An assessment of hydrological models commonly used in NZ

By:

Etienne Gil-Goldsbrough (WSP)

Mark Groves (WSP)

Brendon Schicker (WSP)

Scope

1. Develop a method of scoring the calibration fit of hydrological models
   - Need to assess the **quality of fit**
   - Also assess the **consistency of fit**
   - To do this we built an Excel tool
2. Assess how different models perform in NZ catchments
   - Used 10yrs of real data across three catchments to compare models to observed flow.

Hydrological Models Tested

- Horton (linked to soil moisture deficit (SMD) similar to MIKE / ICM)
- Curve Number (SCS-CN)
- Fixed % run-off

Methodology

- Spreadsheet model of the three hydrological models.
- Consistent run-off routing model and time of concentration (Tc) estimate.
- Fit scored consistently across three calibration events for each river system.
- Best fit obtained using a multi-start non-linear solver.
- Scoring and variance automated.

Assessing Visual Fit

- R^2 can give misleading results – explains variance from a linear regression model – **not goodness of fit!**
- RMSE based on the 'residual' so a better method
  0 = perfect fit. Measures peak error.
- NSE developed for this purpose 1 = perfect fit. Measures overall volume difference.
- Hwang et al recommend combining two methods e.g. NSE and RMSE.
- Above approach tested.

Scoring Methodology

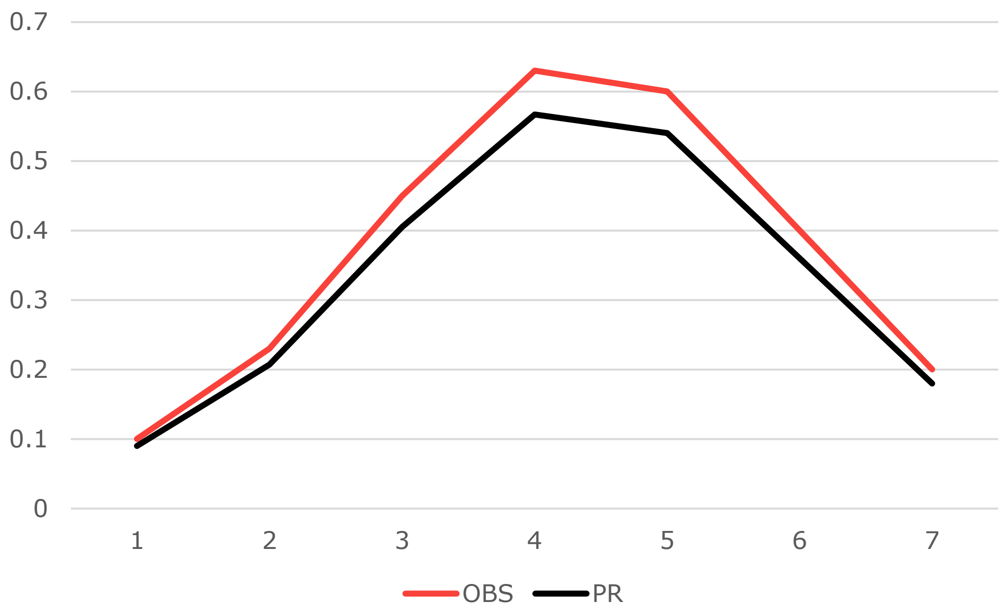
- Nash-Sutcliffe Efficiency (NSE) developed for hydrological calibration. Measures overall fit (and thus volume).
- RMSE as a measure of peak error.
- The two methods are combined for an overall score (volume and peak).

$$NSE - (RMSE \div Q_M)$$

Q_M = *mean flow*

< 0	*= unacceptable*
0.0 - 0.2	*= poor calibration*
0.2 - 0.4	*= moderate calibration*
0.4 – 0.6	*= good calibration*
> 0.6	*= excellent calibration*

Assessing Visual Fit – Off by Constant Factor



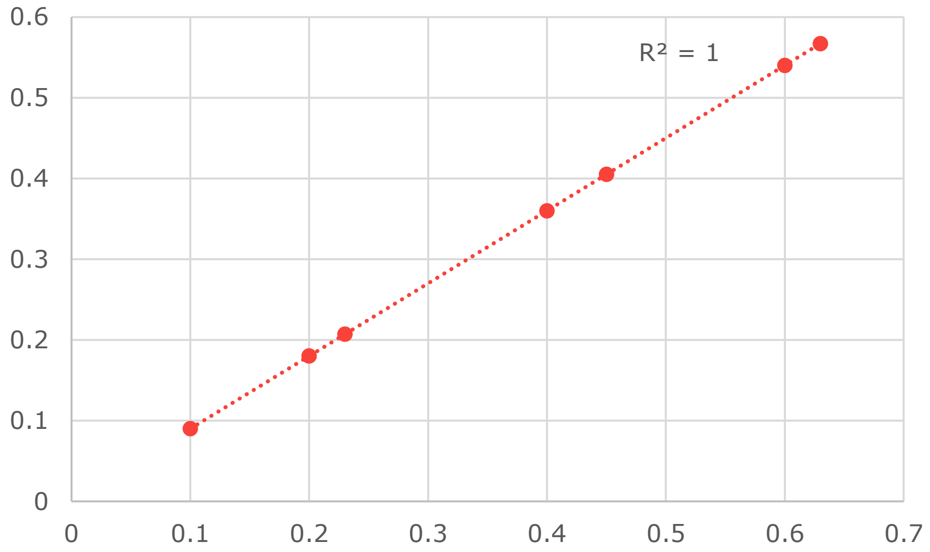
R^2 = 1.0

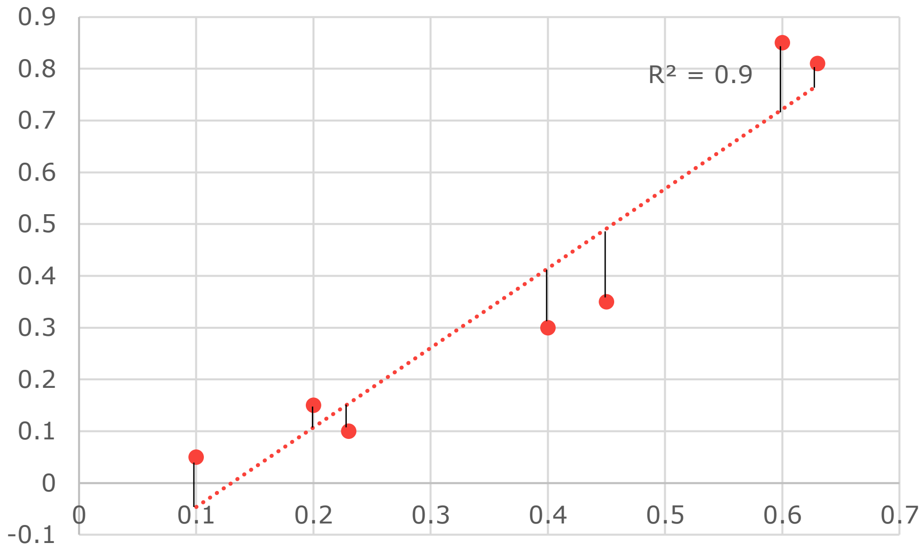
NSE = 0.95

RMSE = 0.11 m^3/s (10%)

Score = 0.84 (excellent fit)

Assessing Visual Fit

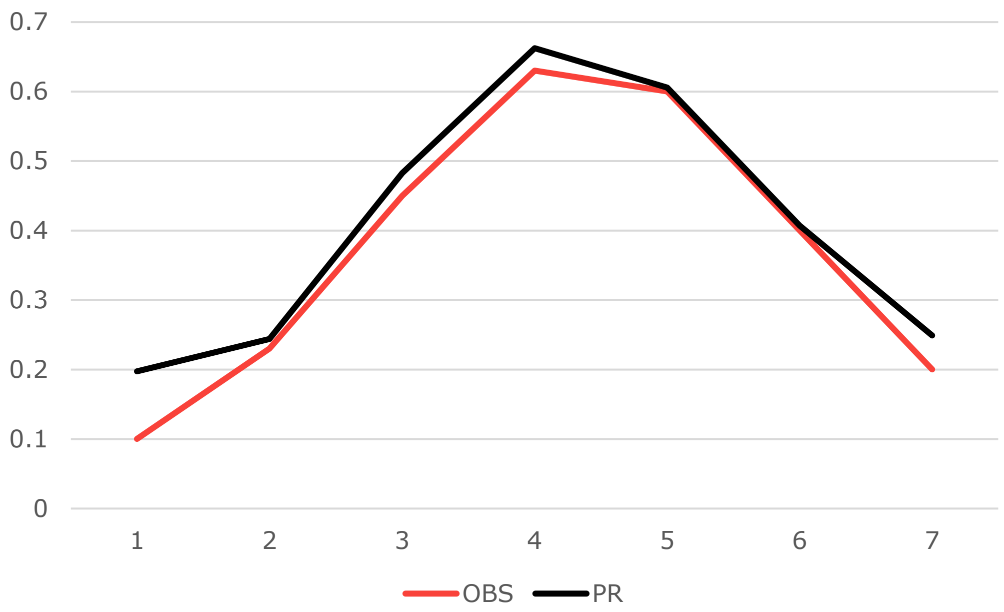




How does $R^2 = 1$ when it is not a perfect model? Because it describes the variance explained by a simple linear regression model and does not need to mean y = x (i.e. observed = predicted)

If a model is out massively, but by a constant % it will have $R^2 = 1$ as the data can be 100% explained by a linear regression model of the form y = mx + c. Hence not a good criteria to use!

Assessing Visual Fit – Good Fit



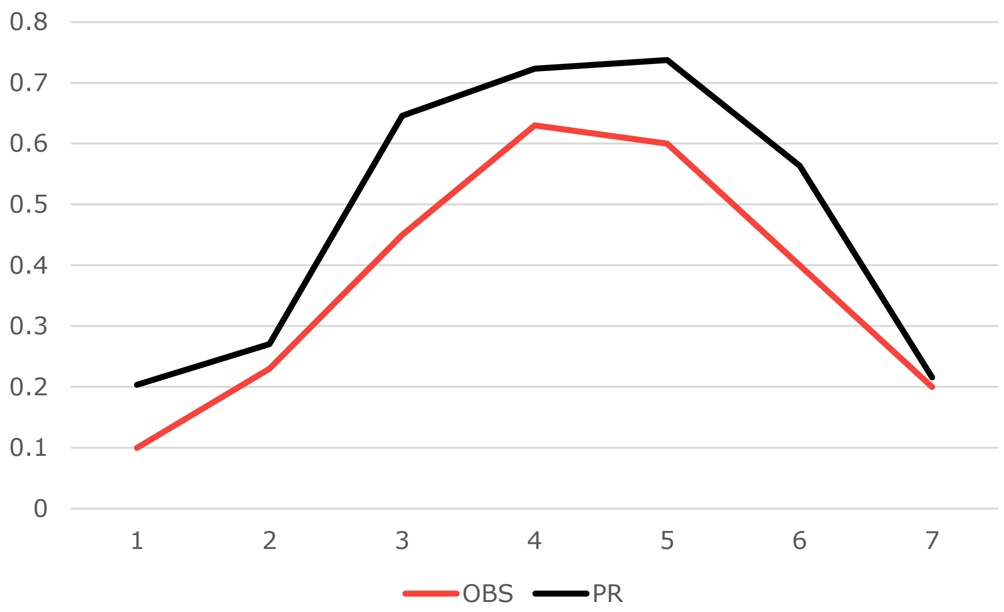
R^2 = 0.98

NSE = 0.88

RMSE = 0.18 m^3/s (15%)

Score = 0.71 (excellent fit)

Assessing Visual Fit – Poor Fit



R^2 = 0.94

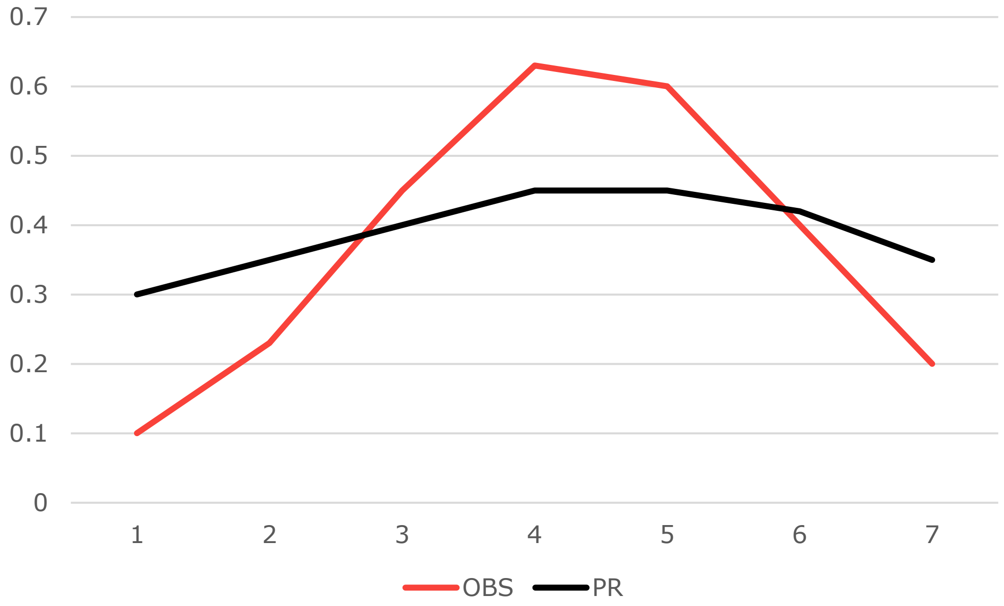
NSE = 0.62

RMSE = 0.31 m^3/s (24%)

Score = 0.31 (poor fit)

R^2 is high, but model is a poor fit for the data.

Assessing Visual Fit – Poor Fit



R^2 = 0.95

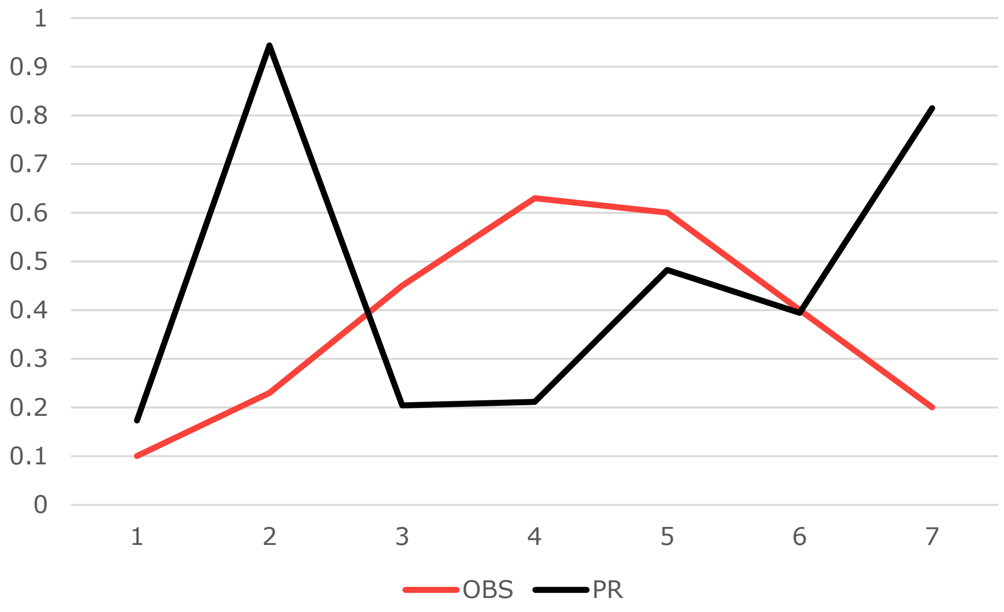
NSE = 0.46

RMSE = 0.37 m³/s (29%)

Score = 0.09 (poor fit)

R^2 is high, but model is a poor fit for the data.

Assessing Visual Fit – Random



R^2 = 0.07

NSE = -3.59

RMSE = 1.08 m^3/s (46%)

Score = -4.67

(no calibration)

Limitations

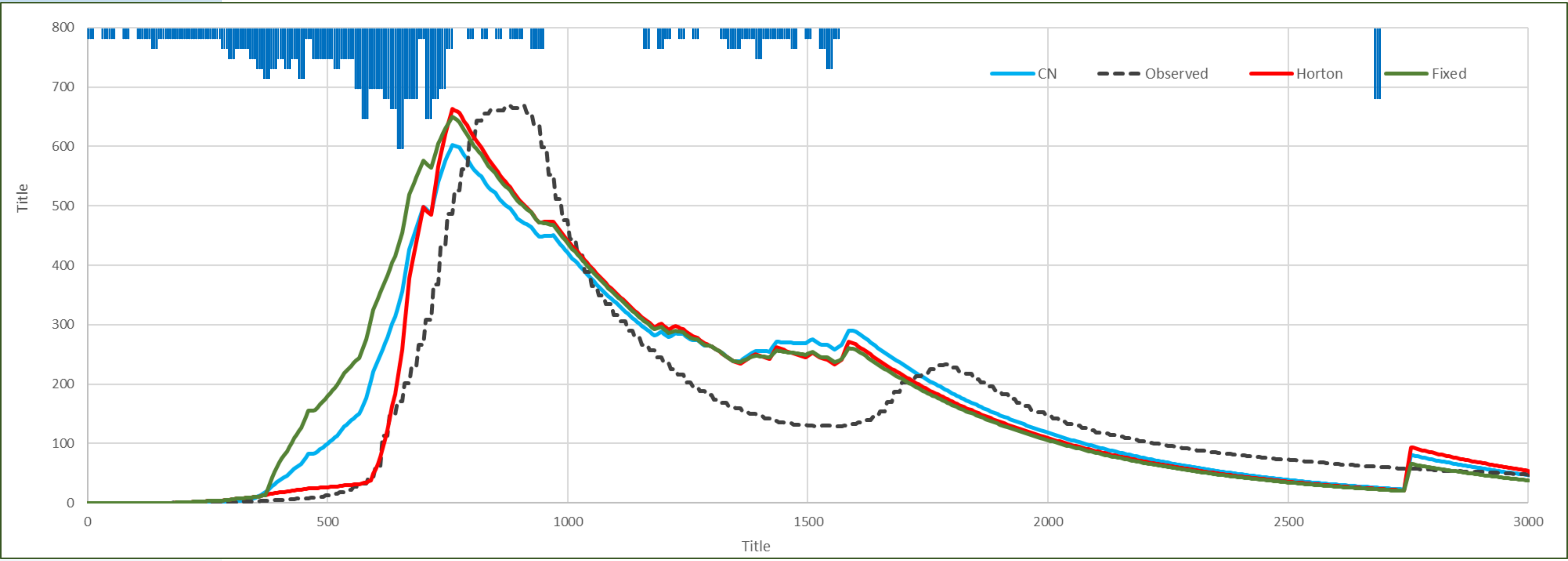
- Spikes (e.g. pumped flow) – smooth data to remove spikes
- Data drop outs – could add logic to equation
- Not many “small” catchments have rain gauges (RG) and flow monitors (FM) with many years of data (for low AEP events need 10+ years of data in ideally 5 minute increments)
- Relying on one RG opens up issues. In practice RG often near FM so not in centre of catchment
- We analysed three events to avoid overfitting to a specific event.

Hutt River

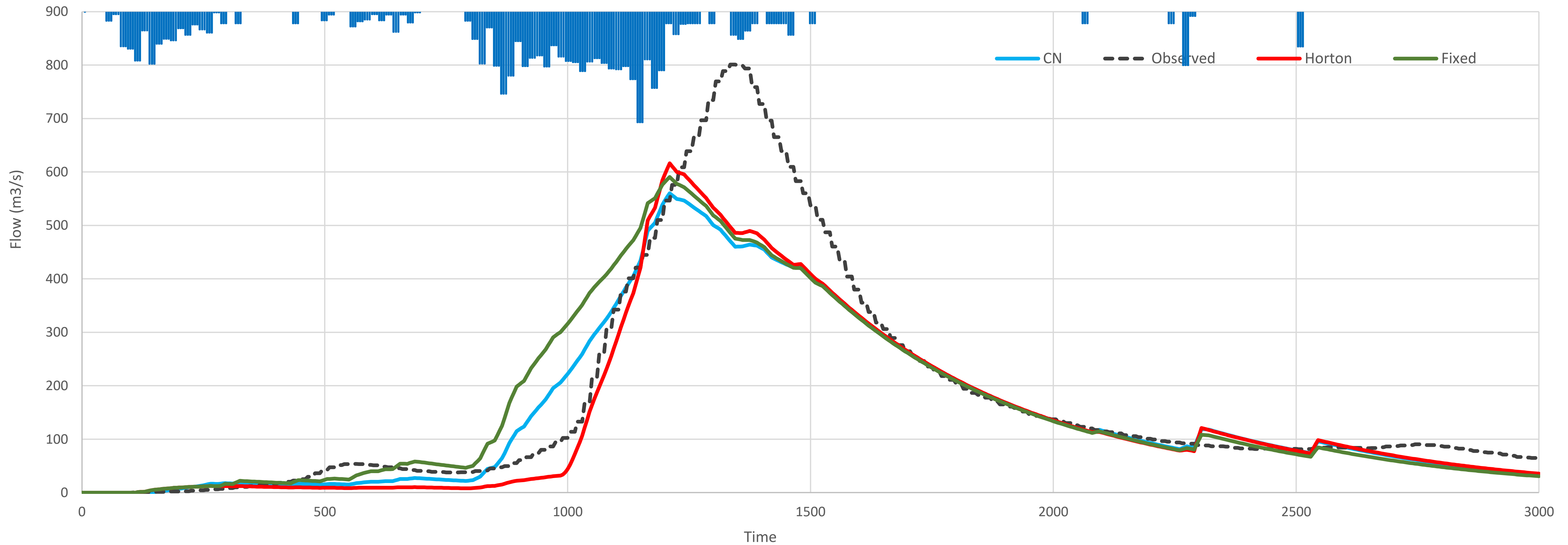
- 450 km^2 (approx.)
- 450 minute Tc
- Flow gauged at Taita Gorge
- Rainfall gauge near flow gauge used

Too large to be represented with one rain gauge and one catchment in practice

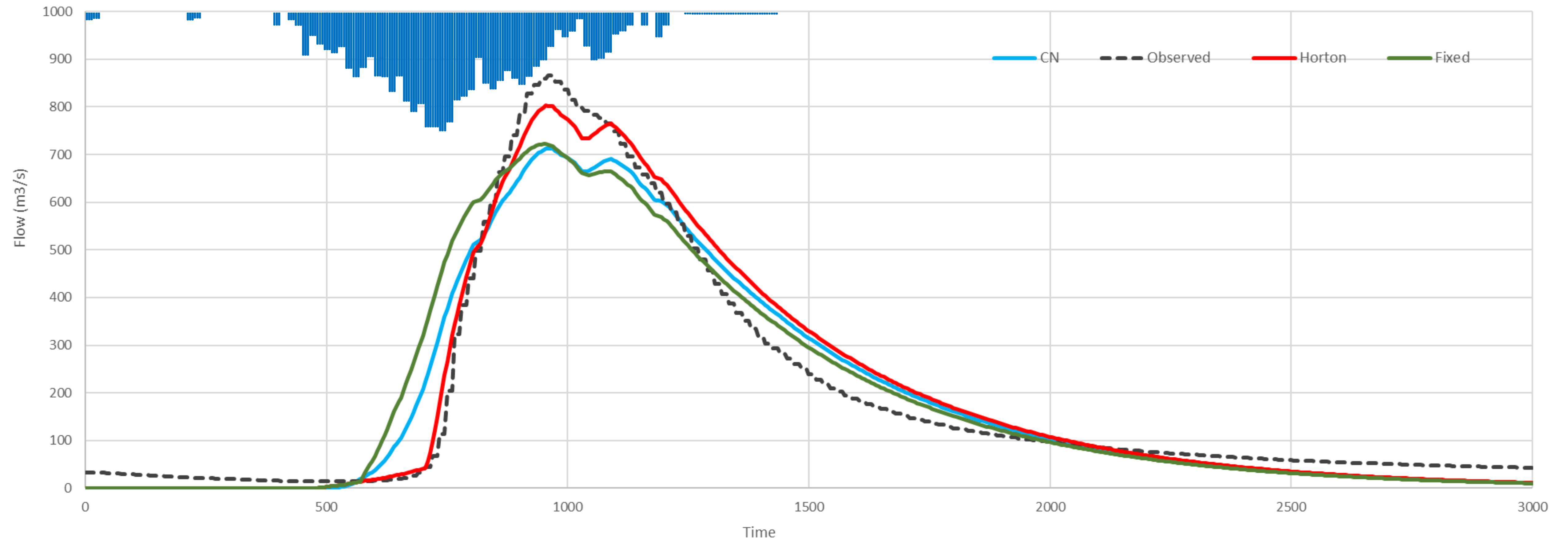
Hutt River – Event 1



Hutt River – Event 2



Hutt River – Event 3



Hutt River – Score

Model Type	Mean Score	Range	Result
CN	0.44	0.35	Good
Horton	0.55	0.33	Good
Fixed	0.30	0.44	Poor

< 0 = unacceptable

0.0 - 0.2 = poor calibration

0.2 - 0.4 = moderate calibration

0.4 – 0.6 = good calibration

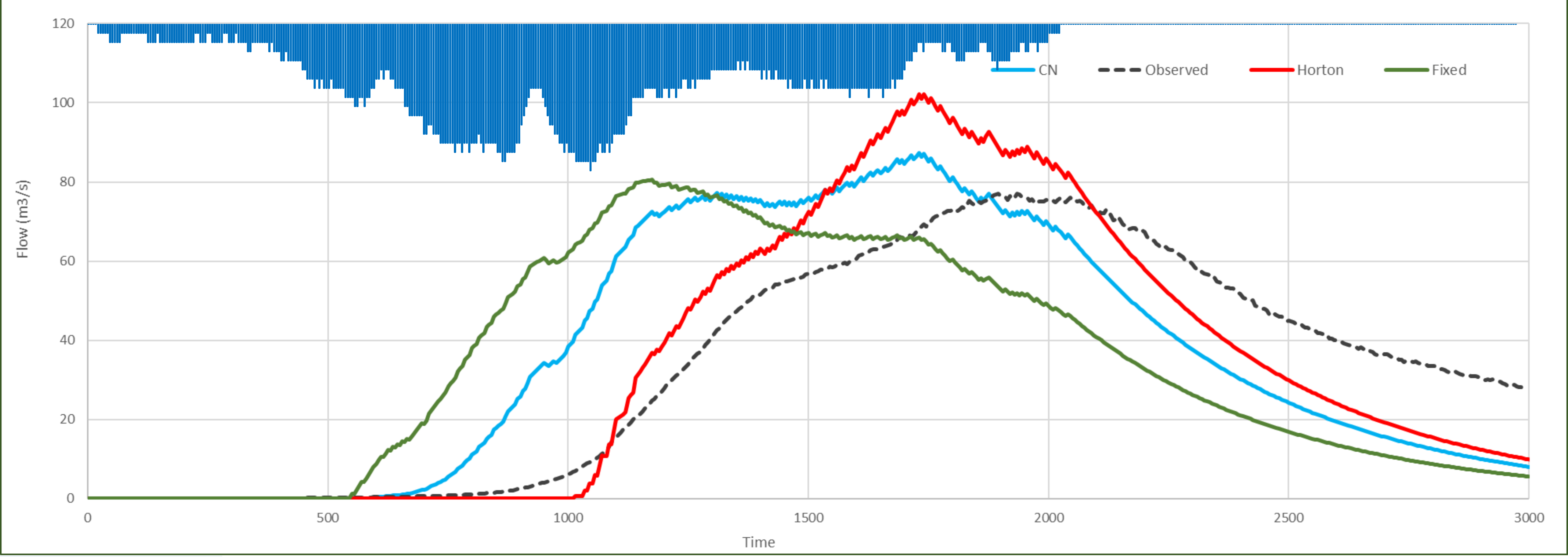
> 0.6 = excellent calibration

Selwyn River

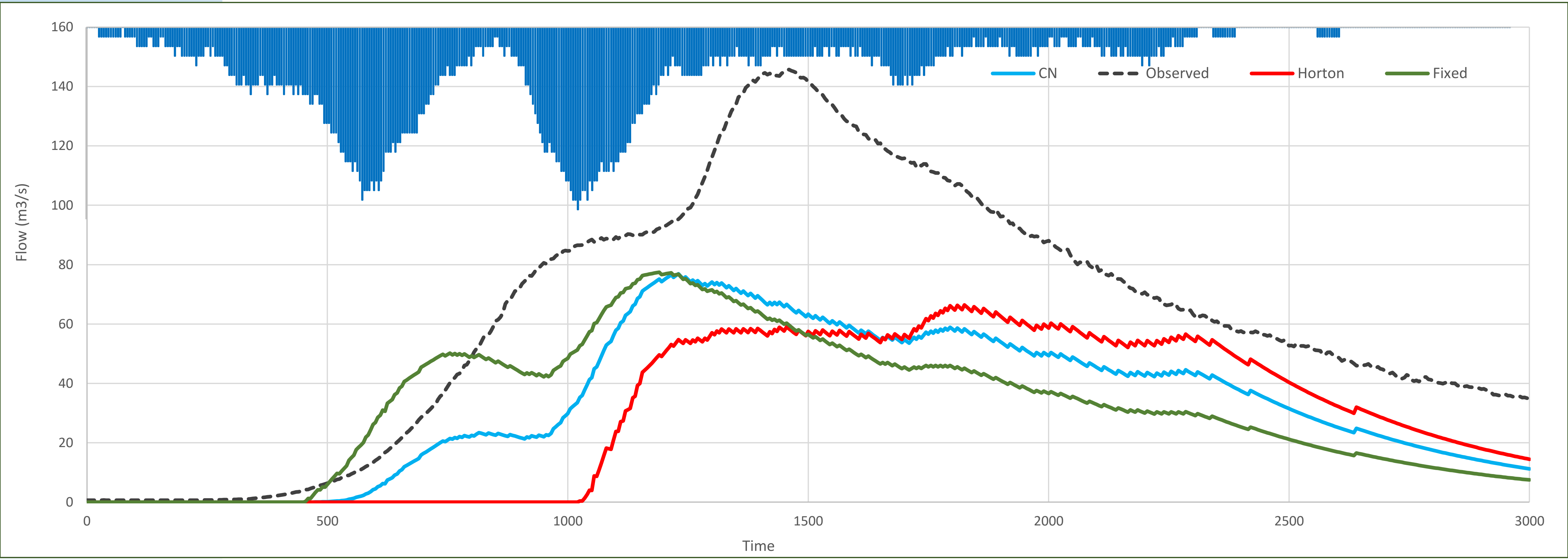
- 164 km^2 (approx.)
- 456 minute Tc
- Flow gauged at Whitecliffs
- Rainfall gauge near flow gauge used

Too large to be represented with one rain gauge and one catchment in practice

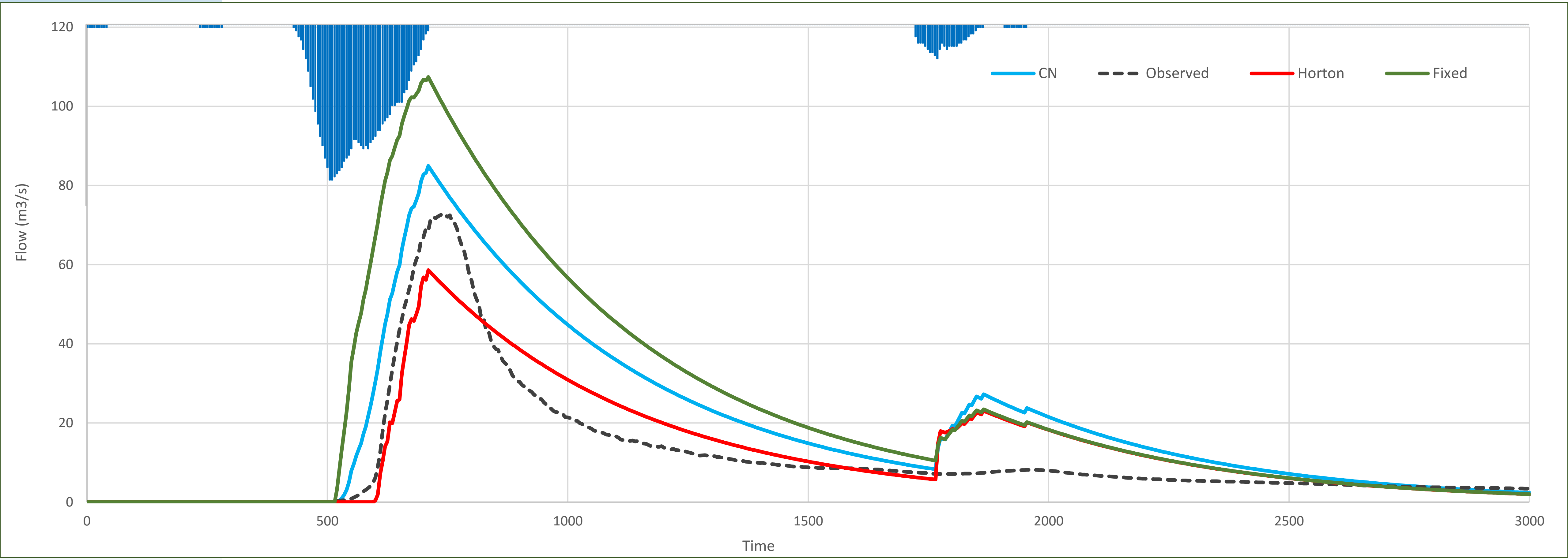
Selwyn at Whitecliffs – Event 1



Selwyn at Whitecliffs – Event 2



Selwyn at Whitecliffs – Event 3



Selwyn at Whitecliffs - Score

Model Type	Mean Score	Range	Result
CN	-0.35	0.39	Unacceptable
Horton	-0.02	1.07	Unacceptable
Fixed	-1.27	1.77	Unacceptable

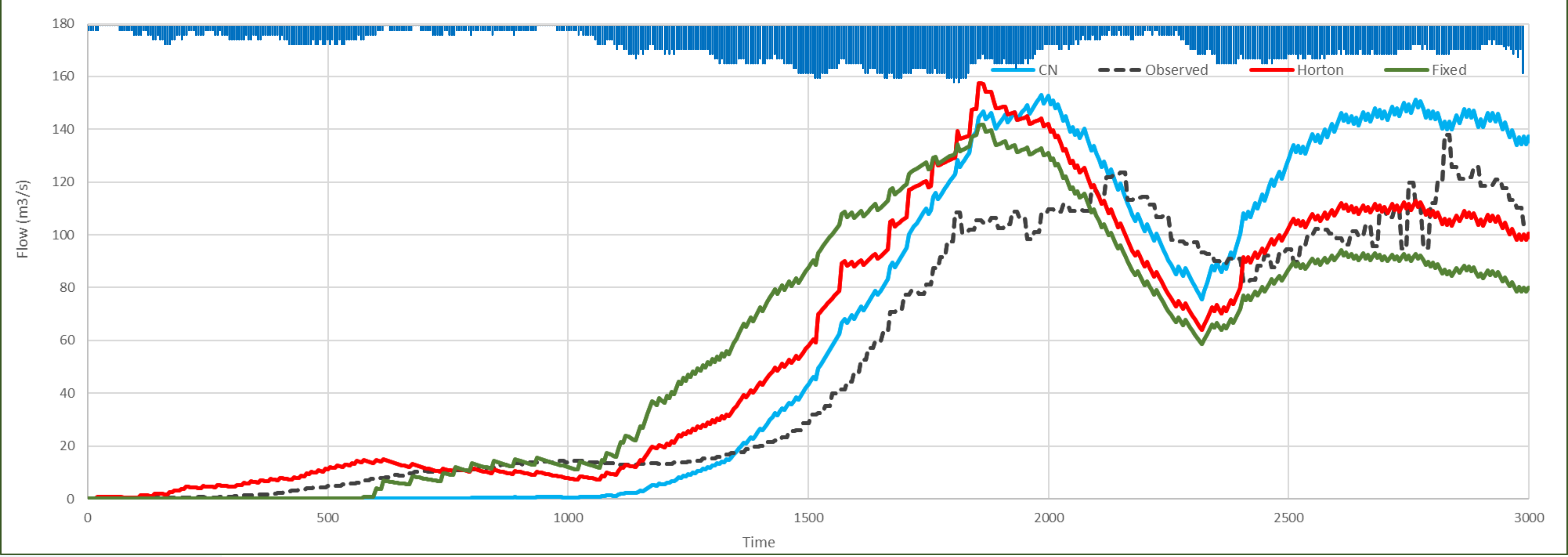
< 0 = unacceptable
0.0 - 0.2 = poor calibration
0.2 - 0.4 = moderate calibration
0.4 – 0.6 = good calibration
> 0.6 = excellent calibration

North Ashburton River

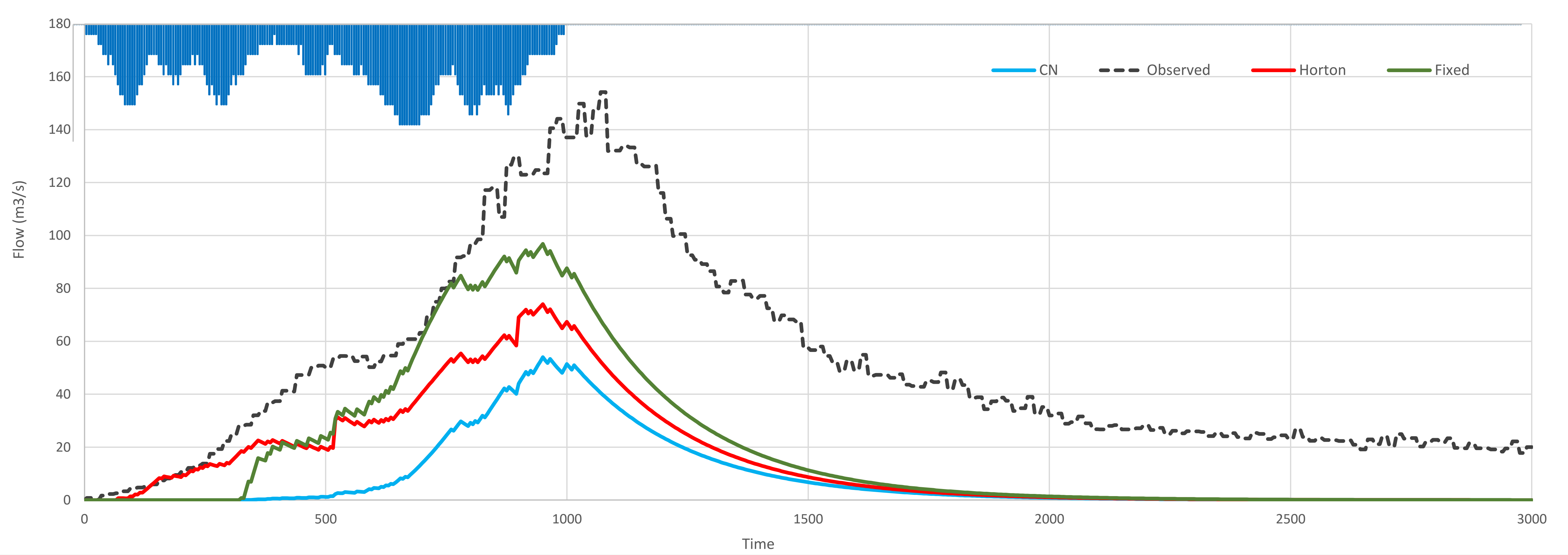
- 276 km^2 (approx.)
- 242 minute Tc
- Flow gauged at Cookies Hut
- Rainfall gauge near flow gauge used

Too large to be represented with one rain gauge and one catchment in practice

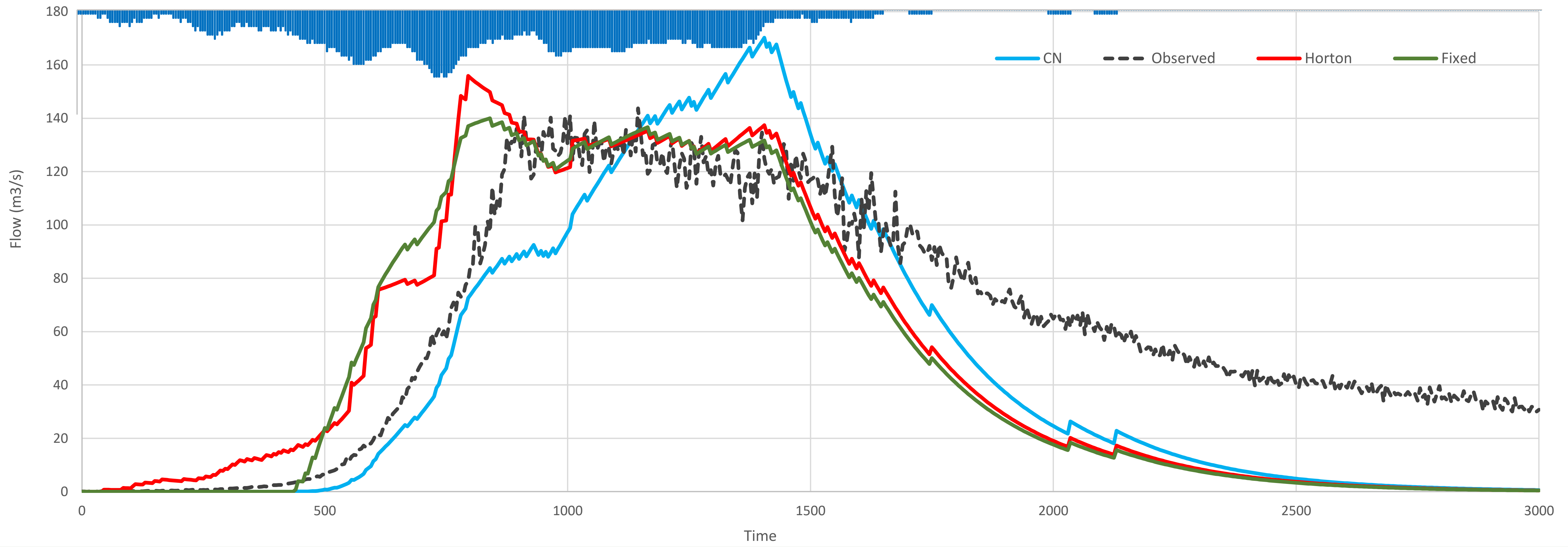
Ashburton River at Cookies Hut - Event 1



Ashburton River at Cookies Hut - Event 2



Ashburton River at Cookies Hut - Event 3



Ashburton River at Cookies Hut - Score

Model Type	Mean Score	Range	Result
CN	-0.37	2.04	Unacceptable
Horton	-0.19	1.49	Unacceptable
Fixed	-0.18	0.80	Unacceptable

< 0 = unacceptable

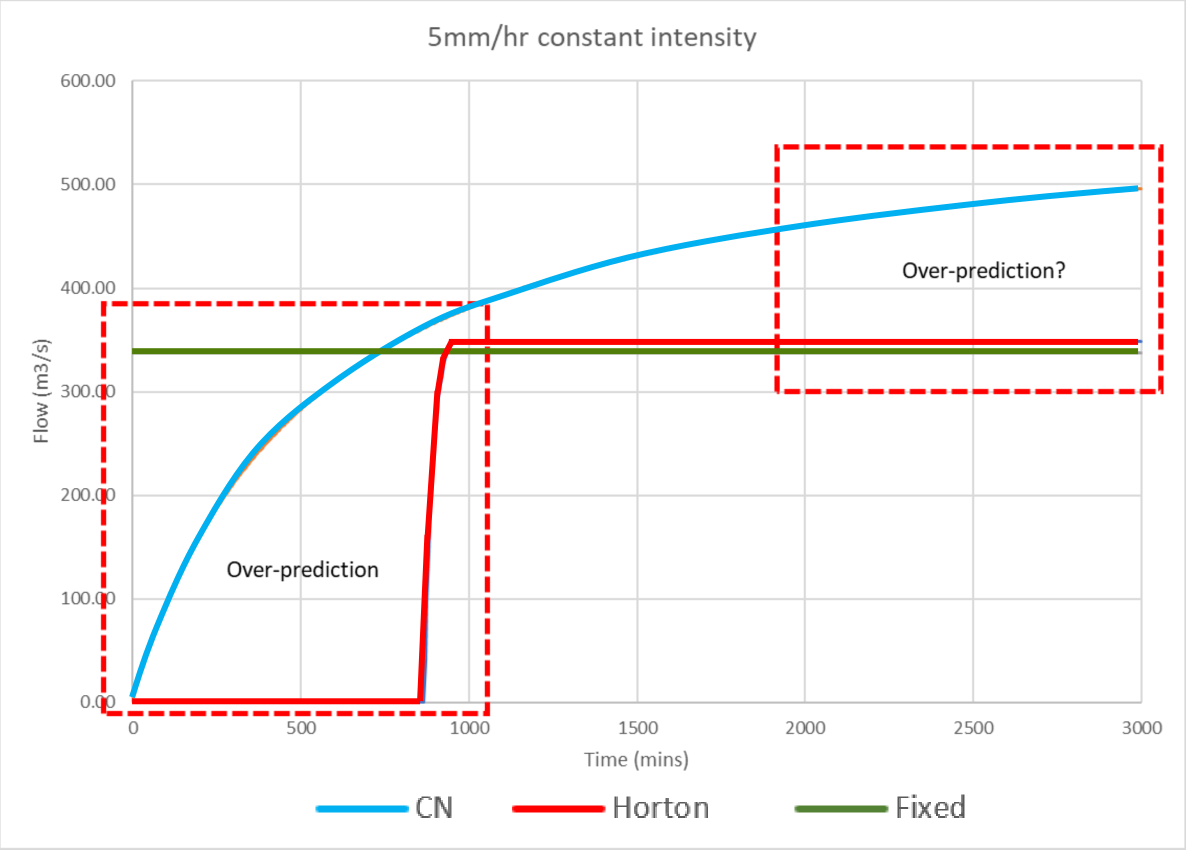
0.0 - 0.2 = poor calibration

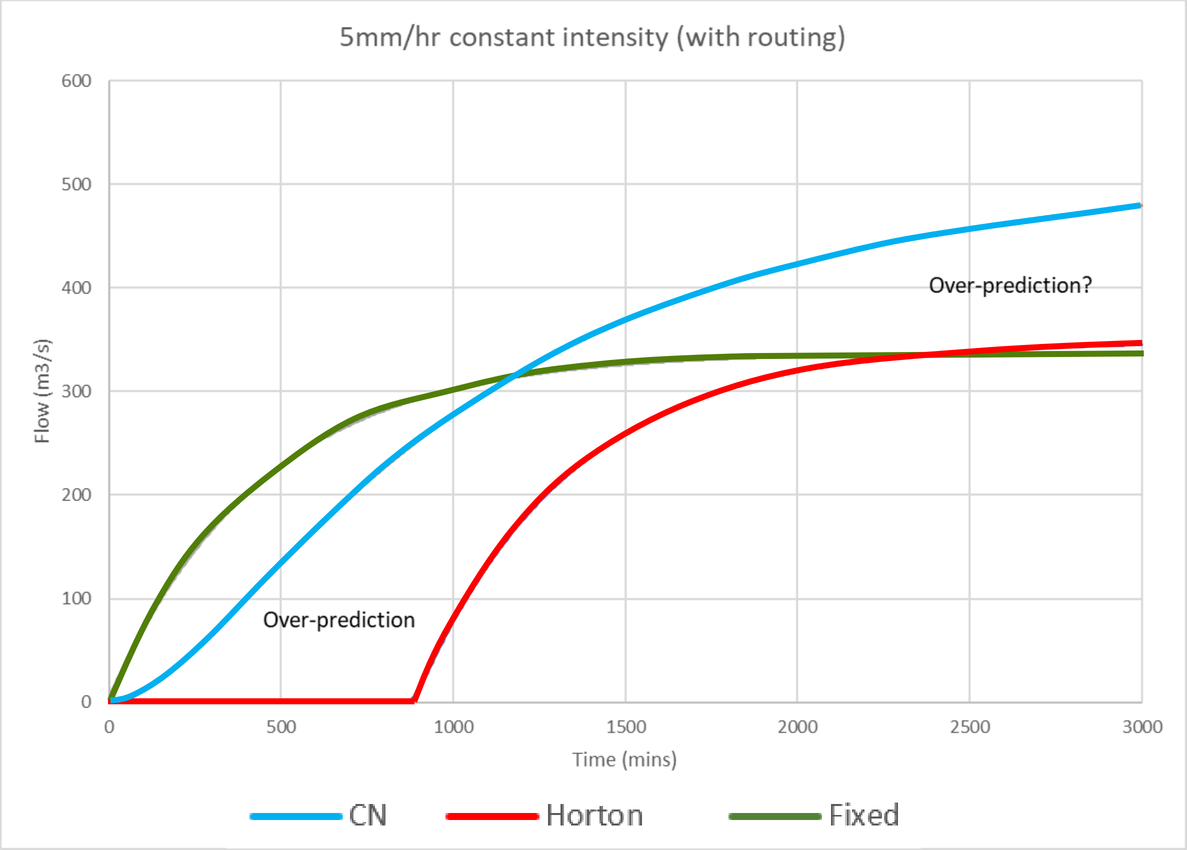
0.2 - 0.4 = moderate calibration

0.4 – 0.6 = good calibration

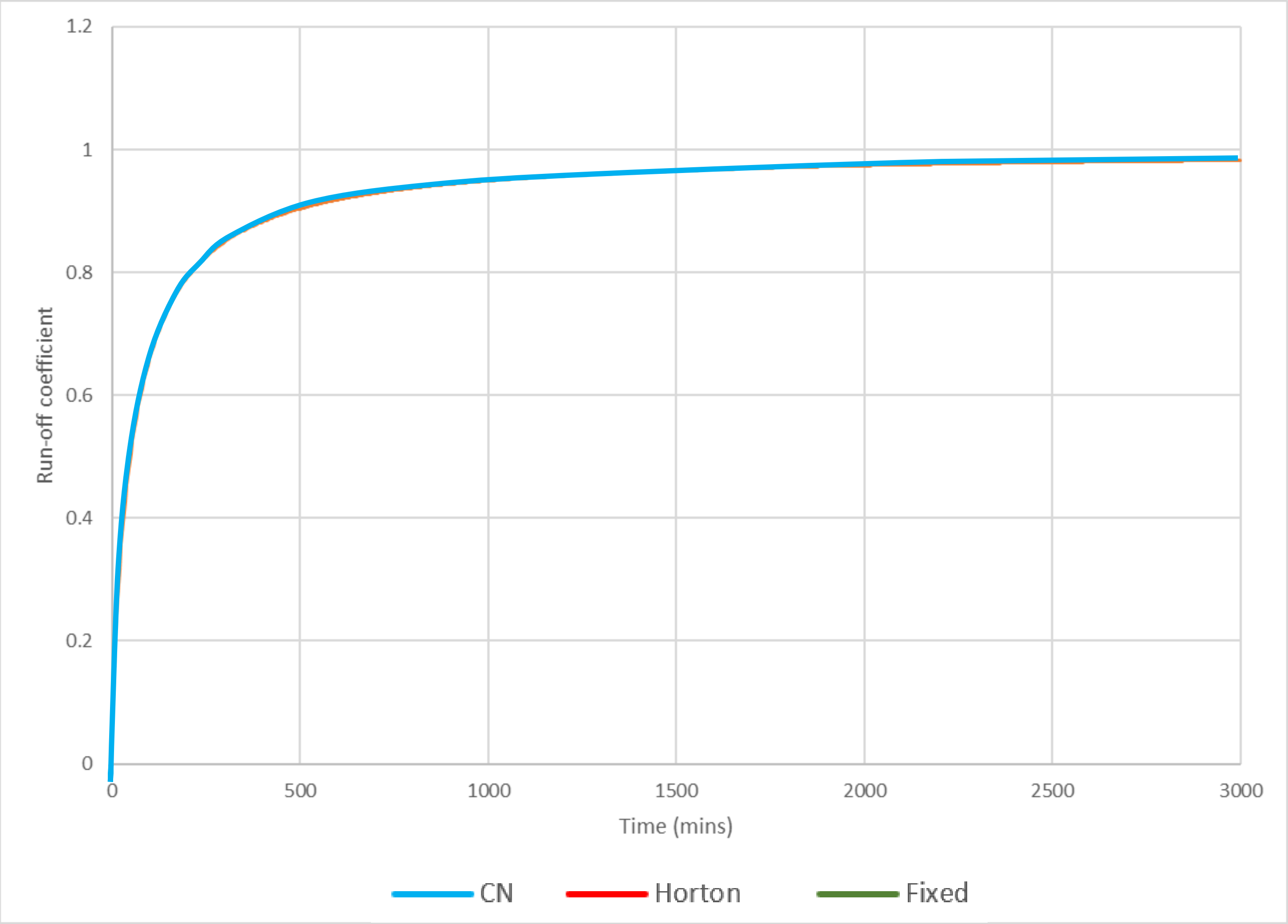
> 0.6 = excellent calibration

Differences (Calibrated to Target)





CN Model



- Run-off increases until unity is achieved.
- Is 0% loss appropriate?
- The NZBC applies 10% loss to impervious surfaces (e.g. roofs)
- 42% is the observed upper bound loss (C = 0.58) (Hutt river example, from observed results)

Conclusions

- Avoid R^2 to assess calibration 'fit'.
- Best to use a combination of RMSE and NSE.
- Using this approach allows for automation in calibration. This was successfully implemented in excel.
- CN performs better than a simple fixed run-off model for two of three catchments tested.
- Horton overall performs the best in terms of fit.
- Horton has the least variance across events for two of three catchments tested.

Thank you

Questions?

Etienne.Gil-Goldsbrough@wsp.com

wsp