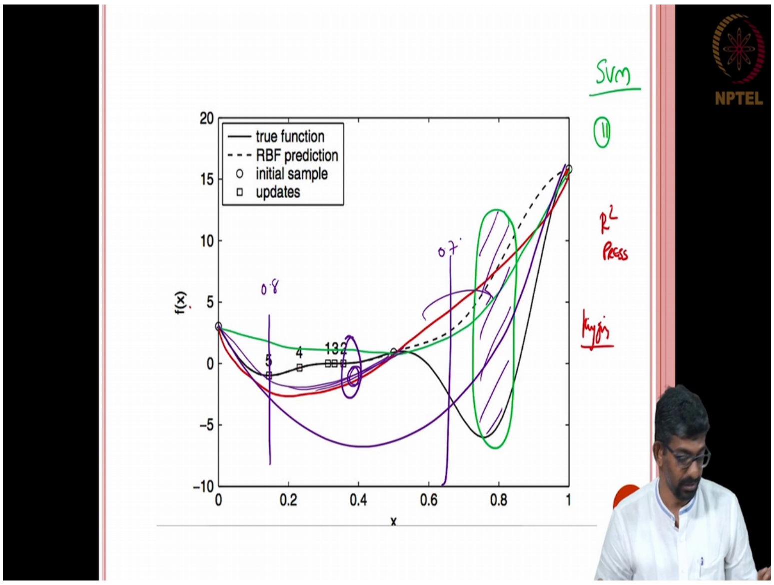
**Surrogates and Approximations in Engineering Design**
**Prof. Palaniappan Ramu**
**Department of Engineering Design**
**Indian Institute of Technology, Madras**

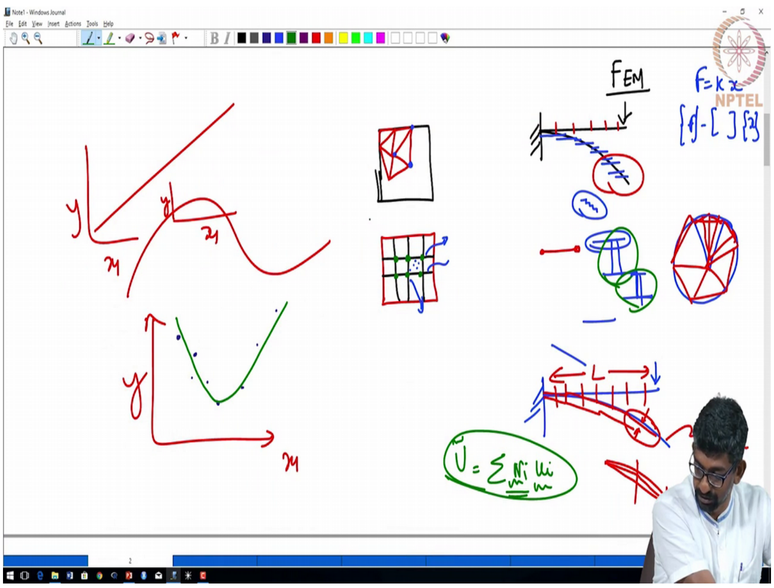
**Lecture - 15**
**Kriging – 2**

(Refer Slide Time: 00:17)

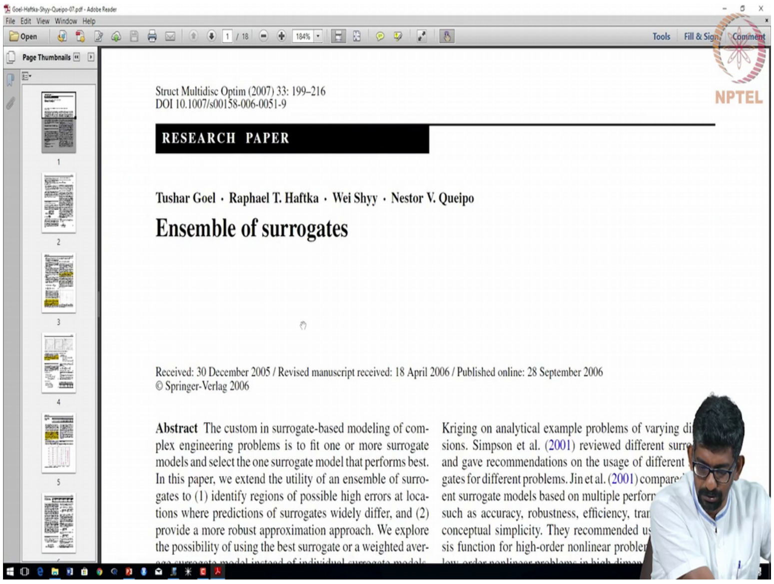



So, in this what I suggest is, I am going to show you a paper that I recommend that you read especially, if you are interested in researching this area.

(Refer Slide Time: 00:23)

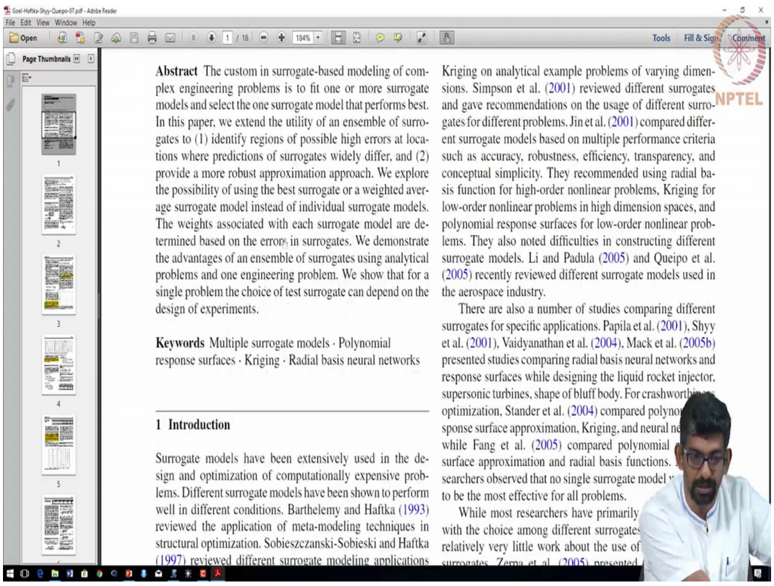


(Refer Slide Time: 00:26)



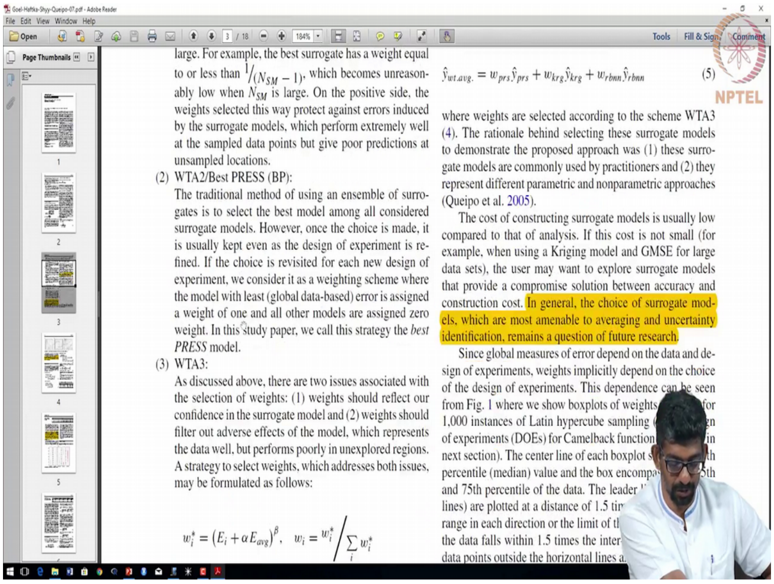

So, this was published in 2007 in the journal of structural multidisciplinary optimization. It was published by Tushar Goel, Raphael Haftka and a couple of other people. Haftka is kind of a guru of surrogates ok.

(Refer Slide Time: 00:43)



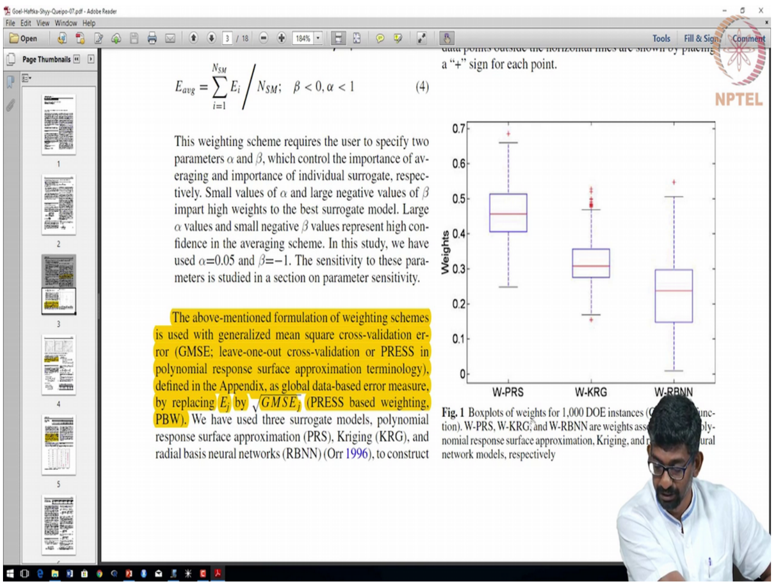

And I guess Abdul Samad also has worked with Haftka. And so you can just take very simple ensemble of surrogates. You can see it has received lot of citations also from some 1000 citations it has received. So, what they are talking about is they will build a weighted average surrogate ok. This w i is a weight for the surrogate that you are talking about. This is again a linear sum if you look at it ok, too many linear sums.

(Refer Slide Time: 01:16)




And there are different schemes that they talk about in terms of the weights. They just generally say E and then at a later point is error they say it is better that you use a presser ok, but you can use any error that is what; that is why they generalized this as E j, but later in the paper they recommend using the presser for doing that. And then they say that you use the best PRESS that is one model and the other one is the weighted average.

(Refer Slide Time: 01:43)

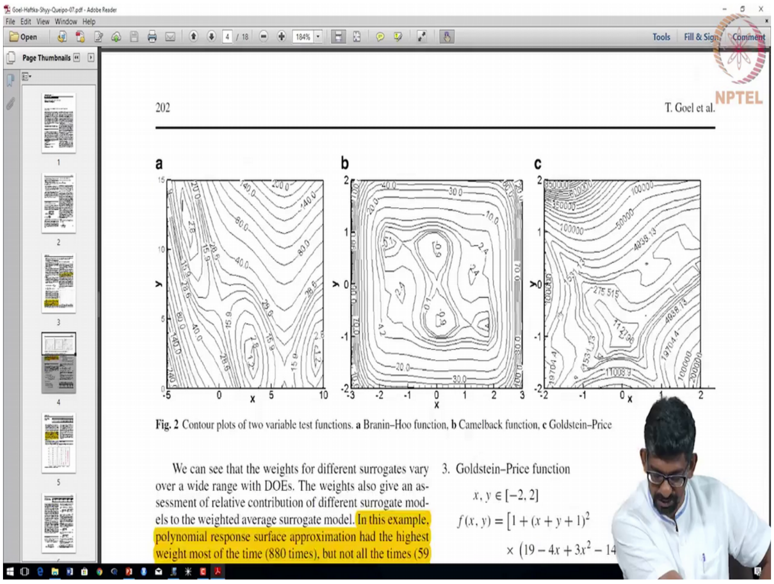


The WTA3 is the weighted average that they say use it this way E i plus alpha E average times beta. The idea is you need to tune your alpha and beta accordingly which was done in a subsequent paper by another author called head Marg Azad.

But here they are taking some beta to be less than 0 and alpha is less than 1. They are taking some specific values of alpha and beta to study. Just to give you an idea; for instance, what they did is they took some function called the camelback function, not the camelback the Branin hoo function. And what they are doing is; they are running this polynomial response Kriging radial basis 1000 times and they checked which model work the best. This is the error I mean this is our weight metric. So, the weights will be given by one of these schemes.

And what they figured out is there was no metamodel that was a runaway winner, you understand what I am saying?

(Refer Slide Time: 02:51)

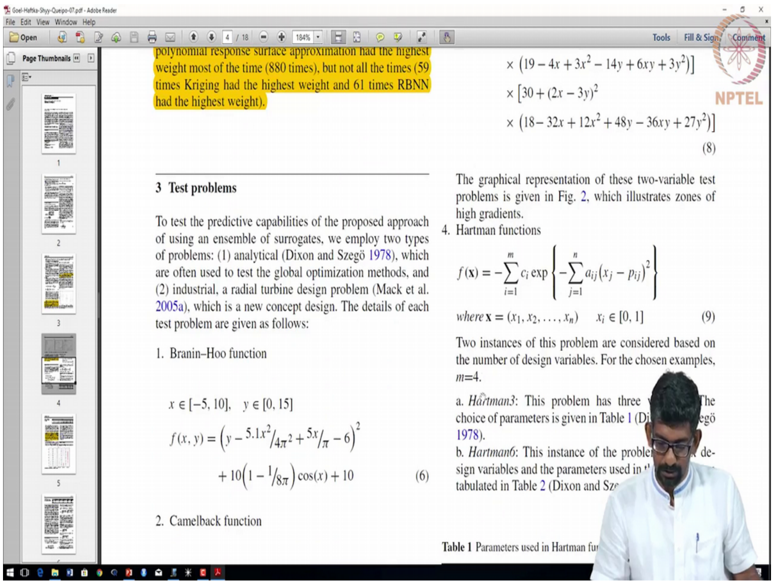



There is some numbers I have captured, PRS had the highest weight most of the time 880 times out of 1000, who knows the one sample that you took could have been that 880 fifth or where PRS was not the best fit.

This is why that random simulations are important, because it is based on DOEs which is again random. So, out of the 1000 DOEs is they did the 880 times PRS was better and then Kriging came only 61 time sorry 59 time well RBRB of took 61 times. This is for a known function, you cannot generate a meaning like now today I do I might get 80, and I might get only 5 times RBF game and then see the remaining times Kriging was better. Because the 1000 DOE is that they created and I created could be different, and this was in Latin hypercube.

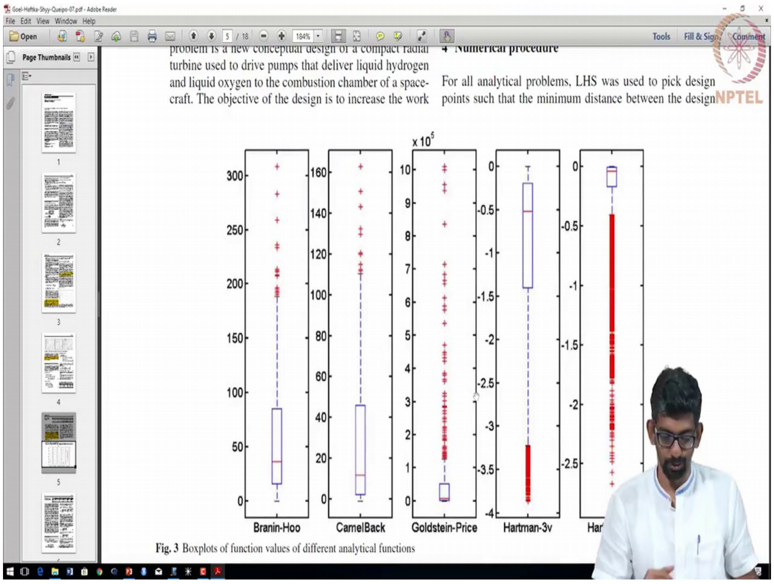
Let us say that you do hammers 3 sequence, you might entirely get a different stuff even these numbers will not be. So, this is a simple x 1, x 2, 2 dimensional problem that we are talking about. So, this is the whole idea is they say that you know no single metamodel is going on; unless let us say that I know this function and I build these approximation over the years, which is what you have what they call subject matter experts. In companies when you go automobile companies aerospace company they have subject matter expert they have very good understanding. In those cases, you know what is a function to be fitted and you can use it, but that is not the case here.

(Refer Slide Time: 04:31)



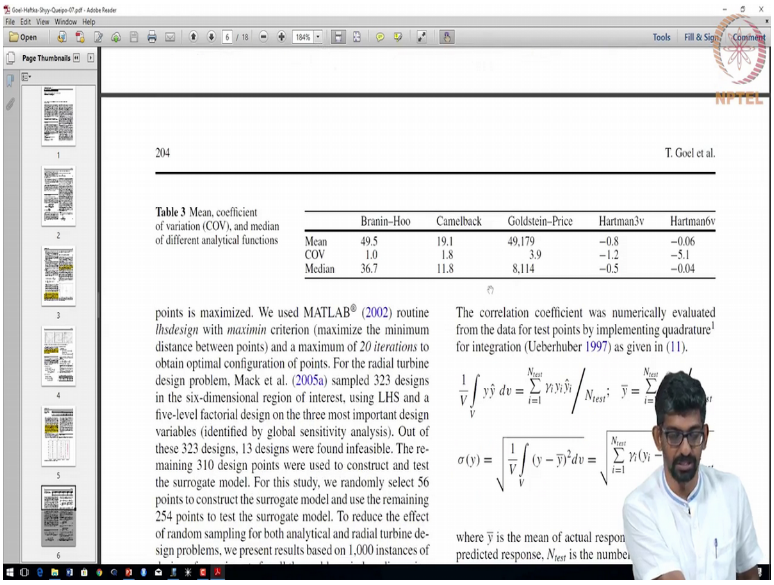

So, then they also give some 6 different test problems of Branin hoo, camelback, Hartman functions.

(Refer Slide Time: 04:40)



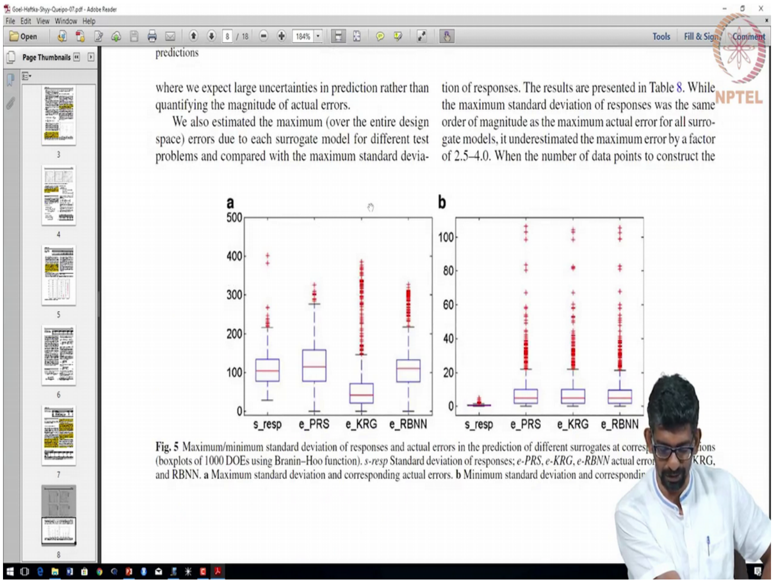

And then it is interestingly what they do is they also show you the variation of the functions themselves. They evaluate the function at these 1000 DOEs for each of them. And then they show you how the Branin hoo function varies it can vary anywhere the value can vary anywhere between 0 to 300 ok.

(Refer Slide Time: 05:03)




So, and then they just show further different functions ok, without any normalization they are trying to run this stuff. And then they are discussing about the prediction matrix, as I pointed out they will use correlation coefficient between the input and the output sorry between the actual and the predicted. They do an RMS error, RMSE we discuss that and the maximum error, you can interesting where are they using the PRESS is discussed this is only for the matrix that they are used. So, this would have been good if it is colour, but it is ok.

(Refer Slide Time: 05:45)




So, what they are doing is; the way they are plotting this one is they are taking a is maximum, they are looking for maximum errors, standard deviation of errors in the prediction.

So, they start with about 12 samples I guess, 20 samples 21 by 21 grid is what they are testamentary is they use about 12 samples for fitting, with the 12 samples for fitting they do 1000 times they repeat this procedure. And then they say that; this is the standard deviation of the responses with respect to the actual errors ok.

If you look at it the maximum standard deviation and then you can see that; you know Kriging performed slightly better than the other 2 guys in terms of the errors, but then you see there are a lot of outliers in Kriging compared to PRS and radial basis.

Similarly, whereas, in this guide where whichever regions the error was minimal you can see that each of them all of them performed very, very similar. There is no variation that is what is captured there is no variation between these performance that is why this; error is the least. Whereas, in this there was error maximum deviation in the function evaluations and then each one predicted something else ok.

So, this is what I meant; whenever there is maximum variation in the predictions it means that there is uncertainty in the design space itself. So, you need more samples to understand what happens there. Whereas, in this case it so happened ok, but please understand that this being more or less the same does not mean that your prediction is good, you might totally be off also ok, but this is guaranteed.

(Refer Slide Time: 07:40)

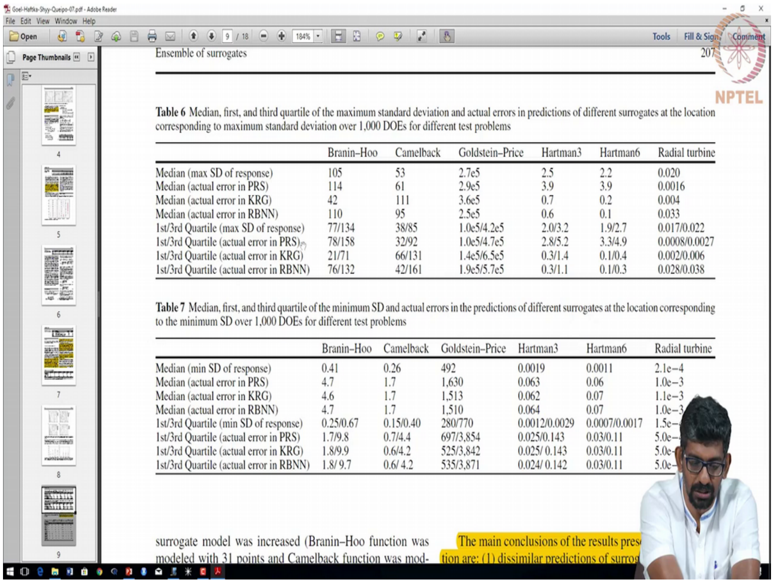

Ensemble of surrogates 207

**Table 6** Median, first, and third quartile of the maximum standard deviation and actual errors in predictions of different surrogates at the location corresponding to maximum standard deviation over 1,000 DOEs for different test problems

| | Branin–Hoo | Camelback | Goldstein–Price | Hartman3 | Hartman6 | Radial turbine |
|---|---|---|---|---|---|---|
| Median (max SD of response) | 105 | 53 | 2.7e5 | 2.5 | 2.2 | 0.020 |
| Median (actual error in PRS) | 114 | 61 | 2.9e5 | 3.9 | 3.9 | 0.0016 |
| Median (actual error in KRG) | 42 | 111 | 3.6e5 | 0.7 | 0.2 | 0.004 |
| Median (actual error in RBNN) | 110 | 95 | 2.5e5 | 0.6 | 0.1 | 0.033 |
| 1st/3rd Quartile (max SD of response) | 77/134 | 38/85 | 1.0e5/4.2e5 | 2.0/3.2 | 1.9/2.7 | 0.017/0.022 |
| 1st/3rd Quartile (actual error in PRS) | 78/158 | 32/92 | 1.0e5/4.7e5 | 2.8/5.2 | 3.3/4.9 | 0.0008/0.0027 |
| 1st/3rd Quartile (actual error in KRG) | 21/71 | 66/131 | 1.4e5/6.5e5 | 0.3/1.4 | 0.1/0.4 | 0.002/0.006 |
| 1st/3rd Quartile (actual error in RBNN) | 76/132 | 42/161 | 1.9e5/5.7e5 | 0.3/1.1 | 0.1/0.3 | 0.028/0.038 |

**Table 7** Median, first, and third quartile of the minimum SD and actual errors in the predictions of different surrogates at the location corresponding to the minimum SD over 1,000 DOEs for different test problems

| | Branin–Hoo | Camelback | Goldstein–Price | Hartman3 | Hartman6 | Radial turbine |
|---|---|---|---|---|---|---|
| Median (min SD of response) | 0.41 | 0.26 | 492 | 0.0019 | 0.0011 | 2.1e−4 |
| Median (actual error in PRS) | 4.7 | 1.7 | 1,630 | 0.063 | 0.06 | 1.0e−3 |
| Median (actual error in KRG) | 4.6 | 1.7 | 1,513 | 0.062 | 0.07 | 1.1e−3 |
| Median (actual error in RBNN) | 4.7 | 1.7 | 1,510 | 0.064 | 0.07 | 1.0e−3 |
| 1st/3rd Quartile (min SD of response) | 0.25/0.67 | 0.15/0.40 | 280/770 | 0.0012/0.0029 | 0.0007/0.0017 | 1.5e− |
| 1st/3rd Quartile (actual error in PRS) | 1.7/9.8 | 0.7/4.4 | 697/3,854 | 0.025/0.143 | 0.03/0.11 | 5.0e− |
| 1st/3rd Quartile (actual error in KRG) | 1.8/9.9 | 0.6/4.2 | 525/3,842 | 0.025/ 0.143 | 0.03/0.11 | 5.0e− |
| 1st/3rd Quartile (actual error in RBNN) | 1.8/ 9.7 | 0.6/ 4.2 | 535/3,871 | 0.024/ 0.142 | 0.03/0.11 | 5.0e− |

surrogate model was increased (Branin–Hoo function was modeled with 31 points and Camelback function was mod-

The main conclusions of the results pres tion are: (1) dissimilar predictions of surrog

So, they test all these things they give you the median plots and all that and they also give you the actual plots here ok.

(Refer Slide Time: 07:43)

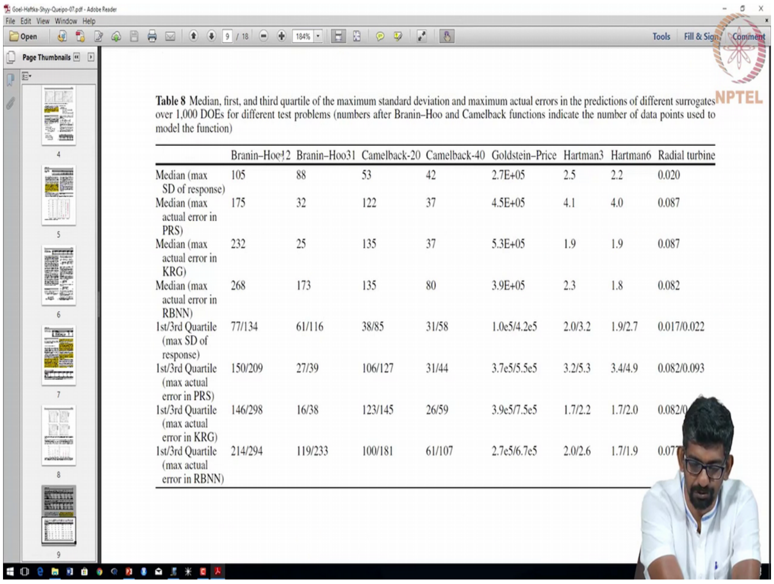
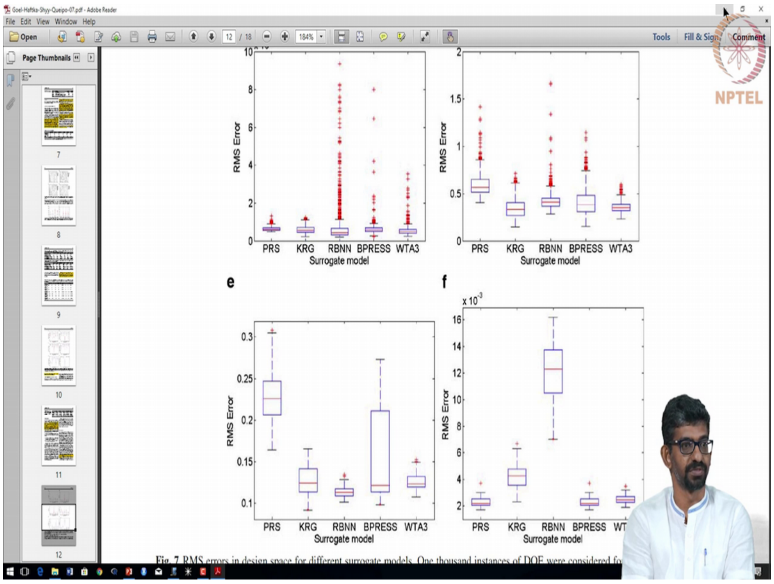


Table 8 Median, first, and third quartile of the maximum standard deviation and maximum actual errors in the predictions of different surrogates over 1,000 DOEs for different test problems (numbers after Branin–Hoo and Camelback functions indicate the number of data points used to model the function)

| | Branin–Hoo12 | Branin–Hoo31 | Camelback-20 | Camelback-40 | Goldstein–Price | Hartman3 | Hartman6 | Radial turbine |
|---|---|---|---|---|---|---|---|---|
| Median (max SD of response) | 105 | 88 | 53 | 42 | 2.7E+05 | 2.5 | 2.2 | 0.020 |
| Median (max actual error in PRS) | 175 | 32 | 122 | 37 | 4.5E+05 | 4.1 | 4.0 | 0.087 |
| Median (max actual error in KRG) | 232 | 25 | 135 | 37 | 5.3E+05 | 1.9 | 1.9 | 0.087 |
| Median (max actual error in RBNN) | 268 | 173 | 135 | 80 | 3.9E+05 | 2.3 | 1.8 | 0.082 |
| 1st/3rd Quartile (max SD of response) | 77/134 | 61/116 | 38/85 | 31/58 | 1.0e5/4.2e5 | 2.0/3.2 | 1.9/2.7 | 0.017/0.022 |
| 1st/3rd Quartile (max actual error in PRS) | 150/209 | 27/39 | 106/127 | 31/44 | 3.7e5/5.5e5 | 3.2/5.3 | 3.4/4.9 | 0.082/0.093 |
| 1st/3rd Quartile (max actual error in KRG) | 146/298 | 16/38 | 123/145 | 26/59 | 3.9e5/7.5e5 | 1.7/2.2 | 1.7/2.0 | 0.082/0 |
| 1st/3rd Quartile (max actual error in RBNN) | 214/294 | 119/233 | 100/181 | 61/107 | 2.7e5/6.7e5 | 2.0/2.6 | 1.7/1.9 | 0.07 |

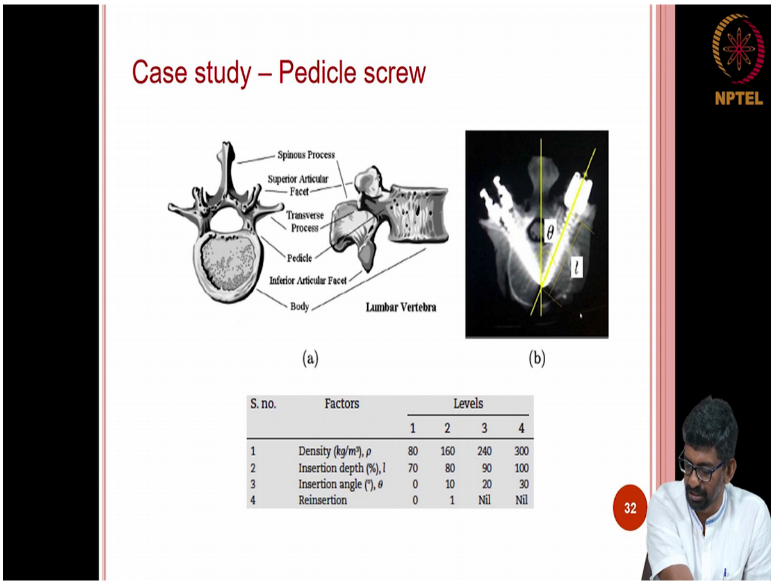
(Refer Slide Time: 07:44)




So, this is a response correlation that they are plotting, you can see how it is varying for each one of them see. I hope you understand a box plot; it is thousand repetitions I am just plotting each one of that the central line is a medial line, this is a 5th percentile, this is a 95th percentile and these are the outliers. It gives you a distribution also.

So, you can this is the interesting paper if you are looking at ensembles. And they suggest that you use a weighted average surrogate or you kind of use a weighted ensemble, unless you have some information on what ensemble to be used in. So, this is one stuff paper that I recommended to you.

(Refer Slide Time: 08:29)




| S. no. | Factors | Levels | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 |
| 1 | Density ($kg/m^3$), $\rho$ | 80 | 160 | 240 | 300 |
| 2 | Insertion depth (%), $l$ | 70 | 80 | 90 | 100 |
| 3 | Insertion angle (°), $\theta$ | 0 | 10 | 20 | 30 |
| 4 | Reinsertion | 0 | 1 | Nil | Nil |

Just discuss a small case study that we did, with that I will wrap it up.

So, we try to apply this idea to a biomedical problem. So, one of the doctors that we work with in CMC Velu wanted to understand for a specific type of degenerative disease and this is called the osteoporotic bones ok. When you grow old the dominant in Indian males there is a condition called osteoporotic bone, which is degenerative.

Generally, your bones are supposed to be generated, but as you grow old they will lose some density and they will become degenerated. So, then what happens is you have some issues; your body weight and your bones needs to realign accordingly and all that, in such cases usually they put some and the bones also become weak so they might break.

So, under osteoporotic conditions when you do a fusion kind of or a graph you put something and then you plate it you screw it. It was not clear whether the regular number of screws that are used on a healthy bone is good enough for an osteoporotic condition also. So, they wanted to understand what is the pull-out strength ok, will this be good enough for it to hold it.

And as you see we really need human bones to test this, but it is not task ok. So, then we can source some caribou bones meaning; bones from the dead body, but that is also a very difficult right like male, that particular age, osteoporotic condition, people should be willing to give the bones specifically for the spinal cord.

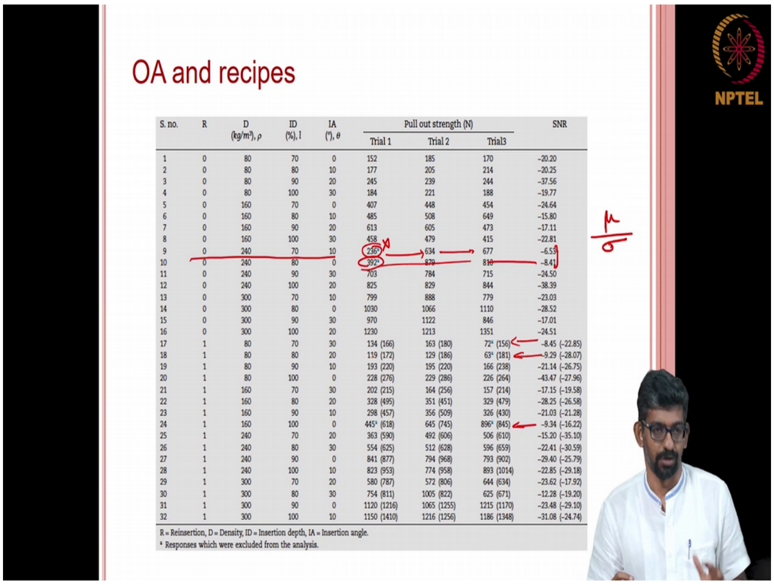
So, it was very expensiveness in that sense, you will have to wait infinitely no you might not be able to get. So, you will finally, after 3 years of wait we were able to get 6 caribou bones to do this study ok, that is all only 6 samples. So, that is the our high fidelity simulation then we use some low fidelity which is the FDA the Federal Drug Agency suggests that some kind of a foam which is equivalent ok.

I do not have the foam thing here there is a foam ok, which by changing the porosity in the foam you can represent the bones. So, they say you can whatever bone related stuff you can do it and this it is an approved test. So, that is large number of simulations that we can do. So, what we do is we mix this information and we build a metamodel. We wanted to give a pull-out strength calculator to be to the doctor.

So, the doctor has some information to begin with which is his input space, density insertion, depth insertion, angle reinsertion you can see what it means. So, this theta is a reinsertion sorry, the insertion angle and l is the insertion depth. Density is the bone density that we are talking about and reinsertion is what happens is they put the screw and then they understand that it is not go on to hold. So, they remove and then they put another screw in the same place, which is slightly longer.

But as you know if you have tried nailing something and removing the nail and then put another nail or a screw in the same spot, it is not going to have the same the hold power. It is not going to have the first time you put you want to put it the right time. So, this information if it is reinsertion means there is no reinsertion 0 means, 1 means there was one reinsertion. So, you can see there are different levels here this they took an orthogonal array to do this, that is a design of experiment these are the different levels.

(Refer Slide Time: 11:59)




| S. no. | R | D (kg/m³), ρ | ID (%), l | IA (°), θ | Pull out strength (N) Trial 1 | Trial 2 | Trial3 | SNR |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 80 | 70 | 0 | 152 | 185 | 170 | −20.20 |
| 2 | 0 | 80 | 80 | 10 | 177 | 205 | 214 | −20.25 |
| 3 | 0 | 80 | 90 | 20 | 245 | 239 | 244 | −37.56 |
| 4 | 0 | 80 | 100 | 30 | 184 | 221 | 188 | −19.77 |
| 5 | 0 | 160 | 70 | 0 | 407 | 448 | 454 | −24.64 |
| 6 | 0 | 160 | 80 | 10 | 485 | 508 | 649 | −15.80 |
| 7 | 0 | 160 | 90 | 20 | 613 | 605 | 473 | −17.11 |
| 8 | 0 | 160 | 100 | 30 | 458 | 479 | 415 | −22.81 |
| 9 | 0 | 240 | 70 | 10 | 236[a] | 634 | 677 | −6.53 |
| 10 | 0 | 240 | 80 | 0 | 392[a] | 879 | 810 | −8.41 |
| 11 | 0 | 240 | 90 | 30 | 703 | 784 | 715 | −24.50 |
| 12 | 0 | 240 | 100 | 20 | 825 | 829 | 844 | −38.39 |
| 13 | 0 | 300 | 70 | 10 | 799 | 888 | 779 | −23.03 |
| 14 | 0 | 300 | 80 | 0 | 1030 | 1066 | 1110 | −28.52 |
| 15 | 0 | 300 | 90 | 30 | 970 | 1122 | 846 | −17.01 |
| 16 | 0 | 300 | 100 | 20 | 1230 | 1213 | 1351 | −24.51 |
| 17 | 1 | 80 | 70 | 30 | 134 (166) | 163 (180) | 72[a] (156) | −8.45 (−22.85) |
| 18 | 1 | 80 | 80 | 20 | 119 (172) | 129 (186) | 63[a] (181) | −9.29 (−28.07) |
| 19 | 1 | 80 | 90 | 10 | 193 (220) | 195 (220) | 166 (238) | −21.14 (−26.75) |
| 20 | 1 | 80 | 100 | 0 | 228 (276) | 229 (286) | 226 (264) | −43.47 (−27.96) |
| 21 | 1 | 160 | 70 | 30 | 202 (215) | 164 (256) | 157 (214) | −17.15 (−19.58) |
| 22 | 1 | 160 | 80 | 20 | 328 (495) | 351 (451) | 329 (479) | −28.25 (−26.58) |
| 23 | 1 | 160 | 90 | 10 | 298 (457) | 356 (509) | 326 (430) | −21.03 (−21.28) |
| 24 | 1 | 160 | 100 | 0 | 445[a] (618) | 645 (745) | 896[a] (845) | −9.34 (−16.22) |
| 25 | 1 | 240 | 70 | 20 | 363 (590) | 492 (606) | 506 (610) | −15.20 (−35.10) |
| 26 | 1 | 240 | 80 | 30 | 554 (625) | 512 (628) | 596 (659) | −22.41 (−30.59) |
| 27 | 1 | 240 | 90 | 0 | 841 (877) | 794 (968) | 793 (902) | −29.40 (−25.79) |
| 28 | 1 | 240 | 100 | 10 | 823 (953) | 774 (958) | 893 (1014) | −22.85 (−29.18) |
| 29 | 1 | 300 | 70 | 20 | 580 (787) | 572 (806) | 644 (634) | −23.62 (−17.92) |
| 30 | 1 | 300 | 80 | 30 | 754 (811) | 1005 (822) | 625 (671) | −12.28 (−19.20) |
| 31 | 1 | 300 | 90 | 0 | 1120 (1216) | 1065 (1255) | 1215 (1170) | −23.48 (−29.10) |
| 32 | 1 | 300 | 100 | 10 | 1150 (1410) | 1216 (1256) | 1186 (1348) | −31.08 (−24.74) |

R = Reinsertion, D = Density, ID = Insertion depth, IA = Insertion angle.
[a] Responses which were excluded from the analysis.

So, we have done about 32 experiments with the foam are these are the different input parameters, here is a pull-out strength ok. Interestingly, this is an experiment it is not a computer experiment. So, for this experimental set up the first one, when I repeated 3 times I get 3 different values, you understand?

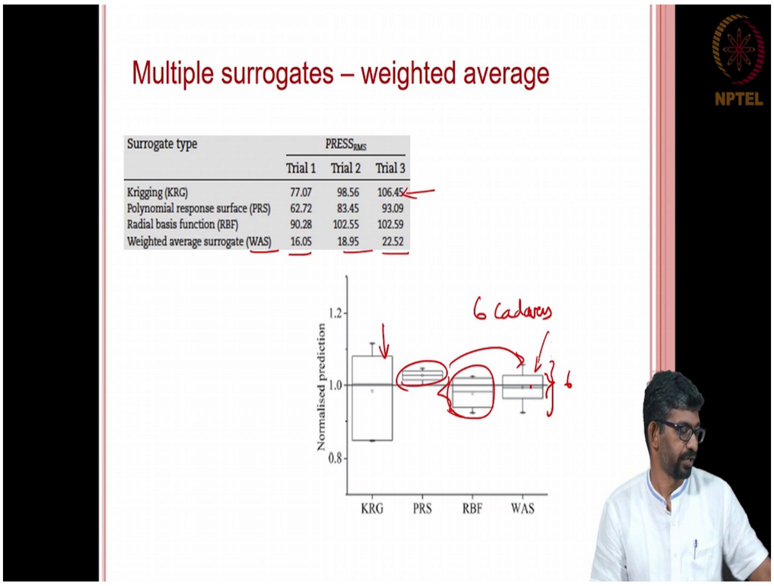
So, which also tells us the foam captures the bone nature I take 18 years old male bone very similar structure, I use another person's bone it will give me 2 different pull out strength that is exactly what this is given ok. So, there should be variability, which is what we have done. And what we did is we use something called an SN ratio Signal to Noise ratio for identifying.

See if you see I do not know whether you are able to see there is a small dot here. We can see 236, 634 and 677 are the other ones other to test whereas, this test gave 236 for this configuration.

So, we know that this guy is an outlier, but then here we can visually do, but when you are giving it to the doctor to do they cannot go and do all these things. So, what we do is we create an SN ratio, SN ratio is signal which is the mean of these information divided by the standard deviation signal to noise. And in this particular case you want the signal to noise to be meaning, your noise should be less then this over all thing will be.

So, if this noise is more you this ratio will be less. So, wherever you get this value to be lesser, then they are all issue prone guys you can see that these were all. Wherever there were less than 10 let us say you put a number on 10 then. So, that is one way of filtering the data; then what he did is he took all these pull out strengths and fitted a metamodel this is what he has done.

(Refer Slide Time: 14:11)




Finally compared it with 6 Cadaver bones predictions so, here is a point PRESS RMS errors with Kriging trial 1 trial 2 trial 3 so 3 respond surfaces ok. Similarly, polynomial respond the R B F and weighted average surrogate. So, you can look at the PRESS error, the weighted average PRESS error was far better than other guys because you want the minimum error, 0 error means that is the best fit.

So, you can see the weighted average had the least error compared to any individual surrogate. Weighted averages you weight and take an average or just take this output this output this output add meaning Kriging PRS RBF and then you average them. What this plot gives this we plot the variations with the respect to the 6 cadaver bones that we tested ok.

So, this is for about 6 what we do is we give the inputs and then we ask our pullout calculator to give out what the pullout strength is. So, was gave a different RBF gave different, PRL gave different, Kriging gave different. We compared it with the actual value from that cadaver and then we take a ratio of that. So, if it is one then my prediction is very close to the actual value.

So, as you can see in this particular stuff Kriging gave a lot of variation. PRS had the least variation compared to even the weighted average surrogate, but then it was way off from the ideal line. And this guy was ok, but he does not have a what we call symmetric distribution this had the median very close to the 1 and then it also had a symmetric distribution.

So, weighted average surrogate was successfully used in this case to give a pullout strength calculator. And currently this is in use basically in a qualitative sense the doctor uses this to understand what is the pull out string and then they make decisions on should they put 2 screws 3 screws or should they use what should be the depth of insertion, accordingly they will choose the pedicle screw to do that ok.

Because pedicle screws are like your shoe sizes ok there are different 2 3 sizes are there they want to design a priori and unless required you do not want to screw further, always they can do a worst case they can screw you know to the deepest, but you do not want to do that you do not want to disturb the nature stuff so it ok. So, with that I guess I am going to wrap this up; unless you have specific questions. If you have specific questions I will take it now, you have any questions? In general, fine.