Improving probabilistic predictions of daily streamflow

Mark Thyer, David McInerney, Dmitri Kavetski

Intelligent Water Decisions Research Group (www.waterdecisions.org)

University of Adelaide

Professor George Kuczera

University of Newcastle

Dr Julien Lerat

Seasonal Streamflow Forecasting, Bureau of Meteorology, Australia

Motivation

- Evaluating uncertainty in hydrological predictions is important for decision making and risk assessment

Aims

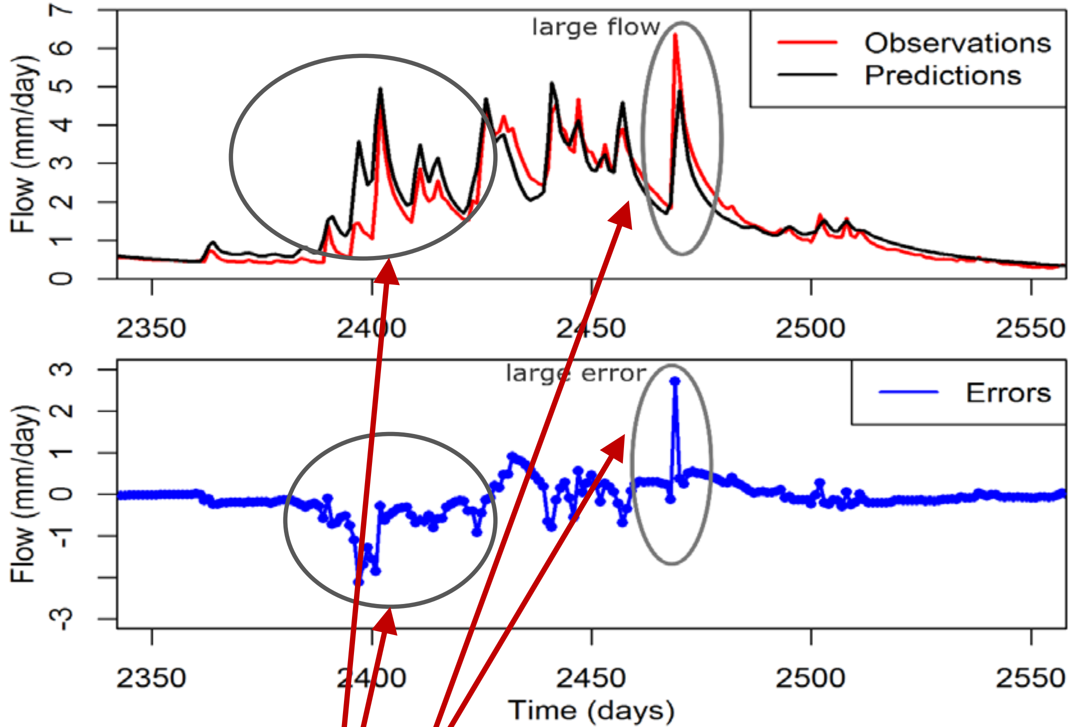
- Improve probabilistic predictions of daily streamflow
- Comprehensive evaluation of approaches for representing predictive uncertainty
- Provide recommendations for researchers and practitioners

Focus

- Aggregated approaches that use residual error models to represent total predictive uncertainty
- More pragmatic than decompositional approaches (e.g. BATEA) that identify individual sources of errors

Challenging features of residuals in hydrology

(Cotter River, Australia)



Streamflow time series

Residual errors time series

Residual = observations-predications

- **Errors are heteroscedastic** (larger errors in large flows)
- **Errors have persistence** (not independent between time steps)
- **Key Challenge:** Identifying residual error models that represent both "features" to achieve reliable and precise probabilistic predictions

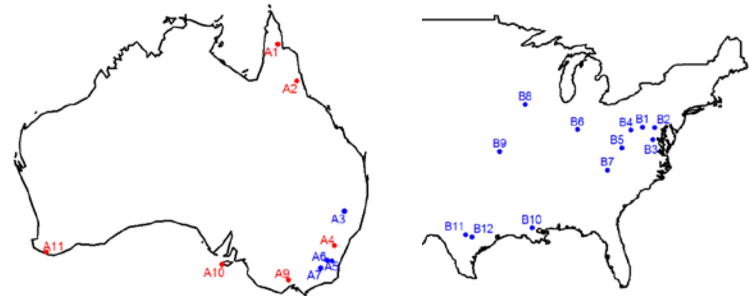
What is the "best" residual error model for making daily streamflow probabilistic predictions?

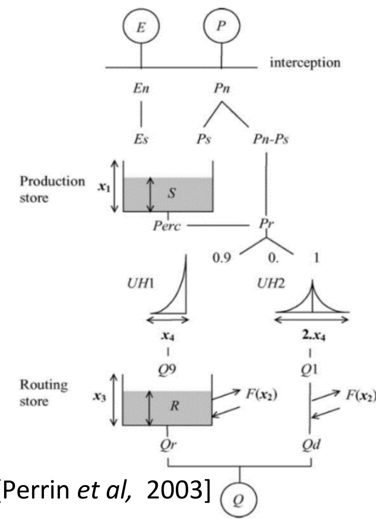
- Research Gap: No study had comprehensively compared the range of residual error models for representing heteroscedasticity in residuals

Residual Error Model	**Description**
No heteroscedasticity	
SLS	Standard least squares (error sd is constant)
Direct approaches for heteroscedasticity	
WLS	Weighted least squares (error sd increases linearly with predictions)
Transformational approaches for heteroscedasticity	
Log	Log transformation
Logsinh	Logsinh transformation (error sd increase "tapers off" with predictions)
BC (inferred λ)	Box-Cox transformation with inferred λ parameter
BC0.2	Box-Cox transformation with fixed λ= 0.2
BC0.5	Box-Cox transformation with fixed λ= 0.5
Reciprocal	Reciprocal transformation

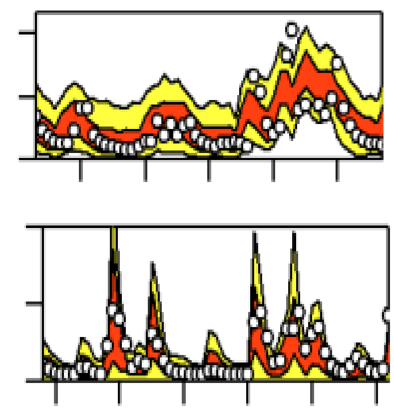
Features of Comprehensive Evaluation

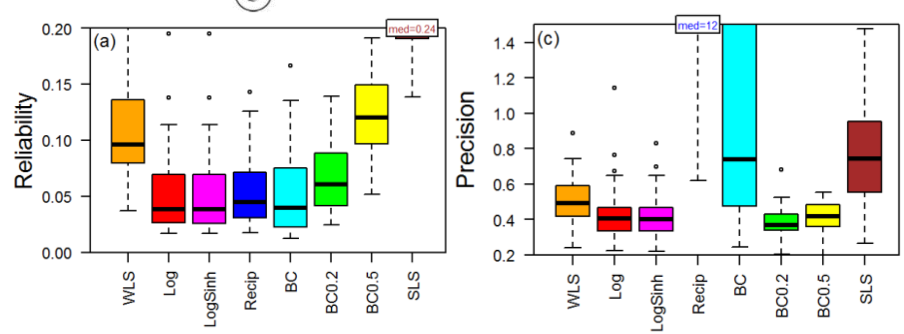
- Improve the robustness of recommendations
- Multiple Catchments
  - 23 climatologically diverse catchments from Australia and USA
- Two Hydrological Models
  - Lumped conceptual models: GR4J and HBV
- Multiple performance metrics
  - Reliability, precision and bias
  - Cross-validation over 10 yr
- Theoretical insights to understand differences in performance
  - Theoretical similarities and differences
  - Synthetic analysis
- McInerney et al (WRR2017)





[Perrin *et al*, 2003]





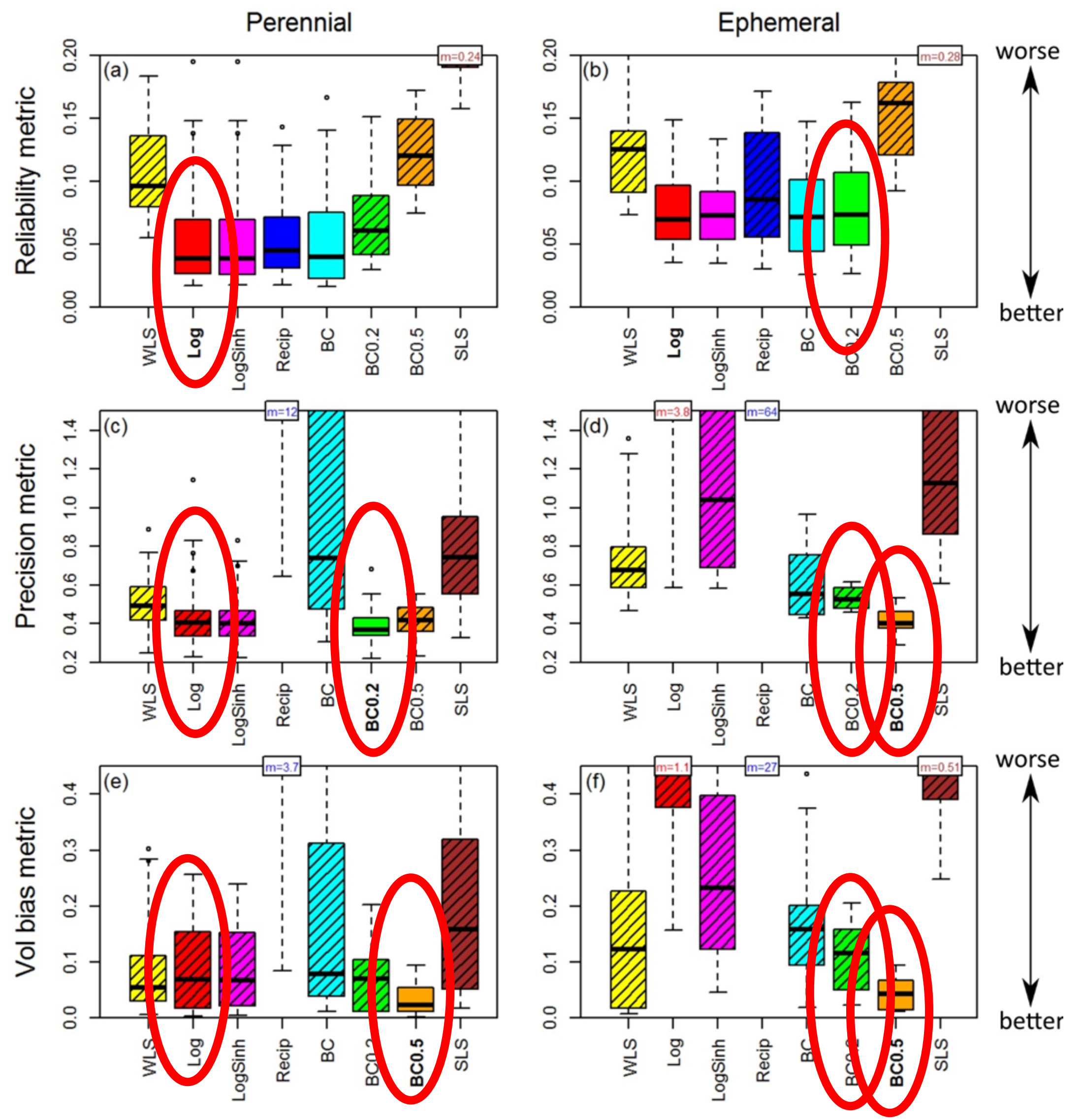
Key Findings: Empirical Results

- Results are dependent on catchment type (perennial/ephemeral)

"Best" Residual Error Models

Model	Outcome
Log	Best reliability in perennial Good precision and bias in perennial
BC0.2	Best precision in perennial Best reliability in ephemeral Good precision and bias in ephemeral
BC0.5	Best bias in perennial Best bias and precision in ephemeral

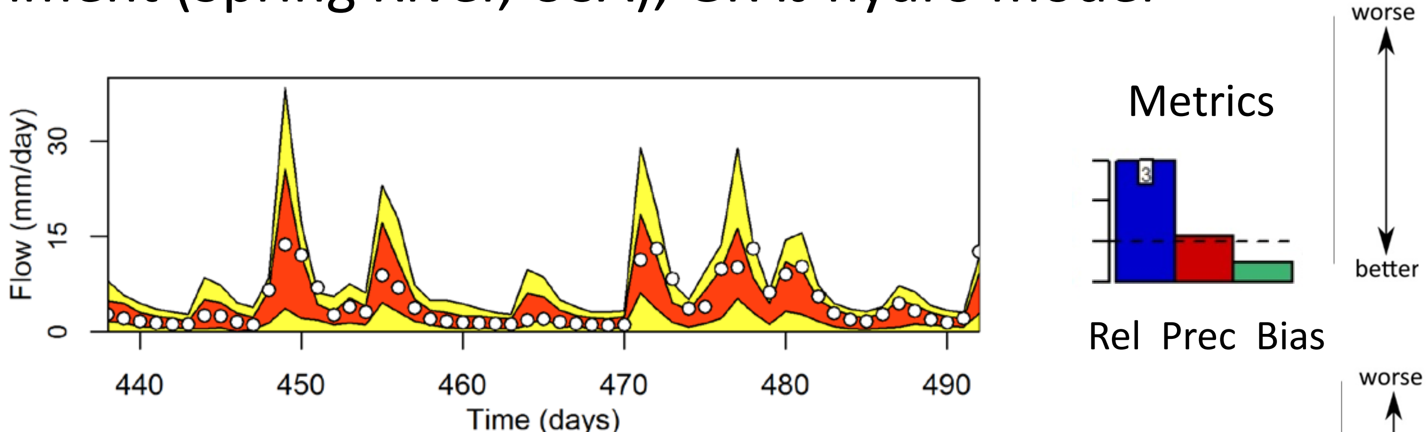
Not Recommended:
- SLS, WLS, Logsinh, BC(inferred λ), Reciprocal
- Either worse reliability, precision or bias or more complex



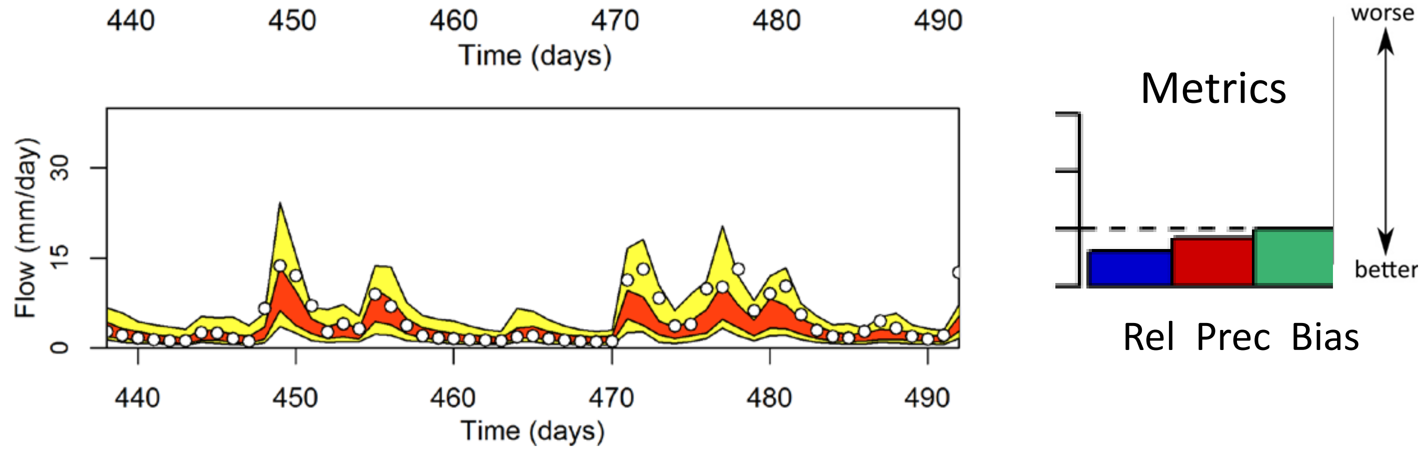
Transformational approaches (Log,BC) outperform direct approaches (WLS)

- Perennial catchment (Spring River, USA), GR4J hydro model

Weighted Least Squares (WLS)



Log transformation

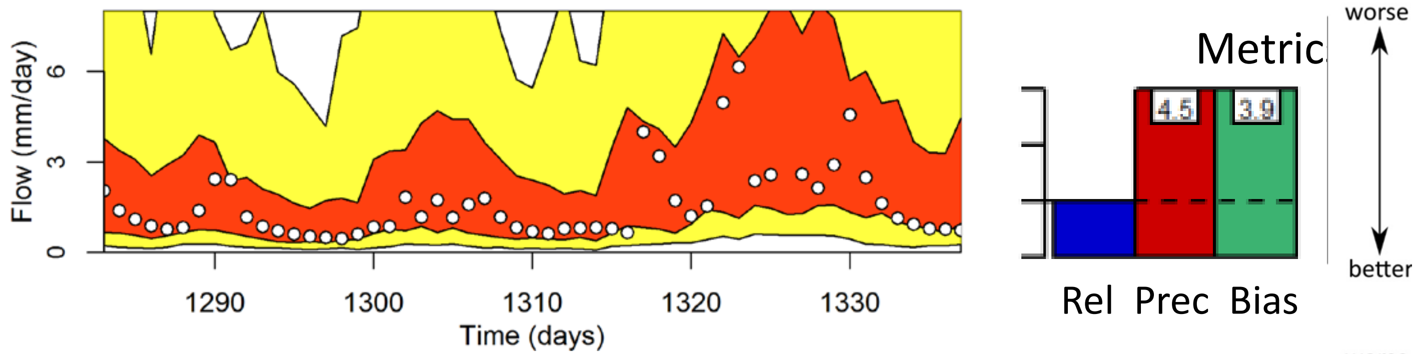


- Log transformation better reliability and precision than WLS
- Theoretical Insight: Transformational approaches (Log and BC) better capture skew and kurtosis in observed residuals than WLS

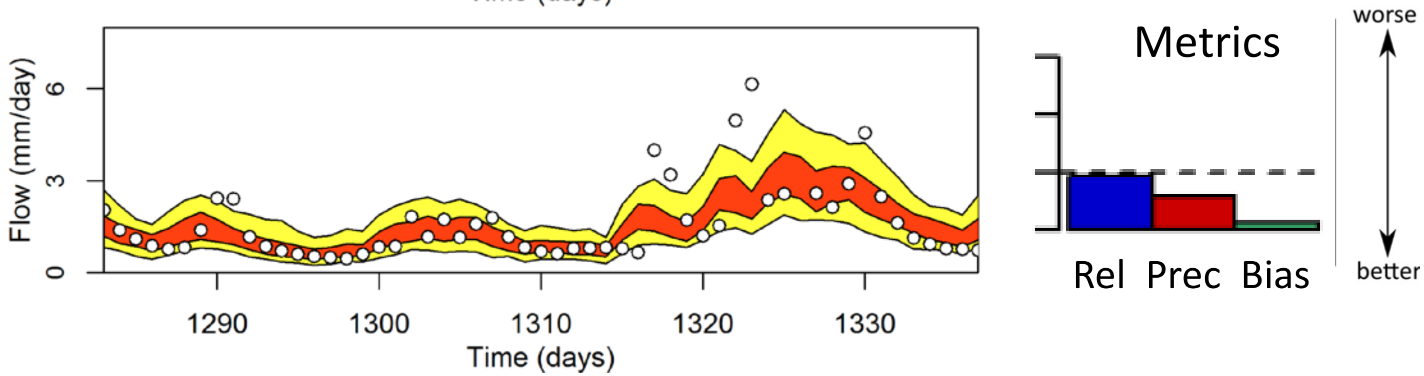
Box-Cox Transformation (fixed lambda) outperforms log transformation in Ephemeral Catchments

- Ephemeral catchment (Rocky River, SA), HBV hydro model

Log transformation



Box Cox transformation (λ=0.2)

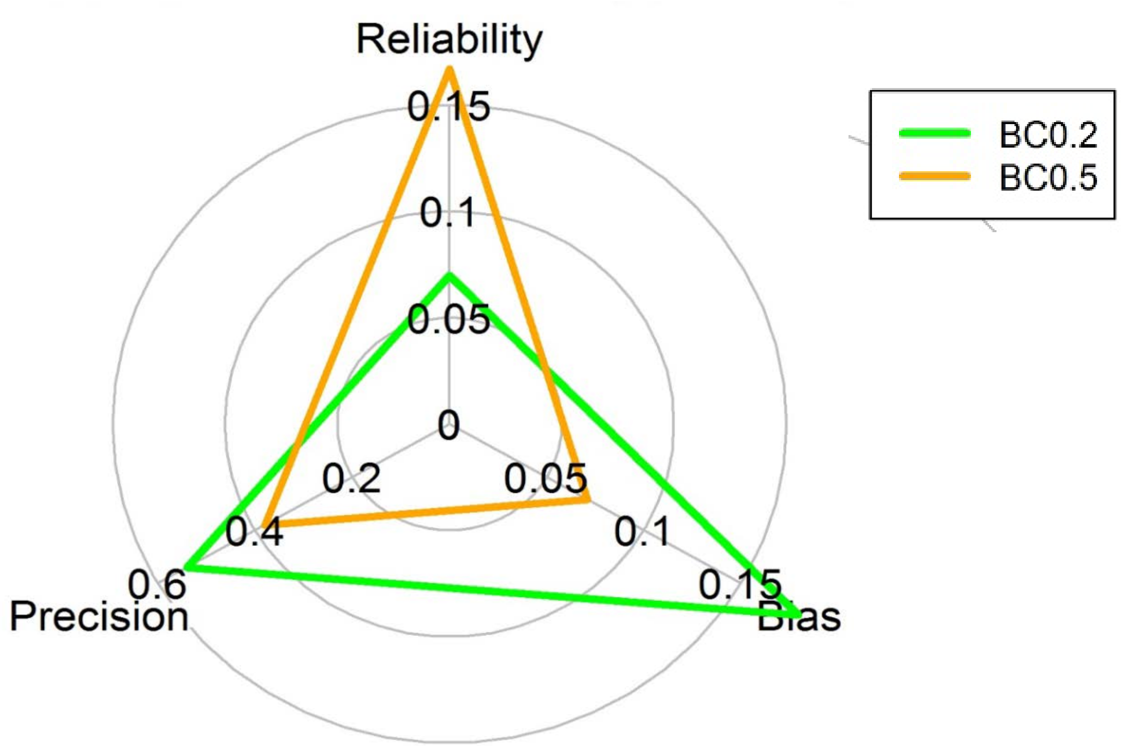


- BC0.2 has similar reliability, but much better precision and bias than log
- Log produces poor precisions (unrealistically large uncertainty) and large bias in ephemeral catchments
- Theoretical Insight: BC transformation better handles zero flows than log in ephemeral catchments

Choose multiple "best" residual error models due to performance trade-offs across multiple metrics

- Not possible to choose a single model that performs best across all metrics

Ephemeral Catchments



- Pareto Optimal Approaches
  - Perennial: Log, BC0.2 and BC0.5
  - Ephemeral: BC0.2 and BC0.5

Broad Recommendations

In **perennial** catchments, use

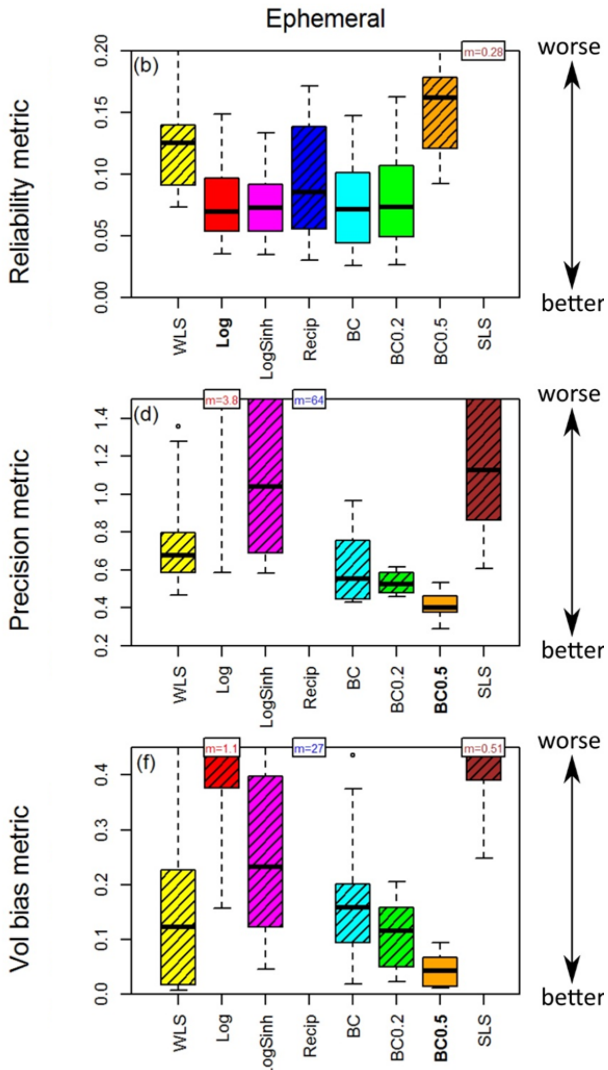
- Log transformation if reliability is important
- Box Cox transformation with fixed λ=0.2 if precision is important
- Box Cox transformation with fixed λ=0.5 if low bias is important

In **ephemeral** catchments, use

- Box Cox transformation with fixed λ=0.2 if reliability is important
- Box Cox transformation with fixed λ=0.5 if precision/bias important

Based on 'median' results across 23 catchments, individual catchment results can differ.

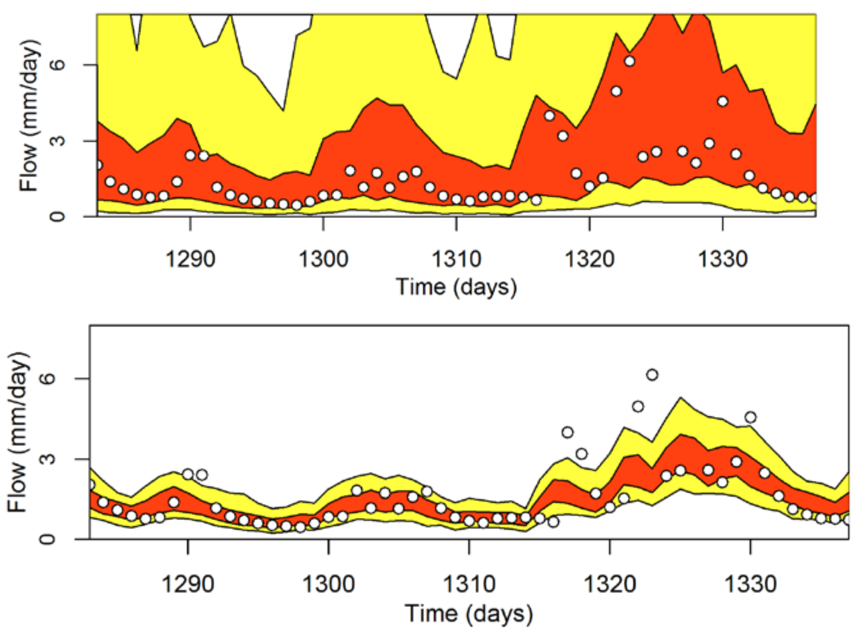
Impact: Significant improvement in probabilistic performance



- Larger impact in ephemeral catchment
- Improved reliability
- Improved precision 105% to 40% of obs streamflow
  - Reduce predictive uncertainty by factor of 2!!!
- Reduced bias from 25% to 4%

Log transformed flows

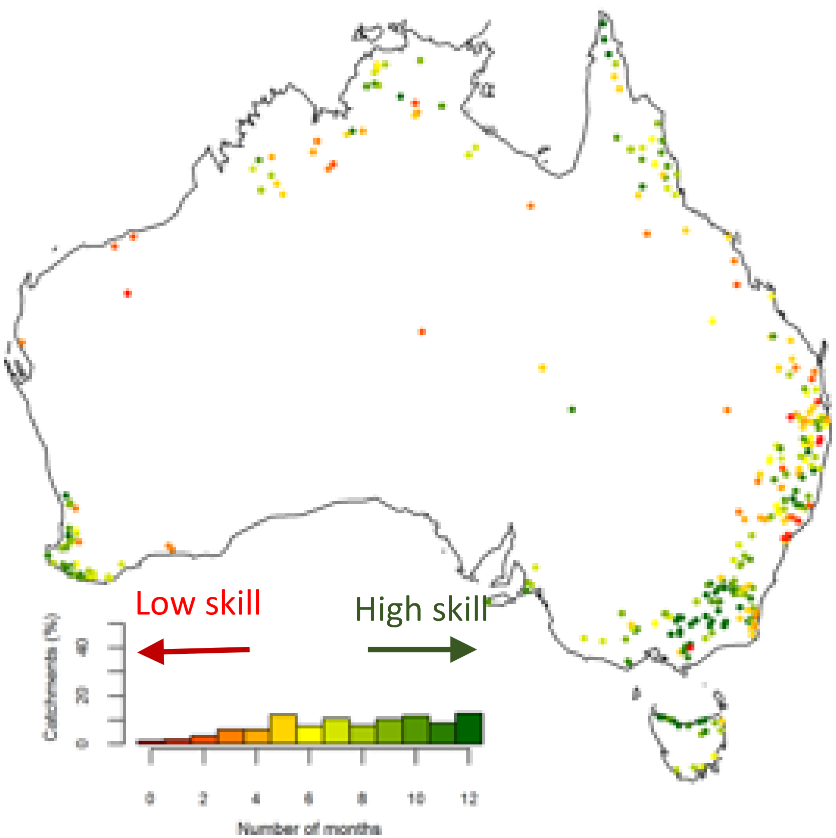
BC0.2 transformed flows



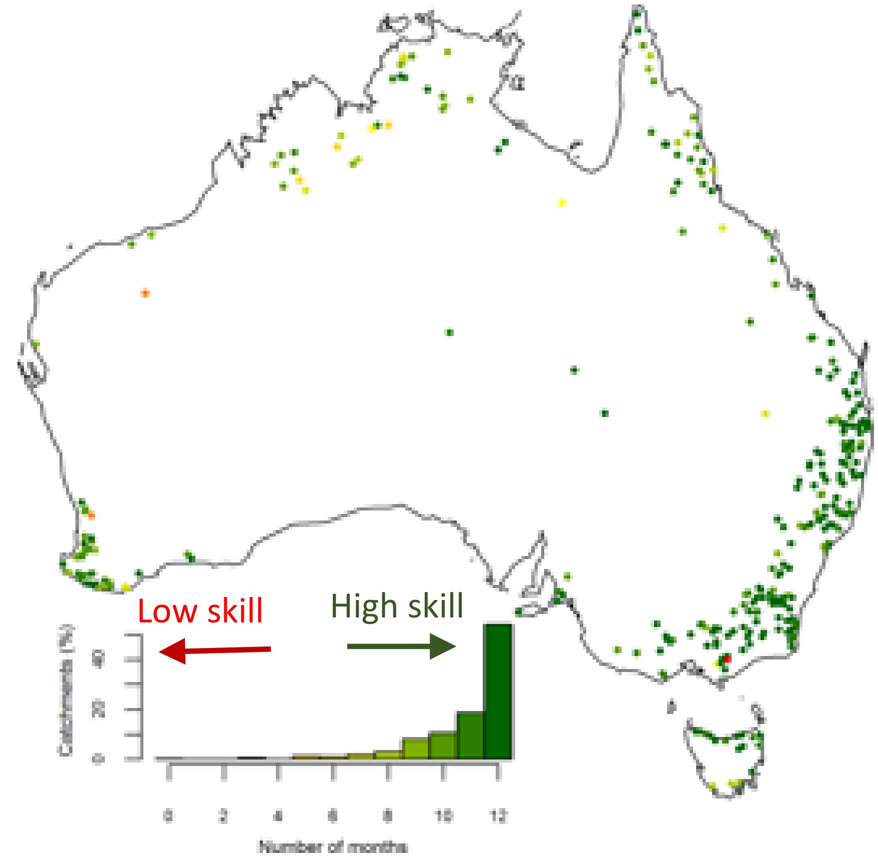
Impacts: Bureau of Meteorology Seasonal Streamflow Forecasts

- Recommendations used to enhance monthly streamflow post-processor

Log/Logsinh



BC0.2



- High Forecast skill: >10 months with reliable forecasts more precise than climatology
- Log/Logsinh: 25-30% sites with high skill
- BC0.2: >80% of sites with high skill
- Preliminary results, subject to peer review (Woldemsekel et al. in prep)

Summary

- Comprehensive evaluation of approaches for predictive uncertainty
  - Eight Residual Error Approaches: Simple=>Complex
  - Multiple catchments/hydro models/performance metrics
  - Theoretical Insights: Understanding reasons for differences in performance
- Broad recommendations
  - “Best” Pareto optimal residual error models in different catchment types
  - significant reductions in predictive uncertainty, while maintaining reliability
- Practical implications: Simplest is often best!
  - Smart use of simple approaches => best predictive performance
  - Simple to implement for reseachers practitioners
- Future research opportunities
  - “Best” residual error model selected from existing approaches
  - Opportunity to improve predictions across flow range, esp near-zero/zero flows

Water Resources Research

RESEARCH ARTICLE

10.1002/2016WR019168

Improving probabilistic prediction of daily streamflow by identifying Pareto optimal approaches for modeling heteroscedastic residual errors

David McInerney[1], Mark Thyer[1], Dmitri Kavetski[1,2], Julien Lerat[3], and George Kuczera[2]

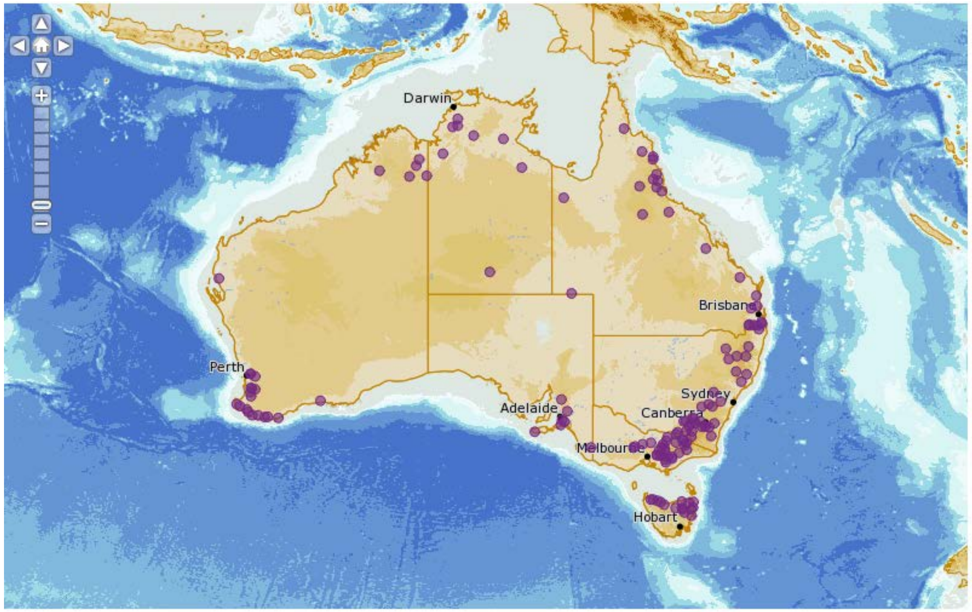
Key Points:

- Choice of heteroscedastic error modeling approach significantly impacts on predictive reliability,

Key Findings: Theoretical Insights

Residual Error Model	Outcome *and Insight*
Log	**Best Reliability in Perennial and Ephemeral** - *Captures heteroscedastity in residuals better than SLS* - *Captures skew and kurtosis in residuals better than WLS* - *Logsinh performance similar to log due to estimated loginsh parameter values*
BC0.2	**Best Precision in Perennial** **-** *BC (inferred λ) has poor precision due to overfitting of low flows* **Better Precision and Bias than log in Ephemeral** - *captures zero flows better than log*
BC0.5	**Best Bias in Perennial** **Best Bias and Precisions in Ephemeral** **Poor Reliability**

Impacts on Forecasting: Bureau of Meteorology Seasonal Streamflow Forecasts



- Seasonal forecasts at ~300 locations
- Used by water managers around Australia
- Based on Statistical and Dynamic Seasonal Streamflow Forecasting System

Dynamic Seasonal Streamflow Forecasting System

Rainfall forecasts (daily)

↓

Rainfall post-processing

↓

Rainfall → Runoff Model + Calibration Approach

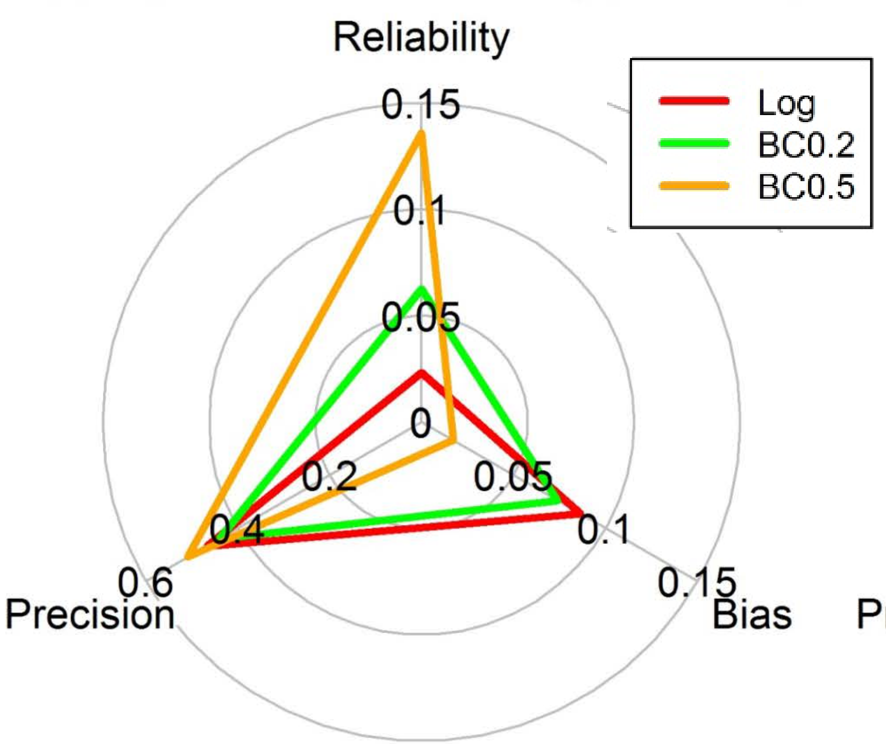
↓

Streamflow post-processing

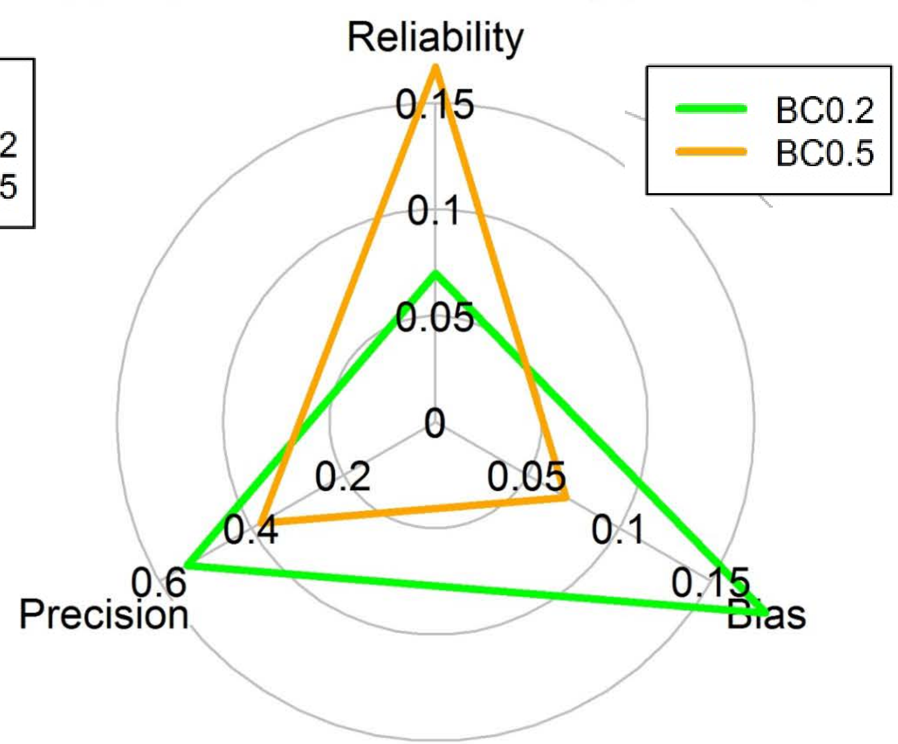
Applied Recommendations to enhanced streamflow post-processor at monthly time scale

Choose multiple "best" residual error models due to performance trade-off's: Pareto optimal approaches

Perennial Catchment



Ephemeral Catchment



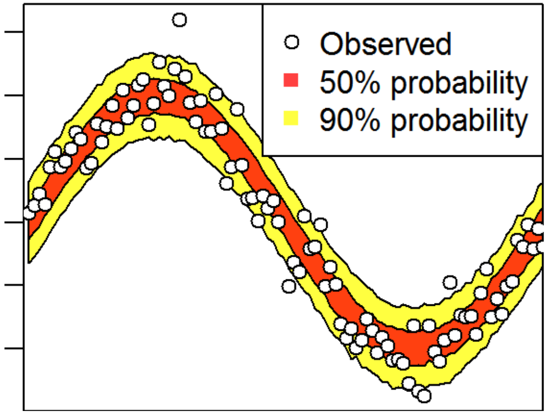
- Not possible to choose a single model that performs best across all metrics
- Pareto Optimal Approaches
  - Perennial: Log, BC0.2 and BC0.2
  - Ephemeral: BC0.2 and BC0.5

Multiple Performance Metrics: What makes good probabilistic predictions?

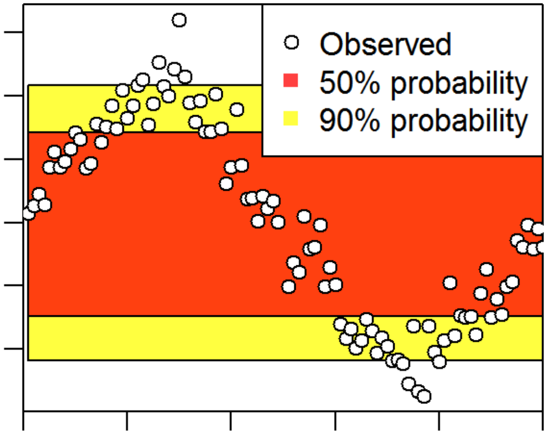
We want predictions that are

- **Reliable:** Predictions statistically consistent with observed data
- **Precise:** Small uncertainty in predictions
- With **low volumetric bias:** total volume from predicted flow matches observations

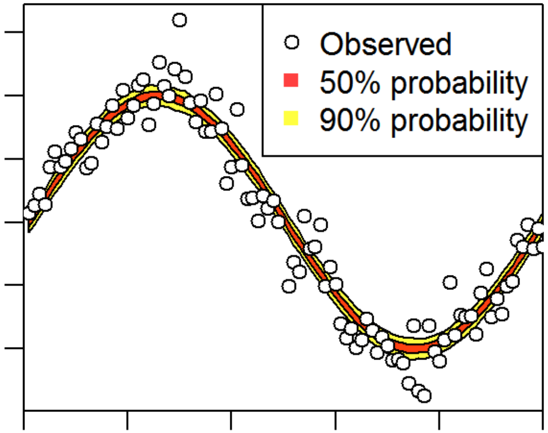
Reliable, precise, unbiased



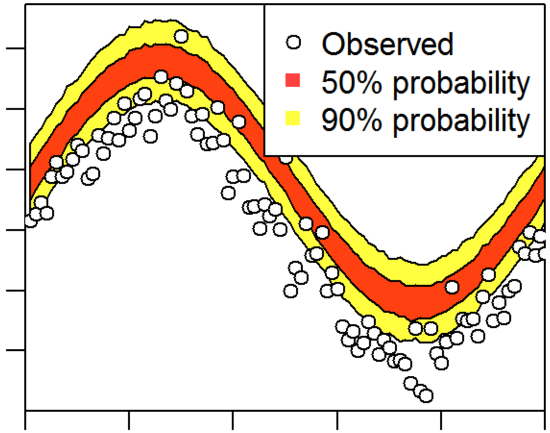
Reliable but imprecise



Precise but unreliable

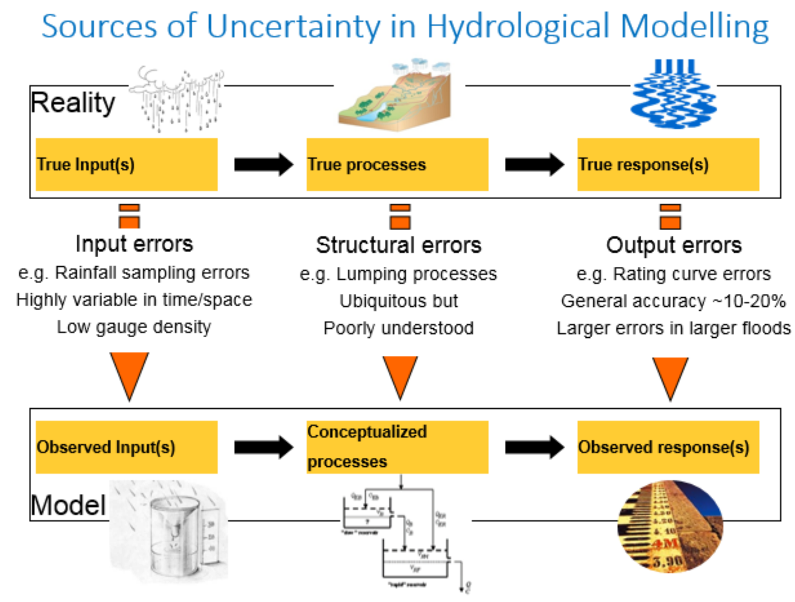


Biased

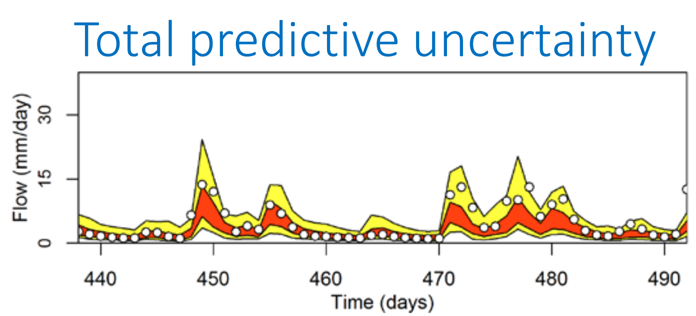


Approaches to modelling uncertainty: Find the right tool for the job

- Decompositional: Estimate individual sources of uncertainty (e.g. BATEA)
  - Diagnose dominant sources of uncertainty
  - Computationally challenging, requires more data and expertise
  - Not really "off-the-shelf" method



- Aggregated: Estimate total uncertainty in predictions
  - Lump all uncertainty into single residual term
  - Common, easy to apply => "off-the-shelf"
  - Unable to estimate the dominant sources



- For decision-making, total predictive uncertainty is of key interest
- Focus: Evaluate residual error models for representing total uncertainty in predictions

Key Findings: Empirical Results: "Best" Residual Error Models

Residual Error Model	Outcome
Log	**Best Reliability in Perennial and Ephemeral**
BC0.2	**Best Precision in Perennial** **Better Precision and Bias than log in Ephemeral**
BC0.5	**Best Bias in Perennial** **Best Bias and Precisions in Ephemeral** **Poor Reliability**