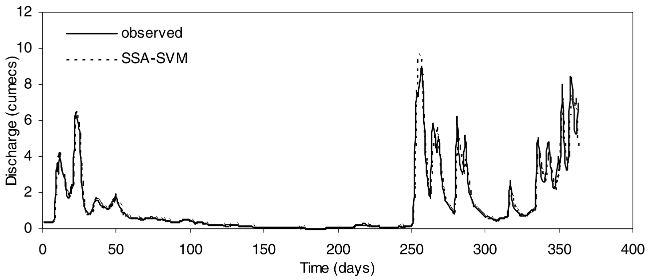
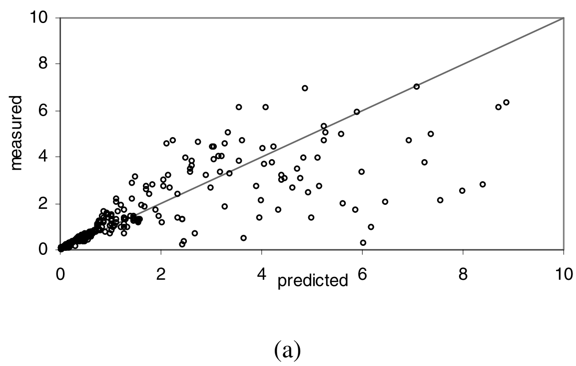


Figure 7 | Comparison between measured and the predicted flows with SSA–SVM: verification.



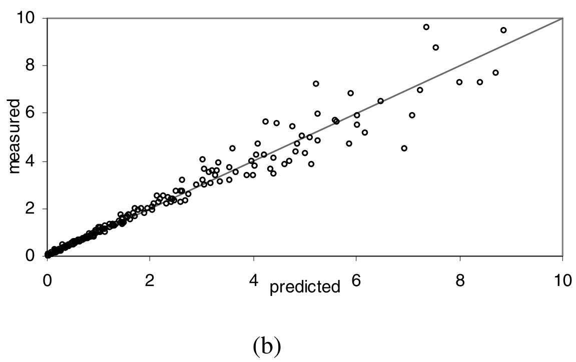


Figure 8 | Scatter plots of verification data for (a) NLP and (b) SSA–SVM.

Figure 8(a, b) show the scatter plots for the verification data with NLP and SSA–SVM, respectively.

CONCLUSIONS

In this study, it has been demonstrated that the proposed approach, SSA–SVM, could yield significantly higher prediction accuracy of hydrologic variables than that of the non-linear prediction (NLP) method. SSA–SVM results in a significant improvement in the case study on Singapore rainfall prediction with a correlation coefficient of 0.70 as opposed to 0.51 obtained by NLP. Similarly, SSA–SVM yields 58.75% improvement (in terms of *RMSE*) over NLP in the runoff prediction for Tryggevælde catchment.

Moreover, the predictions from SVM offer special advantages as compared to other machine learning techniques like ANN. Unlike ANN, SVM does not require the architecture to be defined *a priori*. The structural risk minimization principle gives SVM the desirable property to generalize well in the unseen data. The dual representation offers the unique advantage of ease in dealing with the high-dimensional input vectors without loss of both generalization accuracy and computational efficiency. The optimization problem formulated for SVM is always uniquely solvable and, thus, does not suffer from the limitation of ways of regularization as in ANN, which may lead them to local minima.

ACKNOWLEDGEMENTS

The authors wish to thank Mr Marco Sandri of Centro di Calcolo–C.I.C.A., Università di Verona, Italy for his constructive discussions on SSA. Also, the authors wish to express their sincere thanks to Danish Hydraulic Institute (DHI) and Meteorological Service Singapore for providing the daily runoff data of Tryggevælde catchment and Singapore rainfall data, respectively.

NOTATION

The following symbols are used in this paper:

$R(\beta)$ = actual risk

$R_{emp}(\beta)$ = empirical risk

Ω = confidence interval

h = VC dimension

x = input data