GFIS ECOTWIN

Engineering handbook

Chatake Innoworks Private Limited
SIH engineering programme · 9 September 2026

01 · The complete system

A plant or research record supplies timestamped observations. The gateway validates units and availability. The state engine maps them to a reactor and gas train. Screening rules identify conditions for operator review. The twin visualises this state. An event record preserves what the operator saw and why an advisory appeared. The website contains Overview, Plant lab, Team, Workflow and Library routes with a return to GFIS.

02 · Anaerobic digestion

Hydrolysis breaks larger organic compounds into soluble products. Acidogenesis produces smaller acids and intermediates. Acetogenesis supplies acetate, hydrogen and carbon dioxide. Methanogenesis converts suitable substrates into methane. Changes in loading, temperature and pH can affect the microbial community. The scene communicates this process sequence but does not solve biochemical reaction kinetics.

03 · Methane and total biogas

Biogas includes methane, carbon dioxide and other components. Methane volume can be computed from total gas volume and methane fraction when both refer to the same sample and measurement basis: V_CH4 = V_biogas × x_CH4 / 100. The GFIS API names one field measured_methane_ml; it may represent the source-derived volume calculation. A millilitre research observation must never be shown as cubic metres per hour. The teaching simulation uses m³/h and an explicitly assumed 60% methane fraction.

04 · Research dataset and provenance

Raquel Núñez (2026), A longitudinal dataset of reactor adaptation and operation during semicontinuous anaerobic digestion of black liquor, Mendeley Data, V1, DOI 10.17632/jpnjmzp6bj.1. G-AQUA, Universidad Politécnica de Madrid, Spain, generated the underlying observations. Licence: CC BY 4.0. The two reactors are RBL and SBL. This site bundles processed public GFIS API snapshots, not the original source workbook. The data dictionary and SHA-256 manifest record the transformation and source URL.

05 · Sampling and missing values

Replay timestamps follow the source observations and may be irregular. The slider advances records, not elapsed minutes. A missing pH, alkalinity or methane measurement remains null. Input coverage equals the number of present temperature, pH and gas fields divided by three. Coverage is not a statistically calibrated confidence probability. Missing temperature prevents temperature-specific conclusions.

06 · Temporal model roles

Persistence predicts the next eligible observation with the previous observation. XGBoost can learn nonlinear relationships using past-only lag and rolling features. LSTM encodes ordered input windows into a recurrent hidden state. Neither complexity nor epoch count guarantees improvement. This SIH site computes the persistence diagnostic directly. It links the existing GFIS model workbench for trained-model evidence and does not disguise persistence as XGBoost or LSTM inference.

07 · Evaluation mathematics

For n valid observation/prediction pairs, MAE = sum(|y_i − yhat_i|) / n and RMSE = sqrt(sum((y_i − yhat_i)^2) / n). The replay diagnostic is a next-observation persistence calculation across the selected public snapshot; it is not a held-out training experiment or fixed-horizon forecast. A proper model comparison freezes chronological train, validation and test periods, fits preprocessing within training folds and reports repeated-seed dispersion on identical target pairs.

08 · The 48-hour storage equation

The teaching model uses Δt = 0.25 h, gas-holder capacity K = 100 m³ and initial storage S_0 = 40 m³. Production q depends on an illustrative feed multiplier and temperature/pH response. Available gas A = S_t + q_t Δt. Used gas U = min(demand × Δt, A). Overflow O = max(0, A − U − K). Next storage S_(t+1) = A − U − O. Therefore storage remains within [0,K] and the gas balance closes by construction. Overflow means unserved model capacity, not a real vent command.

09 · Demonstration response functions

The simulation assumes q = 10 f exp(−((T−35)/12)^2) max(0.1, 1−0.65|pH−7.2|) [1+0.1 sin(t/3)], with a minimum efficiency floor of 0.08. f is the dimensionless feed multiplier. Constants are teaching parameters, not calibrated plant coefficients. An overfeeding scenario lowers synthetic pH after hour 12. A heat scenario raises temperature by 7 °C after hour 16. Sensor loss hides temperature after hour 12.

10 · Alert authority

Demonstration screening thresholds are pH 6.5–8.5, temperature 25–40 °C, storage above 90%, and a gas-output drop greater than 40% from the prior record. These are prototype rules and require plant-specific commissioning. An advisory contains a reason and inspection action. It never writes to an actuator. Sensor absence is distinct from a safe reading. Acknowledgement records operator review and does not prove that the underlying condition was resolved.

11 · ADM1 and advanced physics

ADM1 describes anaerobic-digestion kinetics and physicochemical processes through coupled states and rates. A calibrated ADM1 model can supply expected trajectories or residuals. This SIH teaching model is not ADM1. A future physics-loss study must identify measured variables, units, parameter uncertainty and an unmodified ML comparator. A VFA/ALK soft sensor needs paired target measurements and held-out validation; the absence of alkalinity is shown explicitly here.

12 · 3D twin and network

The WebGL scene uses Three.js meshes for the digester, feed tank, gas holder and sensor markers. The current timestamp controls marker colour and gas-holder state. Mouse selection or accessible buttons explain components. This is a visual representation of state, not CFD, a calibrated pressure simulation or proof of plant hardware. The network view shows three hypothetical nodes so the team can discuss roll-up alerts without inventing monitored facilities.

13 · Edge gateway and protocol

Use schema gfis-sih-telemetry-v1 with plant_id, timestamp_utc, sequence, variables, units and sensor_status. Maintain monotonic sequence counters, reject wrong units, bound payload size and retain a local queue during network interruption. The delivered JSON replay provides a software input contract. Real ESP32 acquisition, clock synchronisation and calibration remain team bench tasks.

14 · Event memory and recovery

Event records use schema gfis-sih-event-v1 and retain source class, current record, settings, input coverage, reason, timestamp, engine version and snapshot checksum. Browser-local memory survives a refresh on the same browser. JSON export is the portable copy; there is no team-shared database in this release. Import validates shape and source class before rendering text. Keep exported personal notes out of public hosting.

15 · Deployment and security

The release is designed for the existing gfis.energy CloudFront/S3 site under /sih/. Bundled assets enable local replay without the API. Refresh from AWS fetches the existing public replay endpoint and reports errors without replacing observations with invented data. No new compute, database or AWS account is needed. Research manuscript files and the protected reviewer archive remain outside the public SIH directory.

16 · Technical limits and next experiment

Current software demonstrates state inspection, deterministic scenario response and replay. It does not establish field leak-detection sensitivity, industrial control safety, cross-plant generalisation or validated biological causality. Next, the students should test the sensor bench, collect labelled fault packets, measure response latency, validate out-of-range rejection and complete a supervised plant-specific risk review.

17 · Student ROS2 extension

Sneha and Sakshi share the 3D/ROS2 work using their existing ROS2 experience. Sneha owns meshes, assets and sensor-to-scene mapping. Sakshi owns the proposed replay bridge, message timing and integration tests. The delivered browser uses Three.js; ROS2 integration is an assigned development task and is not yet connected.