

RESUMO

O Manual da Shell para dimensionamento de pavimentos, Shell Pavement Design Manual, foi publicado em 1978. No entanto, o método e os seus procedimentos ainda são utilizados atualmente por muitas administrações rodoviárias. Entre os ábacos incluídos no Manual está o designado Chart RT, destinado a determinar a temperatura equivalente do pavimento para dimensionamento (WMAPT). A temperatura é um fator chave no dimensionamento dos pavimentos betuminosos porque influencia a viscosidade do betume, que tem comportamento termoplástico. Assim, a temperatura equivalente do pavimento tem um impacto substancial nos resultados do dimensionamento. Contudo, não existem expressões matemáticas disseminadas que possam ser utilizadas para substituir analiticamente a utilização do Chart RT. Neste trabalho foi utilizada a Regressão Polinomial Evolutiva (RPE) para gerar modelos matemáticos que possam simular os resultados obtidos do Chart RT. A RPE é uma técnica de modelação híbrida que combina a regressão linear e a procura evolutiva de estruturas de modelos matemáticos. O software, MO-EPR, gera um conjunto de expressões matemáticas que são as melhores no ajuste da informação de treino com diferentes níveis de parcimónia (ou seja, número de coeficientes e grupos adimensionais). Deste modo, um conjunto de dados de treino e dados de verificação foi obtido do Chart RT. Esses dados foram utilizados para gerar um conjunto de modelos utilizando diversas estruturas e funções. Foi avaliado o ajuste dos modelos, realizada uma análise de resíduos e avaliada a sua homocedasticidade. A partir desta análise é proposto um modelo matemático, com níveis adequados de ajuste e parcimónia, para ser utilizado na determinação analítica da temperatura equivalente do pavimento.

Palavras-chave: Regressão Polinomial Evolutiva, Dimensionamento de Pavimentos, Temperatura Equivalente WMAPT.

1 INTRODUCTION

This paper focuses on using Evolutionary Polynomial Regression (EPR), a modelling methodology proposed by Giustolisi and Savic (2006, 2009), to generate an analytical expression to substitute the use of Chart RT. EPR is a hybrid modelling technique that combines linear regression and evolutionary search for mathematical model structures. Thus, EPR works as a two-stage model that (1) searches model structures based on a Genetic Algorithm and (2) estimates their parameters based on linear optimisation (Doglioni et al., 2010). At the end of the run, EPR returns a set of “optimal” model expressions (i.e., the Pareto front of models), which are the best in fitting training data at different levels of parsimony, i.e. number of coefficients and dimensionless groups (Laucelli and Giustolisi, 2011). Therefore, this modelling technique has the advantage of being able to avoid excessive model complexity. Expressions are automatically ranked according to both their goodness of fit and complexity. The use of parsimonious models contributes to a straightforward use of the models

and can give a better physical understanding of the models (Laucelli and Giustolisi, 2011). EPR has been previously used to obtain mathematical expressions for more complex problems such as the Colebrook-White friction factor (Giustolisi et al., 2011) or rubber concrete behaviour (Ahangar-Asr et al. 2011). This work used the implementation of EPR available in the software EPR MOGA-XL (Giustolisi and Savic, 2006); however, some alternatives and improvements of this methodology have been proposed (Li et al., 2023).

Bitumen has a thermoplastic behaviour. Hence, temperature is essential in asphalt pavement performance since it influences the viscosity of the bitumen. Bitumen becomes softer, less viscous at higher temperatures, and stiffer at lower temperatures. Distresses such as permanent deformation on the bituminous mixtures or low-temperature cracking can be associated with high or low temperatures (Gardete et al., 2018; Zaumanis et al., 2018). Because of the change in the viscosity of the bitumen, bituminous mixtures are stiffer, with higher elastic modulus at low temperatures, and more deformable, with lower elastic modulus at higher temperatures. For the same traffic and construction conditions a pavement in a hot zone will need more thickness because it will have higher deformation when subjected to traffic loading (Kumlai et al. 2017, Hasan and Tarefder, 2018). Hence, the in-service temperature of the pavement is a critical factor in flexible pavement design (Sreejith, 2021).

The Shell Pavement Design Manual was published in 1978 (Shell, 1978). The manual includes charts to be used to design flexible pavements and overlays. Whilst the foundation of this design method has some time, the technique and some of its procedures are still currently used (INIR, 2009; Austroads 2017). Among the charts found in the Manual, there are some intended to determine the effective pavement temperature for the design. A chart that is used in the determination of this temperature is Chart RT. This chart allows us to determine the weighted mean annual pavement temperature (WMAPT) as a function of the weighted mean annual air temperature (WMAAT) and the asphalt thickness (Figure 1). Although this chart is still used, an analytical expression is not commonly referred to in literature. Some authors and guides usually refer to an expression for 100 mm of asphalt thickness, Equation 1 (Sreejith, 2021; Austroads 2017).

$$WMAPT = -12.4 + \frac{6.32 \times WMAAT}{\ln(WMAAT)} \quad (1)$$

where, WMAPT = weighted mean annual pavement temperature (°C);

WMAAT = weighted mean annual air temperature (°C).

This work proposes an analytical expression to substitute Chart RT generated using EPR. The aim is to generate a single expression that can be used for the complete range of temperature and asphalt thickness

present in Chart RT. EPR was selected because it creates the structure of the model and number of terms, which are not previously defined by the user. However, the exponents used in the models are chosen from a user-defined set. Also, different models can be generated, considering the fit and parsimony of the models.

2 EVOLUTIONARY POLYNOMIAL REGRESSION

EPR aims to achieve polynomial expressions that could represent a system by searching for the exponents of a polynomial function with a fixed maximum number of terms. During one execution, it returns expressions with increasing numbers of terms up to a limit set by the user to allow the desired number of terms to be selected. The general form of expression used in EPR can be presented as in Equation 2 (Rezania et al., 2008, Giustolisi and Savic, 2006).

$$y = \sum_{j=1}^m F(X, f(X), a_j) + a_0 \quad (2)$$

where, y = estimated value.
 F = expression generated in the process
 X = matrix of input variables
 f = function selected by the user
 m = number of terms of the expression
 a_j = adjustable parameters
 a_0 = optional bias

The first step in the identification of the model structure is to transfer Equation 2 into the vector form, Equation 3 (Rezania et al., 2008, Giustolisi and Savic, 2006):

$$Y_{N \times 1}(\theta, Z) = [I_{N \times 1} Z_{N \times m}^j] \times [a_0 \ a_1 \ \dots \ a_m]^T = Z_{N \times d} \times \theta_{d \times 1}^T \quad (3)$$

where, $Y_{N \times 1}(\theta, Z)$ = is the least-squares (LS) estimate vector of the N target values.
 $\theta_{d \times 1}$ = vector of $d = m + 1$ adjustable parameters a_j ($j = 1:m$) and a_0
 $Z_{N \times d}$ = matrix formed by 1 (unitary vector) for bias a_0 , and m vectors of variables Z_i .
 Z^j = For a fixed j is a product of the independent predictor vectors of inputs.

In general, EPR is a two-stage technique for constructing symbolic models. In the first stage, a standard genetic algorithm (GA) searches for the best form of the function structure, i.e., a combination of vectors of independent inputs. The second stage performs a least square (LS)

regression to find the adjustable parameters, θ , for each combination of inputs. This way, a global search algorithm is implemented for both the best set of input combinations and related exponents simultaneously, according to the user-defined cost function (Rezania et al., 2008, Giustolisi and Savic, 2006).

The matrix of inputs is shown in Equation 4.

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1k} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2k} \\ x_{31} & x_{23} & x_{33} & \cdots & x_{3k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{N1} & x_{N2} & x_{N3} & \cdots & x_{Nk} \end{bmatrix} = [X_1 \quad X_2 \quad X_3 \quad \cdots \quad X_k] \quad (4)$$

where, X = matrix of inputs.
 x_{ij} = candidate variables for matrices Z^j .

Therefore, Z^j in Equation 3 can be write as in Equation 5.

$$Z_{N \times 1}^j = [(X_1)^{ES(j,1)} \cdot (X_2)^{ES(j,2)} \cdot (X_3)^{ES(j,3)} \cdots (X_k)^{ES(j,k)}] \quad \forall j = 1 \cdots m \quad (5)$$

where, $Z_{N \times 1}^j$ = is the j th vector whose elements are products of candidate independent inputs.
 ES = is a matrix of exponents.
 X_i = vectors of inputs.

The aim is to find the matrix $ES_{k \times m}$ of exponents whose elements can assume values within a user-defined set (Rezania et al., 2008, Giustolisi and Savic, 2006). Thus, each exponent in ES corresponds to a value from the user-defined vector EX . Also, each row of ES determines the exponents of the candidate variables of the j th term in Equations 2 and 3.

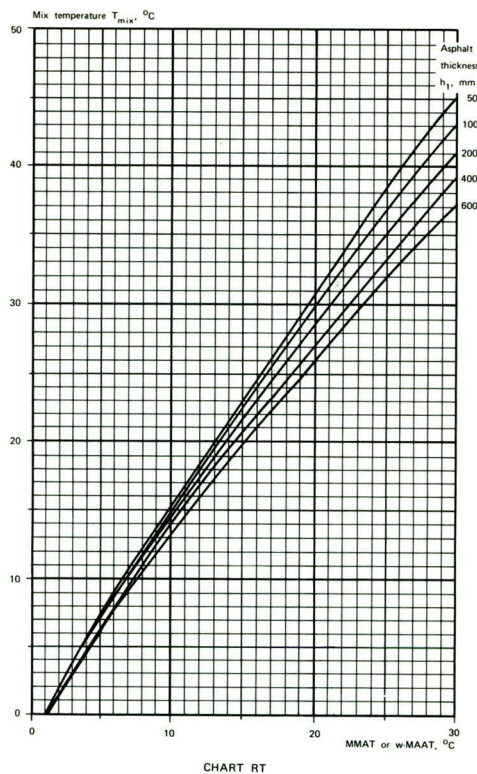
Constants a_i are evaluated using the linear LS method based on the minimisation of the sum of squared errors (SSE) as the cost function, and of the exponents in ES corresponds to a value from the user-defined vector EX . This allows the transformation of the symbolic regression problem into one of finding the best ES , i.e. the best structure of the EPR equation. The global search for the best form of the structure of the equation is performed using a standard GA over the values in the user-defined vector of exponents, EX (Rezania et al., 2008, Giustolisi and Savic, 2006). One of the major advantages is that the user does not have to specify the form of the regression model in advance.

3 MODEL GENERATION

The objective was to achieve a simple mathematical expression with a good fit across the range of values defined in Chart RT. The lines in Chart RT define this range, i.e. weighted mean annual air temperature from 2°C to 30°C and asphalt thickness from 50 mm to 600 mm (Figure 1).

To generate the model, 145 points were determined from the chart and used as training data. The values were taken for each temperature value and for each thickness line. The values were obtained by visual reading. A scan of the chart was made and used to collect the data. Care was made to try and minimize reading errors. Nevertheless, at some degree WMAPT values were affected by reading errors. It should also be mentioned that the training data do not constitute a random data set, but it was used as it gave full coverage of Chart RT.

Figure 1 – Chart RT



Source: SHELL (1978)

It used a set of candidate exponents that could encompass the expected relations between variables; 13 values were used for exponents: 0, -0.25, 0.25, 0.5, -0.5, 1, -1, 1.5, -1.5, 2, -2, 3, -3. It was observed that using opposite sign exponents is unnecessary as EPR can create expressions that could surpass this issue. EPR MOGA-XL has options to include several mathematical functions in the models. These are called inner functions and can be logarithmic, exponential, tangent hyperbolic or secant hyperbolic. One attractive advantage of EPR is that the software

generates the terms, and no previous expression format is needed. The software generates several models according to the user options and gives information on the fit of the models. A specified number of models can be generated with an increasing number of terms; more terms correspond to more complexity but also better fit.

The generated models were assessed for their parsimony and fit. In a first approach for each of the inner function options, the simplest model that included both dependent variables, asphalt thickness and WMAAT, was selected. These models are presented in Table 1. The CoD (Coefficient of Determination) is offered for the training data. CoD represents the proportion of variance explained by the model and measures how well the estimated values match the original values.

Table 1 - Generated models for WMAPT

Model	Expression	CoD
1	$1.8959 \times WMAAT - 0.022212 \times t^{0.25} \times WMAAT^{1.5} - 1.8503$	99.96
2	$1.5142 \times WMAAT - 0.93947 \times \ln(t^{0.25} \times WMAAT^{1.5}) + 3.1382$	98.05
3	$\exp\left(0.042289 \times \frac{WMAAT}{t^{0.25}} - 5.8661 \times \frac{1}{WMAAT^{0.25}} + 5.9148\right)$	99.81
4	$\tan\left(0.0012548 \times \frac{WMAAT}{t^{0.25}} - 1.4194 \times \frac{1}{WMAAT^{1.5}} + 1.5448\right)$	99.59
5	$8.7125 \times WMAAT \times \frac{1}{\ln(t+1)^{0.25} \times \ln(WMAAT+1)} - 9.4722$	99.82
6	$1.5584 \times WMAAT - 1.8184 \times 10^{-5} \times t \times WMAAT^2 - 0.82827$	99.79
7	$1.0601 \times WMAAT \times \exp\left(\frac{1}{t^{0.25}}\right) + 0.091807$	99.74
8	$1.8069 \times WMAAT \times \tanh(WMAAT)^3 - 0.10773 \times t^{0.25} \times WMAAT \times \tanh(t)^{-0.5} \times \tanh(WMAAT)^{-3} + 0.1678$	99.84
9	$1.3997 \times WMAAT + 21635.6341 \times \tanh(t^{0.25} \times WMAAT) - 21635.5145$	97.95
10	$1.3899 \times WMAAT \times \tanh\left(\frac{WMAAT^{1.5}}{t^{0.25}}\right) + 0.32971$	98.02
11	$0.050794 \times t^{0.25} \times WMAAT^{1.5} * \operatorname{sech}(t^{-0.5}) + 8.8794$	74.10
Where: WMAAT - Weighted Mean Annual Air Temperature (°C) t - Asphalt Thickness (mm) CoD - Coefficient of Determination		

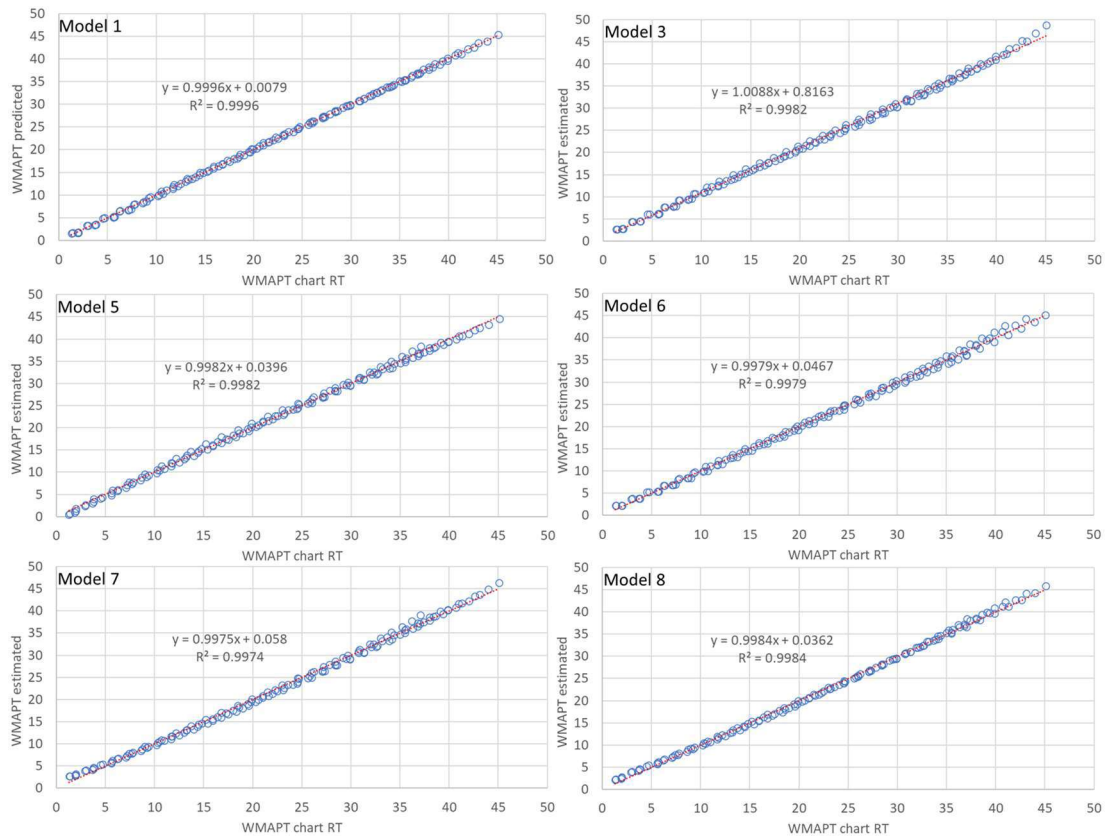
4 GENERATED MODELS ANALYSIS

4.1 Fit analysis

Figure 2 presents the fit of 6 models. The models were selected because of their high COD. Some models achieve a CoD higher than 0.9950,

indicating a good fit to the training data. The highest CoD was 0.9996 for model 1. As a comparison, the model given by Equation 1 achieves a CoD of 0.9977 for the corresponding training data (data for an asphalt thickness of 100 mm).

Figure 2 – Fit for generated Models (1), (3), (5), (6), (7) and (8)



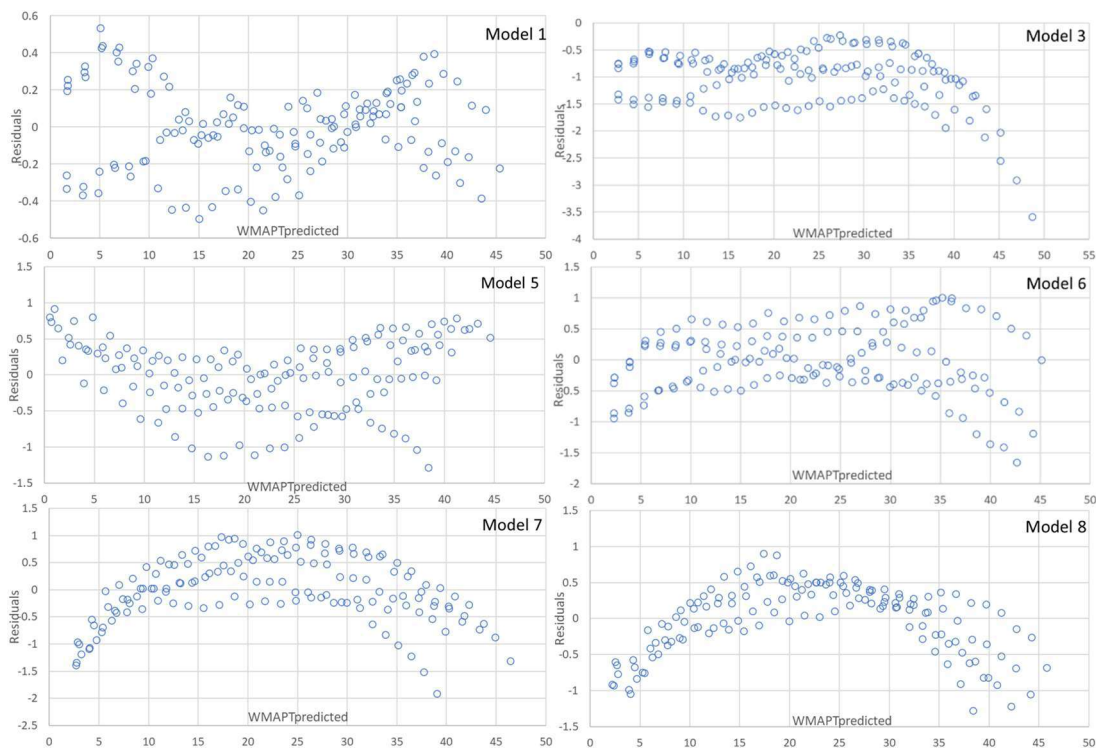
4.2 Residuals Analysis

The gap between the original value of the dependent variable and the estimated value is known as the residual. Residuals should be statistically independently distributed (independence), normally distributed with zero mean (normality), have the same variance (homoscedasticity), and there should be no correlation between independent variables (Multicollinearity) (Suleiman et al., 2015). Thus, residuals should be randomly distributed not showing any pattern; residuals should be small and spread around zero constantly throughout the range (Elsayir, 2019). A graphic analysis can be helpful to find any type of pattern in the residuals. The residuals for the models with the highest CoDs are presented in Figure 3.

Normality was examined graphically using normal probability plots (Q-Q plots). All model residuals followed a normal distribution with some minor deviations at the tails. Residuals are smaller for Model 1, with values mostly lower than $\pm 0.4^\circ\text{C}$ throughout the range. In some models' higher residuals were observed in the lower and higher ranges of temperatures (WMAPT below 10°C and above 35°C). The graphic analysis shows patterns in

residual plots in some models. It can be seen in Figure 3 that residuals from Model 7 and Model 8 show an inverted U shape. Model 3 residuals do not distribute around zero. Models 1, 5 and 6 seem to have a more random distribution. Breusch-Pagan and White's tests were performed to evaluate the homoscedasticity of the residuals (Onifade and Olanrewaju, 2020; Uyanto, 2022). White test was more restrictive and only for Model 5 did the White test indicate homoscedasticity. Breusch-Pagan test indicated homoscedasticity for Model 5 and Model 7.

Figure 3 – Residuals for generated models (1), (3), (5), (6), (7) and (8)



4.3 Testing data

A set of 50 random points was used to analyse the fit of the most promising models (Models 1, 5, 6, 7). This data set was generated as a random data set with WMAAT from 2°C to 30°C and asphalt thickness from 50 mm to 600 mm. This testing data set intended to analyse the fit of the models using intermediate values for the asphalt thickness. This required more care collecting the data, but it also tested the robustness of the models.

Results are presented in Figure 4 and Figure 5. Model 1 achieved higher CoD and lower residuals. Model 5 gave good results, and for common pavement design temperatures (20°C-35°C) the residuals were below $\pm 0.5^\circ\text{C}$. However, outside this temperature interval, residuals are higher. Results for models 6 and 7 show higher residuals.

5 DATA DISCUSSION AND MODEL ANALYSIS

Using data collected from Chart RT (145 points), 11 models were generated using several structures and mathematical functions in EPR.

Most models did exhibit a good fit and low residuals. However, Model 1 presented the best fit overall. Whereas the White test did indicate heteroscedasticity, in a graphic analysis, residuals do not show a clear trend and have a relatively good dispersion throughout the range. The parsimony of the model is acceptable with an expression that does not pose calculation complexity. The testing data confirmed these results. Model 1 also has the advantage of not including less common mathematical functions or functions that could introduce limitations in the domain, leading to computation errors. Therefore, Model 1 is recommended as an expression to analytical substitute Chart RT, Equation 6.

Figure 4 – Fit for testing data for generated models (1), (5), (6), (7)

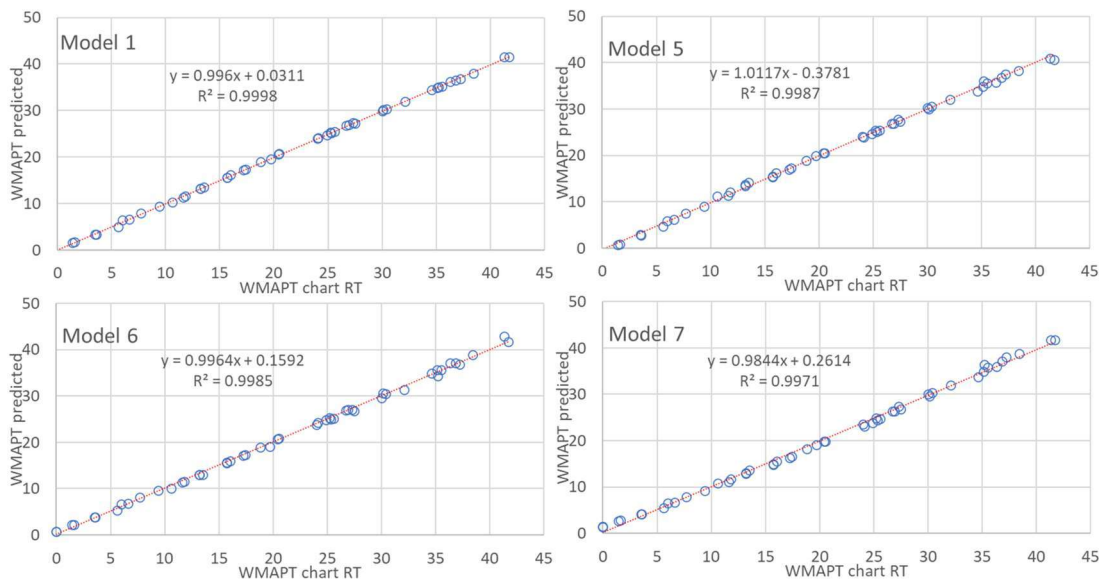
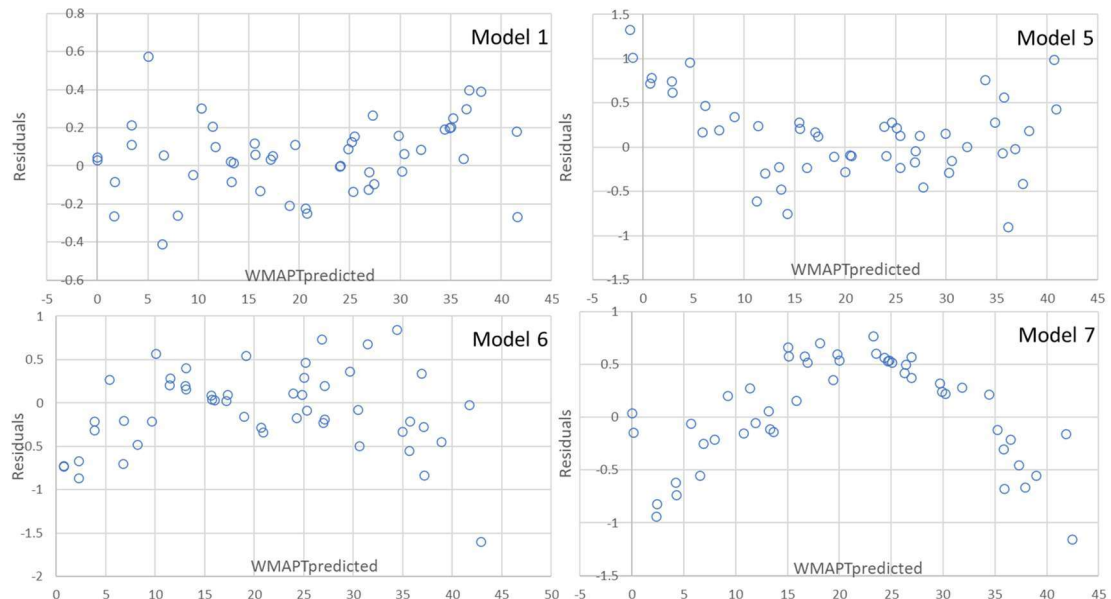


Figure 5 – Residuals for testing data; generated models (1), (5), (6), (7)



$$WMAPT = 1.8959 \times WMAAT - 0.022212 \times t^{0.25} \times WMAAT^{1.5} - 1.8503 \quad (6)$$

where, WMAPT = weighted mean annual pavement temperature (°C);
WMAAT = weighted mean annual air temperature (°C);
t = Asphalt thickness (mm).

As previously mentioned, the software MOGA-XL can create a set of models with increased complexity and better fit. The model generated for the same conditions as Model 1, Equation 6, but adding one step more complexity is represented in Equation 7. This is an example of how EPR generates models increasing the number of terms to achieve better fit. The increase in complexity makes the expression less appealing to use and does not significantly improve the fit (CoD = 0.9996), but it slightly lowers the residuals and can be used if this extra accuracy is needed (for the training data, the average of the absolute values of the residuals decreases from 0.1798 to 0.1299).

$$WMAPT = 1.8468 \times WMAAT - 0.020269 \times t^{0.25} \times WMAAT^{1.5} - 0.0019278 \times t \times WMAAT^{-0.25} - 1.3093 \quad (7)$$

where, WMAPT = weighted mean annual pavement temperature (°C);
WMAAT = weighted mean annual air temperature (°C);
t = thickness (mm).

6 CONCLUSIONS

This work attained a single analytical expression to replace the use of Chart RT. The generated model can be used to analytically substitute the use of chart RT with a satisfactory level of accuracy in a large range of temperatures (WMAAT from 2°C to 30°C) and asphalt thickness (5 mm to 600 mm). EPR generated a set of models from which 11 were selected for further analysis. The ability to create models with increased complexity provides an interesting feature for this task. The proposed model is easy to use and does not introduce complex mathematical functions. The selected model provides a good fit (CoD = 0.9996 for the training data), and the residuals are low and dispersed throughout the range. However, the Breusch-Pagan and White tests did indicate heteroscedasticity for this model. An analysis of the proposed models' residuals shows that in most cases, the expected errors could be acceptable and not higher than what could be expected when reading values in the chart (for the 50 points of the testing data in 33, the error was lower than ±0.2°C). Furthermore, the accuracy of the proposed model is higher on the most common range of temperatures (WMAAT above 10°C and below 35 °C).

ACKNOWLEDGEMENTS

The authors would like to acknowledge all at IDEA-RT for providing the software EPR MOGA-XL and for their inestimable attention and help.

REFERENCES

AHANGAR-ASR, A., FARAMARZI, A., JAVADI, A., GIUSTOLISI, O. Modelling mechanical behaviour of rubber concrete using evolutionary polynomial regression, **Engineering Computations**, Vol. 28 N°4, pp. 492-507. 2011. <https://doi.org/10.1108/02644401111131902>

AUSTROADS. Guide to Pavement Technology Part 2: Pavement Structural Design. Austroads. 2017. ISBN 978-1-925854-69-5

DOGLIONI, A., MANCARELLA, D., SIMEONE, V., GIUSTOLISI, O. Inferring groundwater system dynamics from hydrological time-series data. **Hydrological Sciences Journal**, 55: 4, 593 — 608. 2010. doi: 10.1080/02626661003747556

ELSAIYR, H.A. Residual Analysis for Auto-Correlated Econometric Model. **Open Journal of Statistics**, 9, 48-61. 2019. <https://doi.org/10.4236/ojs.2019.91005>

GARDETE, D., PICADO-SANTOS, L., CAPITÃO, S. Improving bituminous mixture performance by optimizing the design compaction energy – A cost effective approach for better pavements. **Construction and Building Materials**, Volume 190 1173-1181. 2018. <https://doi.org/10.1016/j.conbuildmat.2018.09.169>

GIUSTOLISI, O., SAVIC, D. A symbolic data-driven technique based on evolutionary polynomial regression. **Journal of Hydroinformatics**. 8 (3): 207–222. 2006. <https://doi.org/10.2166/hydro.2006.020b>

GIUSTOLISI, O., SAVIC, D. Advances in data-driven analyses and modelling using EPR-MOGA. **Journal of Hydroinformatics**. 11 (3-4): 225–236. 2009. <https://doi.org/10.2166/hydro.2009.017>

GIUSTOLISI, O., BERARDI, L., WALSKI, T. M. Some explicit formulations of Colebrook–White friction factor considering accuracy vs. computational speed. **Journal of Hydroinformatics**, Vol. 13 401-418. 2011. <https://doi.org/10.2166/hydro.2010.098>

HASAN, A., TAREFDER, R. Development of temperature zone map for mechanistic empirical (ME) pavement design. **International Journal of Pavement Research and Technology**, Volume 11, Issue 1 pp. 99-111. 2018. <https://doi.org/10.1016/j.ijprt.2017.09.012>

INIR I.P. Directivas para a concepção de pavimentos - Critérios de dimensionamento, Instituto de Infra-Estruturas Rodoviárias, Lisboa, 2009.

KUMLAI, S., JITSANGIAM, P., PICHAYAPAN, P. The Implications of Increasing Temperature due to Climate Change for Asphalt Concrete Performance and Pavement Design. **KSCE Journal of Civil Engineering**, 21(4):1222-1234. 2017. DOI 10.1007/s12205-016-1080-6

LAUCELLI, D., GIUSTOLISI, O. Scour depth modelling by a multi-objective evolutionary paradigm. **Environmental Modelling & Software**, Volume 26, Issue 4 pp. 498-509. 2011. <https://doi.org/10.1016/j.envsoft.2010.10.013>.

ONIFADE, O., OLANREWAJU, S. Investigating Performances of Some Statistical Tests for Heteroscedasticity Assumption in Generalized Linear Model: A Monte

Carlo Simulations Study. **Open Journal of Statistics**, 10, 453-493. 2020. <https://doi.org/10.4236/ojs.2020.103029>

LI, Y., LI, M., ZHANG, L. Evolutionary polynomial regression improved by regularization methods. **PLoS ONE**, 18(2): e0282029. 2023. <https://doi.org/10.1371/journal.pone.0282029>

REZANIA, M., JAVADI, A., GIUSTOLISI, O. An evolutionary-based data mining technique for assessment of civil engineering systems. **Engineering Computations: International Journal for Computer-Aided Engineering and Software**, Vol. 25 No. 6 pp. 500-517. 2008. DOI: 10.1108/026444400810891526

SHELL. **Shell Pavement Design Manual—Asphalt Pavements and Overlays for Road Traffic**. Shell International Petroleum Company Ltd., London. 1978

SREEJITH, G. H. Pavement Temperature Data Maps for Worldwide Locations. **International Journal of Pavement Research and Technology**, 15:948–956. 2021. <https://doi.org/10.1007/s42947-021-00064-8>

SULEIMAN, A., ABDULLAHI, U., AHMAD, U. An analysis of residuals in multiple regressions. **International Journal of Advanced Technology in Engineering and Science**, Volume No 03, Special Issue No. 01. 2015.

UYANTO, S. Monte Carlo power comparison of seven most commonly used heteroscedasticity tests, **Communications in Statistics - Simulation and Computation**, 51:4, 2065-2082. 2022. DOI: 10.1080/03610918.2019.1692031

ZAUMANIS, M., POULIKAKOS, L., PARTL, M. Performance-based design of asphalt mixtures and review of key parameters, **Mater. Design**, 141 185–201. 2018. <https://doi.org/10.1016/j.matdes.2017.12.035>.