The National Institutes of Health | Division of Technical Resources | Office of Research Facilities

Design Requirements Manual **News to Use**

May 2026
Vol. 03, No. 05

The formulae $\frac{\partial \rho U_i}{\partial t}+\frac{\partial}{\partial x_j}(\rho U_i U_j)=-\frac{\partial P}{\partial x_i}+\frac{\partial}{\partial x_j}\left(\mu\frac{\partial U_i}{\partial x_j}\right)+g_i(\rho-\rho_0)$ *for building* $\frac{\partial}{\partial x_j}(\rho \overline{U}_i \overline{U}_j)=-\frac{\partial P}{\partial x_i}+\frac{\partial}{\partial x_j}\left(\mu\frac{\partial \overline{U}_i}{\partial x_j}-\rho\overline{u_i' u_j'}\right)+g_i(\rho-\rho_0)$ *state of the art* $\frac{\partial}{\partial x_i}(\rho \overline{U}_i \overline{H})=\frac{\partial}{\partial x_i}\left(\lambda\frac{\partial \overline{T}}{\partial x_i}-\rho\overline{u_i' h'}\right)$ *biomedical research facilities.*

Denoising Data: Utilizing Wavelet-GARCH Technique for Anomaly Detection

Anomaly detection architecture typically uses sensor data and machine learning to analyze how predictor variables relate to a target outcome, but noise can obscure true operating conditions and degrade reliability. To address this, the Data Science and Artificial Intelligence Application Branch (DSAIAB) team developed a generalized monitoring template using discrete wavelet transform (DWT) denoising with Generalized Autoregressive Conditional Heteroskedasticity (GARCH) monitoring that is designed to be adaptable and applied across a wide range of campus sensor data. Originally intended for single-use denoising, the template was adapted for a live chilled water makeup flow anomaly detection application, where improved signal quality enhances model performance. This is critical given the rarity of anomalies (<1%) and their wide range of magnitudes, from catastrophic failures to subtle leaks. This article examines the template's modification, tuning, key considerations, and impact on results to showcase an example of applying the template for a specific use-case.

Template Modification for Application Architecture

To generalize the template's functionality and better support the anomaly detection application, an additional configuration structure was introduced: a list of target signals to be denoised. This allows the processing pipeline to process multiple signals in a single run. Wavelet and GARCH parameters were also updated to accept list-based inputs, enabling systematic evaluation across all signals.

To support anomaly detection, the modified template retains its existing capability to flag GARCH volatility and flatline or stale sensor data, enabling operators to identify and respond to potential anomalies. Additionally, negative values from DWT denoising are replaced with zeros to preserve data integrity and prevent positive value bias.

Mathematical Background

This section outlines the mathematical framework for denoising, GARCH-based flagging, and parameter selection. Denoising is performed using wavelet soft thresholding, which shrinks coefficients toward zero by a threshold lambda:

$$\hat{d}_{j,k} = sign\,(d_{j,k})max\,(\mid d_{j,k} \mid -\lambda, 0), \lambda = \alpha \cdot \hat{\sigma}\sqrt{2log\ N}$$

Following denoising, volatility is modeled using a GARCH(p, q) process to capture time-varying variance:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

If the estimated variance exceeds a threshold value, it is flagged as an anomaly. Residual whiteness evaluates temporal independence by computing the maximum absolute autocorrelation over lags from 1–20:

$$W = \max_{1 \le \tau \le 20} \left| \frac{\sum(r_t - \bar{r})(r_{t+\tau} - \bar{r})}{\sum(r_t - \bar{r})^2} \right|$$

Where lower values indicate less structure, implying more noise removed from the signal (Almeida & Figueiredo, 2013). Spectral flatness then assesses whether the residual is noise-like by comparing the geometric and arithmetic means of its power spectrum (Johnston, 1988):

$$SF = \frac{\exp\left(\frac{1}{M}\sum_{m=1}^{M} log P_r(m)\right)}{\frac{1}{M}\sum_{m=1}^{M} P_r(m)}, P_r(m) = |\mathcal{F}\{r_t\}|^2$$

The value assesses the uniformity of the residual's frequency content by quantifying how tone-like versus noise-like a sound is (Dubnov, 2004), where values near one indicate the removed component exhibits a flat, noise-like spectrum. Computational complexity is approximated as a function of signal length, wavelet filter length determined by wavelet family selection, and decomposition level:

$$\eta = \log\left((N - L_w) * L_w * J\right)$$

These metrics are normalized from zero to one to ensure a consistent scale and are then combined into a weighted composite score for parameter selection:

$$score = \big(0.45 * (1 - SF)\big) + (0.45 * W) - (0.1 * \eta)$$

The National Institutes of Health | Division of Technical Resources | Office of Research Facilities

Design Requirements Manual **News to Use**

May 2026
Vol. 03, No. 05

The formulae $\frac{\partial \rho U_i}{\partial t}+\frac{\partial}{\partial x_j}(\rho U_i U_j)=-\frac{\partial P}{\partial x_i}+\frac{\partial}{\partial x_j}\left(\mu\frac{\partial U_i}{\partial x_j}\right)+g_i(\rho-\rho_0)$ *for building* $\frac{\partial}{\partial x_j}(\rho \bar{U}_i \bar{U}_j)=-\frac{\partial P}{\partial x_i}+\frac{\partial}{\partial x_j}\left(\mu\frac{\partial \bar{U}_i}{\partial x_j}-\rho\overline{u_i'u_j'}\right)+g_i(\rho-\rho_0)$ *state of the art* $\frac{\partial}{\partial x_i}(\rho \bar{U}_i \bar{H})=\frac{\partial}{\partial x_i}\left(\lambda\frac{\partial \bar{T}}{\partial x_i}-\rho\overline{u_i'h'}\right)$ *biomedical research facilities.*

Parameter Tuning for Application Signals

A quantitative approach was initially used to determine appropriate denoising parameters for the target signals. Unlike traditional methods that rely on a known clean reference, this application required blind wavelet parameter selection, where performance is evaluated solely based on the noisy signal and resulting denoised output.

Step 1: Data Preprocessing - The first step ensures numerical stability and consistency by standardizing the data, interpolating missing values, and mean-centering each signal to support spectral and autocorrelation analyses. The signal is then divided into windows to capture local variation in noise and signal structure, improving robustness.

Step 2: Signal Denoising - Following preprocessing, each signal is denoised using a range of parameter combinations selected based on avoidance of overfitting. These combinations included variations in wavelet type, decomposition level, and threshold multiplier. The denoising process is applied iteratively across all possible parameter configurations defined in the search space.

```
wavelets=["haar", "db2", "db4", "sym4", "coif1", "coif3"]
levels=[1, 2, 3, 4, 5, 6, 7, 8]
alphas=np.linspace(0.5, 2.0, 7)
```

Figure 1. Conservatively chosen quantitative criteria to iterate through to avoid overfitting during denoising

Step 3: Denoising Evaluation - Each resulting denoised signal is then evaluated using quantitative criteria designed to assess performance without requiring ground truth reference. The primary metrics are residual whiteness and spectral flatness, both computed on the residual signal.

To compare parameter configurations, all metrics are normalized and combined into a composite score, with residual whiteness and spectral flatness equally weighted and computational complexity assigned a lower weight. This prioritizes denoising performance while accounting for cost. The output display will show the top three scored parameter combinations for each signal.

```
wavelet  level  alpha  whiteness  flatness       eta  runtime_ms  n_windows    score
    db4      6    0.5   0.427584  0.412401 5.292752    0.318325          8 0.664809
    db4      7    0.5   0.436137  0.420888 5.359699    0.411788          8 0.662626
    db4      5    0.5   0.410217  0.391059 5.213571    0.191088          8 0.660980
Best parameters for MUWP1_VFD_FEEDBACK: wavelet='db4', level=6, alpha=0.5, score=0.66480

wavelet  level  alpha  whiteness  flatness       eta  runtime_ms  n_windows    score
   haar      6    0.5   0.124230  0.524775 4.691329    0.205637          8 0.824064
    db2      5    0.5   0.107299  0.517500 4.912966    0.384438          8 0.821010
   haar      5    0.5   0.122663  0.514146 4.612148    0.243763          8 0.820234
Best parameters for Chw_Makeup_Flow: wavelet='haar', level=6, alpha=0.5, score=0.82406
```

Figure 2. Truncated output showing three best scored denoising parameter combinations

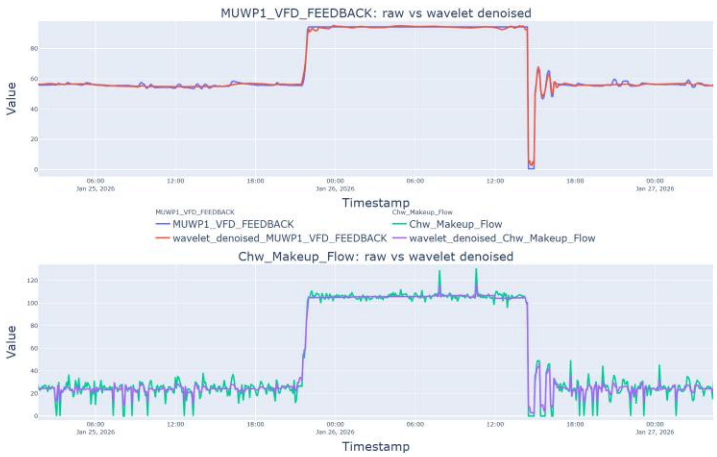

Figure 3. Denoised signals based on quantitatively chosen parameters using residual whiteness and spectral flatness

Step 4: Visual Tuning - Selected parameters are validated through visual inspection of the denoised signals, with additional fine-tuning via threshold multiplier to control denoising strength. Higher values increase smoothing, while lower values preserve more signal detail. This final step ensures that denoising is adequate after conservative quantitative parameter selection.

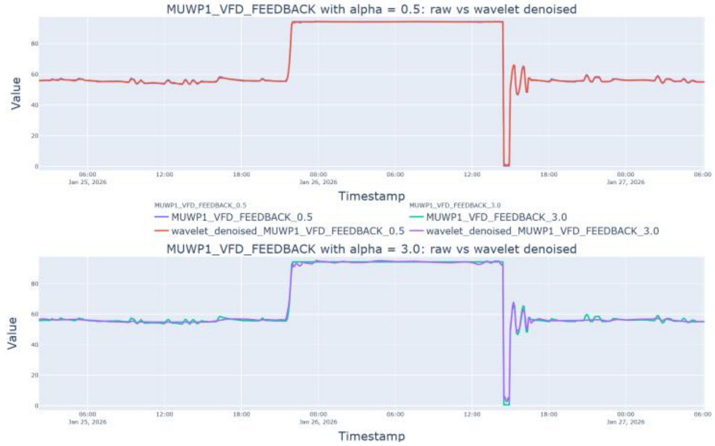

Figure 4. Denoised signal using a threshold multiplier (alpha) of 3.0 vs 0.5 decided via visual tuning of parameter selection

'Design Requirements Manual (DRM) News to Use' is a monthly ORF publication featuring salient technical information that should be applied to the design of NIH biomedical research laboratories and animal facilities. NIH Project Officers, A/E's and other consultants to the NIH, who develop intramural, extramural and American Recovery and Reinvestment Act (ARRA) projects will benefit from 'News to Use'. **Please address questions or comments to:** megan.teague@nih.gov

The National Institutes of Health | Division of Technical Resources | Office of Research Facilities

Design Requirements Manual **News to Use**

May 2026
Vol. 03, No. 05

The formulae $\frac{\partial \rho U_i}{\partial t}+\frac{\partial}{\partial x_j}(\rho U_i U_j)=-\frac{\partial P}{\partial x_i}+\frac{\partial}{\partial x_j}\left(\mu\frac{\partial U_i}{\partial x_j}\right)+g_i(\rho-\rho_0)$ *for building* $\frac{\partial}{\partial x_j}(\rho \overline{U}_i \overline{U}_j)=-\frac{\partial \overline{P}}{\partial x_i}+\frac{\partial}{\partial x_j}\left(\mu\frac{\partial \overline{U}_i}{\partial x_j}-\rho\overline{u_i u_j}\right)+g_i(\rho-\rho_0)$ *state of the art* $\frac{\partial}{\partial x_i}(\rho \overline{U}_i \overline{H})=\frac{\partial}{\partial x_i}\left(\lambda\frac{\partial \overline{T}}{\partial x_i}-\rho\overline{u_i h}\right)$ *biomedical research facilities.*

15-minute Sampled Data Usage (Boundary Effect)

A known limitation of DWT in live denoising is the boundary effect, where the use of insufficient data for convolution-based filtering leads to edge distortion. To compensate, padding or signal extension is used, which can introduce other bias and distortion (Montanari et al., 2015). In streaming applications, this creates right-boundary distortion at the most recent data point and left-boundary artifacts between adjacent windows, compounding errors and reducing denoising effectiveness. In this application, data sampled every five minutes and updated every fifteen minutes requires effective boundary handling to avoid minimal denoising and persistently noisy signals.

To mitigate this issue, the anomaly detection application utilizes a proposed method for 15-minute sample denoising called updated half-edge extension (Yang et al., 2023). This technique is used in the context of sliding window wavelet filtering: it uses the previous window's original data as the extension values during decomposition of the wavelet to reduce the left-boundary effect, then inserts the right-boundary denoised samples. The application follows a similar design by using the most recent 72 hours of data while only appending the most recent 15 minutes of data to the denoised signal. This ensures that the denoised signal is minimally affected by boundary artifacts.

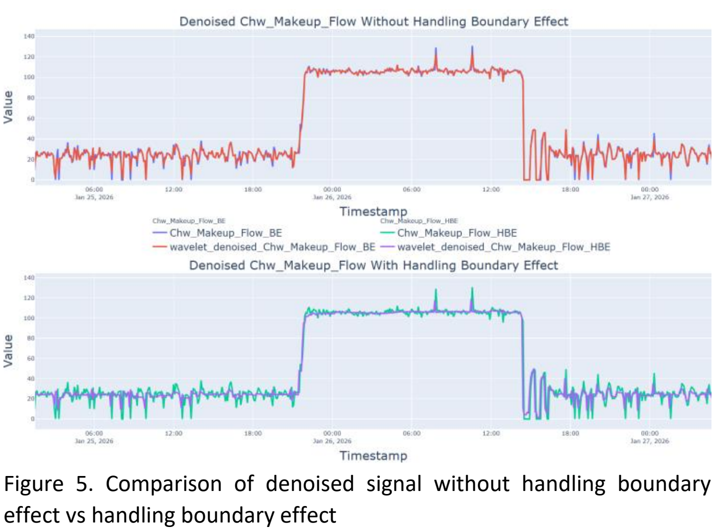

Figure 5. Comparison of denoised signal without handling boundary effect vs handling boundary effect

Conclusion

The Wavelet-GARCH template can be adapted for anomaly detection using 15-minute data ingestion by enabling multi-signal denoising with quantitative parameter selection driven by spectral flatness, residual whiteness, and computational complexity. Negative values generated by wavelet outputs should be replaced with zeros to ensure stable behavior and eliminate potential positive bias. The application utilizes an updated half-edge extension method to reduce boundary effects, improving signal reliability. Together, these modifications create a stable, scalable denoising framework applicable to 15-minute sample anomaly detection.

References

1. Almeida, M. S. C., & Figueiredo, M. A. T. (2013). Parameter estimation for blind and non-blind deblurring using residual whiteness measures. IEEE Transactions on Image Processing, 22(7), 2751–2763.
2. Johnston, J. D. (1988). Transform coding of audio signals using perceptual noise criteria. IEEE Journal on Selected Areas in Communications, 6(2), 314–323.
3. Dubnov, S. (2004). Generalization of spectral flatness measure for non-Gaussian linear processes. *IEEE Signal Processing Letters, 11*(8), 698–701.
4. Montanari, L., Basu, B., Spagnoli, A., & Broderick, B. M. (2015). A padding method to reduce edge effects for enhanced damage identification using wavelet analysis. Mechanical Systems and Signal Processing, 52–53, 264–277.
5. Yang, X., Chen, X., Sun, K., Xiong, C., Song, D., Lu, Y., Huang, L., He, S., & Zhang, X. (2023). A wavelet transform-based real-time filtering algorithm for fusion magnet power signals and its implementation. Energies, 16(10), 4091.

Nomenclature

$\hat{d}_{j,k}$: thresholded detail (wavelet) coefficient
λ: wavelet coefficient threshold
J: final decomposition scale level
α: threshold multiplier
$\hat{\sigma}$: noise standard deviation
σ_t^2: conditional variance in GARCH at time t
ω: GARCH baseline variance
α_i: GARCH impact of past shocks
β_j: GARCH persistence of volatility
ϵ_t: residuals in the GARCH model at time t
p: order of GARCH terms in the model
q: order of ARCH terms in the GARCH model
r_t: denoising residual at time t
$\bar{r}$: mean of the denoising residual
τ: time lag
L_w: wavelet filter length
$P_r(m)$: power spectral density
$\mathcal{F}\{r_t\}$: Fourier transform of denoising residual
N: number of samples
SF: spectral flatness
W: residual whiteness
η: computational complexity
M: frequency bins