# WHAM to POMP 2

## 2025-11-03

## 1. The process equations are nonlinear

WHAM's hidden states evolve through biological laws, not linear algebra.

$$\log N_{a,t} = \begin{cases} \log f(SSB_{t-1}) + \varepsilon_{1,t}, & a = 1, \\ \log N_{a-1,t-1} - Z_{a-1,t-1} + \varepsilon_{a,t}, & 1 < a < A, \\ \log(N_{A-1,t-1}e^{-Z_{A-1,t-1}} + N_{A,t-1}e^{-Z_{A,t-1}}) + \varepsilon_{A,t}, & a = A. \end{cases}$$

Nonlinearities come from:

- **Beverton–Holt recruitment**
  $f(SSB_{t-1}) = \frac{\alpha SSB_{t-1}}{1 + \beta SSB_{t-1}}$
- **Exponential survival**
  $e^{-Z}$
- **Plus-group summation**
  $N_{A-1}e^{-Z_{A-1}} + N_A e^{-Z_A}$

These involve products, ratios, and exponentials, so they cannot be written as a linear matrix equation like $X_t = AX_{t-1} + w_t$.

## 2. The observation equations are also nonlinear

The expected observations depend multiplicatively on the hidden states:

$$\hat{C}_{a,t} = N_{a,t}(1 - e^{-Z_{a,t}})\frac{F_{a,t}}{Z_{a,t}}, \quad \hat{I}_{a,t} = q_i \, s_{a,t} \, N_{a,t} \, e^{-Z_{a,t}f_{t,i}}.$$

Each equation uses exponentials, products, and ratios — all nonlinear in $N_{a,t}, F_{a,t}, M_{a,t}$.
Thus, the observation function $g_\theta(X_t)$ is not linear.

## 3. Random effects are Gaussian on transformed scales

Although the model is nonlinear, WHAM assumes random deviations are Gaussian on **transformed scales**:

- Log-abundance deviations: $\varepsilon_{a,t} \sim \mathcal{N}(0, \Sigma_{NAA})$
- Log-natural mortality deviations: $\delta_{a,t} \sim \mathcal{N}(0, \Sigma_M)$
- Logistic-selectivity deviations: $\zeta_{p,t} \sim \mathcal{N}(0, \Sigma_{Sel})$
- Environmental AR(1): $\eta_t \sim \mathcal{N}(0, \sigma_X^2)$

So, while the process noise is Gaussian, applying nonlinear transformations (log, logistic, exponential) makes the overall system non-Gaussian in its **original space**.

## 4. Observation errors are non-Gaussian likelihoods

- **Catch and index** modeled as log-Normal:

$$\log Y \sim \mathcal{N}(\log \hat{Y} - \tfrac{1}{2}\sigma^2, \sigma^2)$$

- **Age compositions** modeled using a logistic-Normal or multinomial likelihood.

Hence, these are Normal **only on transformed scales**, not in $Y$ itself.

## 5. Why we say "approximately Gaussian"

WHAM uses **Template Model Builder (TMB)**, which applies the Laplace approximation:

$$\int p(X, Y|\theta)\, dX \approx p(Y|\hat{X}, \theta)\, (2\pi)^{k/2} |\Sigma_{\text{post}}|^{1/2}.$$

This means the nonlinear model's posterior of random effects is locally approximated as **multivariate Normal near its mode**.
Thus, WHAM is *approximately Gaussian*: TMB treats random effects as Gaussian when integrating over them, even though the true model is nonlinear.

## Summary Table

| Property | Linear? | Gaussian? | Explanation |
|---|---|---|---|
| **Process dynamics** | Nonlinear | Gaussian errors (log-scale) | Exponentials, Beverton–Holt, logistic selectivity |
| **Observation model** | Nonlinear | Non-Gaussian | Log-Normal, logistic-Normal likelihoods |
| **Inference method** | — | Approx. Gaussian | Laplace approximation in TMB |

*In short, WHAM combines nonlinear biological dynamics with Gaussian random effects and uses approximate Gaussian inference via the Laplace method.*

# 1. POMP General Structure

$$\begin{cases} X_t \sim f_\theta(X_t \mid X_{t-1}) & \text{(process model / hidden state)} \\[2em] Y_t \sim g_\theta(Y_t \mid X_t) & \text{(observation model)} \end{cases}$$

where:
- $X_t$: unobserved true system state at time $t$
- $Y_t$: noisy observation at time $t$
- $f_\theta$: transition density (evolution of hidden states)
- $g_\theta$: observation density (links hidden state to data)
- $\theta$: static parameters

# 2. State Process — Hidden Biological System

$$X_t = \left( N_{1:A,t}, M_{1:A,t}, s_{1:A,t}, X_t^{Ecov} \right)$$

includes: - $N_{a,t}$: numbers at age $a$ - $M_{a,t}$: natural mortality - $s_{a,t}$: selectivity - $X_t^{Ecov}$: environmental covariates (e.g. temperature index)

The system evolves via nonlinear biological laws, with Gaussian random effects in log-space:

$$f_\theta(X_t \mid X_{t-1}) = \mathrm{MVN}\left( h_\theta(X_{t-1}), \Sigma_{\mathrm{proc}} \right)$$

# 3. Nonlinear Mean Function $h_\theta(X_{t-1})$

## (a) Numbers-at-age Process

$$\log N_{a,t} = \begin{cases} \log f(SSB_{t-1}) + \varepsilon_{1,t}, & a = 1, \\\\ \log N_{a-1,t-1} - Z_{a-1,t-1} + \varepsilon_{a,t}, & 1 < a < A, \\\\ \log\left(N_{A-1,t-1}e^{-Z_{A-1,t-1}} + N_{A,t-1}e^{-Z_{A,t-1}}\right) + \varepsilon_{A,t}, & a = A \end{cases}$$

with

$$Z_{a,t-1} = F_{a,t-1} + M_{a,t-1}, \quad f(SSB_{t-1}) = \frac{\alpha SSB_{t-1}}{1 + \beta SSB_{t-1}} e^{\beta_1 X_{t-1}^{Ecov} + \beta_2 (X_{t-1}^{Ecov})^2}$$

Thus,

$$E[\log N_{a,t} \mid X_{t-1}] = h_{a,\theta}^N(X_{t-1})$$

## (b) Natural Mortality Process

$$\log M_{a,t} = \mu_{M_a} + \delta_{a,t}, \quad E[\log M_{a,t} \mid X_{t-1}] = \mu_{M_a}$$

Random deviations $\delta_{a,t}$ have 2D AR(1) covariance across age and year.

## (c) Selectivity Process

$$s_{a,t} = \frac{1}{1 + \exp[-(a - a_{50,t})/k_t]}$$

with:

$$a_{50,t} = l_{a50} + \frac{u_{a50} - l_{a50}}{1 + e^{-(\nu_1 + \zeta_{1,t})}}, \quad k_t = l_k + \frac{u_k - l_k}{1 + e^{-(\nu_2 + \zeta_{2,t})}}$$

and AR(1) deviations $\zeta_{p,t}$ in both parameter ($p$) and year ($t$).

## (d) Environmental Covariate Process

$$X_t^{Ecov} = \mu_X(1 - \phi_X) + \phi_X X_{t-1}^{Ecov} + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_X^2)$$

## 4. Combined Process Model

Combine all components into a single latent vector:

$$\mathbf{X}_t = \begin{bmatrix} \log N_{1,t} \\ \vdots \\ \log N_{A,t} \\ \log M_{1,t} \\ \vdots \\ \log M_{A,t} \\ \zeta_{1,t} \\ \zeta_{2,t} \\ X_t^{Ecov} \end{bmatrix}$$

and mean function:

$$\mathbf{h}_\theta(\mathbf{X}_{t-1}) = \begin{bmatrix} h_\theta^N(\cdot) \\ \mu_M \\ E[a_{50,t}, k_t \mid X_{t-1}] \\ \mu_X(1 - \phi_X) + \phi_X X_{t-1}^{Ecov} \end{bmatrix}$$

Covariance:

$$\Sigma_{\text{proc}} = \begin{bmatrix} \Sigma_{NAA} & 0 & 0 & 0 \\ 0 & \Sigma_M & 0 & 0 \\ 0 & 0 & \Sigma_{Sel} & 0 \\ 0 & 0 & 0 & \Sigma_{Ecov} \end{bmatrix}$$

Each block has correlation across age and time, e.g.

$$\text{Cov}(\varepsilon_{a,t}, \varepsilon_{\tilde{a},\tilde{t}}) = \sigma_a \sigma_{\tilde{a}} \rho_{\text{age}}^{|a-\tilde{a}|} \rho_{\text{year}}^{|t-\tilde{t}|}$$

## 5. Full Process Distribution

$$f_\theta(X_t \mid X_{t-1}) = \mathcal{N}(X_t \mid \mathbf{h}_\theta(X_{t-1}), \Sigma_{\text{proc}})$$

or equivalently:

$$X_t = \mathbf{h}_\theta(X_{t-1}) + \omega_t, \quad \omega_t \sim \mathcal{N}(0, \Sigma_{\text{proc}})$$

## 6. Observation Model

Observations are catches, survey indices, and age compositions:

$$Y_t = (C_{t,i}, I_{t,i}, p_{a,t}^C, p_{a,t}^I)_{i,a}$$

Aggregate catch model:

$$\log C_{t,i} \sim \mathcal{N}\left( \log \hat{C}_{t,i} - \tfrac{1}{2}\sigma^2_{C_{t,i}}, \ \sigma^2_{C_{t,i}} \right)$$

Predicted catch:

$$\hat{C}_{a,t,i} = N_{a,t}(1 - e^{-Z_{a,t}})\frac{F_{a,t,i}}{Z_{a,t}}, \quad Z_{a,t} = F_{a,t} + M_{a,t}$$

For survey indices:

$$\hat{I}_{a,t,i} = q_i s_{a,t,i} N_{a,t} W_{a,t,i} e^{-Z_{a,t} f_{t,i}}$$

and

$$\log I_{t,i} \sim \mathcal{N}\left( \log \hat{I}_{t,i} - \tfrac{1}{2}\sigma^2_{I_{t,i}}, \ \sigma^2_{I_{t,i}} \right)$$

**Age Composition (Logistic-Normal)**

For catch or survey dataset $d$:

$$-\log \mathcal{L} = \frac{n-1}{2}\log(2\pi v) + \frac{1}{2v}\sum_{a=a^*}^{A-1}\left[ \log \frac{p_a^*}{p_{A^*}^*} - \log \frac{\hat{p}_a}{\hat{p}_{A^*}} \right]^2$$

where $v = \tau_d^2/N_{\text{eff},d,y}$.

## 7. Final Joint Observation Density

$$g_\theta(Y_t \mid X_t) = \prod_i \mathcal{N}\left( \log C_{t,i}; \ \log \hat{C}_{t,i} - \tfrac{1}{2}\sigma^2_{C_{t,i}}, \ \sigma^2_{C_{t,i}} \right)$$

$$\times \prod_i \mathcal{N}\left( \log I_{t,i}; \ \log \hat{I}_{t,i} - \tfrac{1}{2}\sigma^2_{I_{t,i}}, \ \sigma^2_{I_{t,i}} \right)$$

$$\times \prod_d \text{LogisticNormal}\left( p_{t,d}^*; \ \hat{p}_{t,d}, \ v_{t,d} \right)$$

## 8. Summary Table

| POMP Component | WHAM Component | Notes |
|---|---|---|
| Latent state $X_t$ | $N_{a,t}, M_{a,t}, s_{a,t}, Ecov_t$ | Hidden biological & environmental states |
| Process model $f_\theta(X_t \mid X_{t-1})$ | Stock–recruitment, survival, 2D AR(1) | Nonlinear Gaussian |
| Observation model $g_\theta(Y_t \mid X_t)$ | Catch, survey, age composition | Lognormal or logistic-normal |
| Parameters $\theta$ | Growth, selectivity, AR(1), variances | Estimated via Laplace (TMB) |
| Process noise | $\varepsilon_{a,t}, \delta_{a,t}, \zeta_{a,t}$ | Gaussian on transformed scales |
| Observation noise | Sampling error | Lognormal or multinomial/logistic-normal |

# Reason 1 — Discrete-Time Markov Process

**WHAM Characteristics:**

- The population evolves in discrete annual time steps.

- The state at year $t + 1$ depends only on the state at year $t$ (Markov property).

- It does not depend on $t$ - $2$, $t$ - $3$, or earlier history.

**Implication:**

POMP is designed for the Markov structure:

$$X_t = f(X_{t-1}, \theta, \varepsilon_t)$$

WHAM implements the exact same form:

$$X_t = f(X_{t-1}, \theta, \varepsilon_t)$$

**Perfect match**

# Reason 2 — Well-defined Process Model

**WHAM has explicit process equations:**

$$R_t = BH(SSB_{t-1}) \cdot e^{\varepsilon_R},$$

$$N_{a+1,t+1} = N_{a,t} \cdot e^{-Z_{a,t}},$$

$$N_{A,t+1} = N_{A-1,t} \cdot S_{A-1} + N_{A,t} \cdot S_A.$$

**Implications:**

- The equations can be directly translated into C code.

- The evolution at each time step is clearly defined.

- The stochastic error term ($\varepsilon_R$) is explicitly included in the model.

**POMP requires:**
> "A function that computes $X_t$ given $X_{t-1}$."

**WHAM provides:** exactly such a function (implemented in TMB C++).

# Reason 3 — Well-Defined Observation Model

The WHAM observation equations are defined as **statistical distributions**,
which can be directly expressed in POMP using equivalent likelihood functions.

**Implication (expanded):**

- **POMP requires the conditional observation density $p(Y_t \mid X_t, \theta)$.**
  In WHAM, this relationship is already fully specified through its statistical likelihood functions for catch, survey indices, and age composition.
  For instance, WHAM models observed catches as lognormal random variables around predicted catches, and age compositions as multinomial (or logistic-normal) draws based on predicted proportions-at-age.
  Therefore, WHAM already provides the exact mathematical form of $p(Y_t \mid X_t, \theta)$ that POMP requires— no additional assumptions or transformations are needed.

- **The Lognormal and Multinomial distributions are standard, well-established probability models.**
  These distributions are widely used in fisheries state-space modeling because they represent realistic sources of stochastic variability:
  the lognormal captures multiplicative process or observation noise (common for quantities like biomass or catch),
  while the multinomial naturally models compositional uncertainty in proportions-at-age or size.
  WHAM's use of these standard distributions ensures that its observation equations are statistically compatible with those expected by POMP.

- **Both frameworks use the same probability density functions (PDFs).**
  WHAM implements these likelihoods internally in TMB via functions such as `dnorm`, `dlnorm`, or `dmultinom` (on the log scale for optimization).
  POMP expresses the same densities explicitly in its `dmeasure` and `rmeasure` functions using the same mathematical formulas—for example, `dlnorm(obs, log(pred), sd, 1)` or `dmultinom(obs, prob=p, log=TRUE)`.
  As a result, given the same states, parameters, and observations, both frameworks evaluate identical likelihood values.
  The difference lies only in *how* they compute or optimize the likelihood—analytically in WHAM (Laplace approximation) versus simulation-based methods in POMP (particle filtering).

**Key insight:**
The observation model is **distribution-based**, not **algorithm-based**.
This means the mathematical structure of the likelihood is identical between WHAM and POMP; only the computational implementation differs.

**Reason 4: States and Observations Are Separable**

**Principle in WHAM**

- **States** $X_t$ are *latent* (not directly observed).

- **Observations** $Y_t$ are noisy functions of the states.

**Example (catch):**

Let $N[a,t]$ be numbers-at-age, $Z_{a,t}$ total mortality, $F_{a,t}^{\text{sel}}$ fishing mortality times selectivity, and waa$_a$ weight-at-age.
The model-predicted catch at time $t$ is

$$\widehat{C}_t \;=\; \sum_a N[a,t] \left(1 - e^{-Z_{a,t}}\right) \frac{F_{a,t}^{\text{sel}}}{Z_{a,t}} \; \text{waa}_a.$$

Observed catch is modeled as a noisy measurement of that deterministic function:

$$C_t^{\text{obs}} \sim \text{Lognormal}\!\left(\log \widehat{C}_t - \tfrac{1}{2}\sigma_C^2, \; \sigma_C\right).$$

**Why this maps to POMP**

- This is exactly a POMP observation equation: $Y_t \sim p\big(g(X_t, \theta)\big)$, where $g$ is the deterministic mapping from states to predictions (here, Baranov catch), and the distribution (here, lognormal) supplies the noise model.

- The *hidden-state + partial-observation* structure is the defining core of POMP, and WHAM naturally follows this structure.

**Reason 5 — Computable Likelihood**

Both WHAM and POMP frameworks ultimately compute the same quantity:
the **marginal likelihood** $p(Y_{1:T} \mid \theta)$.

**WHAM Method**

1. Uses the **Laplace approximation** to integrate out random effects.

2. Computes:

$$L(\theta) \;=\; \int p(Y, X \mid \theta)\, dX \;\approx\; p(Y \mid X^*, \theta)\, (2\pi)^{k/2}\, |H|^{-1/2},$$

where $X^*$ is the mode of the random effects and $H$ is the Hessian matrix of second derivatives.

**POMP Method**

1. Uses **particle filtering (Monte Carlo integration)** to approximate the same integral.

2. Computes:

$$L(\theta) \;=\; \int p(Y, X \mid \theta)\, dX \;\approx\; \frac{1}{N} \sum_{i=1}^{N} w_i,$$

where $w_i$ are the particle weights from sequential importance sampling.

**Interpretation**

- Both frameworks compute the same target — the **marginal likelihood** of the observed data.

- The only difference lies in the **numerical integration method** used:
  - WHAM uses **analytic integration** (Laplace approximation).
  - POMP uses **stochastic integration** (Monte Carlo via particles).

- In both cases, the integral $\int f(x)\, dx$ represents the same quantity; only the computational approach differs.

**Analogy**  To compute an integral $\int f(x)\, dx$:

- **WHAM** $\rightarrow$ uses *analytic approximation* (Laplace).

- **POMP** $\rightarrow$ uses *numerical approximation* (Monte Carlo).

The integral being evaluated is the same — only the **method of approximation** differs.