Uncertainty in prediction and simulation of flow in sewer systems

Models are commonly applied for design of urban drainage systems. Typically they are of deterministic nature although it is well accepted that they only reflect reality approximately. When measurements are available they can be used for calibration of models. However, deviations between model outputs and observations will often remain and should hence be quantified, especially when used for model predictive control. The objective with this thesis has been to quantify and qualify the modelled output uncertainty.

For this purpose a catchment in Ballerup (1,320 hectares) was selected and data included flow from downstream of the catchment, rain measured at two rain gauges and monthly evaporation. The data period covered subperiods of 2007-2010. The catchment area consists of both combined and separated drainage systems and significant infiltration inflow enters the system through permeable surface areas. The simple serial linear reservoir flow routing principle was applied for modelling both the fast rainfall runoff from paved areas and the slow infiltration inflow from permeable areas. The wastewater flow variation was modelled by a harmonic function. Models of different complexity in terms of describing features such as flow constraints, basins and pumps were tested for their ability to describe the output with a time resolution of 15 minutes.

Two approaches to uncertainty quantification were distinguished and adopted, the stochastic and the epistemic method. Stochastic uncertainty refers to the randomness observed in nature, which is normally irreducible due to the inherent variation of physical systems. Epistemic uncertainty on the contrary arises from incomplete knowledge about a physical system. For quantifying stochastic uncertainties a frequentist approach was applied whereas the generalised likelihood uncertainty estimation method (GLUE) was adopted for the epistemic approach. Two different uncertainty estimates were furthermore distinguished: prediction and simulation uncertainty. To quantify the prediction uncertainty the model should accommodate an updating step thereby benefiting from observations and avoid in continuation of the predictions made. The simulation uncertainty on the other hand is calculated from data of a limited measuring campaign and the model does not accommodate a model correction step. The stochastic approach was applied for uncertainty quantification in both prediction and simulation whereas the epistemic uncertainty was assessed only in simulation. A maximum likelihood method was applied for parameter estimation in the stochastic approach, i.e. one optimal parameter set was derived that minimises the errors between model outputs and observations. Conversely in GLUE, parameters are viewed as stochastic variables and many acceptable parameter sets were therefore identified.

The predictive stochastic models were built on stochastic differential equations that include a drift term containing the physical description of the model and a diffusion term describing the uncertainty in the state variables. Additionally the observation noise is accounted for by a separate observation noise term. This approach is also referred to as stochastic grey-box modelling. A state dependent diffusion term was developed using a Lamperti transformation of the states, and implemented to compensate for heteroscedastic state uncertainty and to avoid predicting negative states. A flow proportional observation noise term introduced by a log transform was furthermore used to avoid predicting negative flows. In the simplest stochastic prediction models all parameters were estimated easily; however increasing the deterministic model complexity involved that some of the parameters had to be fixed. The statistical assumptions that require the residuals to correspond to a white noise process were fulfilled for the one-step prediction but beyond the one-step prediction auto-correlated residuals were obtained. The Akaike’s (AIC) and the Bayesian (BIC) information criteria were used to identify preferred models for the one-step prediction whereas a skill scoring criterion addressing both the reliability and the sharpness of the confidence bounds was used when assessing the forecasting performance beyond the one-step. The reliability was satisfied for the one-step prediction but were increasingly biased as the prediction horizon was expanded, particularly in rainy periods.

GLUE was applied for estimating uncertainty in such a way that the selection of behavioral parameter sets continued until a required coverage of observations was obtained (targeting 90%). A likelihood measure were used for ranking the parameter sets and two different ways of drawing parameter sets were tested, a Latin Hypercube Monte Carlo method and a modified Monte Carlo Markov Chain method. When using the stochastic models for simulation, it was found that the simulation uncertainty was best described when estimating parameters by the output error minimisation method. In order to remove the heteroscedastic residuals structure it were necessary to apply a transformation of the observations, however autocorrelation remained in the simulation case. A skill scoring comparison of a simulation and a prediction model showed that a major improvement is gained by updating the model states continuously, i.e. updating of model states results in much lower forecasting uncertainty at shorter prediction steps. In the GLUE methodology there are no requirements to the residuals. Nevertheless the aim is the same as for the stochastic simulation models, namely to cover a proportion of observations consistent with the considered quantile with maximum sharpness, i.e. to minimise the skill score. In one calibration case, even though very broad prior parameter ranges were specified, it was difficult to acquire a 90% coverage of observations and the reliability in rainy periods was much lower than in dry weather. However the GLUE method proved quite consistent in the sense that similar coverage rates were obtained in both calibration and validation periods with the same set of retained parameter sets. A comparison of the stochastic and epistemic approaches to uncertainty evaluation was conducted by comparing the sharpness, the reliability and the skill score on the same set of data. Very similar performance was obtained with the stochastic method as the preferred. The thesis has demonstrated that the statistical requirements to the formal stochastic approach are very hard to fulfill in practice when prediction steps beyond the one-step is considered. Thus the underlying assumption of the GLUE methodology, that uncertainty in modeling and simulation is not only of stochastic nature, seems fairly consistent with the results of this thesis. A major drawback of the GLUE methodology as applied here is the lumping of total uncertainty into the parameters, which entails a loss of physicality of the model parameters. Conversely the parameter estimates of the stochastic approach are physically meaningful. This thesis has contributed to developing simplified rainfall-runoff models that are suitable for model predictive control of urban drainage systems that takes uncertainty into account.