# Spatial Distribution of moisture content in Cement mortar structures using Radial Basis Functions

Layan Kaushik

Indian Institute of Technology (I.I.T.) Roorkee, Roorkee, Uttarakhand, Postal Code: 247667

Professor Bishwajit Bhattacharjee

Indian Institute of Technology (I.I.T.) Delhi, Hauz Khas, New Delhi, Postal Code: 110016

# Spatial Distribution of moisture content in Cement mortar structures using Radial Basis Functions

Layan Kaushik

Indian Institute of Technology (I.I.T.) Roorkee, Roorkee, Uttarakhand, Postal Code: 247667

Professor Bishwajit Bhattacharjee

Indian Institute of Technology (I.I.T.) Delhi, Hauz Khas, New Delhi, Postal Code: 110016

#### **Abstract**

The basic ingredients of all concrete structures are aggregates, cement and water. The relative composition of water in concrete preparation is around 16% of the total ingredients. Workability of concrete depends on the water cement ratio at a certain section, although the mix is considered to be uniform throughout the concrete mix. But at microscopic level the water content in the beam differs axially as well as longitudinally. This uncertainty of uniformness in concrete structures may arise due external factors like exposure to atmospheric moisture and rain. Excess water in concrete structures results into the common phenomenon of corrosion. Many more moisture related problems can be traced to poor decisions in design, construction or maintenance. The core idea of the project is to develop an efficient, accurate and reliable method to estimate the moisture content in laboratory synthesised Concrete mortar blocks. Since direct calculation of water content in a cement concrete element is not possible, the project aims to find a relation between the impedance value of cement concrete block, which can be measured directly, and moisture content in it after a time, 't'. To meet the objective, concepts of RLC circuit is applied to multiple specimens of Concrete blocks synthesised in the laboratory. The blocks are place in a simulator in series with several known resistors and capacitors and tested to in multiple formations. The aim of doing so is to find the impedance value, subsequently cumulative impedance, of the Concrete block by shifting the circuit connection to different positions along its length. A Digital Storage Oscilloscope is in use to collect raw data for the experimentation. The data obtained from such experiments are discrete which is not suitable for the purpose it serves. Thus in order to find a continuous function, which can define the impedance (or potential) at any point along the length of specimen, we take assistance of mathematical tools and concepts of Splines and Radial Basis Functions (RBF). These functions are used to interpolate the collected data between experimentally obtained data sites. We also analyse the methods of choosing a

specific Radial Basis Function to comply the above. The scope of this paper is limited to spatial analysis of the water content in a standard Cement Concrete block.

**Keywords or phrases**: RLC circuit, Splines, Corrosion, Digital Storage Oscilloscope.

#### Abbreviations

#### Abbreviations

RBF	Radial Basis Function
RMS	Root Mean Square
SP	Shape Parameter
MLP	Multi Layered Polynomial
Н	Hypothesis
exp	Exponential
f	Function
cos	Cosine
$\sum$	Summation
ſ	Integral
$\infty$	Infinity
Π	Product

#### 1 INTRODUCTION

# 1.1 Background/Rationale

Concrete mortar is a conglomerate made by mixing aggregates, sand, cement and water in suitable proportions. The resulting composite material is often described by its strength and stiffness. Under different temperatures and drying time, cracks develop on concrete beams, which explain that moisture content fluctuates inside concrete, and this change is not

homogeneous. Since at micro-scale, water distribution is varied, this changes the strength and stiffness to a little extent.

This work consists of supervised data set i.e. the experimental data collected. For a scattered data approximation, a number of methods can be applied. The most common of all is regression. In regression method, a linear relationship is established using the help of least square method algorithm. For example, fitting a straight line: f(x) = ax + b, to a bunch of points,  $\{(x_i,y_i)\}_{i=1}^p$ , is parametric regression because the functional form of the dependence of y on x is given, even though the values of a and b are not. Typically, in any given parametric problem, the free parameters, as well as the dependent and independent variables, have meaningful interpretations. Linear models are simpler to analyse mathematically. In particular, if supervised learning problems are solved by least squares then it is possible to derive and solve a set of equations for the optimal weight values implied by the training set (f(x) = ax + b). The same does not apply for nonlinear models, such as MLPs, which require iterative numerical procedures for their optimisation.

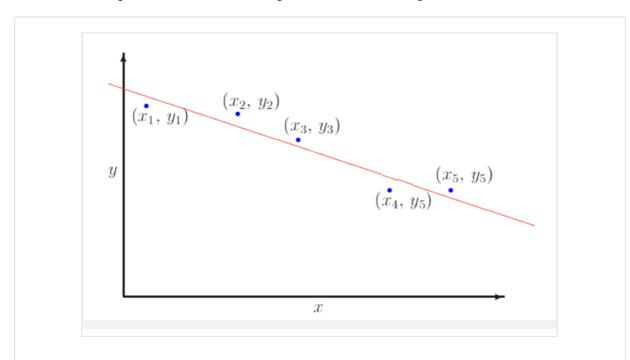


Fig 1 Fitting astraight line to a bunch of points is a kind of parametric regression where the form of the model is known.

The drawbacks of regression is that, all variables aren't linearly related i.e. x may not be linearly related to y. For the same, a solution for non linear regression points out the idea of

basis function. Basis Function is a feature space: define  $\phi(x)$  where  $\phi$  is a non-linear function of x. The model is estimated such that the target is a linear combination of this non-linear functions i.e.  $y(x) = \sum_{j=1}^K w_j \phi_j(x)$ . Neural networks, including radial basis function networks, are nonparametric and non linear models and their weights and other parameters, have no particular meaning in relation to the problems to which they are applied. Estimating values for the weights of a neural network, or the parameters of any nonparametric model, is never the primary goal in supervised learning. The primary goal is to estimate the underlying function, or at least to estimate its output at certain desired values of the input. On the contrary, RBF approximation depends on the centre that is carefully and tactfully chosen by the user using algorithms for perfect fit and can be plotted numerous times. Here, we follow unsupervised non-linear learning of centres followed by pseudo-inverse of the model. RBF can be derived based purely on regularisation:

$$\sum_{n=1}^N (h(x_n)-y_n)^2 + lpha \sum_{k=0}^\infty a_k \int_{-\infty}^\infty \left(rac{d^k h}{dx^k}
ight) dx$$

The above equation is used for smoothest interpolation.

#### 1.2 Statement of the Problems

The spatial distribution of water content in any concrete structure relates to its durability and strength. In our study of cement mortar, we go into depth cross sectional analysis of the mortar sample using concepts of RLC circuit and supervised learning. Furthermore interpolation of the collected data is done using mathematical concepts. For this paper we have used RBF with different shape parameters to interpolate and plot the collected data. Although the water content of cement blocks is generally measured, there isn't a standard method to do so. This research aims to develop an efficient, accurate and reliable method to measure the water content and approximately distribute the data with least error.

### 1.3 Objectives of the Research

#### 1.3.1 Overall objective

The objective of the paper is to develop a standard method to calculate the heterogeneous distribution of moisture content in the cement mortar. This results in efficient designing and proper strengthening of cement mortar which is generally used for making paste for concrete walls and structures.

## 1.4 Scope

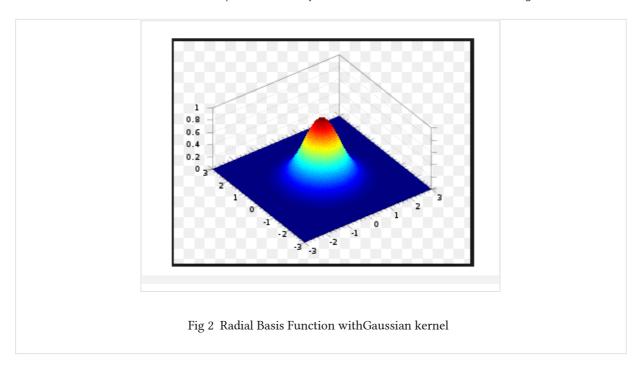
The scope of this paper is to derive a standard and accurate methodology to calculate the distribution of water content in mortar. This distribution will thus be used to design more consistent structures and mortar paste. The scope of the paper is limited to mortar and can be extended to reinforced concrete by further research.

#### 2 LITERATURE REVIEW

# 2.1 Learning Radial Basis Function

#### 2.1.1 Basic RBF Model

Each  $(x_n, y_n)$  (belongs to Data set D, influence a hypothesis h(x) based on  $||X - X_n||$ ). The hypothesis function we choose is affected by the distance of the data set. Key component of RBF being that a point chosen from the data set affects the nearby points more than it affects far away points. Moreover this function is symmetric about a point or a centre which in turn describes the function as only distance dependent.



In case of the figure above, influences of data point  $X_n$  at the top of the Gaussian function influence its nearby points more rigorously than the far away points.

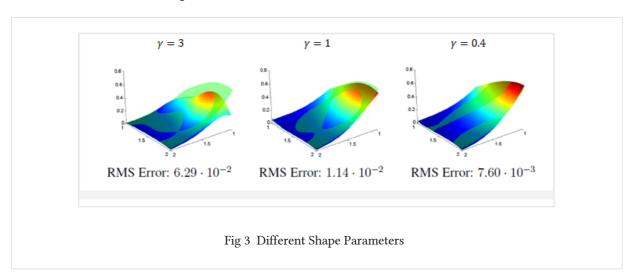
#### 2.1.2 Standard Form of RBF

 $h(x) = \sum_{n=1}^{N} w_n exp \left(-\gamma||X-X-n||^2\right) \ \ (\text{Equation 1}), \text{ is a typical Gaussian function } \\ \text{with Gaussian kernels. The hypothesis, h(x) is defined as the contribution of point X for evaluation the function, according to the data point X_n, from the data set. Every point of the data set has an influence, in turn having a parameter that reflects the value. In the above equation 1, w_n is defined as the weight which tells from the data point propagates its influence. The dependence of distance or inclusion of <math>\|X-X_n\|$  makes the equation radial and the building block of  $\exp\left(-\gamma||X-X_n||^2\right)$  makes it a basis function. More aptly, a radial basis functions are a special class of function. Key feature of radial basis function is that their value depends only on the distance from the origin i.e. it is radially symmetric around some point c called the function centre. Understanding of the radial function can be traced back to Fourier series properties. The superposition principle which is applicable for the infinite sine and cosine functions for successive approximation to common function in Fourier series derives a corollary to prove radial function from functions called splines. Few properties of splines are; 1. It is composed of piecewise polynomials, 2. The national cubic splines

interpolant is the function from Sobolev space  $H^2$  [ a, b] that minimines the semi-norm  $||f^{"}||^2_{L_2[a,b]}$  under the conditions  $f(x_j)=f_j$ ,  $1\le j\le N$ . Extending splines tot the multivariate setting is based on property 1. Property 2 leads to theory of RBF as splines no longer consisting of piecewise polynomials.

#### 2.1.3 Effect of Shape Parameter, $\gamma$

In the hypothesis model, there can be different width of the model of the radial basis function. The width of the function is determined by the relative value of  $\gamma$ . Small value indicates wider Radial basis function shape; larger value indicates sharper shape of the function. The interpolation in case of large  $\gamma$  is poor because the influence of the points on the Gaussian dies out and between two points no inference can be made.



#### 2.1.4 Learning Algorithm

The aim of the learning algorithm is to find the parameters of the RBF and minimize the error of the model with appropriate values of the weights  $w_n$ . We can either perform exact interpolation by using supervised learning i.e. with the help of data set label  $y_n$  or use Lloyd's algorithm to find the centres and the weights. RBF can be related to nearest-neighbour Method, we look for the closest point within the training set to the point we are considering. For a point X, we find  $X_n$  in the data set that is closest to X in Euclidean distance and then inherit the value that point has. The issue of abruptness in nearest neighbour problem is tackled with modification to become K nearest neighbour, i.e. instead of the value of the closest point; we look for K closest points. A similarity based method is adopted, classifying points on the basis of how similar the centres are to points in the training set.

For a Radial Basis Function with K centres, we have N parameters  $w_1, \ldots, w_k$  based on N points in the data set. Here K<N centres  $\mu_1, \ldots, \mu_K$ . A number of important centres are chosen for the data set and have those points influence the neighbourhood around. The RBF now becomes  $h(x) = \sum_{k=1}^K w_k exp(-\gamma||X-\mu_k||^2)$ . Evaluation of x is done using the K centres. According to the distance from that centre, the influence of that particular centre, which is captured by  $w_k$  is contributed. The centre  $\mu_k$  is a representation of a cluster of points from the data set. The centres are chosen such that to minimize the distance between  $x_n$  and the closest centre. This method is K-means clustering. The formalization of the centres is done by splitting the data set  $x_1, \ldots, x_n$  into clusters  $S_1, \ldots, S_K$ ; each cluster have a centre. We then minimize the equation  $\sum_{k=1}^K \sum_{X_n \in S_k} ||X_n - \mu_k||^2$ . We minimize the mean square error at the cluster in terms of Euclidean distance using unsupervised learning techniques.

An iterative approach can also be undertaken for K-means clustering. Lloyd's algorithm which is an iterative approach focuses to minimize the equation  $\sum_{k=1}^K \sum_{X_n \in S_k} ||X_n - \mu_k||^2$  with respect to  $\mu_k$  and  $S_k$  which are the parameters for the algorithm. The way of the algorithm is it fixes one of the parameters in the equation and minimizes the other. Under this algorithm we first take the mean of the data set as the first centre and fix it to minimize the cluster.

Equation  $\frac{1}{|S_k|}\sum_{x_n\in S_k}x_n$  is used for calculating mean of the cluster. The squared error to the mean is smallest of the squared error to any point. That happens to be the closest to the points collectively, in terms of mean squared value is taken. Then freezing the value of  $\mu_k$ , the algorithm creates a new cluster. For cluster we minimize using  $\{x_n:||X_n-\mu_k||\leqslant all||X_n-\mu_k||\}$ , the validity of the condition is checked and inclusiveness of data set  $x_n$  in a cluster  $S_k$  is concluded. This method reduces the in sample error since we use mean and shuffle points among the considered clusters. In shuffling, the points move to the closest centre, therefore the condition value becomes smaller. The convergence of the iterative algorithms happens at the local minimum of the parameters. The centres and clusters are thus obtained by iterative clustering and centring. Centres are vectors of the dimension found without using the label  $y_n$  from the training data set.

#### 2.1.5 Choosing the Weights $w_1, \ldots, w_n$

To find the weight parameter of the radial basis function model, we make the hypothesis  $h(x) = \sum_{k=1}^K w_k exp(-\gamma||X_n - \mu_k||^2)$ , approximately equal to data set label  $y_n$  which consist of N equations and K unknowns; K<N.

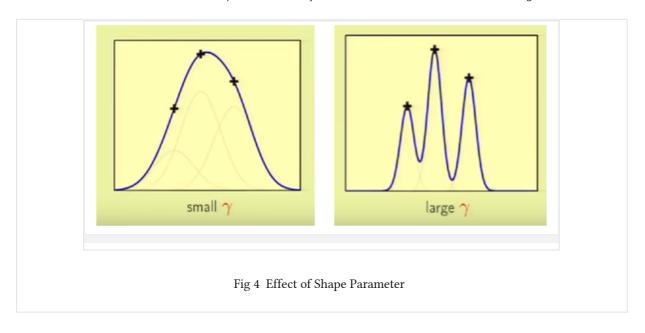
$$h(x) = \sum_{k=1}^K w_k exp(-\gamma ||X_n - \mu_k||^2) pprox y_n$$
 in matrix form is;

$$\begin{bmatrix} exp(-\gamma||X_1-\mu_1||^2) & \cdots & exp(-\gamma||X_1-\mu_K||^2) \\ \vdots & \ddots & \vdots \\ exp(-\gamma||X_N-\mu_1||^2) & \cdots & exp(-\gamma||X_N-\mu_K||^2) \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_K \end{bmatrix} \approx \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$$

The matrix which Gaussian kernel is represented by  $\phi$ , which is multiplied with the weight matrix approximately giving y matrix. For the solution of the matrix we check whether  $\phi^T\phi$  is invertible. If so, the solution is given by;  $w=(\phi^T\phi)^{-1}\phi^Ty$ , which is the pseudo inverse solution of the matrix equation. The value of weight minimizes the mean square distance in the basis function. Although this method doesn't guarantee the correct value at every data point, it is the most suitable unsupervised learning method in Radial basis function to obtain optimum weight vector.

### 2.1.6 Choosing the value of $\gamma$

For the hypothesis  $h(x) = \sum_{k=1}^K w_k exp(-\gamma||X-\mu_k||^2)$ , we treat the shape parameter  $\gamma$  as a parameter. Interpolation of the model depends on the value of  $\gamma$ . It can minimize the in-sample error. For the solution of  $\gamma$ , we use gradient descent approach. In this algorithm, we use an iterative approach called Expectation Maximization algorithm in mixture of Radial basis function kernels. In this we first fix  $\gamma$ , and solve for the weights  $w_1, \ldots, w_k$ . After that, we fix the weights and minimize the hypothesis with respect to  $\gamma$ . The same is iterated until the combination converges to give the values of the parameters i.e. the weights and the shape parameter  $\gamma$ . We can also have different  $\gamma_k$  for each centre  $\mu_k$  and solve by the same iterative approach algorithm.

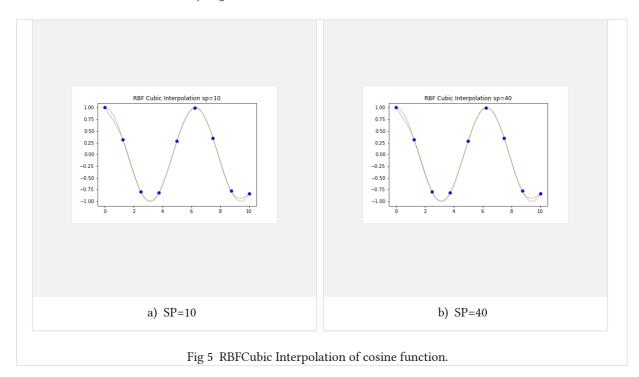


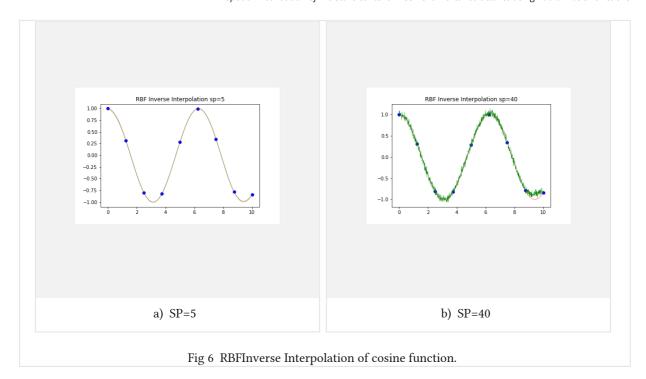
#### 2.1.7 Types of Radial Basis Function

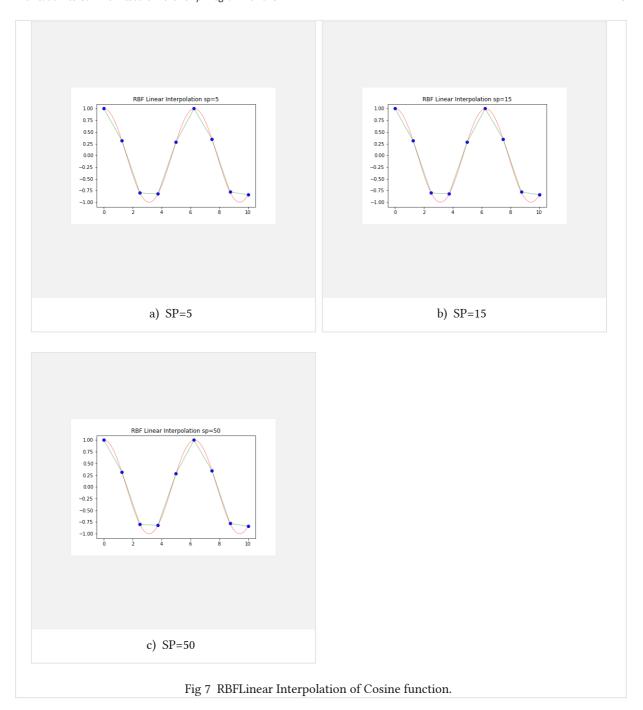
The meaning of radial basis function is a function symmetrical about a point, centre here, and with a Euclidean distance of  $||X-X_n||$ . Any function with can written in this form is a radial Basis Function. For example, a cylindrical function centred at the centre of the base of the cylinder is a radial basis function. Few more examples could include: Gaussian function:  $exp(-\gamma||X-X-n||^2)$ , Thin plate spline:  $||X-X_n||^2 ln(||X-X_n||)$ ; Polyharmonic spline:  $||X-X_n||^k$ , for k=1,3,5; Multiquadric:  $\sqrt{1+\gamma||X-X_n||^2}$ , etc. All basis functions which are radially symmetrical about a centre can be regarded as a Radial Basis Function.

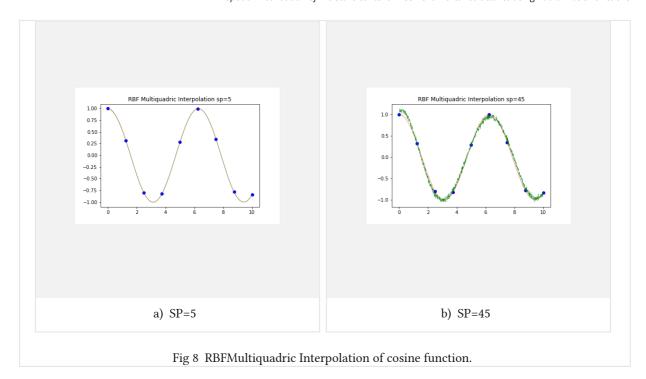
# 2.2 TApplication of Radial Basis Function

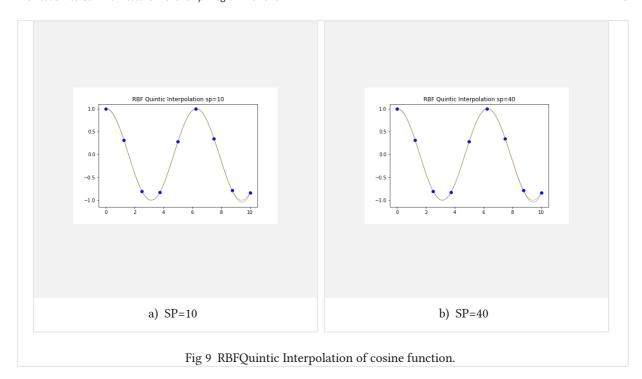
For basis functions like Gaussian function:  $exp(-\gamma||X-X_n||^2)$ , polynomial:  $\sum_{i=0}^n a_i x^i$ ; sine function  $\sin\theta$ , cosine function  $\cos\theta$ , the collected experimental data can be plotted. Using softwares and tools such as MATLAB, Python programming and algorithms studied above, several functions are incorporated to the scattered data using approximation and proper interpolation will be carried out. We have tried to plot and interpolate cosine function data in RBF with several basis functions and different shape parameters.

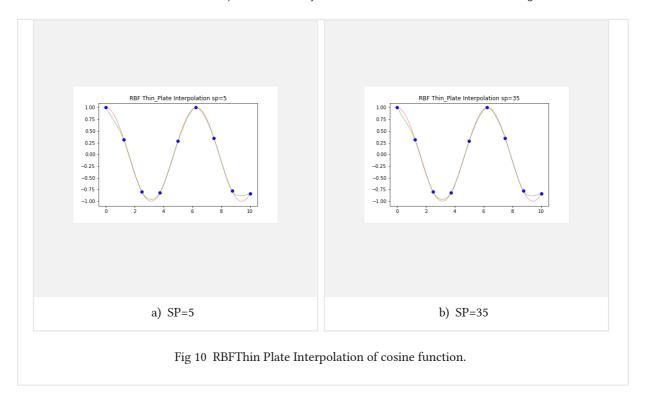












The above interpolations can be used in the collected data of the research but due to time constraint only Gaussian and Inverse Interpolation are used in this paper.

#### 3 METHODOLOGY

#### 3.1 Methods

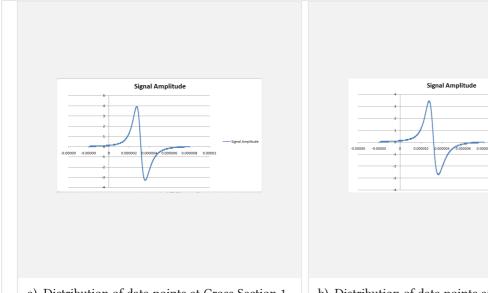
For the study, the concepts of RLC circuit and RBF method were widely used. Since direct calculation of water content in a cement concrete element is not possible, we tried to find a relation between the impedance value of cement concrete block, which can be measured directly, and moisture content in it after a time, 't'. To meet the objective, the concept of RLC circuit is applied to multiple specimens of concrete blocks synthesised in the laboratory. The blocks are placed in a simulator in series with several known resistors and capacitors and tested to in multiple formations. The aim of doing so is to find the impedance value, subsequently cumulative impedance, of the concrete block by shifting the circuit connection to different positions along its length. The signal output generated in our case is the current across a resistance of 1k-ohm. Moreover, for interpolation, mathematical tools like RBFs are used. The basis function radially interpolates the obtained data taking a centre generated by k-

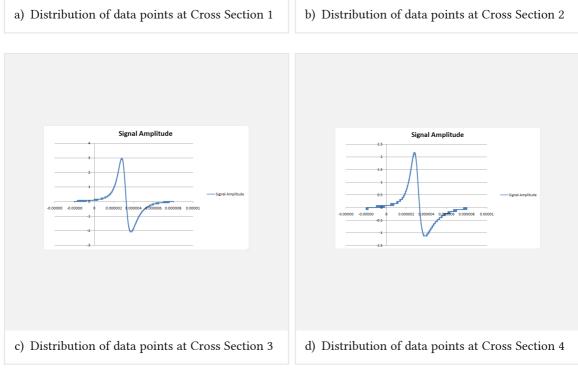
means clustering algorithm. Several observations are made using different basis functions and shape parameters for the data set.

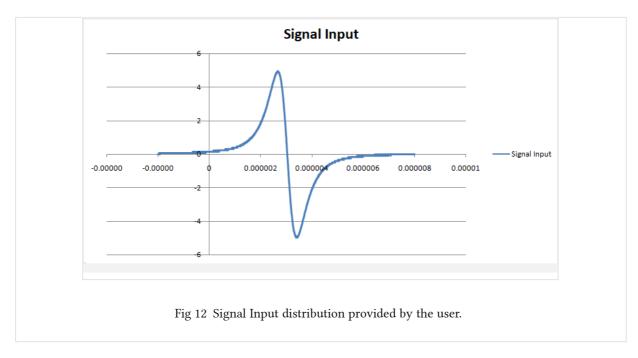
Table 1 Details on the collected data

	A	В
1	Record Length	2500
2	Sample Interval	0.000000004
3	Trigger Point	500
4	Source	CH1
5	Vertical Unit	Volts
6	Vertical Scale	2
7	Vertical Offset	0.32
8	Horizontal Unit	seconds
9	Horizontal Scale	0.000001
10	Pt Fmt	Y
11	Y Zero	0
12	Probe Atten	1

The studies were undertaken in the Civil Department lab of IIT Delhi. A DSO is in use to collect raw data for the experimentation. The data obtained from such experiments are discrete corresponding to seven cross sections. The Signal Output across each cross section is measured after a time interval of 0.0000000004 seconds and a total of 2500 data points are recorded for each cross section. The same procedure is repeated for each cross section eight times keeping the same time interval, as to observe and eliminate errors in the data set. An input signal is provided to the DSO as well. Using it the output signal is generated across the resistance of 1k-ohm to measure current. The data is thus recorded and stored.







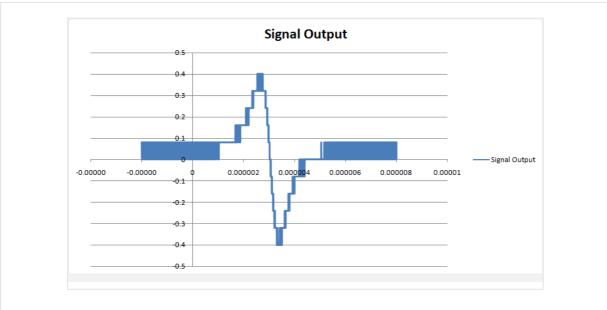


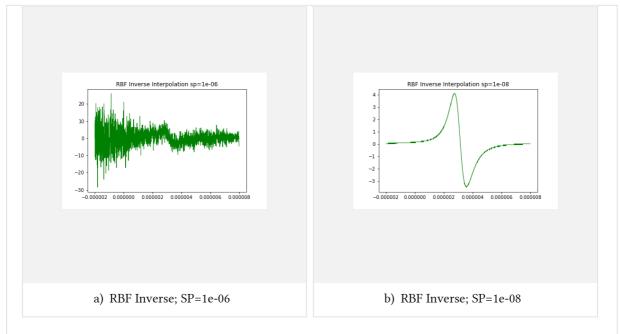
Fig 13 Output Signal Distribution across the resistance of 1k-ohm to measure current on the cement mortar.

Using this standard data distribution and software packages like Matlab and Python, interpolation of the data set is undertaken. This is done in order to find a continuous function, which can define the impedance (or potential) at any point along the length of block. We have

applied Gaussian and Inverse basis function with variable shape parameter to interpolate all the generated data and derive a standard and accurate function to define this model.

#### **4 RESULTS AND DISCUSSION**

Basis functions of Gaussian and Inverse with small shape parameters can be used to get a accurate and refined interpolation. We have observed that for shape parameter 1e-09 and 1e-10, the interpolation is more accurate and smooth.



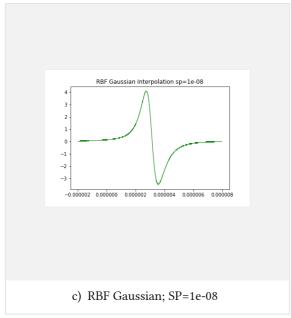


Fig 14 Cross Section 1 Interpolation using Inverse and Gaussian Basis Function with different Shape Parameter.

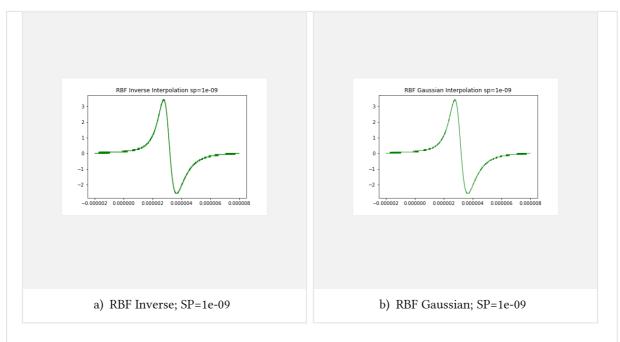
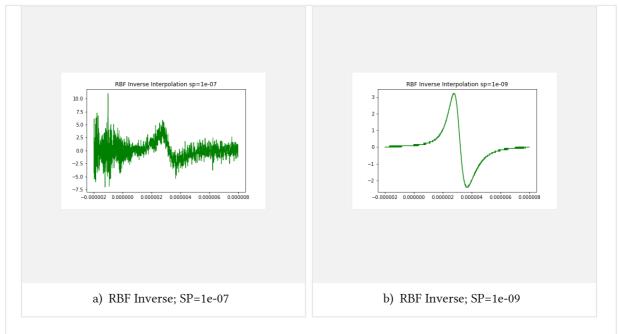


Fig 15 Cross Section 2 Interpolation using Inverse and Gaussian Basis Function with Shape Parameter 1e-09.



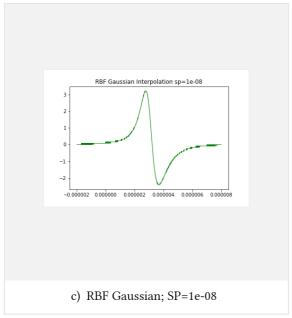


Fig 16 Cross Section 3 Interpolation using Inverse and Gaussian Basis Function with different Shape Parameter.

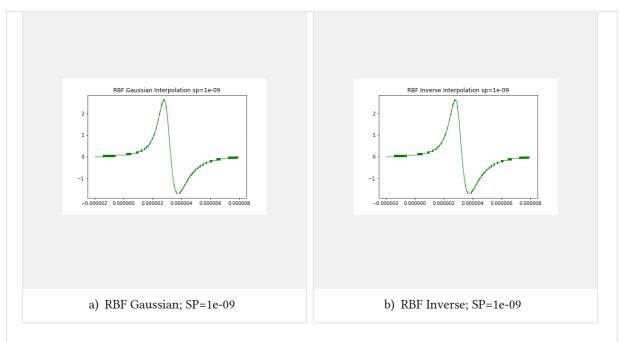
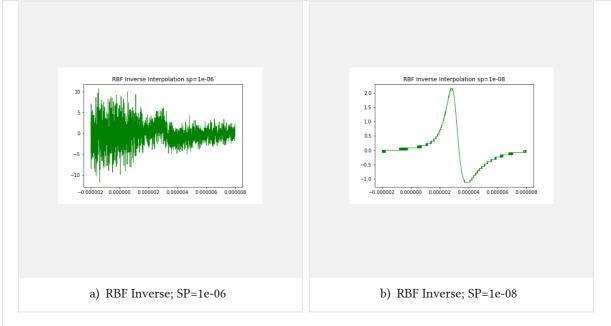


Fig 17 Cross Section 4 Interpolation using Inverse and Gaussian Basis Function with Shape Parameter 1e-09.



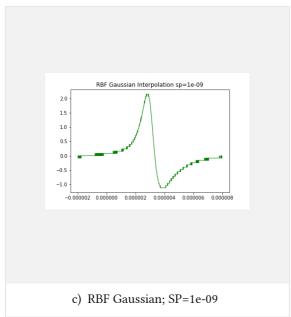


Fig 18 Cross Section 5 Interpolation using Inverse and Gaussian Basis Function with different Shape Parameter.

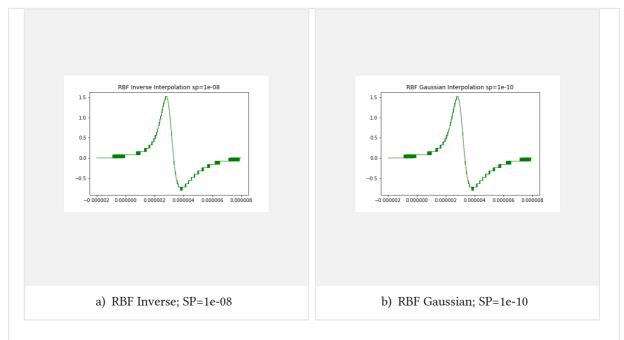
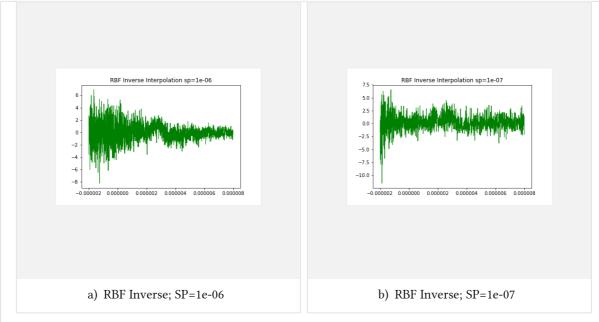
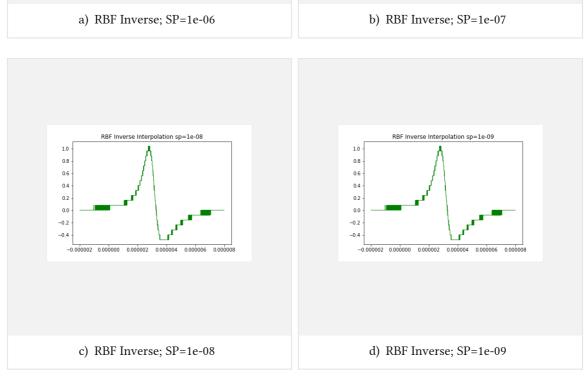
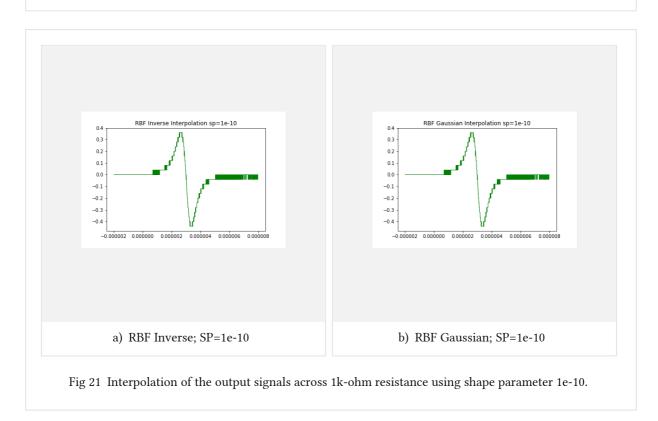


Fig 19 Cross Section 6 Interpolation using Inverse and Gaussian Basis Function with different Shape Parameter.









It is seen that the error and the roughness increases as the shape parameter is increased. In a RBF, therefore, shape parameter can be considered a key part of interpolation to formulate a continuous function for the distribution. Through iterative method, it is found that shape parameters around 1e-09 or smaller gives smooth curves with minimum obvious errors. The

expected interpolation is thus obtained using basis functions (Inverse and Gaussian here) with shape parameter less than 1e-08.

#### 5 CONCLUSION AND RECOMMENDATIONS

In this project, few key concepts of engineering and mathematics were encountered. The heterogeneous distribution of moisture content after curing of concrete mortar directly influences its durability and strength. This quandary was tackled using few engineering and mathematical tools. The concept of RLC and RBF was applied to derive a standard method to measure and plot the heterogeneous distribution of moisture in the concrete block. The interpolated functions provide several advantages. It can help design better structure with

more strength and durability. A standard method is also helpful for designers and engineers to operate and use it whenever required with accuracy. The scope of this paper is limited to concrete mortar and simple beams. Further research can be undertaken to measure water content for Reinforced concrete structures which will provide a holistic approach to the method used to plot water content distribution.

#### **REFERENCES**

- [1] Radial Basis Functions, 15-486/782: Artificial Neural Networks, David S. Touretzky, Fall 2006.
- [2] Introduction to Radial Basis Function Networks, Mark J. L. Orr, Centre for Cognitive Science, University of Edinburgh, 2, Buccleuch Place, Edinburgh EH8 9LW, Scotland, April 1996.
- [3] Choosing Basis Functions and Shape Parameters for Radial Basis Function Methods, Michael Mongillo, October 25, 2011.
- [4] On the selection of the most adequate radial basis function, Humberto Rocha, Portuguese Catholic University, Polo de Viseu, 3504-505 Viseu, Portugal, 10 March 2008.
- [5] Lectures on Radial Basis Functions, Stefano De Marchi, Department of Mathematics "Tullio Levi-Civita", University of Padua (Italy), February 13, 2018.
- [6] An Efficient k-Means Clustering Algorithm: Analysis and Implementation Tapas Kanungo, Senior Member, IEEE, David M. Mount, Member, IEEE, Nathan S. Netanyahu, Member, IEEE, Christine D. Piatko, Ruth Silverman, and Angela Y. Wu, Senior Member, IEEE.
- [7] The use of MRI to observe the structure of concrete, E. Marfisi\*, C. J. Burgoyne\*, M. H. G. Amin† and L. D. Hall†, University of Cambridge

#### **ACKNOWLEDGEMENTS**

For the completion of this paper, I would like express my sincere gratitude to Indian Academy of Sciences (IASc-INSA) for offering me this opportunity to undergo the project. I am deeply indebted and grateful to respected Dr. Bishwajit Bhattarcharjee, Professor, Department of Civil Engineering, Indian Institute of Technology, Delhi, for his critical supervision, crucial inputs and guidance throughout the course of the project.

I would also like thank Mr. Gopinandan Dey, Research Fellow at IIT Delhi, for his distinctive guidance and encouragement in the course of the project.

I would like to extend my special thanks to Mr. Kartikay Kaushik, who has worked as a research fellow alongside me in this project, for his remarkable suggestions and help during the project.

I would like to express my deepest sense of gratitude towards my parents and my friends who have always supported and encouraged me in my life.

#### **APPENDICES**

#### Code for cosine function

```
import numpy as np
from scipy import interpolate
import matplotlib.pyplot as plt
x = \text{np.linspace} (0,10,9)
y = np.cos(x)
xi = np.linspace (0,10,1001)
yi = np.cos(xi)
f = interpolate.Rbf(x,y, function = 'basis function', epsilon = \gamma)
si = f(xi)
plt.plot(x,y,'bo')
plt.plot(xi,yi,color = 'red', linewidth = '0.5')
```

```
plt.plot(xi,si,color = 'green', linewidth = '0.5')  
plt.title('RBF Inverse Interpolation sp = \gamma)  
plt.savefig('rbf_basisfunction_basisfunction_4.png')
```

#### Code for interpolation of data points

```
import numpy as np
from scipy import interpolate
import pandas as pd
import matplotlib.pyplot as plt
matplotlib inline
data= pd.read csv("/home/desktop/Summer Intern/savefile.CSV")
x=data.iloc[:,3].values
y =data.iloc[:,4].values
f=interpolate.Rbf(x,y,function='basisfucntion', epsilon=\gamma)
si = f(x)
plt.plot(x,y,'bo')
plt.plot(x,si,color='green', linewidth='1.0')
```

plt.title('RBF Basis Function sp=  $\gamma$ )

plt.savefig('savefile.png')