Performance Assessment 6.5.12

2025-10-30

Validation of Model Performance

To confirm that the updates introduced in **OverseerFM version 6.5.12** did not affect the overall predictive performance of the N leaching, results were compared against measured leaching data from farmlet trials used in prior validation work. The comparison was conducted for both **version 6.5.11** (pre-update) and **version 6.5.12** (post-update), using identical datasets and statistical performance metrics.

Model Performance Comparison on the farmlet dataset

The two figures below illustrate the relationship between measured and modelled N leaching for versions 6.5.11 and 6.5.12, respectively. The dashed red line represents the 1:1 line, where perfect agreement between modelled and measured values would lie.

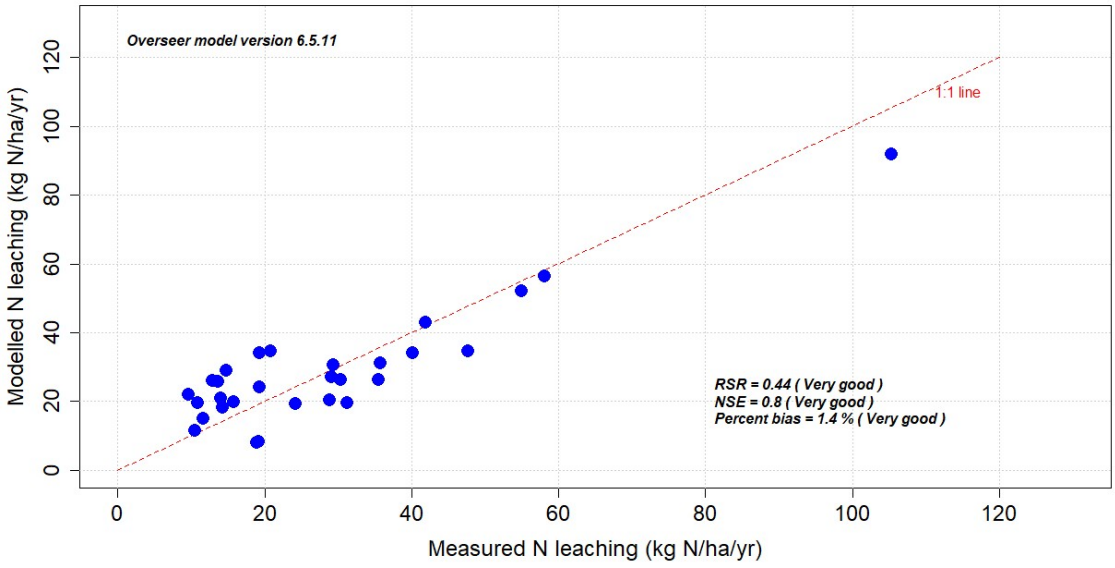

Figure 1. Comparison of measured and modelled N leaching for OverseerFM version 6.5.11.

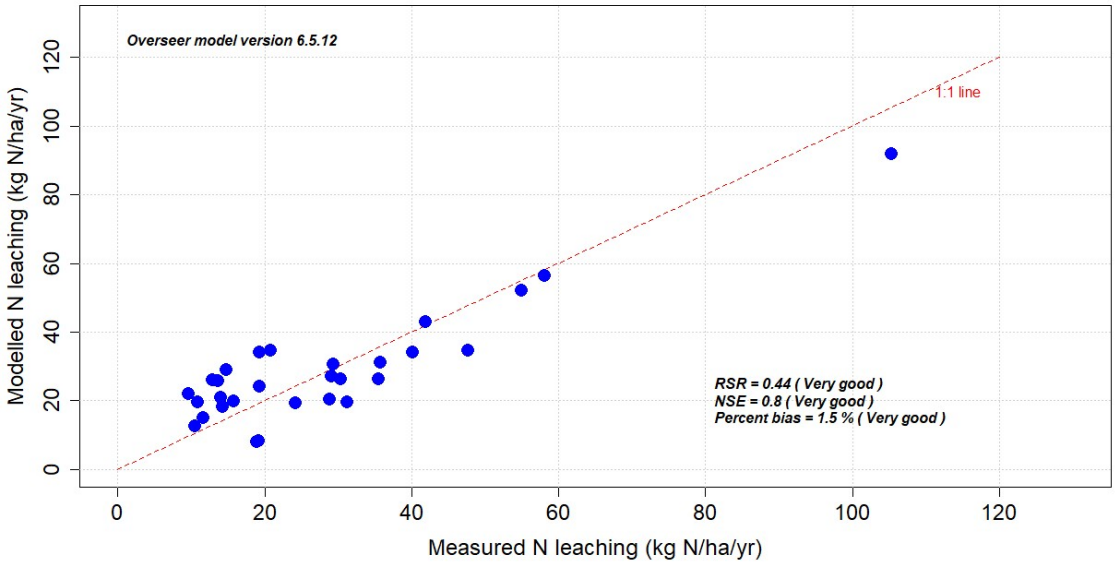

Figure 2. Comparison of measured and modelled N leaching for OverseerFM version 6.5.12.

Performance Statistics:

- **RSR** (Ratio of the Root Mean Square Error to the Standard Deviation of Observations) = 0.44 (**Very good**)
- **NSE** (Nash–Sutcliffe Efficiency) = 0.8 (**Very good**)
- **Percent bias** (Pbias) = 1.4% (v6.5.11) → **1.5% (v6.5.12)** (**Very good**)

Interpretation

The performance metrics demonstrate that the N leaching predictions remain **highly consistent** following the 6.5.12 update. Both model versions yield identical RSR and NSE values, confirming that predictive skill and accuracy have been fully preserved.

The **percent bias (Pbias)** shows a minimal change from **1.4% to 1.5%**, indicating only a **0.1 percentage point increase** in bias toward slightly higher predicted leaching values. This variation is negligible and falls well within the "Very good" performance classification according to the metric.

Origin of the variation

The update to OverseerFM version 6.5.12 is not a trivial revision.

Overseer

The table below lists the blocks from the farmlet dataset where Total N values have changed between versions 6.5.11 and 6.5.12, reflecting the influence of the corrections on block-level N accounting.

Comparison of Total N by Farmlet between OverseerFM versions 6.5.11 and 6.5.12

Farm name	Analysis name	Block name	Total N (6.5.11)	Total N (6.5.12)
Ovr-Butterworth 17-18	Butterworth 17-18	Milking Platform	10,502	10,543
Ovr-Butterworth 17-18	Butterworth 17-18	Effluent area	3,275	3,288
Ovr-Butterworth 17-18	Butterworth 17-18	Yr 2 Kale to grass	1,147	1,137
Ovr-Butterworth 17-18	Butterworth 17-18	Bull	786	788
Ovr-Butterworth 17-18	Butterworth 17-18	Grass to Yr 1 Swedes	4,095	4,096
Ovr-Butterworth 17-18	Butterworth 17-18	Leachate paddocks	367	369
Ovr-Butterworth 18-19	Butterworth 18-19	Milking Platform	10,578	10,618
Ovr-Butterworth 18-19	Butterworth 18-19	Effluent area	3,327	3,340
Ovr-Butterworth 18-19	Butterworth 18-19	Yr 2 Kale to grass	1,112	1,102
Ovr-Butterworth 18-19	Butterworth 18-19	Bull	795	797
Ovr-Butterworth 18-19	Butterworth 18-19	Leachate paddocks	435	436
Ovr-Massey No. 4 Dairy - DCG 60 ha (2009) - Peak cow _	Massey No. 4 Dairy - DCG 60 ha (2009) - Peak cow #	DCG trial	1,623	1,945
Ovr-Massey No. 4 Dairy - DCG 60 ha (2011) - Peak cow _	Massey No. 4 Dairy - DCG 60 ha (2011) - Peak cow #	DCG trial	845	953
Ovr-Telford Con 2012	Telford Con 2012	Main - Turnips	64	73
Ovr-Telford Con 2014	Telford Con 2014	Main - Turnips	24	22

Farm name	Analysis name	Block name	Total N (6.5.11)	Total N (6.5.12)
Ovr-J & C Paterson Trust - 2016-17	Year ending 2016	Pasture	4,679	4,684
Ovr-J & C Paterson Trust - 2016-17	Year ending 2016	Effluent	1,900	1,902
Ovr-J & C Paterson Trust - 2016-17	Year ending 2016	Rolling Pasture OOC	146	147
Ovr-J & C Paterson Trust - 2016-17	Year ending 2016	Turnips	756	755
Ovr-J & C Paterson Trust - 2016-17	Year ending 2016	Ceramic cup paddocks	890	891
Ovr-J & C Paterson Trust - 2017-18	Year ending 2016	Pasture	4,595	4,582
Ovr-J & C Paterson Trust - 2017-18	Year ending 2016	Effluent	1,926	1,919
Ovr-J & C Paterson Trust - 2017-18	Year ending 2016	Rolling Pasture OOC	142	141
Ovr-J & C Paterson Trust - 2017-18	Year ending 2016	Ceramic cup paddocks	670	669
Ovr-J & C Paterson Trust - 2018-19	Year ending 2016	Pasture	4,249	4,235
Ovr-J & C Paterson Trust - 2018-19	Year ending 2016	Effluent	1,922	1,917

Origin of the change for the farm name:

- **Ovr-Butterworth 17-18** and **Ovr-Butterworth 18-19**: Two dairy replacement mobs are now correctly modelled following the correction to mob handling in the Dairy Replacement enterprise. Previously, Overseer sometimes merged Calf and Replacement mobs after weaning, leading to double-counting of animals and nutrient demand, and occasional “Insufficient P” errors. The model now keeps these mobs separate throughout the post-weaning period, preserving correct timing and herd structure defined by the user.
- **Ovr-Massey No. 4 Dairy - DCG 60 ha (2009) - Peak cow** and **Ovr-Massey No. 4 Dairy - DCG 60 ha (2011) - Peak cow**: The change in N loss is due to the correction in how organic fertiliser is modelled. Previously, only the inorganic N fraction was considered, while the organic component was ignored. With version 6.5.12, the

organic N fraction is now correctly included, contributing additional N through decomposition and mineralisation.

- **Ovr-Telford Con 2012** and **Ovr-Telford Con 2014**: Very limited changed.
- **Ovr-J & C Paterson Trust**: Two dairy replacement mobs (calves+replacement) are now correctly modelled following the correction to mob handling in the Dairy Replacement enterprise. Previously, Overseer sometimes merged Calf and Replacement mobs after weaning, leading to double-counting of animals and nutrient demand, and occasional "Insufficient P" errors. The model now keeps these mobs separate throughout the post-weaning period, preserving correct timing and herd structure defined by the user.

Conclusions

- Model performance against measured data remains **very good** and statistically unchanged between versions 6.5.11 and 6.5.12.

- The small change in percent bias (0.1%) indicates **no systematic effect** of the implemented corrections on overall model calibration or accuracy.

- The validation confirms that version 6.5.12 maintains the same predictive capability for nitrogen leaching as the previous version, ensuring continuity and reliability for users and regulatory applications.

Comparison on the Lincoln datasets

The comparison using the Lincoln dataset shows no differences in results between versions 6.5.11 and 6.5.12. This confirms that the recent model updates have no impact on crop block performance, which remains identical to the previous version.

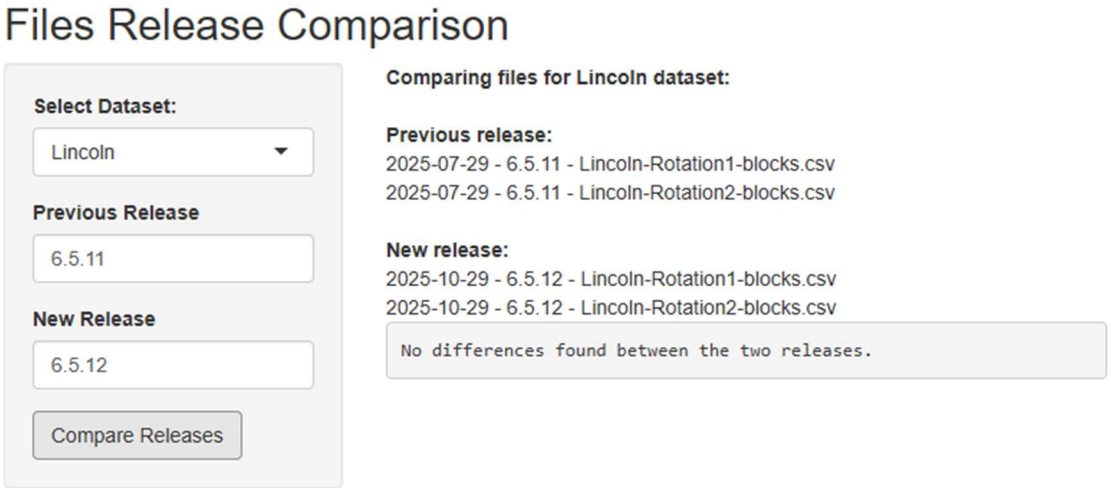

Figure 3. Comparison of Lincoln datasets for OverseerFM versions 6.5.11 and 6.5.12.

Conclusions

The release of OverseerFM version 6.5.12 represents a significant technical update, not a trivial maintenance change. While the overall model performance remains highly stable and validated—showing identical RSR = 0.44 and NSE = 0.8 with only a minor change in percent bias (+0.1 pp), the underlying corrections have meaningful effects at the farmlet and enterprise levels.

Validation against the Lincoln and farmlet datasets confirms that predictive performance for N leaching is statistically identical to version 6.5.11, preserving regulatory and scientific continuity.