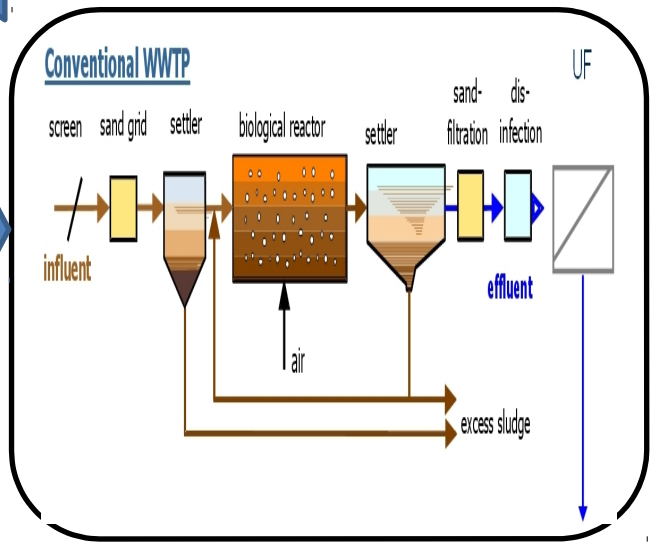
Task 1: Experimental Design

- Carefully design the experiment to produce a good coverage of the input space
- Space filling design
- Generate n x p LHS matrix with uniformly distributed on interval [0,1]

Task 2: LAMMPS

- Run the LAMMPS for fixed number of days to obtain simulation data
- Make repeated run to capture stochasticity
- Parallel configuration setting for HPC



Task 3: Data Processing

- Identify the important inputs
- Summarize the relevant outputs from microscale to mesoscale (floc)
- Characterize the size and shape of the resulting floc or biofilm
- Variable selection
- Output behaviour: stochastic
- multiple outputs; multi-scale emulator

Task 4: Emulation

Consider 2 techniques

- **GP** in form of kriging
- Execute **A two stage**: linear model + GP regression
- Structure identification: mean and covariance functions
- Train the emulator with processed data using R
- Inference: estimating parameters
- Cross-validation: check validity of emulator

Task 5: Outputs

- Emulator algorithm is performed and provided by R routine
- Emulator provides prediction at mesoscale level
- Information is transfer back to LAMMPS as a R codes for scaling up.