Many areas of science and technology dealing with complex processes frequently face the problem of system identification: specifying an optimal model structure which can explain the system behavior in an appropriate manner. In essence, this involves identifying a few meaningful predictors (either as values of previous time lags under assumptions of Markovian dependence, or exogenous variables in systems impacted by external forcings and factors) and formulating a relation between these identified predictors and the system response. Here, we introduce partial information (PI) and residual variance (RV), as two statistical measures for specifying the system using empirical principles, be it a hydrologic system or any natural system that can be expressed in a model form. The PI is a modification of the Partial Mutual Information (PMI) which has been used in hydrological and non-hydrological settings to identify predictor variables, whereas the RV offers a means to specify the partial weightage (equivalent to a classical regression coefficient in a multiple regression framework) when the model is formulated without imposing a defined (linear or nonlinear) structure. In contrast to other algorithms for nonlinear system specification, the proposed framework efficiently provides optimal number of predictors along with the model structure as specified using nonparametric regression and estimation principles. The utility of the approach is demonstrated using synthetically generated linear, non-linear and a high-dimensional dataset, along with a chaotic data set which has been studied extensively for model estimation purposes in the past. Implications of using the proposed framework in a hydrologic setting are discussed.