Table 3. Mean outcomes for NE, LE and BSE

	NE	**LE**	**BSE**
Mean error	0.1882	0.1849	0.1875
Mean bias	0.1358	0.1315	0.1310
Mean variance	0.0525	0.0534	0.0565

Table 4. Win/Draw/Loss: LE vs. alternative

	NE		**BSE**	
	W/D/L	p	W/D/L	p
Error	29/13/14	0.0158	35/2/19	0.0201
Bias	40/13/3	<0.0001	17/2/37	0.0045
Variance	10/14/32	0.0005	44/2/10	<0.0001

alternative algorithms is shown in Table 5. The p value is the outcome of a one-tailed binomial sign test. We assess a difference as significant if $p \leq 0.05$.

Figure 2 graphs the relative error, bias and variance of the three classifiers. The values on the y-axis are the outcome for BSE divided by that for NE. The values of the x-axis are the outcome for LE divided by that for NE. Each point on the graph represents one of the 56 data sets. Points on the left of the vertical line at LE/NE=1 in each subgraph are those for which LE has better results than NE. Points below the horizontal line at BSE/LE=1 indicate that BSE wins in those domains compared with NE. Points above the line $X=Y$ represent that LE has lower values than those of BSE.

Considering first the error outcomes, LE achieves the lowest mean error. The win/draw/loss record indicates that LE has a significant advantage over NE and BSE. However, there is no significant error difference between NE and BSE. From the error graph in Figure 2, we can see that the majority of the points are on the left of the vertical line at LE/NE=1, and above the line $X=Y$. The error ratios of LE and BSE over NE on King-rook-vs-king-pawn (the point at the bottom of the graph) are 0.8478 and 0.6237 respectively. The error ratio of BSE over LE is 0.7357. That is, both LE and BSE reduce the error of NE consider-

Table 5. Win/Draw/Loss: BSE vs. alternative

	NE		**LE**	
	W/D/L	p	W/D/L	p
Error	21/3/32	0.0845	19/2/35	0.0201
Bias	47/2/7	<0.0001	37/2/17	0.0045
Variance	7/3/46	<0.0001	10/2/44	<0.0001

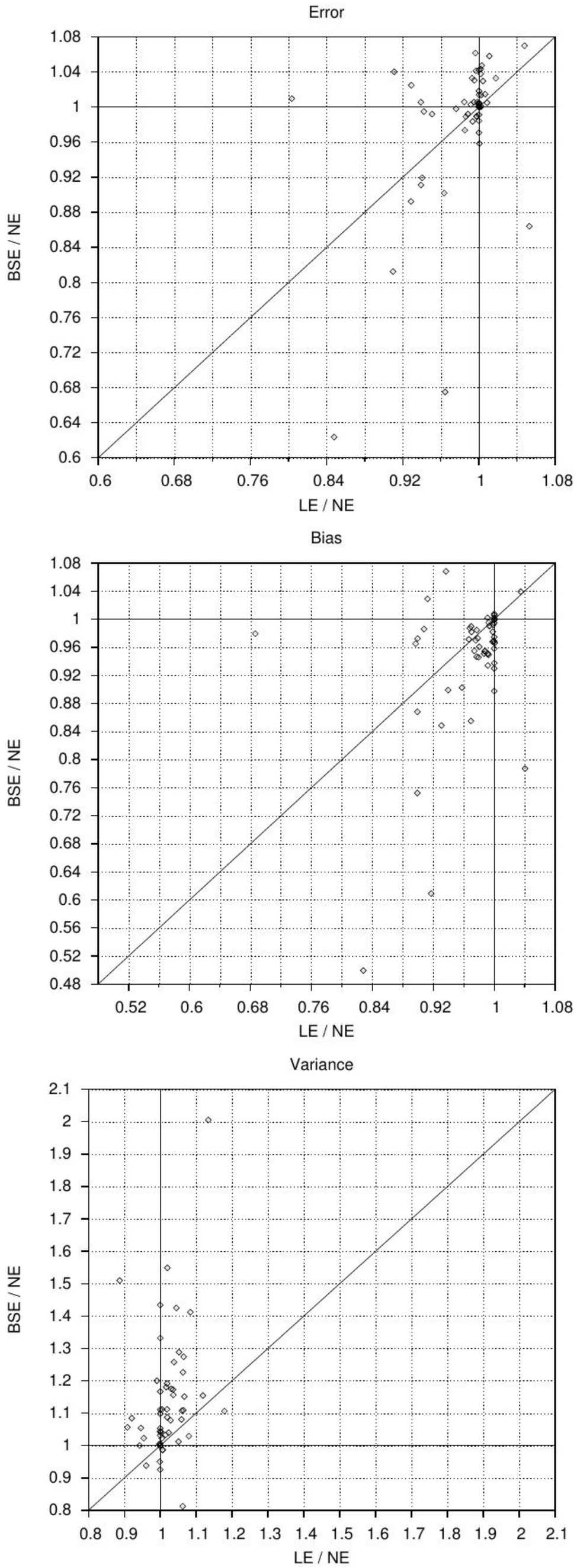

Figure 2. Comparison of Error, Bias and Variance