Study ID	Estimation Technique	Accuracy Measure Used	Estimation Accuracy Achieved	Notes (If Any)
S2	• Least squares regression	NA	NA	• The aim of this paper was to construct a cost estimation model that took into account the "learning factor" involved in getting to grips with an OO framework. • LOCs and OOFPs were used as size metrics. • Least-squares regression was used to generate the learning rate parameter using data from the 5 web-based applications.
S3	NA	NA	NA	• This paper looks at validating metrics proposed to measure Web development effort. • It does not propose or validate an effort prediction model.
S4	• CBR/Analogy	• MMRE • Pred(25)	• MMRE of 52.23% for the LEL group • MMRE of 29.15% for the LEL-G1 dataset. • MMRE of 28% for the LEL-G2 dataset. • MMRE of 40.68% for the HEL group • MMRE of 23.34% for the HEL-G1 dataset. • MMRE of 19.76% for the HEL-G2 dataset.	• Dataset divided into 2 groups: ○ LEL group (low experience) with 41 projects. ○ HEL group (high experience) with 29 projects. • Both LEL and HEL groups later subdivided into 2 homogenous groups:

				○ LEL-G1 (22 projects) and LEL-G2 (19 projects) ○ HEL-G1 (15 projects) and HEL-G2 (14 projects). • No single best CBR configuration: dependent upon dataset.
S5	• CBR/Analogy	NA	NA	• Actual accuracy achieved not explicitly mentioned (nor was the measure of accuracy used).
S6	• Linear regression • Multiple stepwise regression	• Boxplot of residuals	NA	• Stepwise regression underestimated effort whereas linear regression overestimated it. • Stepwise regression showed slightly better results than linear regression.
S7	• CBR/Analogy • Linear regression • Stepwise regression	• MMRE • MdMRE	**LEL Group** • Overall MMRE of 52.23% • LEL-G1 ○ MMRE of 29.15% ○ MdMRE of 11.20% • LEL-G2 ○ MMRE of 28% ○ MdMRE of 16.26% **HEL Group** • Overall MMRE of 40.68% • HEL-G1	• Dataset divided into 2 groups: ○ LEL group (low experience) with 41 projects. ○ HEL group (high experience) with 29 projects. • Both LEL and HEL groups later subdivided into 2 homogenous groups: ○ LEL-G1 (22 projects) and LEL-G2 (19 projects)

			○ MMRE of 23.34% ○ MdMRE of 18.21% • HEL-G2 ○ MMRE of 19.76% ○ MdMRE of 11.51% **Linear Regression** • LEL-G1 ○ MMRE of 49.19% ○ MdMRE of 45.74% • LEL-G2 ○ MMRE of 99.32% ○ MdMRE of 100% • HEL-G1 ○ MMRE of 64.3% ○ MdMRE of 45.07% • HEL-G2 ○ MMRE of 45.89% ○ MdMRE of 54.95% **Stepwise Regression** • LEL-G1 ○ MMRE of 50.83% ○ MdMRE of 36.39% • LEL-G2 ○ MMRE of 55.83% ○ MdMRE of 65.12% • HEL-G1 ○ MMRE of 92.5% ○ MdMRE of 93.04% • HEL-G2 ○ MMRE of 43.8% ○ MdMRE of 43.62	○ HEL-G1 (15 projects) and HEL-G2 (14 projects). • Stepwise regression not consistently better than linear regression. • Estimation by analogy outperforms both regression techniques. • In terms of MMRE, estimation by analogy significantly better than stepwise and linear regression for all datasets apart from LEL-G1. • In terms of MMRE, estimation using stepwise regression was only significantly better than linear regression for dataset LEL-G2. No statistical difference otherwise. • In terms of MdMRE, estimation by analogy statistically superior to both regression techniques for all datasets. • In terms of MdMRE, estimation using stepwise regression was only significantly better than linear regression for dataset LEL-G2. No statistical difference otherwise. • No single best CBR

				configuration: dependent upon dataset.
S8	• CBR/Analogy	• Percentage error between actual and estimated effort	• Percentage error varied from 0% to 33.33%. • Average percentage error of 9.21%	
S9	• Linear regression • Stepwise regression	• Boxplot of residuals	NA	• None of the models produced reasonably accurate estimates of effort. • There was no statistical difference between the various models (based on absolute residuals) regardless of regression technique used or size metric considered.
S10	• Linear regression • Multiple stepwise regression	• MMRE	• Linear regression had a MMRE of 75%. • Multiple stepwise regression had a MMRE of 56%.	
S11	• Linear regression • Multiple stepwise regression	• MMRE	**Linear regression** • Application – 75% • Page – 105% • Media not reused – 34% • Media reused – 110% • Program not reused – 17% • Program reused – 53% **Stepwise multiple regression** • Application – 56% • Page – 108% • Media not reused – 51% • Media reused – 110%	

			• Program not reused – 24% • Program reused – 48%	
S14	• CBR/Analogy	• MMRE • MdMRE • Pred(25)	**Case Study 1** • Weighted Euclidean distance provided the best accuracy when used with 1 analogy. MMRE = 0.1, MdMRE = 0.09, and Pred(25) = 94.12. • Maximum measure provided the worst result regardless of the number of analogies/adaptation techniques used. MMRE from 0.23 to 0.32, MdMRE from 0.15 to 0.34, and Pred(25) from 26.47 to 76.47. • For Unweighted Euclidean, MMRE from 0.12 to 0.15, MdMRE from 0.10 to 0.12, and Pred(25) from 76.47 to 88.24. **Case Study 2** • Weighted Euclidean distance provided the best accuracy when used with 3 analogies and Mean adaptation. MMRE = 0.08, MdMRE = 0.08, Pred(25) = 70.59. • Maximum measure provided the worst result regardless of the number of analogies/adaptation techniques used. MMRE from 0.64 to 0.88, MdMRE from 0.54 to 0.83, and Pred(25) from 5.88 to 26.47. • For Unweighted Euclidean, MMRE from 0.14 to 0.22, MdMRE from 0.11 to 0.18, and Pred(25) from 44.12 to 55.88.	• 2 case studies (34 applications and 25 applications respectively). • Overall estimates better for case study 1 than 2.
S15	• Linear regression • Stepwise regression	• Boxplot of residuals	NA	• None of the models produced reasonably accurate estimates of effort. • There was no statistical difference between the various models regardless of regression technique used or size metric

				considered.
S16	• CBR/Analogy	• MMRE • MdMRE • Pred(25) • Boxplot of residuals	**Requirements and Design** • Unweighted Euclidean distance using the mean of the closest 2 analogies gave the best results (MMRE=0.07, MdMRE = 0.08, Pred(25) = 70.59). • In general predictions obtained using the unweighted Euclidean distance (MMRE from 0.07 – 0.12, MdMRE from 0.08 – 0.09, Pred(25) from 61.76% – 70.59%) were more accurate than those generated using the weighted Euclidean distance (MMRE from 0.08 – 0.15, MdMRE from 0.08 – 0.15, Pred(25) from 55.88% - 67.65%). Also corroborated by boxplot of residuals. • Differences found to statistically significant (at 5%). **Implementation** • Weighted Euclidean distance using the mean of the closest 3 analogies gave the best results (MMRE = 0.08, MdMRE = 0.08, Pred(25) = 70.59%). • In general predictions obtained using the weighted Euclidean distance (MMRE from 0.08 – 0.13, MdMRE from 0.08 – 0.11, Pred(25) from 55.88% - 70.59%) were more accurate than those using the unweighted Euclidean distance. (MMRE from 0.14 – 0.22, MdMRE from 0.11 – 0.18, Pred(25) from 44.12% - 55.88%). Also corroborated by boxplot of residuals. • Differences were statistically significant (at 5%).	• Application Measures did not give better prediction accuracy in any comparison and in some cases generated predictions significantly worse than those using Requirements and Design measures.
S17	• Case Based Reasoning • Multiple Linear Regression • Stepwise	• MMRE • MdMRE • Pred(25) • Boxplots of residuals	**Unweighted Euclidean** • MMRE from 0.11 to 0.15, MdMRE from 0.08 to 0.12, Pred(25) from 76.47% to 91.18%. **Weighted Euclidean**	• For evaluation of CBR configurations, leave 1 out cross-validation used. • For comparisons of techniques 2 types of cross-

	Regression • Classification and Regression Trees		• MMRE from 0.10 to 0.14, MdMRE from 0.09 to 0.12, Pred(25) from 82.35% to 97.06%. **Maximum Distance** • MMRE from 0.23 to 0.32, MdMRE from 0.15 to 0.34, and Pred(25) from 26.47% to 76.47%. • Best overall CBR setup was Weighted Euclidean distance, with 1 analogy, Standardized variables, and Mean adaptation with a MMRE of 0.10, MdMRE of 0.09 and Pred(25) of 94.12. This was then compared with other techniques. **Best CBR** • Split 1: MMRE from 0.09 to 0.16, MdMRE from 0.06 to 0.15, Pred(25) of 63.63% to 90.91%. • Split 2: MMRE from 0.10 to 0.18, MdMRE from 0.09 to 0.15, Pred(25) of 80%. **Linear/Stepwise regression** • Split 1: MMRE from 0.03 to 0.04, MdMRE from 0.02 to 0.04, Pred(25) of 100%. • Split 2: MMRE from 0.02 to 0.05, MdMRE from 0.01 to 0.04, Pred(25) of 100%. **CART** • Split 1: MMRE from 0.13 to 0.22, MdMRE from 0.08 to 0.11, Pred(25) from 81.82% to 90.91%. • Split 2: MMRE from 0.10 to 0.23, MdMRE from 0.07 to 0.10, Pred(25) of 80%.	validation was used. • 3 instances of a 66% split (SP1) and 3 instances of a 86% split (SP2). • Weighted Euclidean showed significantly better results than Unweighted Euclidean when comparing MMRE and paired MRE ($\alpha = 0.01$), but not for MdMRE. Corroborated by boxplot of residuals. • Final models for multiple linear regression and stepwise regression were the same. • Looking at MMRE and MdMRE, Multiple linear and Stepwise regression gave the best prediction accuracy. • T-tests for MMREs and Pred(25) and Mann-Whitney tests for MdMREs showed that regression models generally gave statistically significantly better results than either CBR or CART. • Boxplots of residuals showed that CART gave the worst results. • Boxplots of residuals showed that CBR had better prediction accuracy than the regression models contrary

				to what the MMRE and MdMRE results show.
S19	NA	NA	NA	• Ordinary Least Squares regression used to investigate relationships between dependent variables (information effort, navigation effort, and total design effort) and independent variables • Estimation model not built or evaluated

S21	• CBR
	• Stepwise Regression
	• Multiple Linear Regression
		• MMRE
		• MdMRE
		• Pred(25)
		• Boxplot of residuals
			Adaptation rule not used
			• Unweighted Euclidean: MMRE from 12% - 15%, MdMRE from 10% - 12%, and Pred(25) from 76.47% - 88.24%.
			    ○ Best estimate from 1 analogy: MMRE = 12%, MdMRE = 10%, Pred(25) = 88.24%.
			• Weighted Euclidean with subjective weights: MMRE 10% - 14%, MdMRE 9% - 12%, and Pred(25) 82.35% - 97.06%.
			    ○ Best estimate from 1 analogy: MMRE = 10%, MdMRE = 9%, and Pred(25) = 94.12%.
			• Weighted Euclidean with Pearson's correlation coefficient weights: MMRE 11% - 15%, MdMRE 9% - 13%, and Pred(25) 88.24% - 97.06%.
			    ○ Best estimate from inverse rank weighted mean with 3 analogies: MMRE = 12%, MdMRE = 10%, Pred(25) = 97.06%.
			• Maximum distance: MMRE 23% - 32%, MdMRE 15% - 34%, 26.47% - 76.47%.
			    ○ Best estimate from inverse rank weighted mean with 3 analogies: MMRE = 23%, MdMRE = 16%, Pred(25) = 67.65%.
			
			With adaptation rule
				• Comparison between regression models and best CBR configuration handled by 20 fold cross validation with 66% split.
				• When considering Unweighted and Weighted Euclidean distance measures, not using the adaptation rule generally produced the better result.
				• When considering the Maximum distance measure, using adaptation rules generally produced a better result.

			• Unweighted Euclidean, MMRE 21% - 32%, MdMRE 12% - 19%, Pred(25) = 67.65% - 76.47%.
			  ○ Best result from Median adaptation with 3 analogies: MMRE = 21%, MdMRE = 12%, Pred(25) = 76.47%.
			• Weighted Euclidean with subjective weights: MMRE 21% - 32%, MdMRE 12% - 23%, Pred(25) 55.88% - 76.47%.
			  ○ Best estimate from 1 analogy: MMRE = 21%, MdMRE = 12%, and Pred(25) = 76.47%.
			• Weighted Euclidean with Pearson's correlation coefficient weights: MMRE 24% - 33%, MdMRE 9% - 19%, Pred(25) 58.82% - 73.53%.
			  ○ Best estimate from 1 analogy: MMRE = 24%, MdMRE = 9%, Pred(25) = 73.53%.
			• Maximum distance, MMRE 14% – 20%, MdMRE 10% - 18%, Pred (25) 67.65% - 85.29%.
			  ○ Best estimate from Mean of 3 analogies: MMRE = 14%, MdMRE = 10%, Pred(25) = 85.29%.
			• Best overall CBR configuration when adaptation rule was not used was with Weighted Euclidean distances with Pearson's correlation coefficient weights, with inverse rank weighted mean adaptation with 3 analogies: MMRE = 12%, MdMRE = 10%, Pred(25) = 97.06%.
			• Best overall CBR configuration with adaptation rule was with Maximum distance, with the mean adaptation of the closest 3 analogies: Best estimate from Mean of 3 analogies: MMRE = 14%, MdMRE = 10%, Pred(25) = 85.29%.
			
			Stepwise/Multiple Linear Regression
			• Both achieved identical results.
			• Only considered Page Count, Reused Count and Reused Media Count predictors.
			• Excellent results, superior to those produced by best	

			overall CBR configuration (no adaptation rule, Pearson's Correlation coefficient weights, inverse rank weighted mean of the 3 closest analogies). • MMRE from 1.5% - 3.85%, MdMRE from 0.62% - 3.65%, and Pred(25) of 100%.	
S22	• CBR/Analogy	• MMRE • Pred(25)	**Dataset 1** • Results obtained without adaptation were never statistically significantly better than those obtained using adaptation rules. Without adaptation results: MMRE from 44.01%-775.52%, Pred(25) from 0%-58.33%. • Whenever the use of adaptation rules gave significantly better predictions, it occurred for all 3 types of adaptation rules (no significant difference). • Adaptation rules without weights gave the best predictions: MMRE from 27.24%-514.52%, Pred(25) from 16.67%-66.67%. **Dataset 2** • Not using adaptation rules gave the best results (for most configurations, statistically significantly better results at 95% confidence): MMRE from 27.01%-301.72%, Pred(25) from 2.70%-59.46%. **Feature Subset Selection** • Dataset 1 o Only one feature selected by ArchANGEL's feature subset selection feature so adaptation with weights not necessary. o FSS generally improved prediction accuracy. All statistically significantly better predictions were obtained when FSS was used in conjunction with adaptation rules (no weights); MMRE from 20.14% - 97.08%, Pred(25) from 16.67%-66.67%.	• For comparison of the CBR techniques, leave 1 out cross-validation (jackknife method) was used. • For evaluation of adaptation rules and feature subset selection, leave 1 out cross-validation was also used.

			• Dataset 2 • FSS generally improved prediction accuracy for DS2 as well. However all the statistically significantly better predictions were obtained when FSS was used alone without adaptation rules; MMRE from 15.13%-301.72%, Pred(25) from 2.70%-67.57%.	

S24	• CBR/Analogy
	• Stepwise regression
		• MMRE
		• Pred(25)
		• Boxplot of absolute residuals
			• For CSD and CBR, MMRE ranged from 56.29%-72.14%, and Pred(25) from 20%-30%.
			• For CSD and SWR, MMRE of 108.01% and Pred(25) of 27.5%.
			• For MCD and CBR, MMRE ranged from 151.37%-199.31%, and Pred(25) from 5.63%-20.63%.
			• For MCD and SWR, MMRE of 297.69%, and Pred(25) of 2.5%.
			• Overall best results for MCD were from CBR with 1 analogy (MMRE of 151.37% and Pred(25) of 20.63%.
			• Overall best results for CSD were from CBR with 3 analogies with MMRE of 56.29% and Pred(25) of 30%.
			• CBR with 3 analogies for CSD gave statistically significantly better results than CBR with 1 analogy for MCD.
				• 36 projects used from the Tukutuku database.
				  ○ 24 projects for the multi-company dataset (MCD).
				  ○ 12 projects for the single-company dataset (CSD).
				• All techniques gave better predictions for company-specific data set than for multi-company dataset.
				• All CSD vs. MCD comparison statistically significant apart from CBR using 1 analogy.

S25	• Case Based Reasoning • Stepwise Regression • Classification and Regression Trees • Mean Effort	• MMRE • MdMRE • Pred(25) • Boxplots of residuals • Boxplots of z • MEMRE used in certain circumstan	**CBR** • Maximum distance showed statistically worst results (MMRE from 23%-31%, MdMRE from 15%-34%, Pred(25) from 26.47%-76.47%). • Weighted Euclidean (MMRE from 10%-14%, MdMRE from 9%-12%, Pred(25) from 82.35%-97.06%) showed slightly better estimates than Unweighted Euclidean (MMRE from 12%-15%, MdMRE from 10%-12%, Pred(25) from 76.47%-88.24%). • For both Unweighted and Weighted Euclidean, using the closest analogy produced the best result.	• For identifying best CBR technique, jackknife method used (leave 1 out cross validation). • For comparisons of techniques to types of cross-validation was used. • 3 instances of a 66% split (SP1) and 3 instances of a 86% split (SP2). • Boxplot of residuals suggest

		ces.	• Best overall CBR setup was Weighted Euclidean distance, with 1 analogy, with a MMRE of 10%, MdMRE of 9% and Pred(25) of 94.12%. This was suggested by boxplot of residuals and boxplots of z where the box length and tails were smaller than for the other models. Also the outlier for the Weighted model was less extreme than the outliers for the other models. **Comparisons of Techniques** • All regression models used Page Count, Media Count and Reused Media Count. • Looking at MMRE and MdMRE, Stepwise regression gave the best prediction accuracy (For SP(1): MMRE from 3%-4%, MdMRE from 2%-4%, Pred(25) of 100%. For SP(2): MMRE from 2%-5%, MdMRE from 1%-4%, Pred(25) of 100%). • It also gave superior predictions to using mean total effort. • CART used only 3 measures, Page Count, Media Count, Reused Media Count (as did stepwise regression). • CART results were as follows: SP(1): MMRE from 13%-22%, MdMRE from 8%-11%, Pred(25) from 81.82%-90.91%. SP(2): MMRE from 10%-23%, MdMRE from 7%-10%, Pred(25) of 80%.	that stepwise regression gives the best prediction accuracy for SP1 (statistically significant at 95%) and SP2. • Boxplot of z show a very similar pattern for SP1, however for SP2, some boxplots of CART although having higher spread are not statistically significantly worse than the stepwise regression models. • No statistical significance between CART and CBR results. Boxplots of residuals and z show a difference in spread of distribution between CBR and CART. Boxplots suggest that CART gave better estimates for SP2. • No statistical significance between the results obtained by CART and CBR, and using mean effort. • Most CBR estimates were biased towards underestimation. • Stepwise regression tended to overestimate. • CART tended to overestimate.

S26	• CWADEE method	• Expert evaluation	• 15 out of 22 projects (68.2%) were estimated with good accuracy. • 5 out of 22 projects (22.7%) were estimated with medium accuracy. • 2 out of 22 projects (9.1%) were estimated with poor accuracy.	• Algorithmic model with significant expert input for calibration. • Projects sized using Data Web Points (DWPs), which are similar to Function Points, Object Points and Web Points. • DWPs represent system functionality from the point of view of its data model, and can be obtained early in the development cycle. • Different categories of DWPs. Each category is given a weight when calculating the total number of DWPs and this is handled by an expert. • The CWADEE model uses Cost Drivers taken from the WebMo model. • Good accuracy – estimate similar to expert opinion. • Medium accuracy – estimate close to expert opinion requiring only minor adjustment. • Poor accuracy – estimate not similar to expert estimate.
S27	• Web-COBRA • Ordinary Least Squares	• MMRE • MdMRE • Pred(25)	• For Web-COBRA, MMRE=0.17, MdMRE=0.15, and Pred(25)=0.75. • For OLS, MMRE=0.25, MdMRE=0.23, and Pred(25)=0.67.	• In terms of boxplot of residuals ○ Expert opinion tended to underestimate

	Regression • Expert judgment	• Boxplot of residuals	• For expert judgment, MMRE=0.37, MdMRE=0.36, Pred(25)=0.25.	development effort. This was not seen with the Web-COBRA and OLS model. ○ Web-COBRA presented the smallest interquartile range. 2 outliers present, both being large projects in terms of development effort. ○ OLS regression also had issues with 2 large projects, although these were different from the 2 outliers seen with Web-COBRA. • Web-COBRA proved to be significantly more accurate than expert judgment with a confidence of 99%. • No significant difference between Web-COBRA and OLS, or between OLS and expert judgment.
S28	• Ordinary Least Squares Regression • Expert judgment	• MMRE • MdMRE • Pred(25) • Boxplot of MRE	**OLS regression with FP:** • MMRE = 0.33, MdMRE = 0.33, Pred(25) = 0.42 **OLS regression with WO:** • MMRE = 0.24, MdMRE = 0.23, Pred(25) = 0.67 **Expert judgment** • MMRE = 0.37, MdMRE = 0.36, Pred(25) = 0.25 **Boxplot of MRE values** • OLS with FP boxplot had highest dispersion of MRE values, and largest distribution of MRE values. • OLS with WO and expert opinion box plot similar.	• OLS with WO significantly outperformed OLS with FP at 99% confidence. • No other significant differences found.

S30	• 4 variations of function point counting: ○ IFPUG ○ Estimated NESMA ○ Indicative NESMA ○ Author's simplified method • Latter 3 methods based on IFPUG.	• MRE and MMRE using the IFPUG measurement of functional size as the basis of comparison.	**Estimated NESMA** MRE ranging from 11% to 26% with an MMRE of 18% **Indicative NESMA** MRE ranging from 19% to 73% with a MMRE of 48% **Simplified** MRE ranging from 0% to 10% with an MMRE of 4%.	
S31	• Linear regression	• MMRE • Pred(25)	• MMRE of 0.1151 Pred(25) of 0.8438	
S32	• Manual forward stepwise regression Median effort	• MMRE • MdMRE • Pred(25) Boxplot of residuals	**Overall Model** • MMRE = 1.05, MdMRE = 0.534, Pred(25) = 26.4% • Average expert opinion accuracy (excluding the 13 single-company projects) = 68.3% (underestimate). **Cross-company model on single-model data** • Overall regression model used, with coefficients recalculated after the exclusion of the 13 single company projects. • Estimate accuracy of stepwise regression model superior to that of median model (95% confidence). • SWR MMRE = 0.565, MdMRE = 0.444, Pred(25) = 30.8%. Median model MMRE = 1.169, MdMRE = 0.759, Pred(25) = 15.4%. **Single-company model on its own data** • The single company model had: MMRE = 0.245, MdMRE = 0.234, and Pred(25) = 53.8%. • Significantly superior to cross-company model (95% confidence). • Not as good as expert opinion provided by the single	• Single company model considered total high effort features, average team experience on the development language used, and number of members on the development team. • Single company model evaluated on its own data using leave 1 out cross validation. • Median model used as benchmark.

			company which was a 10% underestimate. **Single Company model on cross-company data** • Poor predictor of cross-company effort. MMRE = 3.57, MdMRE = 0.78, Pred(25) = 15%.	
S34	• Stepwise regression • Case-based reasoning • Median effort • Expert judgment	• MMRE • MdMRE • Pred(25) • Mean absolute residuals • Median absolute residuals	**Overall median estimate** • MMRE=194% **Overall expert judgment** • 61.1% underestimate **Expert judgment on single company data** • 47% underestimate **Overall regression model** • MMRE=99% • MdMRE=70% • Pred(25)=9% • Mean absolute residual=374.9 • Median absolute residual=59.6 **CCM1 (Median model estimates in brackets)** • MMRE=93% (143%) • MdMRE=61% (70%) • Pred(25)=7.1% (7.1%) • Mean absolute residual=25.33 (33.1) • Median absolute residual=21.95 (26.9) **CCM2 (Median model estimates in brackets)** • MMRE=230% (428%) • MdMRE=151% (304%) • Pred(25)=14.3% (7.1%) • Mean absolute residual=55.4 (69.4) • Median absolute residual=54.4 (77.9)	• Median model used as benchmark. **Overall regression model** • Accuracy worse than expert judgment, and not significantly better than median model. **CCM1 and CCM2** • Estimates not significantly better than the corresponding median model estimates (95% confidence). • Both cross company models gave worse predictions than expert judgment. **Within company model evaluated on own data** • Significantly better results than median model. • Better prediction than expert judgment. • Significantly better results than either cross-company model (95% confidence). **Within company model evaluated on cross company**

			Within company model evaluated on own data (Median model estimates in brackets) • MMRE=38% (82%) • MdMRE=38% (61%) • Pred(25)=28.6% (14.3%) • Mean absolute residual=11.2 (30.3) • Median absolute residual=8.36 (15.8) **Within company model evaluated on cross-company data (53 projects)** • MMRE=94% • MdMRE=89% • Pred(25)=3.8% • Mean absolute residual=395.1 • Median absolute residual=88.4 **CBR** • MMRE from 100%-236% • MdMRE from 45%-136% • Pred(25) from 5.7%-25.4% • Mean absolute residual from 35.4-372.01 • Median absolute residual from 31.3-59.5	**data (53 projects)** • Slightly worse accuracy than overall cross-company regression model. **CBR** • Predictions for overall dataset significantly superior to regression model (95% confidence). • Predictions for single company projects using cross company data not significantly different from regression model predictions. • Predictions for single company projects using single company data significantly worse than regression model (95% confidence). • Predictions for cross-company projects using single company data not significantly different from regression model predictions.
S35	• Custom measurement framework	• MRE	• Estimate accuracy for the 3 applications (measured using MRE) were 33.81% (overestimate), 15.25% (underestimate), and 12.2% (underestimate). • Overall accuracy for application suite; MRE of 3.7% (overestimate).	• The Web project is first described in terms of design patterns used. • Each design pattern involves the implementation of one or more functional processes, and hence forms the basis for

				sizing a project using COSMIC Full Function Points. • Size of the project can be adjusted using a "series of multipliers which reflect the complexity of the application or of individual components or processes" termed "Size Factors" • Productivity is determined either from expert judgment (i.e. the members of the development team) or from historical data on projects they have finished. Like the size value, the productivity value can also be adjusted using multipliers termed "Productivity Factors". • The predicted effort required for the Web project can then be calculated by dividing the adjusted size value by the adjusted productivity value. • In order to take into account the probabilistic nature of resource estimation a probability distribution is created for the effort estimate using Monte Carlo simulation.
S37	• Linear regression • CBR/Analogy	• MMRE • Pred(25)	**Test Effort (unit)** • Linear regression model had an overall MMRE of 151.5% and Pred(25) of 23.8 %. • CBR had an overall MMRE of 127% and Pred(25) of 15%.	• Accuracy assessed using leave 1 out cross validation.

			Reliability Rate (unit) • Linear regression model had an overall MMRE of 91.4% and Pred(25) of 0%. • CBR had an overall MMRE of 48.1% and Pred(25) of 52.9%. **Unit Defect Density (unit)** • Linear regression model had an overall MMRE of 77.7% and Pred(25) of 27.2%. • CBR had an overall MMRE of 68.7% and Pred(25) of 75%. **Test Effort (concerns)** • Linear regression model had an overall MMRE of 58% and Pred(25) of 40.9%. • CBR had an overall MMRE of 98.5% and Pred(25) of 45.4%. **Reliability Rate (concerns)** • Linear regression model had an overall MMRE of 286.2% and Pred(25) of 28.5%. • CBR had an overall MMRE of 48.1% and Pred(25) of 52.9%. **Unit Defect Density (concerns)** • Linear regression model had an overall MMRE of 127.8% and Pred(25) of 25%. • CBR had an overall MMRE of 45% and Pred(25) of 45%. **Defect Density (concerns)** • Linear regression model had an overall MMRE of 168% and Pred(25) of 55%.	

			• CBR had an overall MMRE of 72.8% and Pred(25) of 13.6%. **Test Coverage (concerns)** • Linear regression model had an overall MMRE of 7.6% and Pred(25) of 100%. • CBR had an overall MMRE of 20.6% and Pred(25) of 77.2%. **Features Coverage (concerns)** • Linear regression model had an overall MMRE of 55.3% and Pred(25) of 31.8%. • CBR had an overall MMRE of 17.4% and Pred(25) of 81.8%.	

S39	• CBR/Analogy	• MMRE
		• Pred(25)
		• Boxplot of residuals
			Data set 1
			• Best predictions obtained for no adaptation rules or adaptation rules using no weights.
			• These 2 CBR configurations were significantly better than the other configurations at 95% confidence, for most distance and adaptation techniques.
			• In most cases they were not significantly different from each other (also at 95% confidence).
			• The exception to this was when the weighted Euclidean distance with subjective weights was used as the distance measure. In this case not estimates were significantly better when not using adaptation rules than when using adaptation rules with no weights for all but closest case (k=1) and median adaptation (k=3).
			• When no adaptation rules were used, MMRE ranged from 10%-32%, with Pred(25) being 100%.
			• When adaptation rules with no weights were used, MMRE ranged from 14% -33% with Pred(25) being 100%.
				• 2 datasets
				• Honours/Postgraduate students
				• Data Set 1
				  ○ 34 hypermedia applications
				• Data set 2
				  ○ 25 hypermedia
				• CBR configurations evaluated:
				  ○ No feature subset selection
				  ○ Unweighted Euclidean, Weighted Euclidean (subjective and correlation-based weights), Maximum distance, similarity measure.

			• No adaptation rules: ○ Except for the mean of 3 cases, all results obtained using the maximum distance measure were significantly worse than when using the remaining 3 distance measures. ○ Boxplots of residuals did not reveal much difference between the remaining 3 distance measures. ○ Weighted Euclidean distance using Pearson's correlation coefficient, with inverse rank weighted mean adaptation, and 3 analogies is a good candidate for the configuration that gives the best predictions overall based on MMRE (13%), Pred(25) (100%) and boxplot of residuals. • Adaptation rule with no weights ○ Except for the mean of 3 cases for maximum distance there were no statistically significant differences between the different configurations. ○ Best prediction when using the adaptation rule with no weights is the maximum distance using the mean of 3 cases based on MMRE (14%), Pred(25) (100%) and boxplot of residuals. **Data set 2** • Bad accuracy levels obtained for all distances used. • Best results obtained when using the unweighted Euclidean distance with no adaptation rules (MMRE from 74% - 83%, Pred(25) from 56% - 72%). Results statistically significant at 95% confidence. • Boxplot of residuals of the configurations using the unweighted Euclidean distance showed residuals that did not differ very much from each other.	○ Scaling ○ 1,2 and 3 analogies ○ Mean, median and inverse rank weighted mean case adaptation. • Adaptation rules: linear size adjustment with no weights, subjective weights and correlation-based size weights.

S40	• Stepwise regression • Median effort • Expert opinion	• MMRE • MdMRE • Pred(25) • Mean absolute residuals • Median absolute residuals	**Stepwise Regression** • MMRE=0.99 • MdMRE=0.70 • Pred(25)=9% • Mean absolute residual=374.9 • Median absolute residual=59.6 **Expert Opinion** • 61.09% underestimate **Median estimate** • MMRE = 1.94	• Median model used as benchmark. • Accuracy of stepwise regression model not significantly better than median model (95% confidence).
S41	• Linear (OLS) regression	• MMRE • MdMRE • Pred(25) • Boxplot of residuals • Boxplot of z	**Split 1 (C-FFPan)** • MMRE = 0.22 • MdMRE = 0.19 • Pred(25) = 0.64 **Split 2 (C-FFPan)** • MMRE = 0.22 • MdMRE = 0.11 • Pred(25) = 0.73 **Split 3 (C-FFPan)** • MMRE = 0.13 • MdMRE = 0.11 • Pred(25) = 0.73 **Split 4 (C-FFPan)** • MMRE = 0.14 • MdMRE = 0.18 • Pred(25) = 0.70 **Split 1 (C-FFPde)** • MMRE = 0.19 • MdMRE = 0.17	• Applications sized using COSMIC-FFP derivates C-FFPan (obtained from the application analysis documents) and C-FFde (obtained from the application design documents). • Boxplot of residuals and z confirm the results obtained by the summary statistics. Models based on C-FFPde have boxplots with smaller tails and box lengths than the boxplots for C-FFPan models. • With the exception of training set 2, the accuracy results between C-FFPde and C-FFPan are not significant (based on absolute residuals and 95% confidence).

			• Pred(25) = 0.55 **Split 2 (C-FFPde)** • MMRE = 0.19 • MdMRE = 0.08 • Pred(25) = 0.64 **Split 3 (C-FFPde)** • MMRE = 0.10 • MdMRE = 0.10 • Pred(25) = 0.91 **Split 4 (C-FFPde)** • MMRE = 0.11 • MdMRE = 0.05 • Pred(25) = 0.90 **Aggregate Accuracy** • C-FFPan: MMRE = 0.18, MdMRE = 0.15, Pred(25) = 0.70 • C-FFPde: MMRE = 0.15, MdMRE = 0.10, Pred(25) = 0.75	
S42	• Multiple linear regression • Forward stepwise regression • Regression trees • Analogy • Regression tree + linear regression • Regression tree + analogy	• MMRE • MdMRE • Pred(25) • Boxplot of residuals and z	**Length Measures** • Linear regression: MMRE = 0.18, MdMRE = 0.15, and Pred(25) = 0.73. • Stepwise regression: MMRE = 0.28, MdMRE = 0.16, and Pred(25) = 0.53. • Regression tree: MMRE = 0.17, MdMRE = 0.11, Pred(25) = 0.80. • Regression tree + Linear regression: MMRE = 0.17, MdMRE = 0.10, Pred(25) = 0.73. • Analogy: MMRE from 0.21 – 0.31, MdMRE from 0.13 – 0.23, Pred(25) = 0.53 – 0.73. • Regression tree + analogy: MMRE from 0.17 – 0.20,	• 2 linear regression models (for length and functional measures). • 2 stepwise regression models (for length and functional measures). • 2 regression tree models (for length and functional measures). • Unweighted Euclidean distance measure, normalized variables, 1,2

			MdMRE from 0.12 – 0.19, Pred(25) from 0.66 – 0.87 **Functional Measures** • Linear regression: MMRE = 0.20, MdMRE = 0.10, and Pred(25) = 0.73. • Stepwise regression: MMRE = 0.16, MdMRE = 0.09, and Pred(25) = 0.80. • Regression tree: MMRE = 0.20, MdMRE = 0.12, Pred(25) = 0.67. • Regression tree + Linear regression: MMRE = 0.19, MdMRE = 0.20, Pred(25) = 0.66. • Analogy: MMRE from 0.24 – 0.28, MdMRE from 0.08 – 0.17, Pred(25) = 0.53 – 0.73. • Regression tree + analogy: MMRE from 0.25 – 0.31, MdMRE from 0.19 – 0.24, Pred(25) from 0.57 – 0.72	and 3 analogies, and mean/inverse distance weighted mean/inverse rank weighted mean distances. Prediction accuracy assessed using leave 1 out cross validation. • For hybrid models, independent variables used for splitting data set with regression tree modeling used as the basis for secondary model (linear regression or analogy). E.g. The number of server side scripts and applications determined the splitting of the data set when evaluating length measures with regression trees, and this was then used as the independent variable for linear regression. • Stepwise regression produced significantly better results when estimating effort with functional measures as opposed to length measures. • Best analogy results with length measures and functional measures obtained using 3 and 2 analogies respectively, and inverse distance weighted

				mean as the adaptation strategy. • Boxplots of residuals and z used to give insight on the behavior of summary stats like MMRE, MdMRE and Pred(25). • Looking at these, the best results seem to have been obtained with stepwise regression applied on functional measures, resulting in boxplots with the smallest box length and tails, and residuals that are symmetrically distributed with the median adjacent to 0. ○ Based on boxplots of z, all but 1 analogy model underestimates effort
S43	• Fuzzy analogy	• MMRE • Pred(0.20)	• Different membership functions were investigated. • Trapezoidal functions were found to be better than triangular functions which were found to be better than Gaussian functions. • Accuracy was dependant on linguistic quantifier α. Accuracy increases monotonously with α. • All 3 membership function shapes had a Pred(0.20)>=70 with appropriately high levels of α. • For Trapezoidal functions, an α value of 25 or greater was required with MMRE values ranging from 78.36 to 58.60 and Pred(0.20) ranging from 71.70 to 84.91. • For Triangular functions, an α value of 30 or greater was required with MMRE values ranging from 73.17	• CBR and fuzzy logic • To build fuzzy representations for the 9 software attributes Fuzzy C-Means clustering technique in conjunction with a Real Coded Genetic Algorithm was used. • FCM was used to generate fuzzy sets for each attribute. ○ RCGA was used to generate membership

			to 62.23, and Pred(0.20) ranging from 73.58 to 79.25. • For Gaussian functions, an α value of 50 or greater was required with MMRE values ranging from 70.17 to 69.22, and Pred(0.20) ranging from 71.70 to 73.58.	functions for the fuzzy sets.
S44	• Stepwise regression • Mean model • Median model	• MMRE • MdMRE • Pred(25)	**SWR** • MMRE = 129% • MdMRE = 73% • Pred(25) = 17.24% **Mean** • MMRE = 4314% • MdMRE = 1413% • Pred(25) = 6.89% **Median** • MMRE = 663% • MdMRE = 149% Pred(25) = 3.44%	• All model validation done with 1-fold cross-validation with a 66% split. • Mean and median models used as benchmarks. •
S45	• Fuzzy least-squares regression	• NA	• NA	• Projects sized using COSMIC-FFP version 2.0. • Accuracy of model not evaluated.

S46	• CBR/Analogy
	• Median model
	• Mean model
		• MMRE
		• MdMRE
		• Pred(25)
			Using OO-HFP functional size measure
			• **CBR**
			  ○ MMRE = 3514.4%
			  ○ MdMRE = 1763.2%
			  ○ Pred(25) = 25%
			• **Mean**
			  ○ MMRE = 1532.5%
			  ○ MdMRE = 786.4%
			  ○ Pred(25) = 0%
			• **Median**
			  ○ MMRE = 156.51%
			  ○ MdMRE = 94.8%
			  ○ Pred(25) 0%
				• CBR config: unweighted Euclidean similarity measure, 1 analogy, no adaptation rules used.
				• Mean and median models used only for comparisons (benchmarking) purposes.
				• OO-HFP and the Tukutuku size measures were compared.
				• No significant difference in accuracy between using OO-HFP or Tukutuku size measures.

			Using Tukutuku size measures
			• **CBR**
			  ○ MMRE = 3614.4%
			  ○ MdMRE = 1864.2%
			  ○ Pred(25) = 25%
			• **Mean**
			  ○ MMRE = 1532.5%
			  ○ MdMRE = 786.4%
			  ○ Pred(25) = 0%
			• **Median**
			  ○ MMRE = 156.51%
			  ○ MdMRE = 94.8%
			• Pred(25) = 0%
				• No significant difference between CBR (regardless of size measure) and Mean and Median estimates.

S47	• NA	• NA	NA	• Ordinary Least Squares regression and Robust regression used to investigate relationships between dependent variables (information effort, navigation effort, presentation effort and total design effort) and independent variables (see metrics discussed below). • Various metrics (e.g. size measures, complexity measures, data cohesion measures, and reuse measures) involved with the information and navigation models of the W2000 design notation were considered. • Both OLS and robust regression also used to

				investigate if there was a statistical association/correlation between estimated effort for the various W2000 models and actual effort. • Estimated effort was also compared to actual effort for the various W2000 models and learning effort to see if students tended to underestimate (which they did in most cases).
S48	• Forward manual stepwise regression • Case-based reasoning • Mean • Median	• MMRE • MdMRE • Pred(25) • Boxplot of absolute residuals	**Tukutuku** • SWR – MMRE = 0.18, MdMRE = 0.14, Pred(25) = 0.73 • CBR – MMRE = 0.16, MdMRE = 0.12, Pred(25) = 0.87 **Web Objects** • SWR – MMRE = 0.17, MdMRE = 0.11, Pred(25) = 0.80 • CBR – MMRE = 0.21, MdMRE = 0.11, Pred(25) = 0.80 **Length** • SWR – MMRE = 0.12, MdMRE = 0.11, Pred(25) = 0.87 • CBR – MMRE = 0.18, MdMRE = 0.12, Pred(25) = 0.87 **Functional** • SWR – MMRE = 0.23, MdMRE = 0.21, Pred(25) = 0.73 • CBR – MMRE = 0.14, MdMRE = 0.11, Pred(25) = 0.93	• Overall predictions obtained using CBR were superior to those using SWR. • Based on MMRE, MdMRE, and Pred(25), the best results for CBR and SWR were obtained using functional and length measures respectively. • Using absolute residuals, SWR with length measures and Web Objects produced significantly superior results to SWR with functional measures, but similar results to SWR with Tukutuku measures. • With CBR, no significant differences amongst the 4 size measures using absolute residuals. • SWR with length measures

			• SWR – MMRE = 0.18, MdMRE = 0.14, Pred(25) = 0.73 • CBR – MMRE = 0.16, MdMRE = 0.12, Pred(25) = 0.87 **Mean** • MMRE = 0.34, MdMRE = 0.27, Pred(25) = 0.47 **Median** • MMRE = 0.33, MdMRE = 0.24, Pred(25) = 0.60	produced significantly superior results to both mean and median estimation. • CBR with Tukutuku measures produced significantly superior results to mean estimation. • CBR presented significantly superior results than SWR with functional measures. ○ Based on boxplot of absolute residuals, best results were obtained for SWR with length measures and CBR with functional measures, as confirmed by other accuracy measures.
S49	NA	NA	NA	• Study designed to look at the relationship between certain metrics and maintenance effort. • No model to estimate maintainability evaluated
S50	• Fuzzy Radial Basis Function Neural Networks	• MMRE • Pred(25)	**Fuzzy C-Means** • Accuracy dependent on the number of clusters and the width of the Gaussian kernels. • 2 width factors available; Haykin and Saha. • For Haykin: MMRE from around 45% to 0% decreasing as the number of clusters increases from 25 to 50. Pred(25) from around 60% to around 95% increasing as the number of clusters increases from	• 3 layer neural network: ○ Input layer – 9 neurons corresponding to 9 variables characterizing Web project. ○ Middle layer; number of neurons variable

			25 to 50.
			• For Saha: MMRE from around 40% to 0% decreasing as the number of clusters increases from 25 to 50. Pred(25) from around 55% to around 100% increasing as the number of clusters increases from 25 to 50.
			
			Hard C-Means
			• C-means classification to minimize objective function J used.
			• MMRE decreases from 200% to 0% as the number of clusters increases from 25 to 50 (there is a sizable spike in MMRE as the number of clusters transitions from 30 to 35).
			  ○ Pred(25) increases from around 25% to around 100% as the number of clusters increases from 25 to 50.
				determined by clustering algorithm.
				  ○ Output layer consisting of a single neuron.
				• Fuzzy C-means and hard C-means clustering algorithms used

S51	• Bayesian networks • Case-based reasoning • Stepwise regression • CART	• MMRE • MdMRE • Pred(25)	**BN** • MMRE = 34.3%, MdMRE = 27.4%, Pred(25) = 33.3%. **SWR** • MMRE = 94.8%, MdMRE = 100%, Pred(25) = 6.7%. **CBR** • MMRE from 134.7 – 203%, MdMRE from 85 – 91.7%, Pred(25) = 13.3% (for all 3 configurations). **CART** • MMRE = 690.4%, MdMRE = 83.2%, Pred(25) = 20%.	**Bayesian Network Model** • Hybrid structural development and parameter estimation. Data-driven model validation. • Hugin Expert used for data-driven processes. **CBR** • Euclidean distance similarity measure, mean adaptation, 1-3 analogies, feature subset selection (using Spearman's rank correlation), no adaptation rules used. • CBR-Works **CART** • Max tree depth of 10, minimum number of cases

				in parent node – 5, minimum number of cases in child nodes – 2. • Least squared deviation used as measure of impurity. • Small Risk Error (SRE) was set to a minimum of 90%. • All models built using the same population of 120 projects and evaluated on the same test set of 30 projects. • BN predictions statistically superior to other methods. • No statistical difference between remaining methods. • Boxplots of residuals showed that BN-based residuals were the smallest with the most compact distribution.
S52	• Bayesian networks • Stepwise regression	• MMRE • MdMRE • Pred(25) • MEMRE • MdEMRE • Boxplot of residuals Boxplot of z	**BN** • MMRE = 34.26%, MdMRE = 27.42%, Pred(25) = 33.33%, MEMRE = 228.41%, MdEMRE 35.83% **SWR** • MMRE = 94.75%, Median = 100%, Pred(25) = 6.67%, MEMRE = 31.12%, MdEMRE = 35.83%	• For the SWR model 24 projects in the validation set had estimated efforts of 0. Therefore MEMRE and MdEMRE measures were only calculated for the remaining 6 cases. For this reason absolute residuals were focused on. • BN accuracy statistically significantly superior to

				mean and median model accuracy (with 95% confidence). • Boxplot of residuals similar.
S53	• Bayesian networks • Mean model • Median model	• MMRE • MdMRE • Pred(25)	**BN** • MMRE = 34.3%, MdMRE = 27.4%, Pred(25) = 33.3%. **Mean** • MMRE = 1106.31%, Median = 252.36%, Pred(25) = 6.67% **Median** MMRE = 132.76%, MdMRE = 85.9%, Pred(25) = 10%	• Hybrid structural development and parameter estimation. Data-driven model validation. • Hugin Expert used for data-driven processes. • Mean and median models used as a benchmark. • BN accuracy statistically significantly superior to mean and median model accuracy (with 95% confidence). • No significant difference between mean and median models. • Looking at boxplot of residuals, the boxplot for the BN model has a lower median than the boxplots for the mean and median models, and fewer (and not as extreme) outliers.
S54	• Case-based reasoning • Stepwise regression • CART • Mean model • Median model	• MMRE • MdMRE • Pred(25) • Boxplot of residuals	**SWR** • MMRE = 94.8%, MdMRE = 100%, Pred(25) = 6.7%. **CBR** • MMRE from 134.7 – 546.3%, MdMRE from 85 – 99.3%, Pred(25) = 10 - 13.3%. **CART** • MMRE = 690.4%, MdMRE = 83.2%, Pred(25) = 20%.	**CBR** • Euclidean distance similarity measure, mean and inverse rank-weighted mean adaptation, 1-3 analogies, feature subset selection (using Spearman's rank correlation), no adaptation

			Median • MMRE = 132.8%, MdMRE = 85.9%, Pred(25) = 10% **Mean** • MMRE = 1106.2%, MdMRE = 252.3%, Pred(25) = 6.7%	rules used. **CART** • Max tree depth of 10, minimum number of cases in parent node – 5, minimum number of cases in child nodes – 2. • Least squared deviation used as measure of impurity. • Small Risk Error (SRE) was set to a minimum of 90%. • The mean and median effort models were used for benchmarking purposes. • All models built using the same training set of 120 projects and evaluated on the same test set of 30 projects. • No statistical difference between the different CBR configurations, SWR and CART. • SWR, CBR (all configs) and CART statistically superior to the mean model, but not the median model. • Boxplot of residuals similar except for CBR with 3 analogies using the inverse rank weighted mean

				adaptation technique. This had the highest median but the most compact distribution. • All boxplots shared the same set of 4 outliers.
S55	• Stepwise Regression • Case-based reasoning • Mean model • Median model	• MMRE • MdMRE • Pred(25)	**Cross-company SWR model on single company dataset** • MMRE = 85.86%, MdMRE = 100%, Pred(25) = 6.67% **Cross-company CBR model on single company dataset** • MMRE = 92.54%, MdMRE = 93.13%, Pred(25) = 0% **Single-company SWR model on single company dataset** • MMRE = 19.51%, MdMRE = 15.44%, Pred(25) = 73.33% **Single-company CBR model on single company dataset** • MMRE = 15%, MdMRE = 15%, Pred(25) = 80% **Single-company Mean model** • MMRE = 31.64%, MdMRE = 25.61%, Pred(25) = 46.67% **Single-company Median model** • MMRE = 32.25%, MdMRE = 23.3%, Pred(25) = 66.67% •	• Mean and median single company models used as a benchmark for comparison. **CBR** • Unweighted Euclidean distance measure, mean adaptation using the closest 2 analogies. **SWR** • 8 projects with large residuals were moved because they destabilized the cross-company model. • Leave 1 out cross validation use to verify model accuracy on single-company data. • Cross-company SWR model significantly worse than mean and median model (95% confidence). • Single-company SWR model significantly better than cross-company SWR model, but not the median and mean models. • Boxplot of residuals show

				that the spread of the distribution for the cross-company SWR model is much wider and that most of the values are much greater. • Cross-company CBR model not significantly different from cross-company SWR model. • Single-company CBR model not significantly different from single-company SWR model. • Cross-company CBR model significantly worse than single-company CBR model. • Single-company CBR model not significantly different from mean or median model.
S56	• A combination of expert judgment, WebMo and neural networks	• NA	• NA	• Project size is first estimated using WebMo • This size is then adjusted using a neural network. The neural network has 10 input parameters (9 for the WebMo cost drivers and 1 for bias). • Training of the neural network is done using estimates of development effort (in person months) obtained by expert judgment.

				• The values of the weights (for the various cost drivers) and bias are adjusted at each training step until the change in values for all the weights falls below a certain threshold level. • Adjustment in weights determined by a calculation that involves the estimation error (the difference between the estimate provided by expert judgment and the automated estimated). No accuracy metrics specified.
S57	• The authors proposed estimating maintenance effort by first sizing the system, then determining what proportion of the size units are affected by the maintenance task in question (i.e. the one to be estimated). • Size was measured in object-points, function points, and number of	NA	NA	• 8 complexity measures including class hierarchy, data usage, coupling and cohesion. • 8 quality measures including modularity, portability, flexibility and maintainability. • Model not evaluated.

	statements. • Raw size measures adjusted by degree of complexity and quality. • Measuring the proportion of the size units affected (“maintenance impact”), the software entities involved need to be identified, as well as the relationships these entities have with dependent entities. • Lastly the error reports and change requests that initiate the maintenance task are also considered.			
S58	• Content Management System Effort Estimation Model (CMSEEM)	• MMRE • Pred(25)	• Projects were divided into 4 types based on their size and total/build ratio. • MMRE ranged from 0.09 to 0.15. • Overall MMRE of 0.12 • Pred(25) > 80%	• Algorithmic estimation model. • Extension of COCOMO 2.0 model • Sizes applications using Object Point Analysis
S62	• Linear regression	• MMRE • MdMRE	**Linear regression with COSMIC** • MMRE = 0.22	• 2 models. • Effort was the dependent

		• Pred(25) • Boxplot of absolute residuals	• MdMRE = 0.11 • Pred(25) = 0.75 **Linear regression with Web Objects** • MMRE = 0.14 • MdMRE = 0.06 • Pred(25) = 0.75	variable, and project size in either Web objects or COSMIC function points the independent variables. • Median of WO boxplot closer to 0 than COSMIC boxplot. • Tails and box length for WO boxplot slightly smaller than for CFP boxplot. • No significant difference between the 2 estimation models.
S63	• WEBMO	MMRE	Effort • Business – MMRE = 99.88% • E-commerce – MMRE = 99.96% • Financial/trading – MMRE = 99.96% • Information portal – MMRE = 99.76% • Information utilities – MMRE = 99.31% Duration • Business – MMRE = 98.03% • E-commerce – MMRE = 98.85% • Financial/trading – MMRE = 99.16% • Information portal – MMRE = 97.45% • Information utilities – MMRE = 96.17%	•
S64	• Radial Basis Function Neural Networks	• MMRE • Pred(25)	**C-Means clustering** • Cluster coherence can be measured via either the objective function *J* or the Dunn index D_1. Aim is to either minimize *J* or maximize D_1. • MMRE-J ranged from around 75% to close to 0% depending on the number of clusters (the higher the number of clusters the lower the MMRE). • MMRE-D_1 ranged from around 60% to around 10%	• Two clustering algorithms considered: C-means and APC-III.

			depending on the number of clusters (the higher the number of clusters the lower the MMRE). • Pred(25)-J ranged from around 70% to around 95% depending on the number of clusters (the higher the number of clusters the larger the Pred(25)). • Pred(25)-D_1 ranged from around 60% to close to 90% depending on the number of clusters (the higher the number of clusters the larger the Pred(25)). **APC-III** • Number of clusters using the APC-III algorithm dependent on the value of parameter α. The higher the α value, the smaller the number of clusters. • α values ranged from 0.2 to 0.7 (with the corresponding number of clusters ranging from 52 to 34). • MMRE ranged from around 0% to around 80% depending on the value of α (the higher the α value, the higher the MMRE). • Pred(25) ranged from 100% to around 58% depending on the value of α (the higher the α value, the lower the Pred(25)).	
S65	• Bayesian Networks • Stepwise Regression • Case-based Reasoning • Mean model • Median model	• MMRE • MdMRE • Pred(25) • MEMRE • MdEMRE Boxplots of residuals and *z*	**Validation Set 1** • Hybrid BN, CBR (1, 2 and 3 analogies) and MSWR all performed statistically better than mean based estimated effort. • Only MSWR performed statistically better than median based estimated effort. MSWR outperformed all techniques (MMRE = 1.5, MdMRE = 0.64, Pred(25) = 23.08, MEMRE = 1.36 and MdEMRE = 0.64). • The Hybrid BN performed significantly better or similar to all methods apart from MSWR (MMRE = 1.9, MdMRE = 0.86, Pred(25) = 15.38, MEMRE = 13.06 and MdEMRE = 2.38).	**Bayesian Networks** • 2 data-driven BNs, and 2 hybrid BNs. • Hugin tool used. • Data-driven structural development done on training set of 130 randomly selected projects. • 2 training sets -> 2 data driven BNs. • Single expert built structure fit to two training sets to

			• CBR 2 and 3, and both BNs performed similarly to median based estimated effort. • CBR 1 showed significantly worse prediction accuracy than median based estimated effort (MMRE = 5.27, MdMRE = 0.97, Pred(25) = 7.69, MEMRE = 31.70, MdEMRE = 3.43). • Similar trends observable using boxplot of residuals. • Using boxplot of z-values, it was the hybrid BN that significantly outperformed all other techniques. • Only CBR1 and the hybrid BN outperformed median based effort estimation. **Validation Set 2** • All techniques performed significantly better than Mean estimated effort when looking at residuals. • Only MSWR was significantly more accurate than median based estimated effort (MMRE = 0.73, MdMRE = 0.66, Pred(25) = 10.77, MEMRE = 2.86 and MdEMRE = 1.21). • Median based estimated effort was significantly superior to all other techniques apart from MSWR. (MMRE = 4.95, MdMRE = 0.89, Pred(25) = 15.38, MEMRE = 4.62 and MdEMRE = 0.78). • The data-driven BN performed better than the hybrid BN. Its performance was similar to those of all CBR-based predictions. (MMRE = 4.09, MdMRE = 0.96, Pred(25) = 1.54, MEMRE = 7.90 and MdEMRE = 0.93). • The hybrid BN was significantly worse than all techniques apart from mean based estimation. (MMRE = 27.95, MdMRE = 5.31, Pred(25) = 3.08, MEMRE = 1.34 and MdEMRE = 0.90). • Similar trends observable with boxplot of residuals. • Similar trends also observable with boxplot of z values, except CBR1 in addition to MSWR presented	generate the 2 hybrid BNs. • Equal-frequency intervals discretisation was used, and 5 intervals generated. • None of the BNs were optimized. • Parameter Estimation data driven using EM-learning algorithm. • Predictive accuracy used to validate BNs. • 66:33 split used for validation with 65 projects. • 2 validation sets used. • Estimated effort calculated using point forecast. **Stepwise Regression** • 2 regression models built. • Same 2 training sets used. • 2 validation sets of 65 projects. • Variables with more than 40% of their values equal to 0 or missing were excluded. • Transformation required due to variables having non-normal distributions. • High influence data points removed to improve model stability. **Case-based Reasoning** • CBR-Works tool used. • Unweighted Euclidean distance similarity measure

			significantly superior predictions as compared to median based prediction.	• 1,2, and 3 analogies • Mean adaptation • Feature subset manually decided using Spearman's rank correlation test. • All 195 cases used as case base. • 2 validation sets of 65 projects used. • Mean and median single company models used as a benchmark for comparison.

S66	• Bayesian networks
	• Manual stepwise regression
	• Case-based reasoning
	• Mean model
	• Median model
		• MMRE
		• MdMRE
		• Pred(25)
		• MEMRE
		• MdEMRE
		• Boxplot of residuals
		Boxplot of z
			Validation Set 1
			• Mean
			  ○ MMRE = 30.35, MdMRE = 3.99, Pred(25) = 15.38%, MEMRE = 1.07, MdEMRE = 0.91
			• Median
			  ○ MMRE = 5.02, MdMRE = 0.93, Pred(25) = 9.23%, MEMRE = 4.43, MdEMRE = 0.94
			• BNAuPo
			  ○ MMRE = 13.97, MdMRE = 2.57, Pred(25) = 4.62%, MEMRE = 0.78, MdEMRE = 0.81
			• BNHyPo
			  ○ MMRE = 36.00, MdMRE = 4.90, Pred(25) = 7.69%, MEMRE = 1, MdEMRE = 0.93
			• BNAuHu
			  ○ MMRE = 7.65, MdMRE = 1.67, Pred(25) = 7.69%, MEMRE = 1.07, MdEMRE = 0.76
			• BNHyHu
			  ○ MMRE =1.90, MdMRE = 0.86, Pred(25) = 15.38%, MEMRE = 13.06, MdEMRE = 2.38
			• MSWR
				BN
				• 8 BN models were made using 2 BN tools; Hugin and Powersoft.
				• 2 datasets of 130 projects were used for both the structural development and parameter estimation phases. These 2 datasets were randomly selected.
				• The remaining 65 projects were used for model validation.
				• 4 of these 8 BNs were data driven (2 datasets and 2 tools).
				• The remaining 4 models were hybrid BNs. The structural development phase was handled by an

			○ MMRE = 1.50, MdMRE = 0.64, Pred(25) = 23.08%, MEMRE = 1.36, MdEMRE = 0.64 • CBR ○ MMRE = 5.06 – 5.27, MdMRE = 0.87 – 0.97, Pred(25) = 7.69% - 10.77%, MEMRE = 3.59 – 31.70, MdEMRE = 0.81 – 3.43 **Validation Set 2** • Mean ○ MMRE = 27.94, MdMRE = 5.31, Pred(25) = 3.08%, MEMRE = 1.34, MdEMRE = 0.90 • Median ○ MMRE = 4.95, MdMRE = 0.89, Pred(25) = 15.38%, MEMRE = 4.62, MdEMRE = 0.78 • BNAuPo ○ MMRE = 14.93, MdMRE = 6.46, Pred(25) = 0%, MEMRE = 0.94, MdEMRE = 0.90 • BNHyPo ○ MMRE = 37.31, MdMRE = 8.05, Pred(25) = 1.54%, MEMRE = 1.14, MdEMRE = 0.93 • BNAuHu ○ MMRE = 4.09, MdMRE = 0.96, Pred(25) = 1.54%, MEMRE = 7.90, MdEMRE = 0.93 • BNHyHu ○ MMRE = 27.95, MdMRE = 5.31, Pred(25) = 3.08%, MEMRE = 1.34, MdEMRE = 0.90 • MSWR ○ MMRE = 0.73, MdMRE = 0.66, Pred(25) = 10.77%, MEMRE = 2.86, MdEMRE = 1.21 • CBR • MMRE = 4.46 – 6.73, MdMRE = 0.84 – 0.92, Pred(25) = 7.69% - 15.38%, MEMRE = 13.26 – 21.81, MdEMRE = 0.89 – 0.95	expert, and the resulting model was used for the remaining 2 KEBN stages which were automated using both datasets and tools. **MSWR** • 2 regression models built, 1 for each dataset (see above). • Models validated on the remaining 65 projects. • 6 of the 19 variables excluded for regression analysis in accordance to assumptions underlying any regression technique. • For dataset 1, 13 high influence projects were removed to improve model stability. • For dataset 2, 9 high influence projects were removed to improve model stability. **CBR** • Feature selection using Spearman’s rank correlation test to select features only significantly associated with total effort. • Unweighted Euclidean distance similarity measure.

				• 1,2 and 3 analogies with mean adaptation. • Mean and Median models used for benchmarking. **Validation Set 1** • Based on absolute residuals, MSWR significantly outperformed all other techniques. It was also the only technique to significantly outperform median based estimation. • BNHyHu presented similar or significantly better accuracy (Mean, CBR1, and BNHyPo) than all techniques except MSWR. • Most techniques presented significantly superior estimation to mean based estimation. • BNHyPo presented the worst predictions out of all techniques including mean based estimation. • Based on z-values however, it was BNHyHu that out performed all techniques including median based prediction. **Validation Set 2** • Based on absolute

				residuals, best predictions obtained by (in descending order), MSWR, CBR3, and CBR2/CBR3/BNHyHu. • Only MSWR presented significantly superior accuracy to median based estimation. • Most techniques presented significantly superior accuracy to mean based estimation. • BNHyPo presented the worst predictions out of all techniques including mean based estimation. • Best BN model was BNAuHu, whose estimates were significantly superior to all other BN models and similar to all CBR predictions. • Significance tests using z values similar to those using absolute residuals. MSWR and CBR1 both presented accuracy significantly superior to median-based predictions. BNAuHu was not the best BN model presenting significantly worse predictions than BnAuPo.
S67	• Stepwise regression	• MMRE • MdMRE	**Baseline SWR** • MMRE = 120.74% MdMRE = 56.54% Pred(25) =	• Mean and median single company models used as a

	• Case-based reasoning
	• Mean model
	• Median model
		• Pred(25)	18.07%
			
			Baseline CBR
			• MMRE = 111.93% MdMRE = 45.45% Pred(25) = 31.33%
			
			Baseline Mean
			• MMRE = 7094.89% MdMRE = 1394% Pred(25) = 3.61%
			
			Baseline Median
			• MMRE = 461% MdMRE = 98.21% Pred(25) = 4.82%
			
			CCM1
			• MMRE = 85.86% MdMRE = 100% Pred(25) = 6.67%
			
			CC-CBR
			• MMRE = 92.54% MdMRE = 93.13% Pred(25) = 0%
			
			CC-Mean
			• MMRE = 86.54% MdMRE = 88.49% Pred(25) = 0%
			
			CC-Median
			• MMRE = 98.74% MdMRE = 98.92% Pred(25) = 0%
			
			CCM2
			• MMRE = 60.35% MdMRE = 55.28% Pred(25) = 6.67%
			
			SCM
			• MMRE = 15.93% MdMRE = 14.68% Pred(25) = 86.67%
			
			SC-CBR
			• MMRE = 15% MdMRE = 15% Pred(25) = 80%
				benchmark for comparison.
				
				CBR
				• Unweighted Euclidean distance measure, mean adaptation using the closest 2 analogies.
				• 3 CBR configurations evaluated.
				  ○ A base-line CBR configuration using all 83 projects evaluated using leave 1 out cross-validation.
				  ○ A CBR configuration that uses the cross-company dataset as a case base and whose accuracy is evaluated on the 15 single company projects.
				  ○ A CBR configuration that uses the single-company dataset as a case base evaluated using leave 1 out cross-validation.
				
				SWR
				• 4 stepwise regression models were built.
				  ○ A baseline model built using the entire dataset of 83 projects. 9 high influence projects were

			SC-Mean • MMRE = 31.64% MdMRE = 25.61% Pred(25) = 46.67% **SC-Median** • MMRE = 32.25% MdMRE = 23.3% Pred(25) = 66.67% **SCM on CC data** • MMRE = 2.61908E+12% MdMRE = 5668.56% Pred(25) = 0% **SC-CBR on CC data** • MMRE = 14430.99% MdMRE = 5146.52% Pred(25) = 1.47% **SC-Mean on CC data** • MMRE = 31208.52% MdMRE = 8781.81% Pred(25) = 5.88% **SC-Median on CC data** • MMRE = 32542.41% MdMRE = 9160.36% Pred(25) = 5.88%	removed from this model to improve stability. Accuracy assessed using leave 1 out cross validation. ○ A cross-company model built by recalibrating the baseline model whilst omitting single-company data (CCM2). Validated on the single-company dataset. ○ A cross-company model built from scratch using only the 68 projects from the cross-company dataset (CCM1). Also evaluated on the single company data. 8 high influence projects were removed from this model to improve stability. ○ A single company model built using the single company dataset (SCM). Evaluated using leave 1 out cross validation. • Baseline SWR superior to baseline mean and median techniques (95%

				confidence). • Baseline CBR similar to baseline SWR. • CCM1 accuracy similar to mean and median techniques. Boxplot of residuals show that CCM1 plot has lower median than mean and median plots suggesting slight (not significant) superiority. • CC-CBR similar in accuracy to CCM1 • CCM2 accuracy similar to mean and median techniques. CCM2 plot has much lower median than mean and median plots suggesting slightly superior predictions (not significant) by CCM2. • SCM similar in accuracy to single company mean and median techniques. SCM boxplot has lower median than mean and median plots suggesting slightly superior predictions. • SC-CBR similar in accuracy to SCM. • CCM1 significantly less accurate than SCM, but not significant difference between CCM2 and SCM. • CC-CBR significantly less

				accurate than SC-CBR. • SCM poor predictor for cross-company data; significantly worse than single company mean and median technique applied to cross-company data. • SC-CBR not significantly difference from SCM when applied to cross-company data.
S69	• Content Management System Effort Estimation Model (CMSEEM)	• MMRE • Pred(25)	• Projects were divided into 4 types based on their size and total/build ratio. • MMRE ranged from 0.13 to 0.24. • Overall MMRE of 0.16. • Pred(25) > 80%	• Algorithmic estimation model. • Extension of COCOMO 2.0 model • Sizes applications using Object Point Analysis

S70	• Custom	• MRE
		• Pred(25)
			• Summary stats for MRE given with along with Pred(25)
			• FP
			  ○ MMRE = 0.49
			  ○ MdMRE = 0.19
			  ○ Pred(25) = 0.6
			• WO
			  ○ MMRE = 0.61
			  ○ MdMRE = 0.36
			  ○ Pred(25) = 0.4
				• Size measure in function points or Web objects used in conjunction with a productivity coefficient (determined on the basis of tech used for project).

S71	• Support Vector Regression (SVR)	• MMRE • MdMRE • Pred(25) • MEMRE • MdEMRE • Boxplot of residuals	**Normalization of variables** • MMRE from 0.81 (Gaussian kernel) to 2.08 (sigmoid kernel). • MdMRE from 0.73 (polynomial kernel) to 0.85 (sigmoid and Gaussian kernels). • Pred(25) from 0.08 (linear and sigmoid kernels) to 0.17 (polynomial and Gaussian kernels).	• 4 kernel functions evaluated: linear, polynomial, Gaussian and sigmoid. • 2 techniques to normalize features evaluated: normalization and

		• Boxplot of z	• MEMRE from 2.63 (polynomial kernel) to 14.5 (Gaussian kernel). • MdEMRE from 0.71 (polynomial kernel) to 4.27 (Gaussian kernel) **Log transformation of variables** • MMRE from 0.78 (linear kernel) to 1.84 (sigmoid kernel). • MdMRE from 0.38 (linear kernel) to 0.85 (sigmoid and Gaussian kernels). • Pred(25) from 0.11 (Gaussian kernel) to 0.32 (linear kernel). • MEMRE from 0.79 (linear kernel) to 13.44 (sigmoid kernel). • MdEMRE from 0.53 (linear kernel) to 2.46 (sigmoid kernel)	logarithmic transformation. • These 8 configurations evaluated using hold-out cross validation: 130-65 training/test split. Projects randomly selected. • With variable normalization, best overall results obtained with polynomial kernel. Confirmed with boxplots which were less skewed with a median closer to 0. • With log transformation, best overall results obtained with the linear kernel. Results confirmed by boxplots where the box and tails for the boxplot of the linear kernel are less skewed than the boxplots for the other kernels with a median closer to 0. • Linear kernel with log transformation presents a better MMRE, MdMRE, Pred(25), MEMRE, and MdEMRE than the polynomial kernel with normalization. Difference is statistically significant ($p < 0.01$) with regards to absolute residuals. • Linear kernel/log transformation results compared with results

				obtained from previous study on same dataset using MSWR, CBR, BN, Mean and Median (see S54 – dataset 1). SVR results with linear kernel/log transformation provided significantly superior results to all other techniques.
S72	• Support Vector Regression • Manual Stepwise Regression • Case-based Reasoning • Bayesian Networks	• MMRE • MdMRE • Pred(25) • MEMRE • MdEMRE • Boxplots of residuals and *z*	**Validation Set 1** • Hybrid BN, CBR (1, 2 and 3 analogies) and MSWR all performed statistically better than mean based estimated effort. • Only MSWR performed statistically better than median based estimated effort. MSWR outperformed all techniques (MMRE = 1.5, MdMRE = 0.64, Pred(25) = 23.08, MEMRE = 1.36 and MdEMRE = 0.64). • The Hybrid BN performed significantly better or similar to all methods apart from MSWR (MMRE = 1.9, MdMRE = 0.86, Pred(25) = 15.38, MEMRE = 13.06 and MdEMRE = 2.38). • CBR 2 and 3, and both BNs performed similarly to median based estimated effort. • CBR 1 showed significantly worse prediction accuracy than median based estimated effort (MMRE = 5.27, MdMRE = 0.97, Pred(25) = 7.69, MEMRE = 31.70, MdEMRE = 3.43). • Similar trends observable using boxplot of residuals. • Using boxplot of z-values, it was the hybrid BN that significantly outperformed all other techniques. • Only CBR1 and the hybrid BN outperformed median based effort estimation. **Validation Set 2** • All techniques performed significantly better than Mean estimated effort when looking at residuals.	**SVR** • 6 configurations; 3 dealing with effort as the dependent variable and 3 dealing with the inverse of effort as the dependent variable. • For effort the 3 configurations are: ○ No transformation of variables. ○ Normalization of features. ○ Log transformation of the effort and features. • For inverse effort the 3 configurations are: ○ Normalization of features. ○ Log transformation of features. ○ Log transformation of effort and features. • All 6 configurations applied to 4 different kernels: linear,

			• Only MSWR was significantly more accurate than median based estimated effort (MMRE = 0.73, MdMRE = 0.66, Pred(25) = 10.77, MEMRE = 2.86 and MdEMRE = 1.21). • Median based estimated effort was significantly superior to all other techniques apart from MSWR. (MMRE = 4.95, MdMRE = 0.89, Pred(25) = 15.38, MEMRE = 4.62 and MdEMRE = 0.78). • The data-driven BN performed better than the hybrid BN. Its performance was similar to those of all CBR-based predictions. (MMRE = 4.09, MdMRE = 0.96, Pred(25) = 1.54, MEMRE = 7.90 and MdEMRE = 0.93). • The hybrid BN was significantly worse than all techniques apart from mean based estimation. (MMRE = 27.95, MdMRE = 5.31, Pred(25) = 3.08, MEMRE = 1.34 and MdEMRE = 0.90). • Similar trends observable with boxplot of residuals. • Similar trends also observable with boxplot of z values, except CBR1 in addition to MSWR presented significantly superior predictions as compared to median based prediction.	polynomial, Gaussian and Sigmoid. **MSWR** • See S59 **CBR** • See S59 **BN** • See S59 **Mean and median** • See S59 **Comparisons** Comparisons between SVR and other techniques done using two 130 (training) to 65 (test) splits.
S74	• Linear regression	• MMRE • MdMRE • Pred(25) • Boxplot of absolute residuals	• MMRE = 0.17 • MdMRE = 0.11 • Pred(25) = 0.80	• Web applications sized using COSMIC-FFP. • Regression models evaluated using leave 1 out cross-validation. • A comparison was made against a previous study with academic data (see S35) • Boxplot of residuals has a smaller box length and tails as compared to the boxplot

				obtained in the previous study. Both boxplots have an outlier and close medians. • No significant difference between estimation accuracy for the 2 studies.
S75	• Web-COBRA • Mean model • Median model	• MMRE • MdMRE • Pred(25)	**Web-COBRA (with Web-Objects)** • MMRE = 0.14, MdMRE = 0.11, Pred(25) = 0.87 **Mean model** • MMRE = 0.34, MdMRE = 0.27, Pred(25) = 0.47 **Median model** ○ MMRE = 0.33, MdMRE = 0.24, Pred(25) = 0.60	• Web applications sized with Web Objects. • Monte Carlo simulation used to obtain a probability distribution for the effort required for the new project. • Mean value of distribution used as estimated effort value. • Mean and median models used as a benchmark for comparison.
S76	• Manual stepwise regression • CBR/Analogy • Regression trees in conjunction with CBR or MSWR • Mean model • Median model	• MMRE • MdMRE • Pred(25) Boxplot of residuals	**Size Measures set 1** • MSWR: MMRE = 12%, MdMRE = 11%, Pred(25) = 87% • CBR (only best results provided): $CBR_{2C\text{-}FSS}$: MMRE = 19%, MdMRE = 9%, Pred(25) = 87%. $CBR_{2C\text{-}PC}$: MMRE = 25%, MdMRE = 17%, Pred(25) = 67%. • RT with CBR (only best result provided): $RT+CBR_{2C\text{-}FSS}$: MMRE = 11%, MdMRE = 9%, Pred(25) = 93%. • RT with MSWR: results not provided, “did not provide good models”. **Size Measures set 2** • MSWR: MMRE = 23%, MdMRE = 21%, Pred(25) = 73% • CBR (only best results provided): $CBR_{2B\text{-}FSS}$: MMRE =	**MSWR** • MSWR model built using data from all 15 projects. **CBR** • ANGEL tool used. • Unweighted Euclidean distance measure. • 1, 2 and 3 analogies. • Mean, inverse distance mean and inverse rank weighted mean adaptation. • FSS feature of ANGEL tool used. As an alternative features were selected using Pearson’s Correlation.

			11%, MdMRE = 10%, Pred(25) = 93%. CBR_{2A-PC}: MMRE = 24%, MdMRE = 17%, Pred(25) = 87%. • RT with CBR (only best result provided): $RT+CBR_{1A-FSS}$: MMRE = 19%, MdMRE = 20%, Pred(25) = 66%. • RT with MSWR: results not provided, “did not provide good models”. • Mean: MMRE = 0.34, MdMRE = 0.27, Pred(25) = 0.47 Median: MMRE = 0.33, MdMRE = 0.24, Pred(25) = 0.60	**Regression Trees** • Regression trees were used in conjunction with CBR and MSWR. • Regression tree technique used to split dataset into 2 based on a particular variable (number of server-side scripts for size measure set 1, and number of external references for size measure set 2). • CBR (or MSWR) used for effort estimation based on the remaining variables and split dataset. • Mean and median based effort estimation used as a benchmark. • All techniques evaluated using leave one out cross validation. • MSWR results significantly (residuals) better with size measure set 1 than 2. MSWR was the only model for which the difference in accuracy between the 2 sets of size measures was significant. • For CBR, A = mean, B = inverse distance mean, and C = inverse rank weighted

				mean. FSS = Feature subset selection, and PC = Pearson's Correlation. • Results with FSS better than results with PC, but only significantly better with size measure set 2. • For size measure set 2, CBR with FSS provided significantly better results than MSWR, CBR with PC and RT with CBR. • For size measure set 1, MSWR provided significantly better results than CBR with PC, and RT with CBR with FSS also provided significantly better results than CBR with PC. • In comparisons with mean and median estimation, for set 1, MSWR and RT with CBR with FSS provided significantly superior estimates to both mean and median estimation. • For set 2, CBR with FSS, and RT with CBR with FSS provided significantly superior estimates to both mean and median estimation. • Results confirmed by boxplots of absolute residuals.

S77	• Web-COBRA in conjunction with COSMIC • Mean model • Median model	• MMRE • MdMRE • Pred(25)	• MMRE of 0.11 • MdMRE of 0.10 • Pred(0.25) of 0.93 • Results obtained were significantly superior to both mean (MMRE of 0.34, MdMRE of 0.27 and Pred(0.25) of 0.47) and median (MMRE of 0.33, MdMRE of 0.24 and Pred(0.25) of 0.60) based estimation. • Practical significance also assessed using effect size which showed that the improvement in accuracy provided by Web-COBRA with COSMIC was practically significant as well.	• Leave one out cross validation used. Mean and median models used as a benchmark for
S80	• Linear regression • CBR/Analogy	• NA	• NA	• Paper describes how the proposed model would be used and its results interpreted. Evaluation of model not done.
S81	• Stepwise regression • Mean • Median	• MMRE • MdMRE • Pred(25)	**SWR** • MMRE – 129% • MdMRE – 73% • Pred(25) – 17.24% **Mean** • MMRE – 4314% • MdMRE – 1413% • Pred(25) – 6.89% **Median** • MMRE – 663% • MdMRE – 149% • Pred(25) – 3.44%	• Manual stepwise regression • 1 fold cross validation used – 66% training set, 34% test set
S82	• Linear regression • Mean model • Median model	• MMRE • MdMRE • Pred(25) • Boxplot of absolute residuals	**OO-HFP** • MMRE = 0.10, MdMRE = 0.08, Pred(25) = 0.93. **FPA** • MMRE = 0.11, MdMRE = 0.10, Pred(25) = 0.87 **Mean** • MMRE = 0.47, MdMRE = 0.28, Pred(25) = 0.49	• Effort is the dependent/response variable, and size either in OOHFP or FPA the independent/predictor variable.

			Median • MMRE = 0.35, MdMRE = 0.23, Pred(25) = 0.62	• OO-HFP measured automatically using the VisualWADE tool. • FPA "measured manually by a certified function point specialist using IFPUG counting practices manual". • Mean and median effort estimation used as a benchmark for comparisons. • Models validated using leave 1 out cross-validation. • SPSS15.0.1 used for regression model and other statistical analysis. • OO-HFP counted automatically using VisualWADE tool with the requirements specification. • FPA counted manually at the implemented Web application level (so that the FPA count would be "as accurate as possible for comparison purposes"). • Boxplots of residuals confirm results of summary statistics. Absolute residuals when using FPA are larger, with the boxplot presenting a slightly larger box when compared to OO-HFP. • OO-HFP presents significantly better prediction accuracy than FPA in terms

				of absolute residuals (95% confidence).
S83	• Bayesian Network	• NA	• NA	• Aim of this research was to aggregate causal maps from 6 Web companies to look at predictors that have a causal effect upon Web effort estimation. • Causal maps qualitative part of Bayesian Network. • All causal maps expert driven
S84	• Support Vector Regression	• MMRE • MdMRE • Pred(25) • MEMRE • MdEMRE • Boxplot of residuals • Boxplot of z	**Dataset 1** • MMRE from 0.59 to 1.989, MdMRE from 0.339 to 0.663, Pred(25) from 0.234 to 0.391, MEMRE from 0.689 to 1.397, MdEMRE from 0.365 to 0.735. **Dataset 2** • MMRE from 0.856 to 2.733, MdMRE from 0.455 to 0.954, Pred(25) from 0.200 to 0.400, MEMRE from 0.498 to 2.855, MdEMRE from 0.410 to 0.716. •	• Linear and RBF kernels evaluated. • No and logarithmic preprocessing of data investigated. • Tabu search used to optimize SVR parameter settings. • These results were compared to the use of SVR with Tabu search optimizing parameter settings (S72). • For both validation sets SVR and TS+SVR produced comparable results with the linear kernel. • For both validation sets, when the RBF kernel was used TS+SVR provided better results, significantly so without data preprocessing.

				• For both validation sets when using Tabu search, best results obtained with logarithmic preprocessing, RBF kernel and Tabu search optimization (significantly better based on absolute residuals). Confirmed by boxplot of residuals. • Results obtained for TS+SVR also compared against other effort estimation techniques (S59). For both validation sets TS+SVR significantly superior to all other techniques including mean and median effort estimation (based on absolute residuals). Confirmed by boxplot of residuals.

S85	• Tabu search	• MMRE
		• MdMRE
		• Pred(25)
		• MEMRE
		• MdEMRE
		• Boxplot of residuals
			Validation set 1
			• MMRE objective function
			  ○ MMRE = 0.75, MdMRE = 0.76, Pred(25) = 0.14, MEMRE = 6.31, MdEMRE = 0.51
			• MdMRE objective function
			  ○ MMRE = 1.37, MdMRE = 0.68, Pred(25) = 0.29, MEMRE = 1.77, MdEMRE = 0.51
			
			Validation set 2
			• MMRE objective function
			  ○ MMRE = 0.80, MdMRE = 0.66, Pred(25) = 0.18, MEMRE = 4.72, MdEMRE = 1.42
			• MdMRE objective function
				• Effort estimation treated as an optimization problem.
				• Effort model described by equation derived from going through the search space using Tabu search which generates values for factor coefficients, equation constant and operators (+, -, *, /).
				• Two objective functions evaluated: accuracy as measured my MMRE or

			• MMRE = 0.99, MdMRE = 0.49, Pred(25) = 0.31, MEMRE = 1.81, MdEMRE = 0.54	MdMRE. • Tabu search using MdMRE as the objective function (TS2) produced the best results based on accuracy measures (apart from MMRE) for both validation sets. • Difference between results significant (based on absolute residuals with 95% confidence). • Confirmed by boxplot of residuals where median for TS2 closer to 0 and box length and tails less skewed. • Tabu search compared with results obtained using other estimation techniques (S54). • Predictions obtained with TS2 significantly superior to all other techniques with regards to the second validation set (based on absolute residuals with 95% confidence).. • Predictions obtained with TS2 significantly superior to all techniques apart from MSWR with regards to the first validation set (based on absolute residuals with 95% confidence).

				• Only MSWR and TS2 provided significantly superior accuracy than median effort estimation (based on absolute residuals with 95% confidence).
S87	• WEBMO+ • VPM+	• MMRE	Effort • WEBMO+ MMRE = 5.0% • VPM+ MMRE = 7.7% Duration • WEBMO+ MMRE = 10.0% • VPM+ MMRE = 16.02%	• WEBMO+ is an adaptation of WEBMO; an algorithmic effort estimation technique which is an extension of COCOMO 2.0 specifically for Web projects (hypermedia and applications). ○ Uses an estimate of the source lines of code based on the number of external use cases to size web applications. • VPM+ is an extension of VPM (Vector Prediction Model). Software size is measured as the magnitude of a 2 dimensional vector of project complexity and functionality. • In the extension magnitude is calculated using Web Objects.
S89	• Radial Basis Function Neural Networks	• MMRE • Pred(25)	**C-Means Sigma-Saha** • MMRE = 34.04% • Pred(25) = 64.15% **C-Means Sigma Haykin**	• 3 variants tested • 2 fuzzy clustering techniques based on the C-means algorithm.

			• MMRE = 37.34% • Pred(25) = 71.70% **K-Means** • MMRE = 79.81% • Pred(25) = 52.83%	• 1 hard clustering algorithm, K-means • Accuracy obtained dependent on the number of hidden neurons used in RBFN. Accuracy improves as number of hidden neurons increase. 30 used in this case. • Different width values (dependent on clustering algorithm) also effect accuracy.
S90	• Support vector regression • Manual stepwise regression • Case-based reasoning	• MMRE • MdMRE • Pred(25) • MEMRE • MdEMRE • Boxplot of residuals • Boxplot of z	**Split 1** • MMRE from 0.591 to 2.423. • MdMRE from 0.411 to 0.777. • Pred(25) from 0.188 to 0.344. • MEMRE from 0.603 to 4.49. • MdEMRE from 0.467 to 0.781. **Split 2** • MMRE from 0.678 to 4.38*10E5. • MdMRE from 0.36 to 0.826. • Pred(25) from 0.077 to 0.415. • MEMRE from 0.506 to 8.609. • MdEMRE from 0.41 to 0.886. **Split 3** • MMRE from 0.645 to 1.7*10E5. • MdMRE from 0.319 to 0.844. • Pred(25) from 0.138 to 0.415. • MEMRE from 0.824 to 5.778. • MdEMRE from 0.326 to 1.051.	**SVR** • 18 separate configurations investigated. • No transformation of features, normalization of features and log transformation of features. • No kernel transformation (linear), polynomial and RBF kernels. • Effort and inverse effort response/dependent variables. • Choice of parameters determined by semi-automatic approach investing a wide range of values for the variables using leave one out cross validation on the training set. **Stepwise Regression**

			Test Set 1	• 3 regression models built.

Test Set 1

	MMRE	MdMRE	Pred(25)	MEMRE	MdEMRE
SVR with C3(R)	0.59	0.41	0.34	0.60	0.47
MSWR	1.50	0.64	0.23	1.36	0.64
CBR1	5.27	0.97	0.08	31.70	3.43
CBR2	5.06	0.87	0.11	3.59	0.81
CBR3	5.63	0.97	0.09	4.17	0.88
Median Effort	4.84	0.92	0.09	4.48	0.94
Mean Effort	29.28	3.94	0.08	1.07	0.91

Test Set 2

	MMRE	MdMRE	Pred(25)	MEMRE	MdEMRE
SVR with C3(R)	0.91	0.36	0.42	0.51	0.41
MSWR	0.73	0.66	0.11	2.86	1.21
CBR1	4.46	0.92	0.08	21.81	0.95
CBR2	6.73	0.89	0.15	15.65	0.90
CBR3	6.09	0.84	0.09	13.26	0.89
Median Effort	4.21	0.85	0.14	5.57	0.86
Mean Effort	26.21	4.71	0.05	1.39	0.90

Test Set 3

	MMRE	MdMRE	Pred(25)	MEMRE	MdEMRE
SVR with C3(R)	1.10	0.33	0.42	1.00	0.35
MSWR	1.24	0.51	0.26	1.50	0.54
CBR1	9.68	0.95	0.08	12.76	0.96
CBR2	7.06	0.91	0.08	5.13	0.88
CBR3	16.41	1.90	0.14	2.33	0.93
Median Effort	2.63	0.91	0.19	6.24	0.86
Mean Effort	14.45	3.74	0.03	1.47	0.91

•

- 3 regression models built.
- Variables with more than 40% of their values equal to 0 or missing were excluded.
- Transformation required due to variables having non-normal distributions.
- High influence data points removed to improve model stability.

Case-based Reasoning

- CBR-Works tool used.
- Unweighted Euclidean distance similarity measure
- 1,2, and 3 analogies
- Mean adaptation
- Feature subset manually decided using Spearman's rank correlation test.
- All 195 cases used as case base.

SVR configurations

- When variables are not transformed best results obtained from Linear and Polynomial kernels (based on summary measures) for the 1st set, and the Polynomial kernel for the 2nd and 3rd sets.
- When variables were normalized Linear kernel produced better results

				(based on summary measures) than the other 2 kernels for the 1st set, but the Polynomial and RBF kernels produced better results for the 2nd and 3rd set respectively. • When variables are log transformed, the best results in terms of summary measures obtained using the RBF kernel for all three sets (although only statistically significant for the 1st and 2nd sets). • With inverse effort and no transformation the best results obtained by the RBF kernel. • With inverse effort and normalization the best results were obtained by the RBF and Polynomial kernels for the 1st and 2nd sets respectively (based on summary measures), with RBF providing the best results for the 3rd set (apart from MMRE). • With inverse effort and log transformation the best results were given by the RBF kernel. • The best overall results were obtained by using

				effort, log transformation of variables and the RBF kernel for the first 2 sets confirmed by boxplots. For the 3rd set RBF also produced the best results although there were other comparable configurations. • For the 1st test case the above RBF configuration provided better results than all other configurations. • For the 2nd and 3rd test cases, the above RBF configuration provided better results for all configurations except normalization of features/effort/polynomial kernel, normalization of features/inverse effort/RBF, and log transformation of features/inverse effort/RBF. However in terms of summary measures the above configuration provides better results than all configurations apart from normalization of features/effort/polynomial kernel in the third set. **Comparison of best SVR config with MSWR and CBR** • Best SVR config as stated above is taken to be Log

<table>
<tr><td></td><td></td><td></td><td></td><td>transformation/Effort/ RBF kernel.
• All predictions obtained with SVR were significantly superior than those obtained with all other techniques using all 3 sets.
• SVR and MSWR were the only 2 estimation techniques to provide significantly better estimates than median effort estimation.
• MSWR significantly outperformed all 3 CBR configurations.
• No difference found between the 3 CBR configurations.</td></tr>
<tr><td>S91</td><td>• SVR with Tabu search (to configure parameters)
• Stepwise regression (manual)
• CBR
• Mean
• Median</td><td>• Median of Absolute Residuals (MdAR)</td><td><table>
<tr><td>TUKUTUKU repository</td><td></td><td></td><td></td><td></td></tr>
<tr><td>DocProNo</td><td>26</td><td>41</td><td>33</td><td>43</td></tr>
<tr><td>DocProYes</td><td>33</td><td>54</td><td>63</td><td>56</td></tr>
<tr><td>Enhancement Projects</td><td>17</td><td>24</td><td>24</td><td>18</td></tr>
<tr><td>MetricNo</td><td>36</td><td>57</td><td>53</td><td>57</td></tr>
<tr><td>MetricYes</td><td>12</td><td>24</td><td>15</td><td>15</td></tr>
<tr><td>NewProjects</td><td>47</td><td>49</td><td>60</td><td>57</td></tr>
<tr><td>ProImprYes</td><td>27</td><td>29</td><td>48</td><td>36</td></tr>
<tr><td>ProImprNo</td><td>41</td><td>52</td><td>52</td><td>56</td></tr>
</table></td><td>• SVR + TS is compared against:
○ SVR with parameters chosen randomly
○ SVR with default Weka parameters
○ SVR with grid search parameters
• RBF Kernel employed for SVR along with logarithmic pre-processing.
• CBR with feature selection and no feature selection investigated.
• 1,2 and 3 analogies considered for adaptation
• 10 fold cross validation</td></tr>
</table>

			38 64 (k=2) 40 (k=2) 287 61 63 64 (k=2) 65 (k=2) 533 81 24 50 (k=3) 27 (k=3) 192 25 65 60 (k=2) 50 (k=2) 532 90 14 21 (k=1) 24 (k=2) 189 164 56 57 (k=3) 43 (k=3) 548 87 39 34 (k=3) 28 (k=3) 164 53 49 56 (k=3) 54 (k=1) 646 84	used. • SVR+TS was significantly superior to SVR+Rand for all Tukutuku datasets. • SVR+TS was significantly superior to SVR+Weka for all but 2 Tukutuku datasets. • SVR+TS was significantly superior to SVR+Grid for all but 1 Tukutuku dataset. • SVR+TS was significantly superior to SWR for all Tukutuku datasets. • SVR+TS was significantly superior to CBR (feature selection) for all Tukutuku datasets. • SVR+TS was significantly superior to CBR (no feature selection) for all but 1 Tukutuku dataset. • SVR+TS was significantly superior to mean and median effort estimation for all Tukutuku datasets.
S92	• Web component model	• MMRE • Pred(25)	• MMRE = 0.07398 • Pred(25) = 0.8	• Extension of Cocomo II model. • Web application sized with function points plus the 4 additional Web specific size measures. • Size along with cost drivers, category of application and language expansion factor

				used to calculate effort.
S93	• Ordinary least squares regression. • Web-COBRA • CBR • Mean • Median	• MMRE • MdMRE • Pred(25)	**OLSR** - Web Objects • MMRE = 0.21 • MdMRE = 0.15 • Pred(25) = 0.7 - Function Points • MMRE = 0.46 • MdMRE = 0.28 • Pred(25) = 0.40 **Web COBRA** - Web Objects • MMRE = 0.18 • MdMRE = 0.12 • Pred(25) = 0.80 - Function Points • MMRE = 0.29 • MdMRE = 0.25 • Pred(25) = 0.50 **CBR** - Web Objects • MMRE = 0.22 • MdMRE = 0.12 • Pred(25) = 0.70 - Function Points • MMRE = 0.49 • MdMRE = 0.17 • Pred(25) = 0.60 **Mean** - Web Objects • MMRE = 0.63	• Hold out cross validation performed. ○ Training set - 15 apps ○ Test set – 10 apps ○ Projects in test set developed after projects in training set. • Applications sized using Web Objects and Function Points. • ANGEL tool used for CBR ○ 1, 2 and 3 analogies ○ Mean, inverse distance and inverse rank weighted mean of k analogies used to generate estimate. • Statistical differences evaluated using T-test or Wilcoxon signed rank test (if data not normal). • r value used to look at “size” of effect. r = 0.2 (small), r = 0.5 (medium), r = 0.8 (large). • For OLSR, WO significantly better than FP. • For Web COBRA, WO significantly better than FP. • For CBR best results obtained using 2 analogies

			• MdMRE = 0.37 • Pred(25) = 0.40 **Median** • MMRE = 0.68 • MdMRE = 0.34 • Pred(25) = 0.40	with mean adaptation. • While WO results better, not significantly so. • No significant differences between the three methods using WO. • All WO and FP results significantly better than mean and median estimates.
S96	• Custom – company specific productivity coefficient used to calculate effort, using project size.	• MRE • Pred(25)	• Summary stats were given for MRE including MMRE and MdMRE **FP** • MMRE = 0.49 • MdMRE = 0.19 • Pred(25) = 62% **WO** • MMRE = 1.23 • MdMRE = 0.66 • Pred(25) = 40% **RWO** • MMRE = 0.45 • MdMRE = 0.19 • Pred(25) = 58%	• Revised Web Objects • Uses size measures (described below) and their associated complexity weighting. • Once size has been calculated, this is used in conjunction with level of reuse and productivity of tools to classify the application. Classification system derived from domain expert. • The classification is used to assign weights that are used in conjunction with application size to estimate effort using a company specific productivity coefficient. • Effort also estimated using FP and WO as size measures
S97	• Bayesian Network	• NA	• NA	• Structural development

				done by domain expert. • Information for parameter information either elicited from DE or obtained from past projects. • Model validation handled by walkthrough and predictive accuracy • Case study considers the creation of a Bayesian Network using a DE. • The model was validated using the 22 past projects. • Estimate with highest probability in effort distribution (outcome of the BN) is compared to actual effort required. • If they do not match, BN recalibration was performed. Happened only 1 time out of 22 cases.
S98	• Factor analysis with structural equation modeling to determine factors related to maintenance and how they relate to each other. • Maintenance estimation obtained using either maximum likelihood or	• MMRE	• Maximum likelihood MMRE = 47.58%. • Bayesian analysis MMRE = 44.08%	• Predictors obtained from literature. • Grouped into 4 categories using factor analysis. • These 4 categories used to model maintenance cost using structural equation modeling. • Effort predicted by model using either maximum likelihood or Bayesian analysis

	Bayesian analysis			• Mention of accuracy also made where accuracy = 100 – MMRE (in %) • Training set – 192 projects, test set – 30 projects.

E1	• Average unit cost model
	• Linear regression
	• Non-linear regression
	• Multiple regression (linear and multiplicative)
		• MMRE
		• Pred(25)
			Average unit cost model
			• Full sample: MMRE = 81.3%, Pred(25) = 20%.
			• When sample subdivided based on programmer experience (into Advanced, intermediate and novice):
			  ○ Advanced (3 projects): MMRE = 35.6%, Pred(25) = 0%.
			  ○ Intermediate (6 projects): MMRE = 28.7%, Pred(25) = 33.3%.
			  ○ Novice (6 projects): MMRE = 111.6%, Pred(25) = 16.66%.
			
			Simple Linear Regression
			• Model built using functional size (in csfu) as the independent variable.
			• R^2 of 0.38, accuracy not given.
			
			Non-Linear Regression
			• Model built using functional size (in csfu) as the independent variable.
			• Logarithmic transformation used.
			• R^2 of 0.716, with an MMRE of 26.66%, and a Pred(25) of 26.66%
			
			Multiple Regression
			• First evaluated using functional size and staff experience (3 level categorical variable).
			  ○ 2 dummy variables designed to represent categorical variable.
			  ○ Linear model
			  ○ R^2 of 0.62 with coeffficients for variables
				○ MMRE referred to in this paper as MRE (mean relative error).

			representing staff experience not being significant. No accuracy results provided ("makes no positive contribution over the linear regression model with a single variable). • Evaluated next using functional size and project difficulty (3 level categorical variable). ○ 2 dummy variables designed to represent categorical variable. ○ Linear model ○ R^2 of 0.66 with coefficients for variables representing project difficulty not being significant. No accuracy results provided. • Lastly evaluated using functional size and a 2 level categorical variable representing both staff experience and project difficulty. ○ Multiplicative model. • MMRE of 45%, and Pred(25) of 53%.	
E2	• Stepwise multiple regression	• MMRE • Pred(25)	Stepwise multiple regression achieved a MMRE of 56% and a Pred(25) of 25%.	•
E3	• Use Case Points (UCP)	• NA	○ NA	• Estimated test effort of 367 man days compared to the actual effort of 390 man days.
E4	• NA	• NA	• NA	• Exploratory study; design to investigate a series of hypotheses empirically. • Ordinary Least Squares regression used to investigate relationships between dependent variables and independent variables. • Findings were:

				○ Median of actual effort to build the information model is significantly higher than the median of the actual effort to build the navigation model. ○ For each effort category, median of actual effort is significantly higher than median of estimated effort (i.e. tendency for underestimation). ○ There is a positive correlation between estimated effort and actual effort. Estimated effort can be used as a predictor for actual effort. ○ A significant relationship does not exist between the subject's technical knowledge and estimated learning

				effort, actual learning effort, estimated effort, and actual effort. ○ There is a significant relationship between a subject's knowledge and the difference between actual and estimated effort. ○ There is a negative correlation between a subject's average grade and the estimated learning effort. There is also no correlation between a subject's average grade and actual learning effort. ○ In other words subjects with a higher proficiency tended to believe that they would have an advantage in learning activities that did not materialize.

				○ There is a significant positive correlation between average grade and actual effort. ○ There is a positive correlation between estimated learning effort, and actual learning effort, but a poor R^2 indicates that estimated learning effort would not be a good predictor. ○ For each effort category there is a positive correlation between learning effort and the corresponding modeling effort (e.g. estimated navigation learning effort and estimated navigation effort). All R^2 values are low indicating that the learning effort would be a poor predictor of

				modeling effort. • No correlation between self-assessment of assignment and actual marks obtained.
E5	• Generalised linear model (GLM)	• NA	• NA	
E6	• Generalised linear model (GLM)	• NA	• NA	
E7	• Bayesian Network	• NA	• NA	• Expert elicited BN Data from 8 past projects used to calibrate BN model during model validation phase.