6595 and 23.74, respectively. The R^2 is significantly higher than those for the previous models (0.4858). Significant observation was also made by the decrease of the F value from 101.02 for PM to 23.74 for FM. Fig. 4 shows a plot of experimental versus predicted M_R/P_a values and a comparison of the predicted M_R values against deviatoric stress for specimens MA-3B, NO-7A, and OS-1B for FM. As expected, the FM predicted the resilient modulus values of specimen NO-7A very closely, while the prediction for specimen MA-3B is much worse. Furthermore, because of the improvement in R^2 and F values both the goodness of fit of the model and the significance model the observations between the experimental and predicted values are closer. It may therefore be assumed at this point that since the F-value is 23.74 and it is the lowest F value, the FM is the most significant statistical model for the development dataset.

Other Models

Other complex regression models, such as separate slope, mixture surface, and homogeneity of slope models were also considered in the present study (Table 4). Separate slope method is used to model the influence of the predictors while mixture surface designs are identical to factorial regression designs. In general, homogeneity of slope models is used to test the whether the predictors interact in influencing responses. Even though the R^2 of these models increased to as high as 0.9430, the F values also increased indicating that these complex models were not significant for the development dataset.

Evaluation of Models

Based on the R^2 and F values for the evaluation dataset (Table 4), the SBM performs the worst; the FM performs the best while MRM and PM are intermediate. Furthermore, since the second order PM is a special case of MRMs and it performs better than the MRM, the two statistical models considered for further evaluation are the second order PM and the full FM.

Furthermore, the evaluation dataset were separated into soils from Woodward County and Rogers County. Separate comparisons were made for the Woodward County soils (WOE-4B) and the Rogers County soils (ROE-20B), and a comparison was made for both counties together (henceforth called "combined evaluation dataset"). This provides different views on the prediction quality and the importance of datasets on statistical analysis [24, 29]. Additionally, a comparison was made between the differences in the R^2 values of the development dataset and the evaluation dataset.

Evaluation of Factorial Model

The R^2 value of the combined evaluation dataset was only 0.3634. Fig. 5 shows a comparison of the experimental and predicted M_R/P_a values for the combined evaluation dataset. Even though the overall R^2 value for the development dataset was 0.6595, it dropped

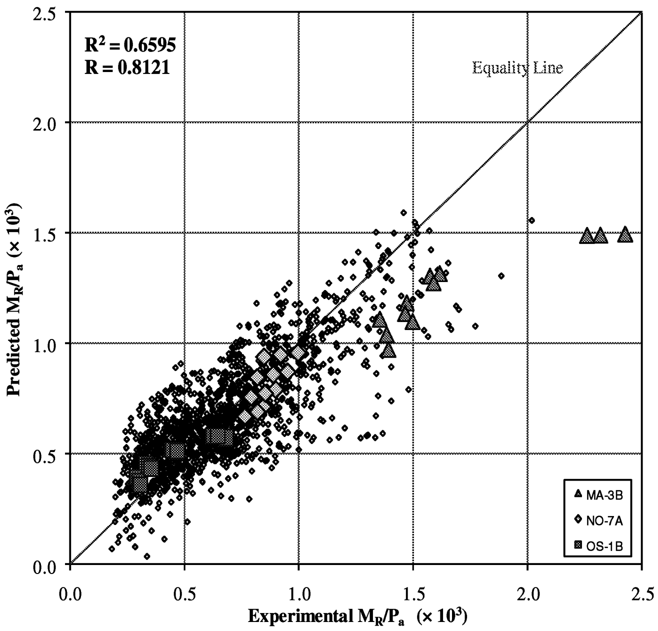

Fig. 5. Comparison of Experimental and Predicted M_R/P_a for Combined Evaluation Dataset: Factorial Model.

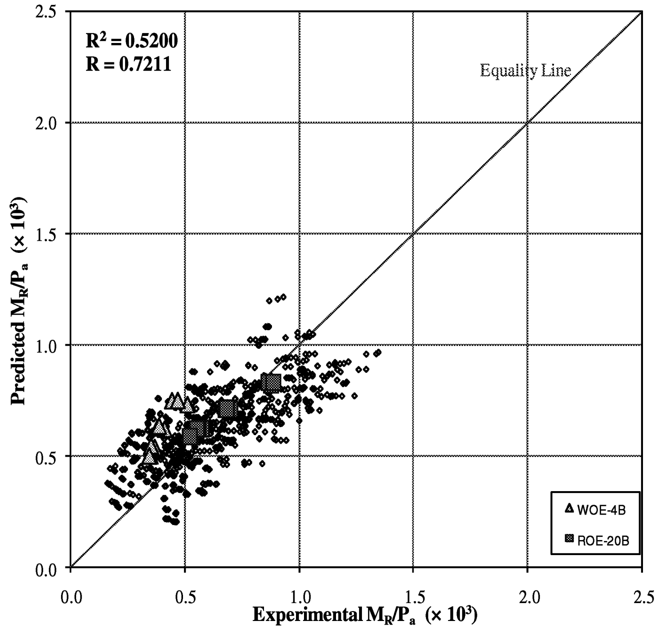

Fig. 6. Comparison of Experimental and Predicted M_R/P_a for Combined Evaluation Dataset: Polynomial Model.

Table 5. Summary of the Statistical Modeling Results Using Evaluation Dataset.

Statistical Model	Combined	Woodward County	Rogers County
	R^2	R^2	R^2
Stress-Based	0.3569	0.5776	0.3666
Multiple Regression	0.5403	0.8077	0.5370
Polynomial	0.5200	0.6212	0.5523
Factorial	0.3634	0.0962	0.4021

significantly to 0.3634 for the evaluation dataset (Table 5). The soils from Woodward County (WOE-4B) have the worst predictions among all the statistical models with a R^2 value of 0.0962. The full FM considered here contains 128 terms in the function and it may be considered as the most complex function among the four statistical models. Therefore, it is possible that the FM over-fitted the development dataset and caused a poor prediction in the evaluation dataset [24, 32]. In the case of present dataset, it appears that the full FM has created a condition known as too much wiggle [30, 31]. Too much wiggle occurs when the equation has too many terms and tries to fit to as many data point as possible. The percent difference in the R^2 between the development dataset and the Woodward County and Rogers County evaluation datasets are 85% and 39%, respectively.

Evaluation of Second Order Polynomial Model

The second order PM predicted the M_R/P_a values with an R^2 value of 0.5200. A plot of the experimental and predicted M_R/P_a values is given in Fig. 6. The results show that the Woodward County (WOE-4B) and the Rogers County soils (ROE-20B) have R^2 values of 0.6212 and 0.5523, respectively. The R^2 values for Woodward and Rogers Counties were approximately 27% and 6.2% higher than the R^2 value for the development dataset. This indicates that the second order PM is capable of predicting the M_R values of the Woodward and Rogers County soils reasonably well, as compared to other models.

Summary and Conclusions

Several statistical models were developed in this study. These models included: stress-based, multiple regression, polynomial, and factorial. Based on the R^2 and F values, the second order polynomial and factorial models were further considered for the evaluation dataset. An evaluation of the two models indicated that for the combined development and evaluation datasets a second order polynomial is a good statistical model for evaluating M_R from the selected routinely determined properties.

The following conclusions drawn from the present study were summarized below:

1. The stress-based model was found to have a low R^2 value (0.3226), with significant scatters in the back-prediction of the development dataset. The F value for this model was found to be 253.37, indicating that a more complex model may be needed in correlating M_R with the selected model parameters.
2. The R^2 and F values for the multiple linear regression model were found to be 0.4357 and 165.88, respectively, indicating a significant improvement over the stress-based model.
3. A second order polynomial multiple regression model was developed for the development dataset. The R^2 and F values for this model were found to be 0.4858 and 101.02, respectively. These values indicate a slight improvement over the multiple linear regression model. Although the effect of this improvement was not visually noticeable for the overall development dataset, the improvement was evident through the improvement in the R^2 values for the evaluation dataset.
4. One of the most complicated models considered in this study was a full factorial model. This model had 128 terms, and the

R^2 and F values for the development dataset was found to be 0.6595 and 23.74, respectively. Based on these values and not considering the evaluation dataset, this model appeared to be the best statistical model.

5. Based on the R^2 and F values second order polynomial and factorial model were selected for further evaluation using the evaluation dataset. Factorial model showed the worst predictions for soils from Woodward County with a R^2 value of 0.0962.
6. Second order polynomial showed the R^2 values for the evaluation dataset (Roger County (0.5523), Woodward County (0.6212), and combined dataset (0.5200)) were relatively high, indicating that this model is comparatively a good model for the combined datasets (development and evaluation).
7. The second order polynomial model developed in this study is expected to be useful in the Level 2 and Level 3 designs of pavements in Oklahoma.

References

1. American Association of State Highway and Transportation Officials (AASHTO), (1986). *Standard Specifications for Transportation Materials and Methods of Sampling and Testing*. Washington, D. C., USA.
2. Kim, D. and Stokoe II, K.H., (1992). Characterization of Resilient Modulus of Compacted Subgrade Soils Using Resonant Column and Torsional Shear Tests, *Transportation Research Record*, No. 1369, pp. 83-91.
3. George, K.P., (1992). Resilient Testing of Soils Using Gyratory Testing Machine, *Transportation Research Record*, No. 1369, pp. 63-72.
4. Kim, D., Kweon, G., and Lee, K., (1997). Alternative Method of Determining Resilient Modulus of Compacted Subgrade Soils Using Free-Free Resonant Column Test, *Transportation Research Record*, No. 1577, pp. 62-69.
5. American Association of State Highway and Transportation Officials (AASHTO), (2004). *Guide for Mechanistic-Empirical Design of new and rehabilitated pavement structures*, Final Report prepared for National Cooperative Highway Research Program (NCHRP), Transportation Research Board, National Research Council, Washington, D.C., USA.
6. Dunlap, W.S., (1963). A Report on a Mathematical Model Describing the Deformation Characteristics of Granular Materials, *Technical Report 1, Project 2-8-62-27, TTI*, Texas A&M University, Texas, USA.
7. Seed, H.B., Mitry, F.G., Monosmith, C.L., and Chan, C.K., (1967). Prediction of Pavement Deflection from Laboratory Repeated Load Tests, *NCHRP Report 35*, Transportation Research Board, Washington D.C., USA.
8. Moossazadeh, J.M. and Witczak, M.W., (1981). Prediction of Subgrade Moduli for Soil that Exhibits Nonlinear Behavior, *Transportation Research Record*, No. 810, pp. 9-17.
9. May, R.W. and Witczak, M.W., (1981). Effective Granular Modulus to Model Pavement Response, *Transportation Research Record*, No. 810, pp. 1-9.
10. Yau, A. and Von Quintus, H.L., (2002). Study of LTPP Laboratory Resilient Modulus Test Data and Response Characteristics, *Publication FHWA-RD-02-051*, FHWA, U.S. Department of Transportation.
11. National Cooperative Highway Research Program (NCHRP), (2003). Procedure for Resilient Modulus of Unstabilized Aggregate Base and Subgrade Materials, *Project 1-28A*, Transportation Research Board, Washington D.C., USA.
12. Hopkins, T.C., Beckham, T.L., Sun, L., and Pfalzer, B., (2004). Kentucky Geotechnical Database, *Publication KTC-03-06/SPR-177-98-1F*, University of Kentucky Transportation Center, College of Engineering, Lexington, Kentucky, USA.
13. Rahim, A.M. and George, J.P., (2004). Subgrade Soil Index Properties to Estimate Resilient Modulus. *Transportation Research Board 2004 Annual Meeting*, CD-ROM, Washington D.C., USA.
14. Khazanovich, L., Celauro, C., Chadbourn, B., Zollars, J., and Dai, S., (2006). Evaluation of Subgrade Resilient Modulus Predictive Model for Use in Mechanistic-Empirical Pavement Design Guide, *Transportation Research Record*, No. 1947, pp. 155-166.
15. Li, D.Q. and Selig, E.T., (1994). Resilient Modulus for Fine-grained Subgrade Soils, *Journal of Geotechnical Engineering, ASCE*, 120(6), pp. 939-957.
16. Gomes, A.C. and Gillett, S., (1996). Flexible Pavement: Resilient Behavior of Soils, Balkema, Rotterdam, ISBN 90 54 10 5232.
17. Paute, J.L. and Hornych, P., (1996). Flexible Pavement: Influence of Water Content on the Cyclic Behavior of a Silty Sand, Balkema, Rotterdam, ISBN 90 54 10 5232.
18. Rada, C. and Witczak, W.M., (1981). Comprehensive Evaluation of Laboratory Resilient Moduli Results for Granular Material, *Transportation Research Record*, No. 810, pp. 23-33.
19. Raad, L., Minassian, G.H., and Gartin, S., (1992). Characterization of Saturated Granular Bases under Repeated Loads, *Transportation Research Record*, No. 1369, pp. 73-82.
20. Zaman, M.M., Chen, D.H., and Laguros, J.G., (1994). Resilient Moduli of Granular Materials, *Journal of Transportation Engineering, ASCE*, 120(6), pp. 967-988.
21. Drumm, E.C., Boateng-Poku, Y., and Pierce, T.J., (1990). Estimation of subgrade resilient modulus from standard tests. *Journal of Geotechnical Engineering,* 116(5), pp. 774-789.
22. Lee, W., Bohra, N.C., Altschaeffl, A.G., and White, T.D., (1997). Resilient Modulus of Cohesive Soils, *Journal of Geotechnical and Geoenvironmental Engineering, ASCE*, 123(2), pp. 131-136.
23. Ebrahimi, A., (2006). Regression and Neural Network Modeling of Resilient Modulus Based on Routine Soil Properties and Stress States. Ph.D. Dissertation, University of Oklahoma, Norman, Oklahoma, USA.
24. Montgomery, D.C., Peck, E.A., and Vining, G.G., (2006). Introduction to Linear Regression Analyses (Wiley Series in Probability and Statistics, John Wiley and Sons, Inc., New Jersey, USA.
25. Federal Highway Administration (FHWA), (2002). Study of LTPP Laboratory Resilient Modulus Test Data and Response

Characteristics, *FHWA-RD-02-051*, Office of Engineering Research and Development, McLean, Virginia, USA.

26. Tarefder, R.A., White, L., and Zaman, M., (2005). Neural Network Model for Asphalt Concrete Permeability. *Journal of Materials in Civil Engineering, ASCE*, 17(1), pp. 19-27.
27. StatSoft, Inc. (2006). Electronic Statistics Textbook, Tulsa, Oklahoma, USA.
28. Carmichael, R.F. III. and Stuart, E., (1978). Predicting resilient modulus: A study to determine the mechanical properties of subgrade soils, *Transportation Research Record*, No. 1043, pp. 20-28.
29. Myers, R.H., Montgomery, D.C., and Vining, G.G., (2001). Generalized Linear Models: With Applications in Engineering and the Sciences, John Wiley and Sons, Inc., New Jersey, USA.
30. Sokal, R.R. and Rohlf, F.J., (1995). Biometry: The Principles and Practice of Statistics in Biological Research, W.H. Freeman and Co., New York, USA.
31. Fernandez-Juricic, E., Sallent, A., Sanz, R., and Rodriguez-Prieto, I, (2003). Testing the Risk-disturbance Hypothesis in a Fragmented Landscape: Non-linear Responses of House Sparrows to Humans, *The Condor*, Vol. 105, pp. 316-326.
32. Hill, T. and Lewicki, P., (2006). STATISTICS Methods and Applications, StatSoft, Tulsa, Oklahoma, USA.

Appendix: Equation of Factorial Model

$$
\begin{aligned}
M_R/P_a = {} & 13.2514795 - 438.31923*w - 13.311426*\gamma_d + 2.41669221*PI + 27.2918109*P_{200} - 35.722370*U_c - 11.240229*\sigma_d + 85.5626222*\theta + \\
& 278.441637*w*\gamma_d - 16.168513*w*PI - 1.0992907*\gamma_d*PI + 14.4631546*w*P_{200} - 9.7399663*\gamma_d*P_{200} - 1.7766282*PI*P_{200} + \\
& 332.908859*w*U_c + 23.0157253*\gamma_d*U_c - .01410921*PI*U_c + 38.7543639*P_{200}*U_c + 799.099567*w*\sigma_d + 31.3535709*\gamma_d*\sigma_d - \\
& 3.3563515*PI*\sigma_d - 69.037455*P_{200}*\sigma_d - 47.303113*U_c*\sigma_d - 279.93578*w*\theta - 38.643251*\gamma_d*\theta - 1.0518368*PI*\theta - \\
& 84.485170*P_{200}*\theta - 29.795776*U_c*\theta - 106.57364*\sigma_d*\theta + 9.15134484*w*\gamma_d*PI - 28.679440*w*\gamma_d*P_{200} + 21.1200089*w*PI*P_{200} \\
& + .662040611*\gamma_d*PI*P_{200} - 200.60637*w*\gamma_d*U_c + 2.75148494*w*PI*U_c - .02351306*\gamma_d*PI*U_c - 244.11931*w*P_{200}*U_c - \\
& 25.359274*\gamma_d*P_{200}*U_c - 1.2799791*PI*P_{200}*U_c - 588.87015*w*\gamma_d*\sigma_d + 15.1565883*w*PI*\sigma_d + .481801078*\gamma_d*PI*\sigma_d - \\
& 218.18881*w*P_{200}*\sigma_d + 10.1638379*\gamma_d*P_{200}*\sigma_d + 2.36460454*PI*P_{200}*\sigma_d - 45.492450*w*U_c*\sigma_d + 12.7119527*\gamma_d*U_c*\sigma_d + \\
& 3.45466718*PI*U_c*\sigma_d + 73.1347282*P_{200}*U_c*\sigma_d + 115.117756*w*\gamma_d*\theta + 7.15819491*w*PI*\theta - .16922596*\gamma_d*PI*\theta + \\
& 319.715817*w*P_{200}*\theta + 37.2140312*\gamma_d*P_{200}*\theta + .456277589*PI*P_{200}*\theta + 87.1036047*w*U_c*\theta + 12.8403258*\gamma_d*U_c*\theta \\
& + .470678587*PI*U_c*\theta + 17.7135930*P_{200}*U_c*\theta + 117.516534*w*\sigma_d*\theta + 45.1075781*\gamma_d*\sigma_d*\theta + 4.48108267*PI*\sigma_d*\theta + \\
& 123.422308*P_{200}*\sigma_d*\theta + 69.9906302*U_c*\sigma_d*\theta - 11.825725*w*\gamma_d*PI*P_{200} - 2.1155135*w*\gamma_d*PI*U_c + 149.381038*w*\gamma_d*P_{200}*U_c \\
& + .461733179*w*PI*P_{200}*U_c + .843999848*\gamma_d*PI*P_{200}*U_c - 3.1294192*w*\gamma_d*PI*\sigma_d + 277.350974*w*\gamma_d*P_{200}*\sigma_d - \\
& 14.172917*w*PI*P_{200}*\sigma_d + .525633666*\gamma_d*PI*P_{200}*\sigma_d + 98.7215822*w*\gamma_d*U_c*\sigma_d - 15.964519*w*PI*U_c*\sigma_d - \\
& 1.1448209*\gamma_d*PI*U_c*\sigma_d - 169.47515*w*P_{200}*U_c*\sigma_d - 22.778234*\gamma_d*P_{200}*U_c*\sigma_d - 2.0574402*PI*P_{200}*U_c*\sigma_d - \\
& 1.1780048*w*\gamma_d*PI*\theta - 136.19640*w*\gamma_d*P_{200}*\theta - 5.5904934*w*PI*P_{200}*\theta + .582825521*\gamma_d*PI*P_{200}*\theta - 32.417984*w*\gamma_d*U_c*\theta - \\
& 4.4154766*w*PI*U_c*\theta - .03434428*\gamma_d*PI*U_c*\theta - 35.024894*w*P_{200}*U_c*\theta - 5.1994910*\gamma_d*P_{200}*U_c*\theta + .197336556*PI*P_{200}*U_c*\theta + \\
& 12.9337904*w*\gamma_d*\sigma_d*\theta - 14.637167*w*PI*\sigma_d*\theta - 1.4228429*\gamma_d*PI*\sigma_d*\theta - 175.26170*w*P_{200}*\sigma_d*\theta - 51.332213*\gamma_d*P_{200}*\sigma_d*\theta - \\
& 4.4923407*PI*P_{200}*\sigma_d*\theta - 150.22381*w*U_c*\sigma_d*\theta - 29.716170*\gamma_d*U_c*\sigma_d*\theta - 2.2982062*PI*U_c*\sigma_d*\theta - 70.820468*P_{200}*U_c*\sigma_d*\theta + \\
& 0.00000000*w*\gamma_d*PI*P_{200}*U_c + 0.00000000*w*\gamma_d*PI*P_{200}*\sigma_d + 5.73119283*w*\gamma_d*PI*U_c*\sigma_d + 0.00000000*w*\gamma_d*P_{200}*U_c*\sigma_d + \\
& 9.80008993*w*PI*P_{200}*U_c*\sigma_d - .05009087*\gamma_d*PI*P_{200}*U_c*\sigma_d + 0.00000000*w*\gamma_d*PI*P_{200}*\theta + 1.92867824*w*\gamma_d*PI*U_c*\theta + \\
& 0.00000000*w*\gamma_d*P_{200}*U_c*\theta + 1.30268945*w*PI*P_{200}*U_c*\theta - .38745801*\gamma_d*PI*P_{200}*U_c*\theta + 3.36590974*w*\gamma_d*PI*\sigma_d*\theta + \\
& 0.00000000*w*\gamma_d*P_{200}*\sigma_d*\theta + 11.4488598*w*PI*P_{200}*\sigma_d*\theta + 1.17782173*\gamma_d*PI*P_{200}*\sigma_d*\theta + 31.7684321*w*\gamma_d*U_c*\sigma_d*\theta + \\
& 5.85679607*w*PI*U_c*\sigma_d*\theta + .607338016*\gamma_d*PI*U_c*\sigma_d*BS + 128.506165*w*P_{200}*U_c*\sigma_d*\theta + 26.8574361*\gamma_d*P_{200}*U_c*\sigma_d*\theta + \\
& 1.63623868*PI*P_{200}*U_c*\sigma_d*\theta + 0.00000000*w*\gamma_d*PI*P_{200}*U_c*\sigma_d + 0.00000000*w*\gamma_d*PI*P_{200}*U_c*\theta + \\
& 0.00000000*w*\gamma_d*PI*P_{200}*\sigma_d*\theta + 0.00000000*w*\gamma_d*PI*U_c*\sigma_d*\theta + 0.00000000*w*\gamma_d*P_{200}*U_c*\sigma_d*\theta + \\
& 0.00000000*w*PI*P_{200}*U_c*\sigma_d*\theta + 0.00000000*\gamma_d*PI*P_{200}*U_c*\sigma_d*\theta - 4.6975464*w*\gamma_d*PI*P_{200}*U_c*\sigma_d*\theta
\end{aligned}
$$