Réseau laitier canadien
Canadian Dairy Network

MS test software:
Applications to Canadian data

Peter Sullivan,

Canadian Dairy Network

Réseau laitier canadien
Canadian Dairy Network

Introduction

- **Use of the software**
  - Flexibility and ease of use
  - Pre-processing requirements
  - Information / results generated
  - Post-processing opportunities
- **Impacts of ignoring genotypes in EBV calculation (i.e. GPS bias) on validation results**

Réseau laitier canadien
Canadian Dairy Network

Use of the software

Flexibility and ease of use

- **Easy to use**
- **Flexible to different breeds, traits, population size, with suitable guidance in the documentation**
- **Documentation is efficient, being very helpful but not longer than necessary**
- **Runs efficiently... ~ 1 hour for all breeds and traits**

Réseau laitier canadien
Canadian Dairy Network

Use of the software

pre-processing requirements

- Need to create input files for each breed
- Can have 1 file with all traits OR by groups of traits OR separate files for each trait
- Input data file(s) include for each animal: pedigree, ebv (E) and reliabilities (R):

an sire dam E_a E_s E_d R_a R_s R_d

- Input parameter file specifies the name of data file to use, which traits within the file should be tested, and the range of years to include in the test

Réseau laitier canadien
Canadian Dairy Network

Use of the software

Information / results generated

- **Output information is divided among multiple files**

  ***.out, *.dat, *.summary, <console>**

- **Provides yearly estimates of: genetic variance, average MS, PEV(MS)**
- **Tests for years with outlier variance**
  - Outlier if H0 rejected for a given year
- **Tests for non-zero trend in variance**
  - H0 rejected if BOTH statistical AND practical tests are FAILS
- **Warnings and good guidance if insufficient data**

Réseau laitier canadien
Canadian Dairy Network

Use of the software

Post-processing opportunities

- **Approximate PEV(MS) provided in the output, which is otherwise difficult to obtain without this software**
- **The subset of provided bulls actually used for the test (e.g. PEV is low enough) are identified**
- **User can modify input data to test impacts on validation test results**
  - Improve understanding of current EBV
  - Develop better EBV systems
- **Genomic Selection**
  - Test years prior to GPS (e.g. Bulls born 1998-2008)
  - Modify EBVs to see impact of GPS ★

Réseau laitier canadien
Canadian Dairy Network

Genomic Selection for positive Mend. Sampling

- Selection intensity for genomic young bulls has increased dramatically... 1 in 20 (5%) chosen for use in AI
- Positive within-family deviations for important traits at the time of GPS selection (i.e. before AI use) ... $MS_{GPS}>0$ for 70-100% of GPS bulls (CAN,NLD,DFS data)
- After AI, the MS deviations from EBV (ignoring genotypes) of these same bulls are regressed strongly towards 0 ... $MS_{EBV}>0$ ~> 50%
- MS_{GPS} is very high, while MS_{EBV} is much closer to zero
- Recent trends in EBV are much lower due to this downward bias of MS for GPS bulls (b_{EBV} ~50% * b_{GEBV})

Réseau laitier canadien
Canadian Dairy Network

Future EBV will not be biased by Genomic PreSelection (GPS)

- **Genetic evaluation centres will develop methodology to reduce/eliminate GPS bias in EBV of dairy cattle**
- **Corrections to remove GPS bias will likely increase %MS>0 ~> 100% and will increase b_{EBV} ~> b_{GEBV}**
- **How will this affect results from Interbull Validation test 4, which tests the following hypothesis?**

H0: For each year, MS ~ N(0,V_0)

- **Since GPS creates truncated distribution of MS, so H0 is dramatically unrealistic... Accepting H0 means:**
  1. GPS effects are small, or power of the test is weak.
  2. EBV are biased / do not reflect the true effects of GPS

Réseau laitier canadien
Canadian Dairy Network

What might future EBV look like if NOT biased by GPS?

- Consider changes that would increase b_{EBV} ~> $2*b_{EBV}$ and increase %MS>0 from 50% ~> 100%, while not changing the EBV of very best GPS bulls (e.g. 3 st. Deviation above the average of birth year cohorts)

$$Top_0 = a + b_{EBV} * yr + 3\sigma_e$$

$$b_{EBV} *= 1.25, 1.50, 1.75, 2.00$$

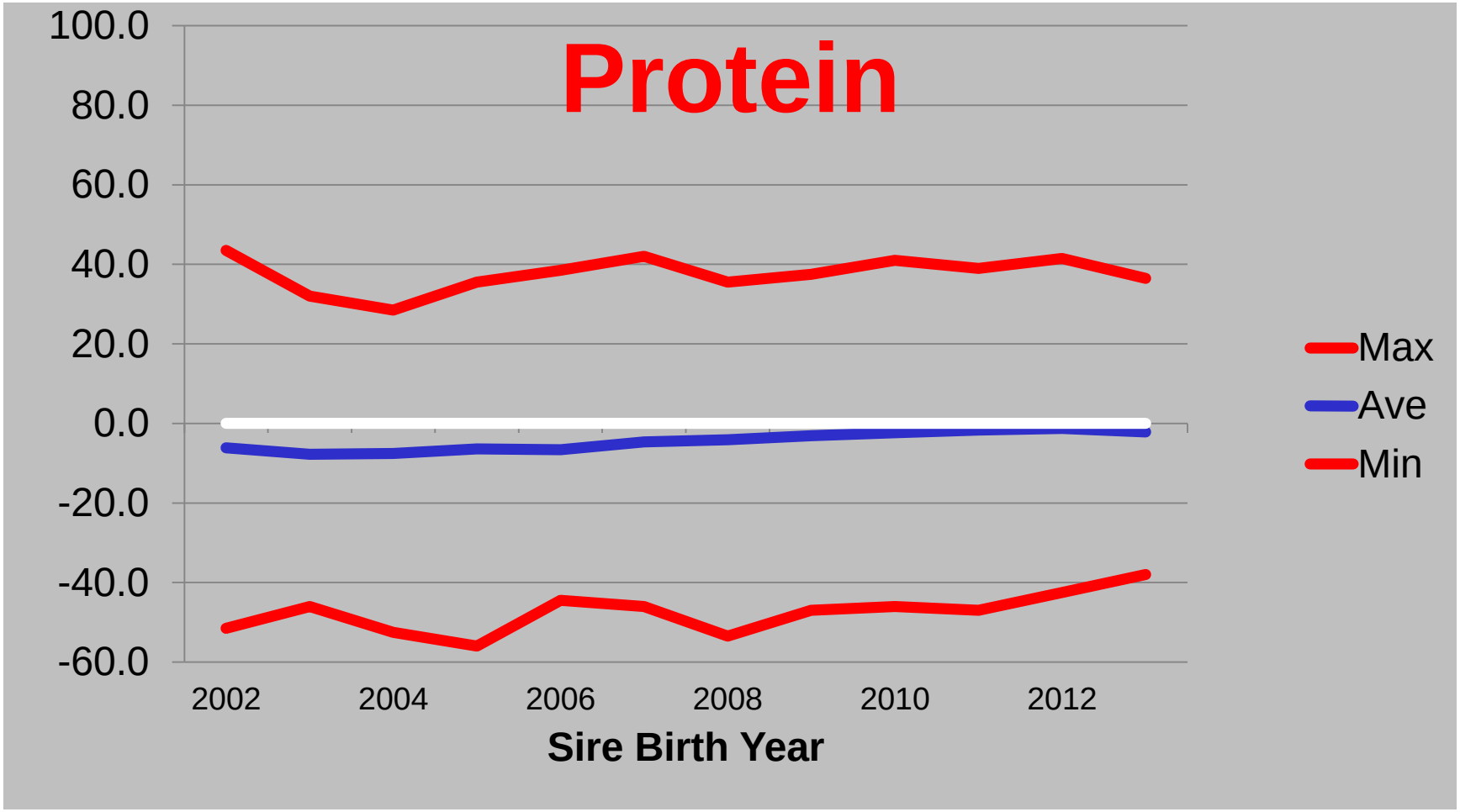
$$Top_1 = a + b_{EBV} * yr + 3\sigma_1$$

Solve for σ_1, such that $Top_1 = Top_0$

Scale EBV-E(EBV) by σ_1 / σ_e

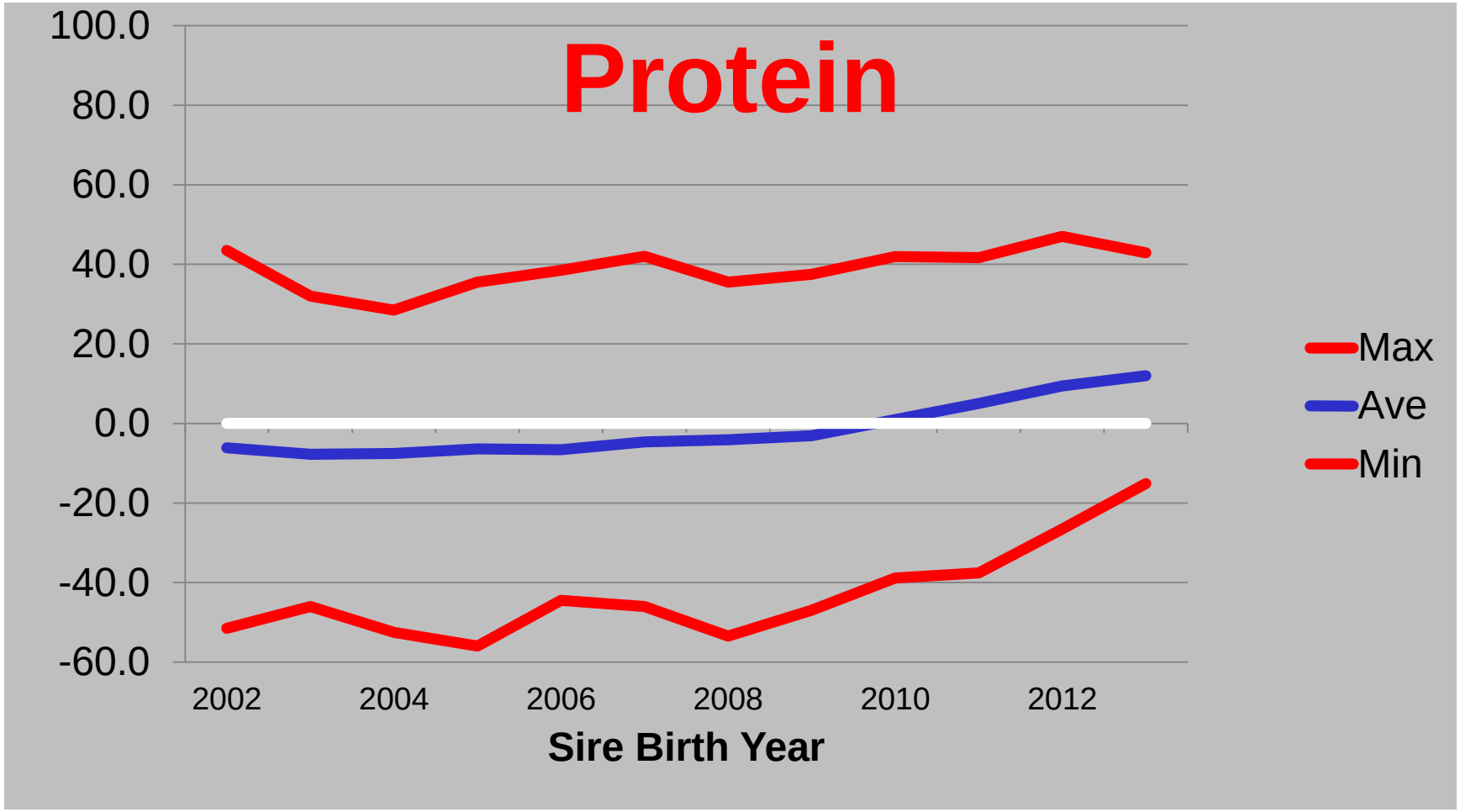
Réseau laitier canadien
Canadian Dairy Network

Range of MS_{EBV} with EBV ignoring genotypes (GPS bias)



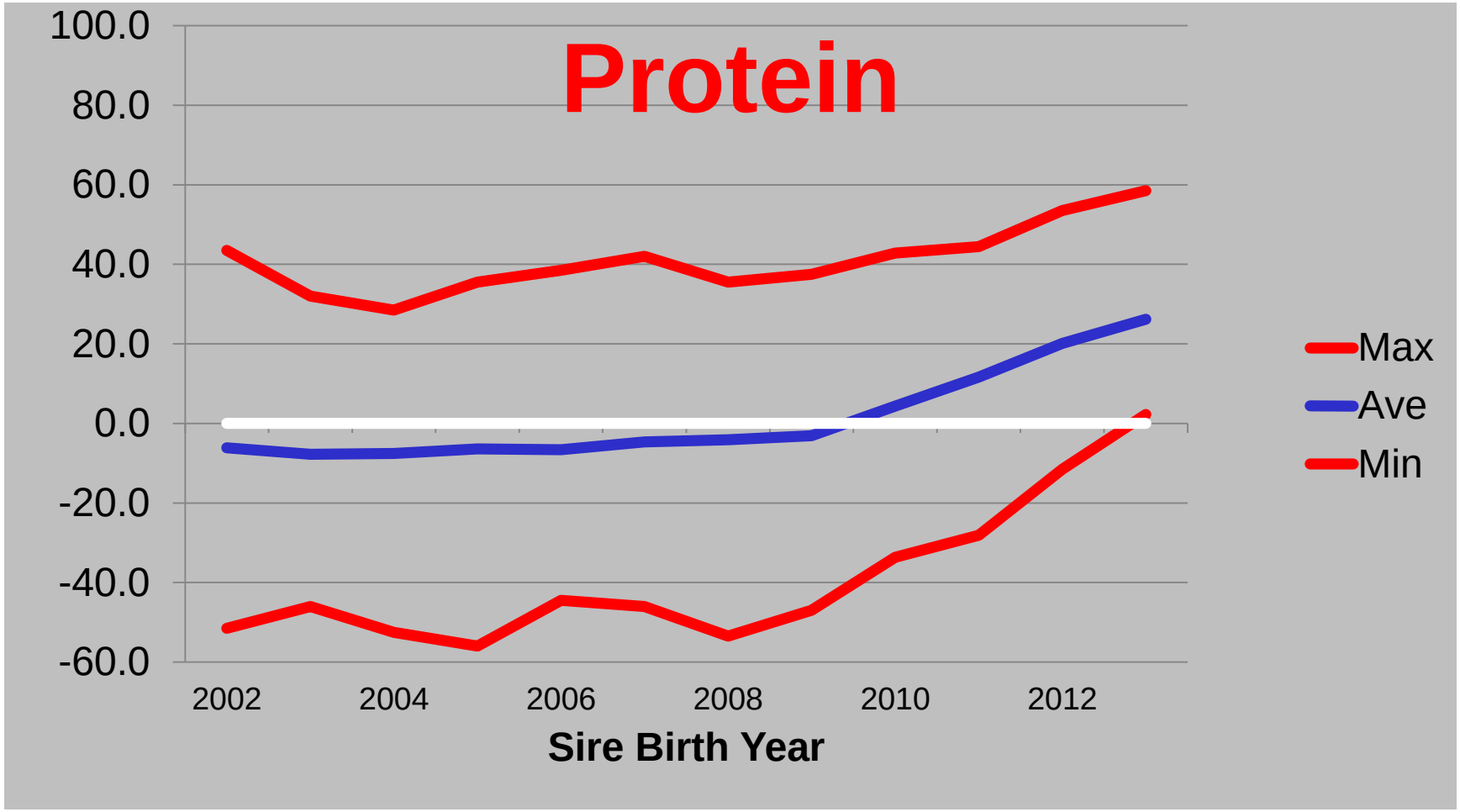
Réseau laitier canadien
Canadian Dairy Network

Range of MS_{EBV} with EBV adjusted for GPS: b_{EBV}*1.50



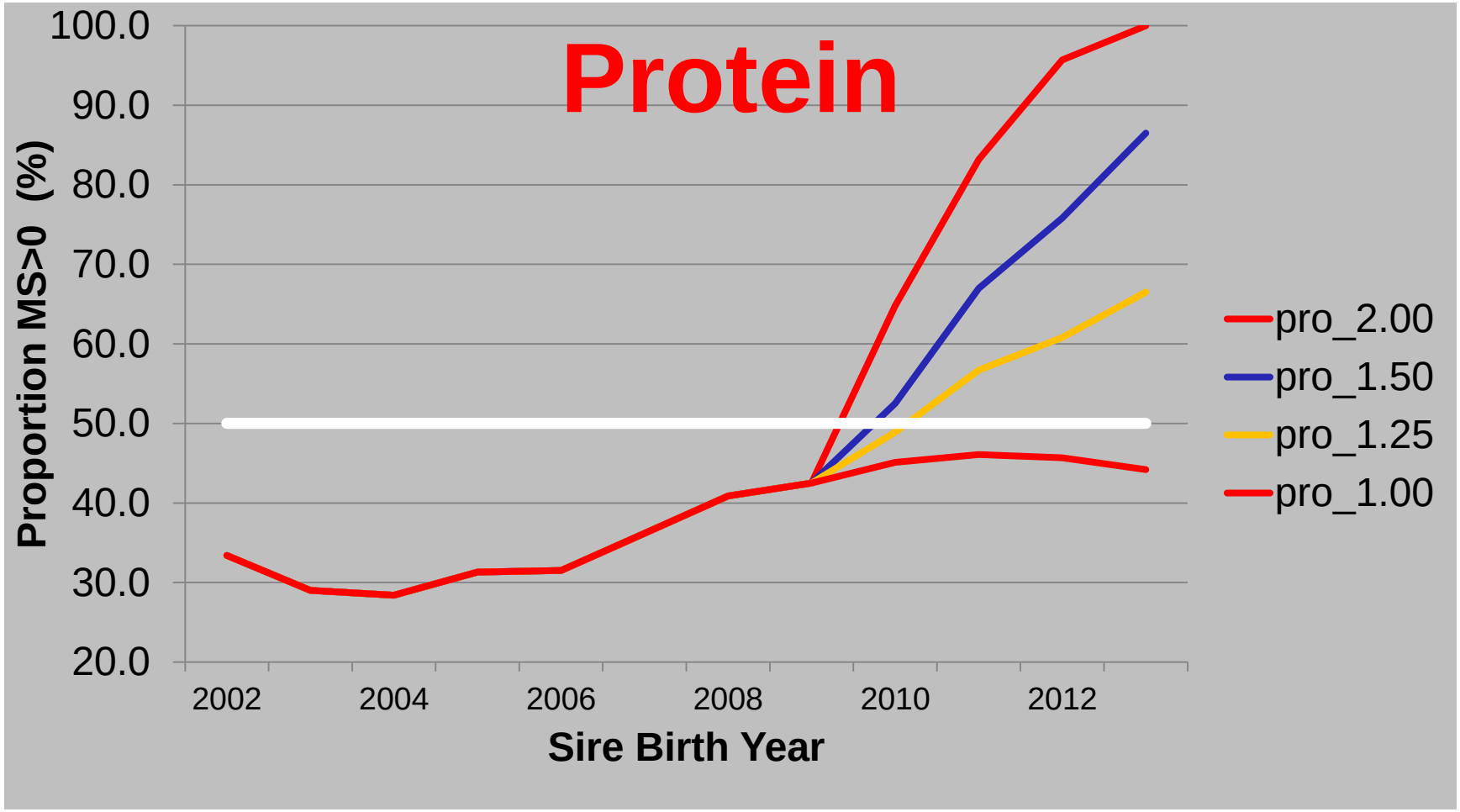
Réseau laitier canadien
Canadian Dairy Network

Range of MS_{EBV} with EBV adjusted for GPS: b_{EBV}*2.00



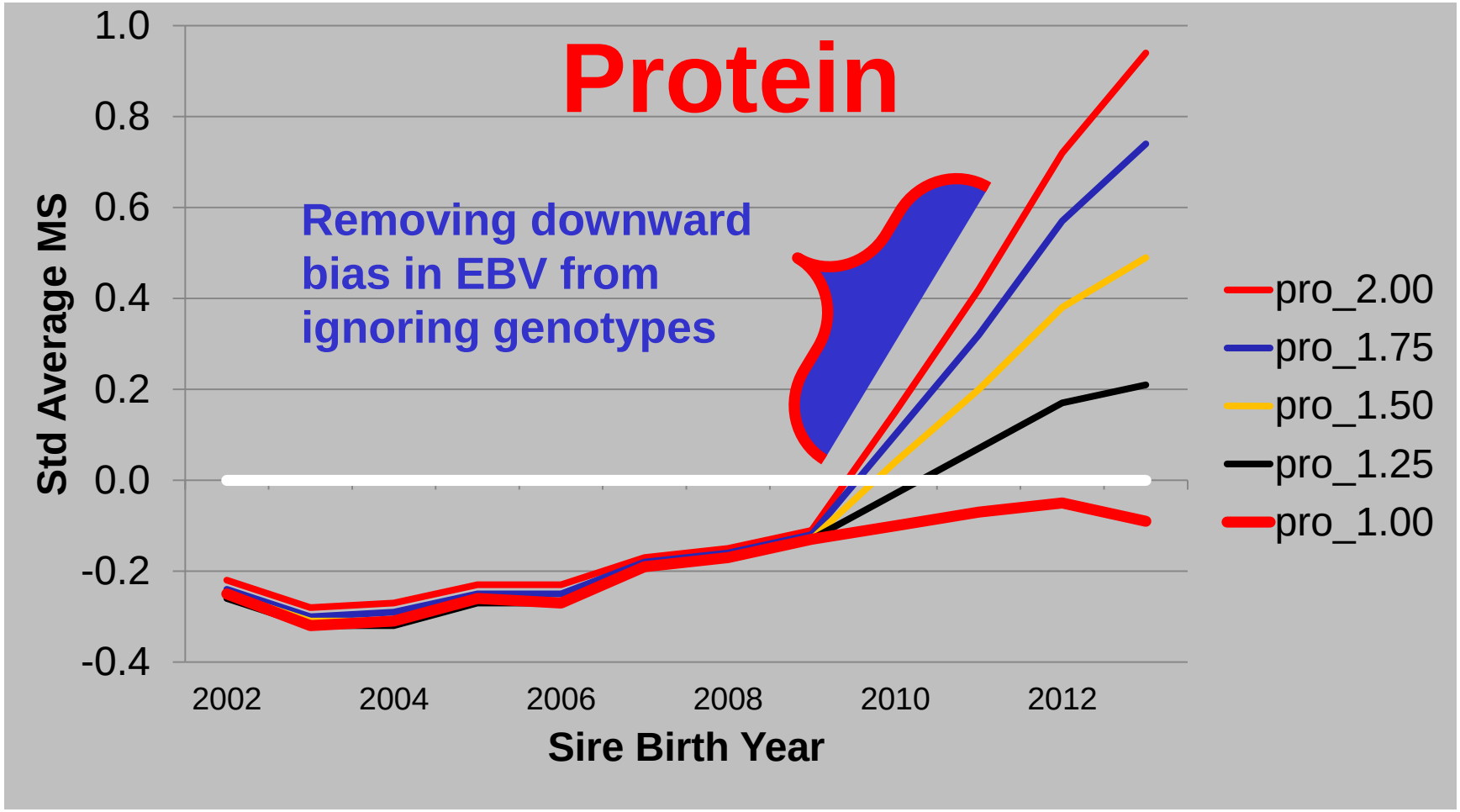
Réseau laitier canadien
Canadian Dairy Network

GPS bulls are the best full-sibs: move %MS>0 from 50% ~> 100%



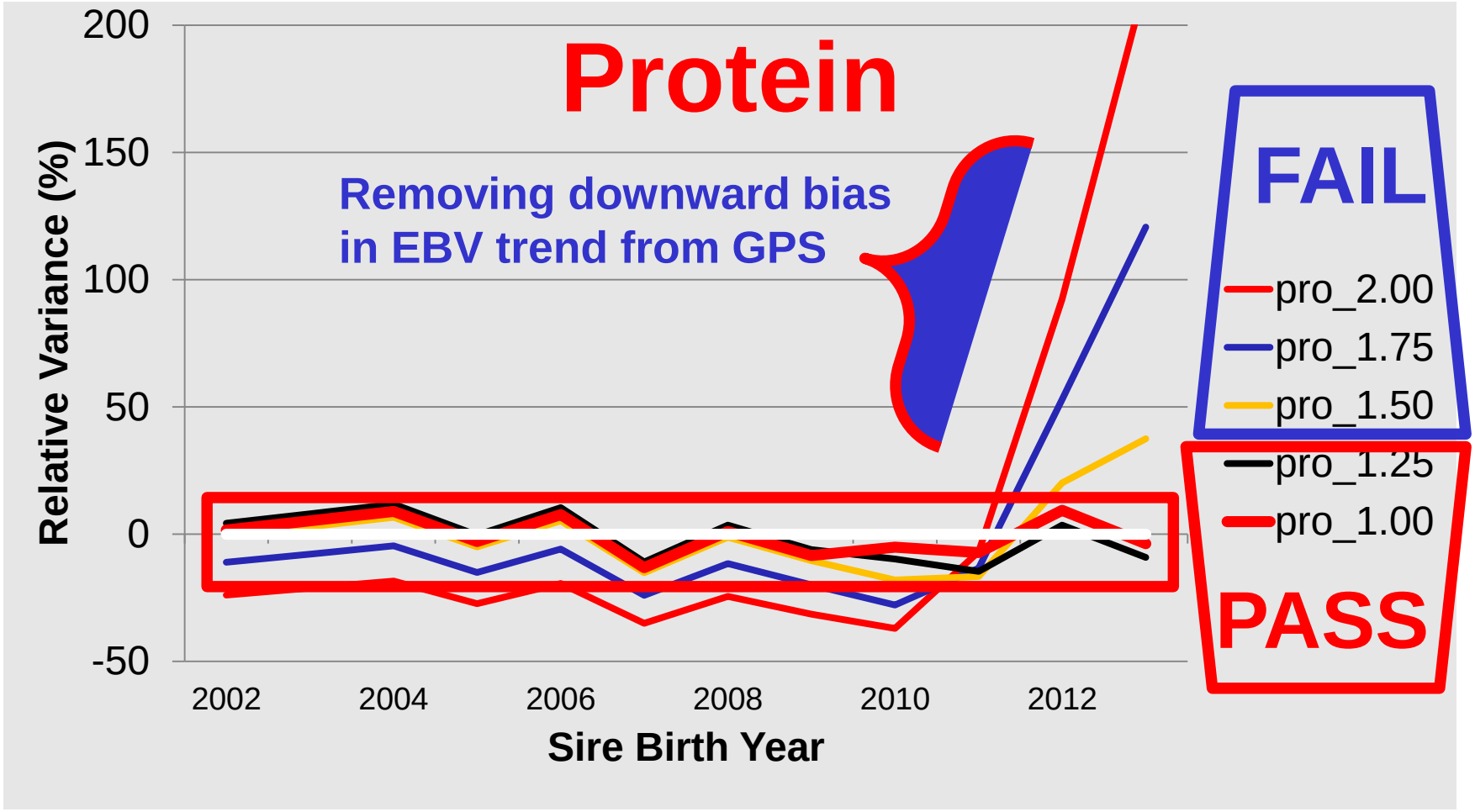
Réseau laitier canadien
Canadian Dairy Network

GPS affects on ITB validation Test4 (Trends in average MS)



Réseau laitier canadien
Canadian Dairy Network

GPS affects on ITB validation Test4 (Trends in variance)



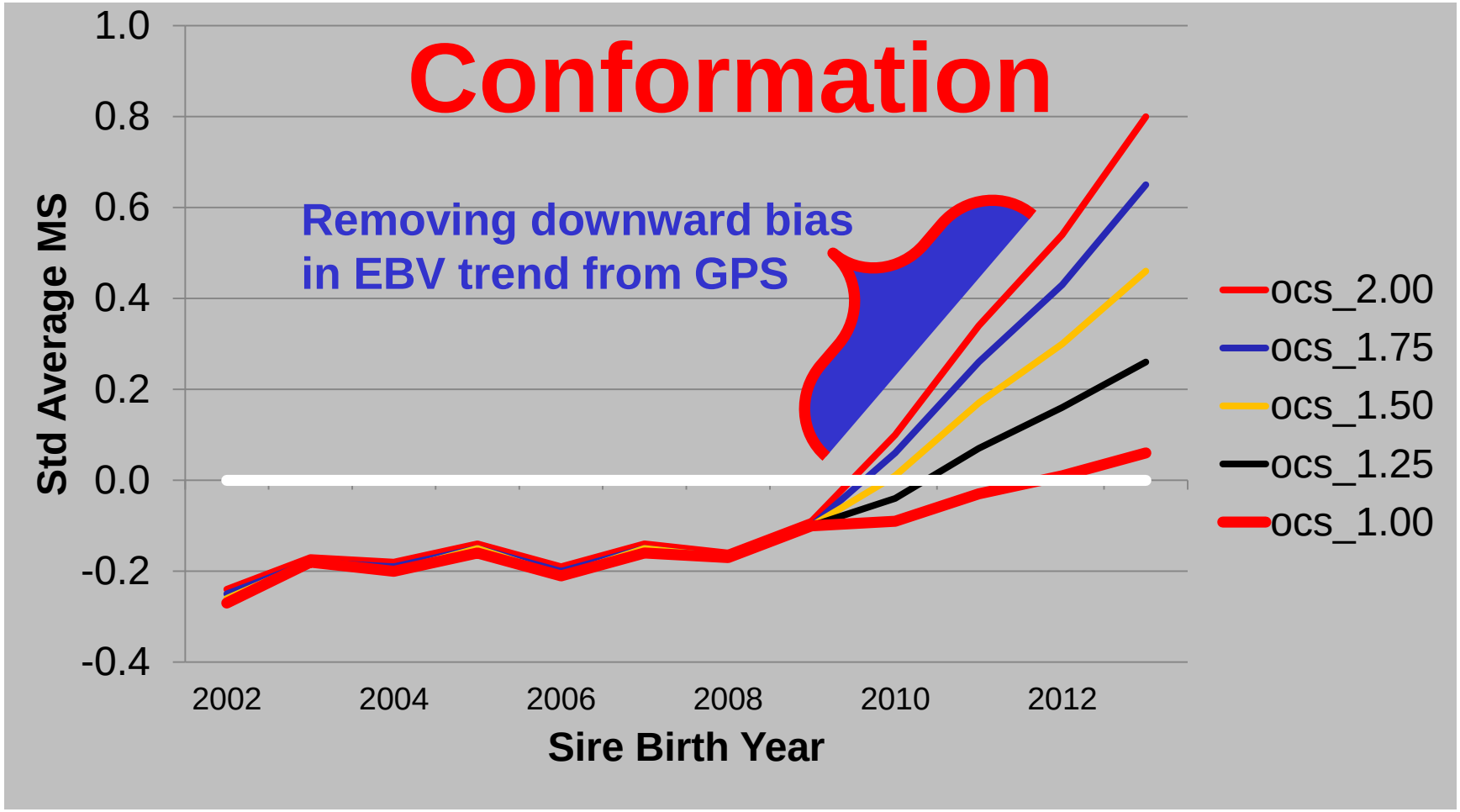
Réseau laitier canadien
Canadian Dairy Network

MS Variance Test
Protein

- **If EBV without genotypes include GPS bias, then the biased EBV will PASS the variance trend test (ITB test 4)**
- **If EBV are corrected to account for GPS, then the unbiased EBV will FAIL the variance trend test (ITB test 4)**
  - Intensity of GPS is **HIGH** for **Protein**

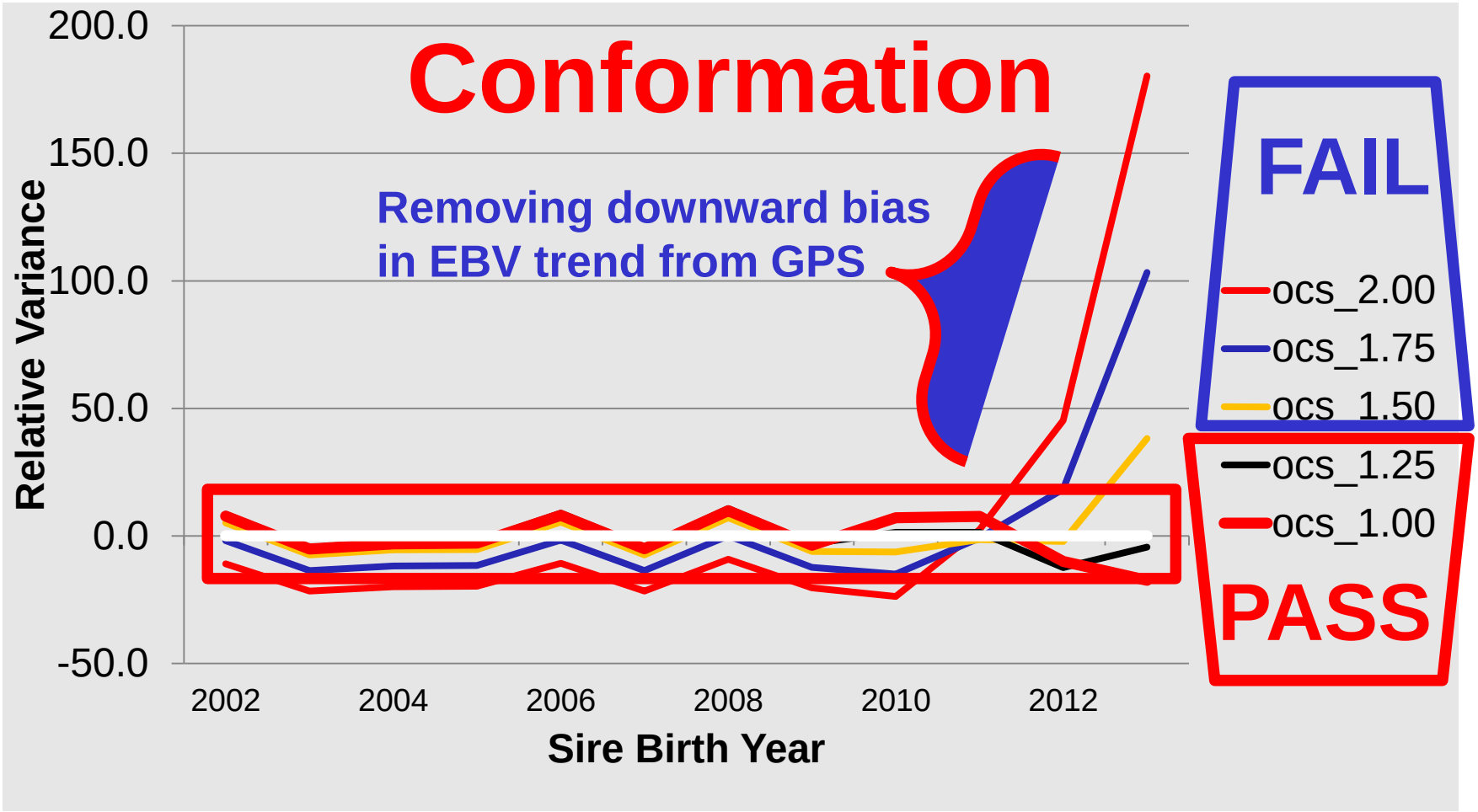
Réseau laitier canadien
Canadian Dairy Network

GPS affects on ITB validation Test4 (Trends in average MS)



Réseau laitier canadien
Canadian Dairy Network

GPS affects on ITB validation Test4 (Trends in variance)



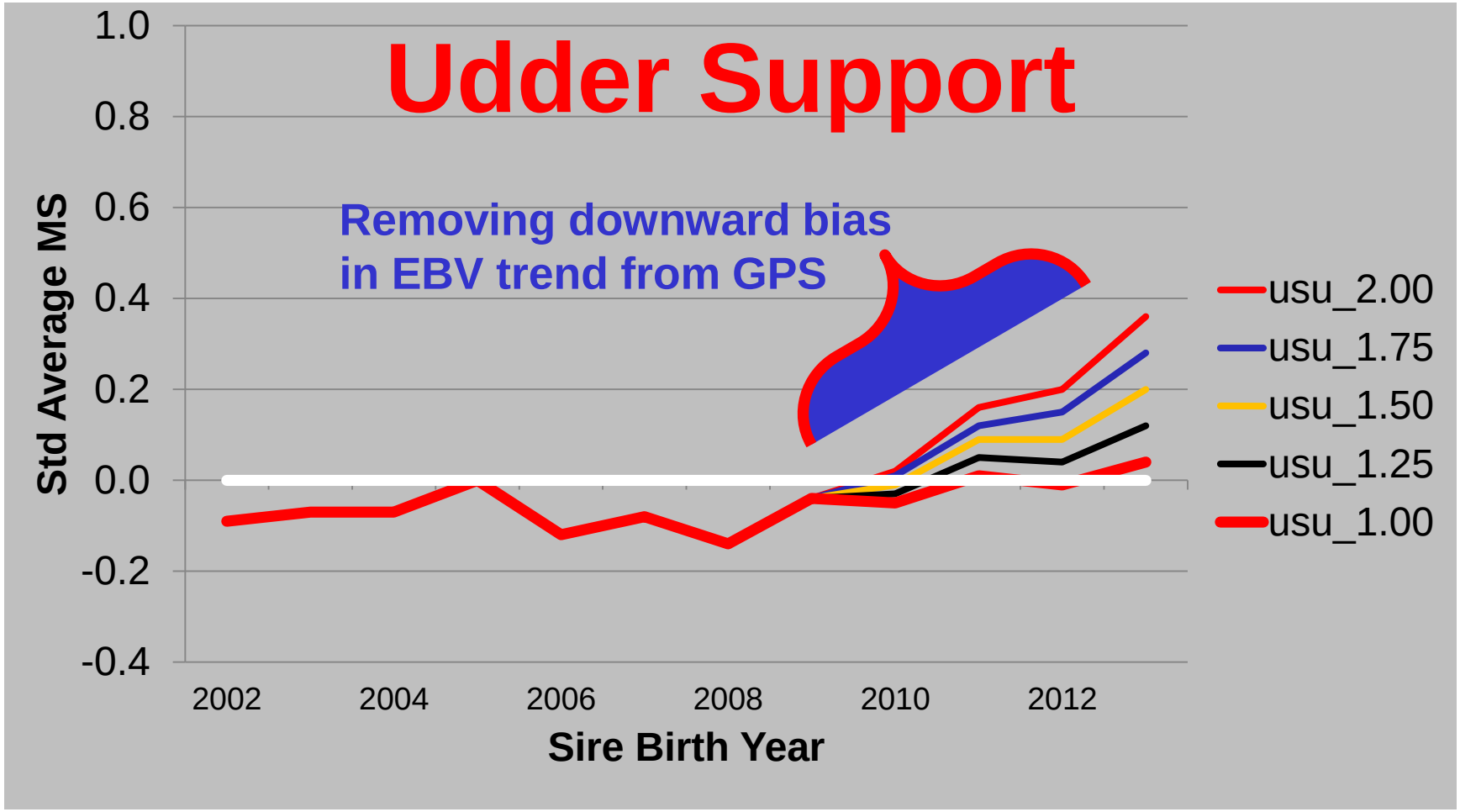
Réseau laitier canadien
Canadian Dairy Network

MS Variance Test
Conformation

- **If EBV without genomics include GPS bias, then the biased EBV will PASS the variance trend test (ITB test 4)**
- **If EBV are corrected to account for GPS, then the unbiased EBV will FAIL the variance trend test (ITB test 4)**
  - Intensity of GPS is **HIGH** for **Conformation**

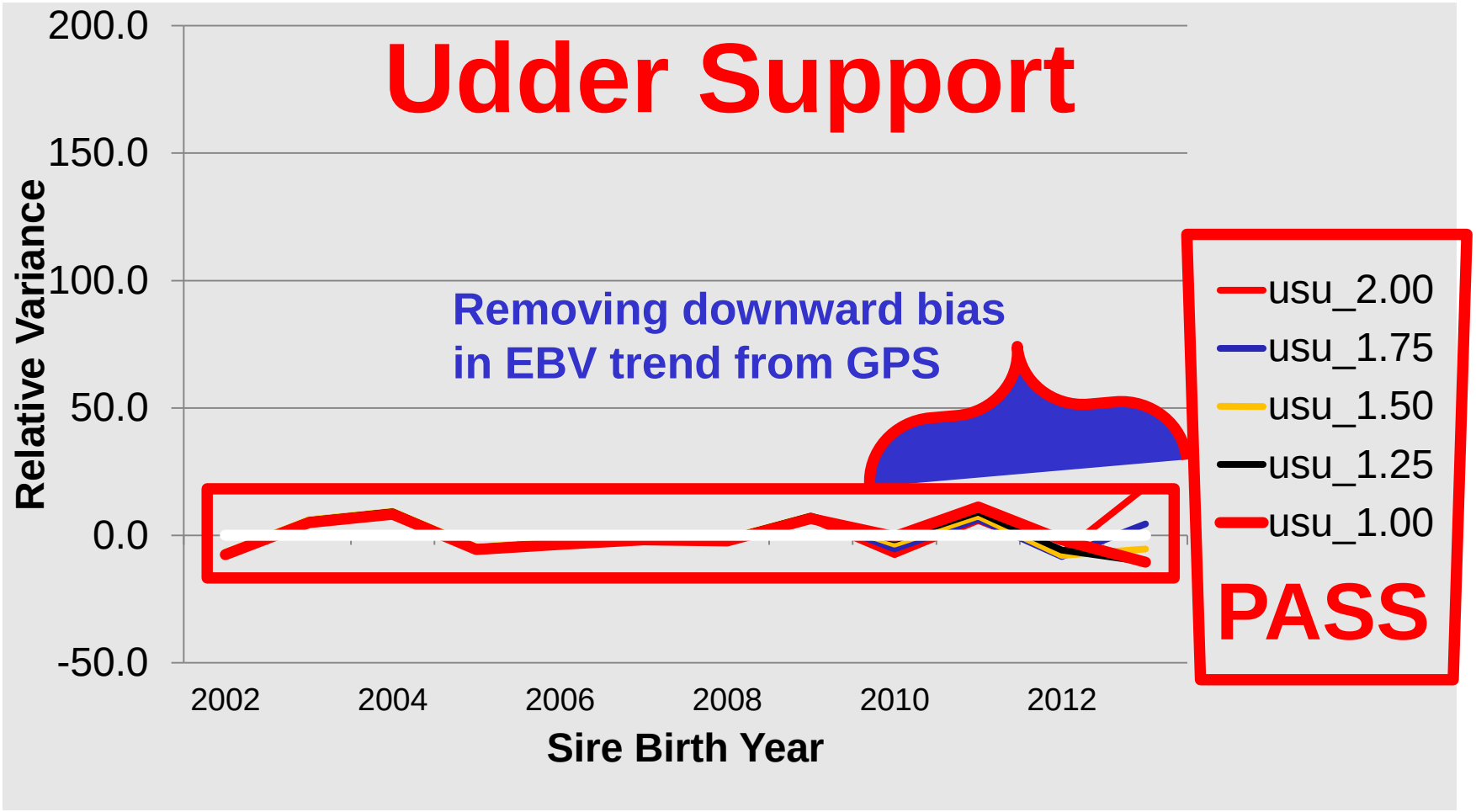
Réseau laitier canadien
Canadian Dairy Network

GPS affects on ITB validation Test4 (Trends in average MS)



Réseau laitier canadien
Canadian Dairy Network

GPS affects on ITB validation Test4 (Trends in variance)



Réseau laitier canadien
Canadian Dairy Network

MS Variance Test
Udder Support

- **If EBV without genomics include GPS bias, then the biased EBV will PASS the variance trend test (ITB test 4)**
- **If EBV are corrected to account for GPS, then the unbiased EBV will PASS the variance trend test (ITB test 4)**
  - Intensity of GPS is **LOW** for **Udder Support**

Réseau laitier canadien
Canadian Dairy Network

Summary

- **Software is flexible, efficient, easy to use**
- **Results of the validation test are NOT RELEVANT to decide inclusion/exclusion of national EBV as input to MACE**
  - GPS-Biased EBV pass the test
  - Unbiased EBV fail the test
- **Software is a very useful tool for "GE systems mining"**
  - Good quality information/data is generated, to allow detailed tracing and post-analysis of within and between family estimates of genetic differences... To improve understanding of GE systems behaviours and for developing improved systems