The development of prognostic tools for MV cable circuits

S. Christou, P. L. Lewin, J. A. Pilgrim and S. G. Swingler

School of Electronics and Computer Science
University of Southampton
Southampton, United Kingdom
e-mail: sc10g09@soton.ac.uk

Abstract— **Cable failures are disruptive, costly to repair and have a serious impact on customer confidence. Thus developing a reliable on-line prognostic tool is of a great interest. An experimental setup has been created to develop a new prognostic thermal model for MV underground cables. This paper introduces a thermal prognostic simulation model based on Support Vector Regression Algorithm which predicts the likely temperature along the cable thirty minutes into the future and is able to detect temperature anomalies which can indicate upcoming failures.**

Keywords—Condition Monitoring; Support Vector Regression; Diagnostic; Insulation System.

I. INTRODUCTION

Cables start to age the moment they are put into operation. The lifetime of an aging cable is affected by a combination of mechanical, thermal and electrical stresses as well as environmental conditions [5].

Partial discharge (PD) monitoring tools can detect the defect prior to the failure of insulation [1]. Hydro Quebec Research Institute [2] and KEMA [3] have developed on-line monitoring systems which are able to detect and localize PD in cable and cable joints and to indicate the upcoming failures by spotting the abnormal PD activity. Another way of condition monitoring is the monitoring of the temperature on the joint of underground cable. Hydro Quebec developed an online monitoring software which recorded temperatures of the cable joints and indicating the health of the joints [4].

All the above condition monitoring systems have been designed to monitor and indicate just one particular factor of probable cable failures. Therefore it is not possible just from this monitored data alone to determine the remaining lifetime of the circuit before catastrophic failure occurs. Thus development of multi-factor online condition monitoring and associated prognostic indicators is of interest.

This paper investigates a method of developing a prognostic capability for evaluation of the health and long term performance of aging distribution cable circuits. Developing such a prognostic model will significantly improve the prognosis accuracy, allowing the targeting of maintenance and reduction of in-service failures. Real-time measurements of weather conditions, partial discharge activity, surface temperature by the cable joint ,strain/stress of

cable joint and cable loading taken close to underground cables can be used to update simulation models giving a more accurate representation of the state of the cable.

At the present stage a thermal prognostic simulation model has been established using the Support Vector Regression (SVR) Algorithm which predicts the likely temperatures along the cable 30 minutes into the future, according to weather conditions and known loading. Anomalies of temperature measurements along the cable compared to predicted temperatures can indicate possible degradation activity in the cable. An experimental surface trough has been set up where operation of power cables is simulated with a control system which is able to model any cable loading. The surface temperature of the cable is continuously monitored as well as the weather conditions such as solar radiation, wind speed, humidity, rainfall and air temperature. The experiment is used to investigate the feasibility of different approaches for monitoring underground cables, where hot-spots are more likely to occur, such as joints and terminations. It is assumed that an increase of the local cable temperature is indicative of accelerated aging of the cable insulation due to thermomechanical, electrical and environmental factors.

II. EXPERIMENTAL SETUP

An experimental surface trough consists of 3.95 m trench filled with small grain sharp sand. Simulated cables are constructed by a heater tape wrapped around a length of aluminum pipe. The tape has a nominal power of 400 W, thus for a pipe of 3.95 m the maximum heat loss is approximately 100 W/m. The surface temperature of the cable is continuously monitored using 15 thermocouples (T/C) placed on the cables at various positions within the trench as its shown in Figure 1. All the thermocouples are connected to the Campbell Scientific datalogger CR1000 via a 32 channel multiplexer. The datalogger record the temperatures every 1 minute giving an average value every 30 minutes.

A weather station is installed on the roof of Tony Davies High Voltage Laboratory to collect measurements of weather conditions. The TABLE I. below shows the weather condition factors that are monitored by the Campbell Scientific datalogger CR10.
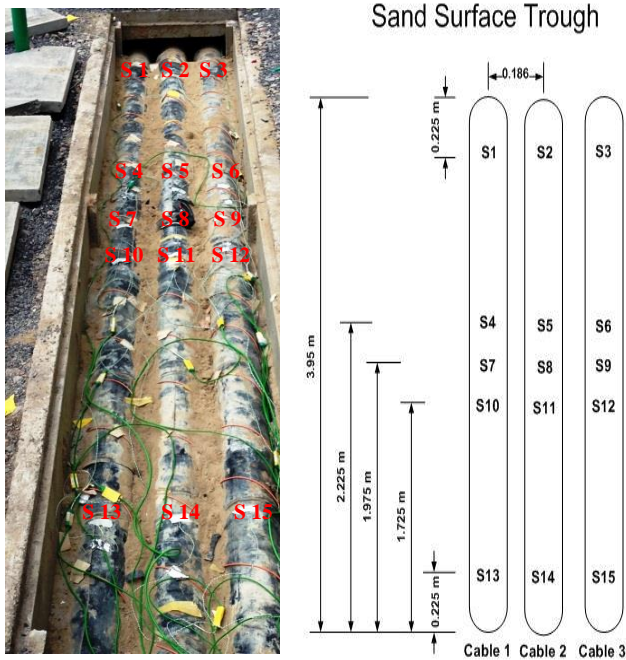
Fig. 1 Layout of T/Cs in the Sand Surface Trough.

TABLE I.   RECORDINGS OF THE WEATHER DATA

Factors	Unit	Measurement Interval	Averaging Interval
Air Temperature	˚C	5 min	30 min
Solar Radiation	W/m²	20 sec	30 min
Wind Speed	m/s	20 sec	30 min
Wind Direction	°	20 sec	30 min
Relative Humidity		5 min	30 min
Total Rainfall	mm	30 min	N/A

A. SVR Input Experiment

During this experiment a constant power demand profile of 100W was loaded in the simulated cables producing an output heat loss of 25 W/m. During the loading condition artificial hot-spots are introduced to the cable, by an external heat source, in order to investigate if the thermal prognostic simulation model is sensitive enough to detect the existence of the temperature anomalies along the cable, produced by the Hot-Spots. Three different hot-spot power profiles of 5, 10,15W were performed.

The experiment investigates if it is possible to monitor specific areas of an underground cable, where Hot-Spots are more likely to occur, such as joints and terminations.

The external heat source is monitored by T/C S8 which is positioned in the middle of the Cable 2 as is shown in Fig1. T/C S2 is used as an input variable for SVR training as well as the Solar Radiation, the Air Ambient Temperature and the load profile demand of the experiment.

III.   THERMAL PROGNOSTIC SIMULATION MODEL BASED ON SUPPORT VECTOR REGRESSION (SVR) ALGORITHM

A thermal prognostic simulation model has been developed using the SVR Algorithm. The model predicts the temperatures along the cable 30 minutes ahead, in line with weather conditions and known loading. Abnormal temperature measurements along the cable compared to predicted temperatures can be a sign of upcoming degradation activity in the cable.

A. Support Vector Regression

Support Vector Regression is a machine learning algorithm which uses a non-linear mapping to transform data into a high dimensional feature space where linear regression is performed [6]. SVR has the advantage that not only fits the training data in order to find a solution but also keep the highest possible degree of generality when new unseen data are introduced [7].

The ε-SVR is used to predict the values by solving the following optimization problem [8]:

$$\text{minimize} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*)$$

$$\text{subject to} \quad \begin{cases} y_i - w^T\varphi(x_i) - b \le \varepsilon + \xi_i \\ w^T\varphi(x_i) - y_i + b \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases} \quad (1)$$

where x_i is a feature vector of the input space with dimension N, y_i is the output value to be estimated, b is a parameter of bias, w controls the smoothness of the model, φ(x) is a function used to map training samples x_i to a higher dimensional feature space, penalty factor C controls the trade-off between complexity of the function and the frequency with which errors are allowed and ξ_i and ξ_i^* are slack variables which compute the error for overestimating and underestimating the true output value y_i. The parameter ε determines the maximum deviation from the target output value y_i, Fig. 2 shows the situation graphically.
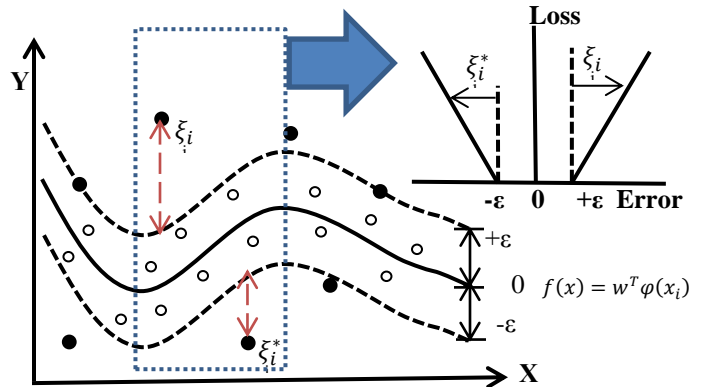

Fig. 2 The parameters for the Support Vector Regression Algorithm.

This above optimization problem can be transformed into the dual problem, as mentioned in [9] and its solution is given by:

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(x_i, x) \quad (2)$$

where the dual variables α_i and α_i^* are subject to constrains $0 \leq \alpha_i$, $\alpha_i^* \leq C$. The kernel function $K(x_i, x)$ is used to avoid the complicated calculations of the inner products of the mapped inputs in the feature space φ(x).

B. SVR Develompent

A thermal prognostic simulation model based on the Support Vector Regression (SVR) Algorithm predicts the likely temperatures along the cable 30 minutes ahead, taking into account weather conditions and known loading. Weather data and cable surface temperature data were taken from 30th of October 2013 until 12th of December 2013. The LIBSVM toolbox [10] in Matlab has been used for the development of the thermal prognostic model.

1) Kernel Selection

The kernel function is used to allow operation in the input space instead of the potentially high dimensional feature space. Hence as was mentioned before the inner product does not need to be computed in the feature space. There are different types of kernel functions such as Gaussian Radial Basis Function, Exponential Radial Basis Function, Multi-Layer Perceptron, Fourier Series, Splines and etc. There is not a straightforward way to select a kernel function yet, but the most effective for general use has been found to be the Gaussian Radial Basis Function (Gaussian-RBF) [11]. Hence here we use the Gaussian-RBF which is defined as:

$$K(x_i, x) = e^{(-\gamma \, \|(x_i - x\|^2\,)} \qquad \gamma > 0 \qquad (3)$$

where parameter γ controls the width of the Gaussian-RBF kernel function of the model.

2) Data Preparation

The input data for SVR training have been prepared and separated as is shown in TABLE II.

TABLE II.   INPUT DATA FOR SVR

Input	Variables	Description
1-126	Solar Radiation	Mean Solar radiation captured every 30 min
127-252	Air Ambient temperature	Mean Air temperature captured every 30 min
253-378	T/C S2 (see Fig.1)	Temperature of T/C S2, captured every 30 min

The input data for the SVR is the Solar Radiation, Air Ambient Temperature and the Temperature of the distant T/C S2. The load demand input was not taken into the account for this model, as the constant value of it does not add any feature to the model but only increases the complexity of the whole system. The T/C S2 was chosen as it is one of the most distantly located T/C on the same cable where the heat source lies.

Input data has to be scaled before being imported in the SVR in order to avoid features with larger numeric values being dominant to those with smaller values. Furthermore scaling makes the computation easier during the application of the kernel function. In this model the data were scaled from [0,1] as follows:

$$Xscale = \frac{Xi - Xmin}{Xmax - Xmin} \qquad (4)$$

where X is the original data, $Xscale$ the scaling data, $Xmax$ and $Xmin$ are the maxium and minimum values of X.

3) Grid-Search and Cross-Validation

The penalty factor C and the kernel parameter γ has to be decided from the user of the LIBSVM toolbox in order for the regression model to be developed. The aim is to identify a good pair of C and γ so that the classifier can accurately predict unknown data. For that reason the data are divided into two groups: a) 30/10/2013 - 20/11/2013 training data and b) 20/11/2013 – 23/11/2013 testing data. To identify the most appropriate C and γ two algorithms are combined together, Cross-Validation and Grid-Search. The training data are split into K segments with equal size in K-fold cross-validation and then each segment is predicted by the remaining K-1 segments sequentially. For each of iterations of Grid-Search, a combination of C and γ values are tested and the one with minimum Cross-Validation error is selected. C and γ pairs are change based on the grid-search algorithm which suggest exponentially growing sequence for C (2^{-5}, 2^{-3}... 2^{15}) and γ (2^{-15}, 2^{-13}... 2^5).

IV. RESULTS

Based on the SVR input experiment, the thermal prognostic model was developed. The T/C S2, Solar Radiation and Air Ambient Temperature are the input parameters to the SVR model which predicts the temperature of the T/Cs buried in the sand surface trough. Initially the model is tested at unknown data from 20/11/2013 – 23/11/2013 and the corresponding Mean Absolute Percentage Error (MAPE) is calculated as defined by:

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{y_a(i) - y_p(i)}{y_a(i)} \right| \qquad (5)$$

where y_a is the real value of T/Cs temperature, y_p is the predicted value of T/Cs temperature and n is the total number of samples. This stage is very crucial as it calculates the prediction accuracy of the model obtained from the unknown data and reflects the performance on classifying and independent data set.

After the confidence gained from the selection of the appropriate C and γ and the corresponding MAPE at the test data, a new experiment was implemented as described in Section II.A. Fig. 3-6 show prediction made for the unknown data for T/C S11 and S15 and Fig. 7-10 show the prediction made while the Hot-Spots were enabled. T/C S11 and T/C S15 were selected because the first one is close to the hot-spot location, 0.25 m, and the second one is located farthest from the hot-spot, 1.5 m away.
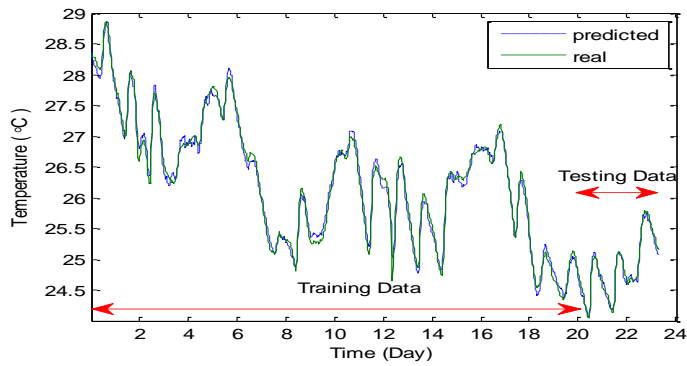
Fig. 3 Predicted temperature on unknown Testing Data and real temperature of T/C S11, MAPE=0.64 % .
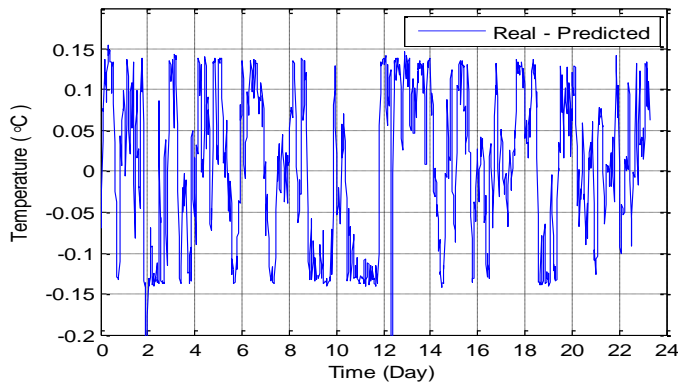

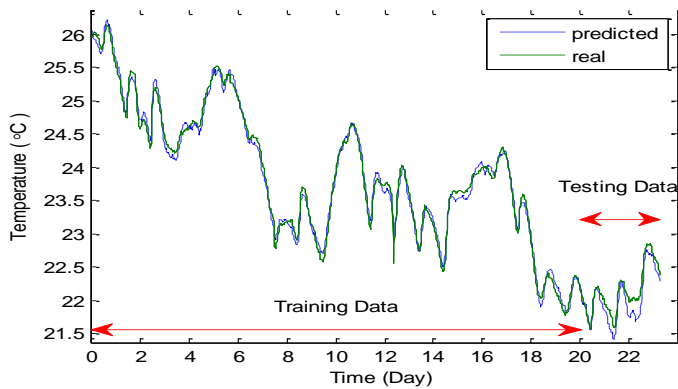Fig. 4 Temperature prediction error for T/C S11 for Testing Data.

Fig. 5 Predicted temperature on unknown Testing Data and real temperature of T/C S15, MAPE=1.86 % .
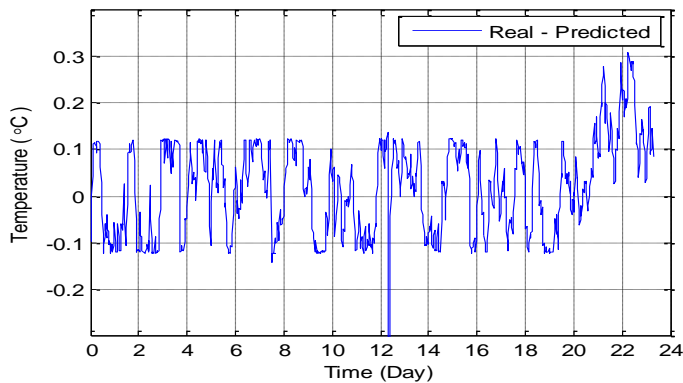

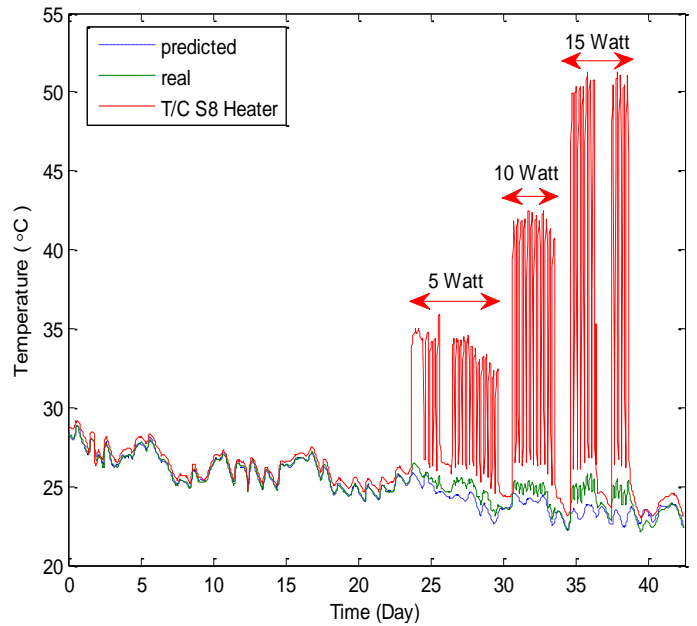Fig. 6 Temperature prediction error for T/C S15 for Testing Data.

Fig. 7 Predicted temperature and real temperature for T/C S11 under the influence of Hot-Spots, MAPE=4.45 % .
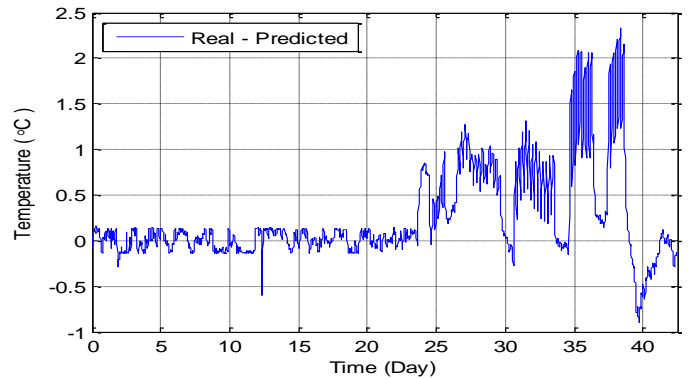

Fig. 8 Temperature prediction error for T/C S11 under the influence of Hot-Spots.
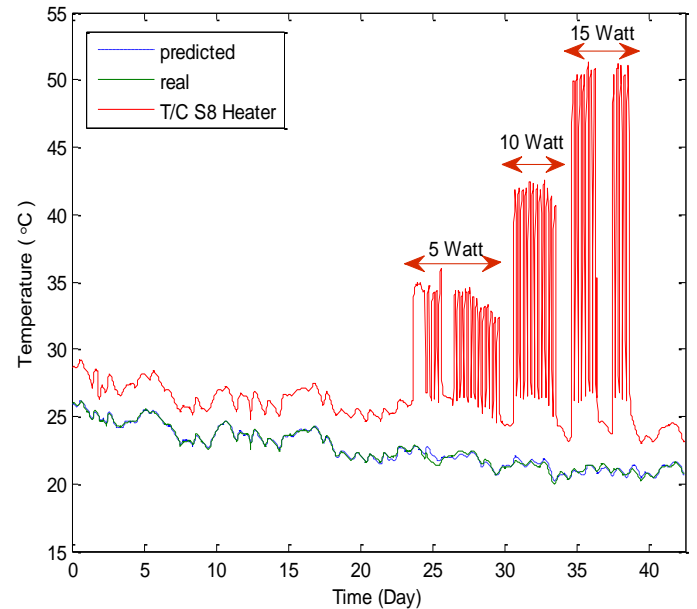

Fig. 9 Predicted temperature and real temperature for T/C S15 under the influence of Hot-Spots, MAPE=1.78 % .
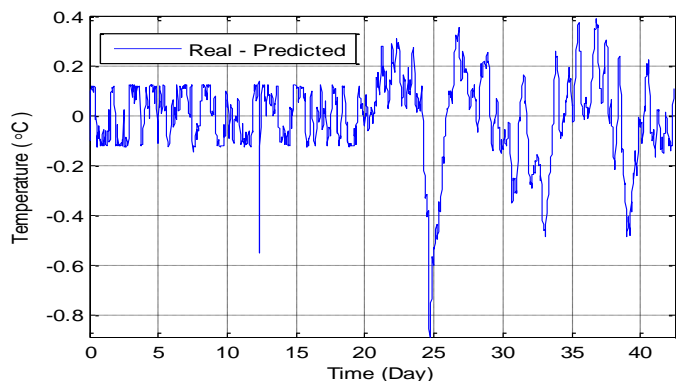
Fig. 10 Temperature prediction error for T/C S15 under the influence of Hot-Spots.

The Fig. 11 shows the distribution of the maximum temperature prediction error along the cable.
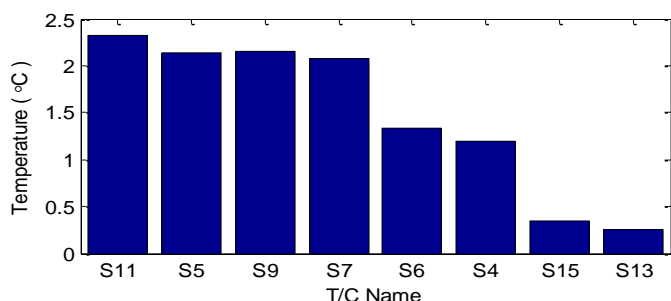

Fig. 11 Maximum temperature prediction error.

V. DISCUSSION

The prediction made for the unknown testing data of T/Cs S11 and S15 shows good results with prediction MAPE less than 2% as it can been seen from Fig. 3-6. After the introduction of the hot-spots from the day 23 until day 39 of the experiment, the MAPE for T/C S11 increased to 4.45%, which is almost 7 times greater than before. A better understanding of the influence of the hot-spots on the T/C S11 can be seen in Fig. 7-8. The temperature prediction error for T/C S11 is increased when the Hot-Spots are on. When the hot-spots are off, on the days 26, 30, 34, 37, 39 and onwards, it can be seen that the predicted temperature matches the real temperature. Hence the T/C S11 is affected by the hot-spots. The MAPE of T/C S15 during the introduction of the Hot-Spots was 1.78%, which is close to its previous testing value of 1.86%. Thus the prediction error of the T/C S15 is within the same boundaries as it was before and is not affected by the introduction of the hot-spots as it can be seen in Fig. 9-10. In the event when the real temperature is higher than predicted temperature, positive part of temperature prediction error, it can be assumed as indication of upcoming failure of the system.

Fig.11 shows the distribution of the maximum temperature prediction error along different locations on the cables. The T/Cs close to the hot-spots such as S11, S5, S7 and S9 have the highest temperature prediction error. The influence of the hot-spots, produced by the heater positioned above T/C S8, is

reduced for the T/Cs located farther such as T/Cs S6 and S4. However T/Cs S6 and S4 can still identify the anomalies of the temperature produced by the hot-spots despite the fact they have a smaller temperature prediction error than the T/Cs S11, S5, S7 and S9. The most distant T/Cs such as S15 and S13, positioned 1.5 m away from the heater location, are not capable to detect temperature anomalies of the system as the sensitivity of the T/Cs is around ± 0.5 °C. Hence it can be concluded that the most effective distance for the thermal prognostic simulation model to identify the location of the Hot-Spots is in the radius of up to 0.30 m.

VI. CONCLUSION

In this paper a novel thermal prognostic tool has been developed using SVR, to identify anomalies of cable temperature 30 min in to the future. The prediction output can be used to identify a possible sign of upcoming degradation activity in the cable. It has been proven that it is possible to monitor specific areas of an underground cable, where hot-spots are more likely to occur, such as joints and terminations. The use of such systems enables a smarter way of prognostic condition monitoring in which you measure less and model more. The use of such models will enable the power network operators to maximize asset utilization and minimize constraint costs in the system.

REFERENCES

[1] J. Densley, "Ageing mechanisms and diagnostics for power cables - an overview," IEEE Electrical Insulation Magazine, vol. 17, no. 1, pp. 14–22, Jan. 2001.

[2] R.-N. Wu and C.-K. Chang, "The Use of Partial Discharges as an Online Monitoring System for Underground Cable Joints," IEEE Transactions on Power Delivery, vol. 26, no. 3, pp. 1585–1591, Jul. 2011.

[3] P.C.J.M. Wielen and E.F Steennis, "Experiences with Continuous Condition Connections Monitoring of In-Service MV Cable Connections," pp. 1–8, 2009.

[4] D. Fournier et al., "Detection, Localization and Interpretation of Partial Discharge in the Underground Distribution Network at Hydro-Quebec,''CIRED 2005,Turin,6-9 June 2005.

[5] D. Fournier and N. Amyot, "Diagnostic of Overheating Underground Distribution cable joints," CIRED 2001, 18-21 June 2001.

[6] V. Vapnik, "The Nature of Statistical Learning Theory," Springer,1995.

[7] E.G. Ortiz-García,et al., "Accurate local very short-term temperature prediction based on synoptic situation Support Vector Regression banks," Atmospheric Research, Volume 107, pp. 1-8, April 2012.

[8] A.J. Smola, B. Schölkopf, " A tutorial on support vector regression,'' Stat. Comput., pp. 199–222, 14 (3) (2004)

[9] V. Vapnik, "Statistical Learning Theory," New York:Willey,1998.

[10] C. Chang and C. Lin. "LIBSVM: a Library for Support Vector Machines", April 2005, http://www.csie.ntu.edu.tw/cjlin/libsvm.

[11] Bernhard Scholkopf and Alexander J. Smola. "Learning with Kernels: Support Vector Machines, Regularization, Optimisation and beyond." The MIT Press, 2002.