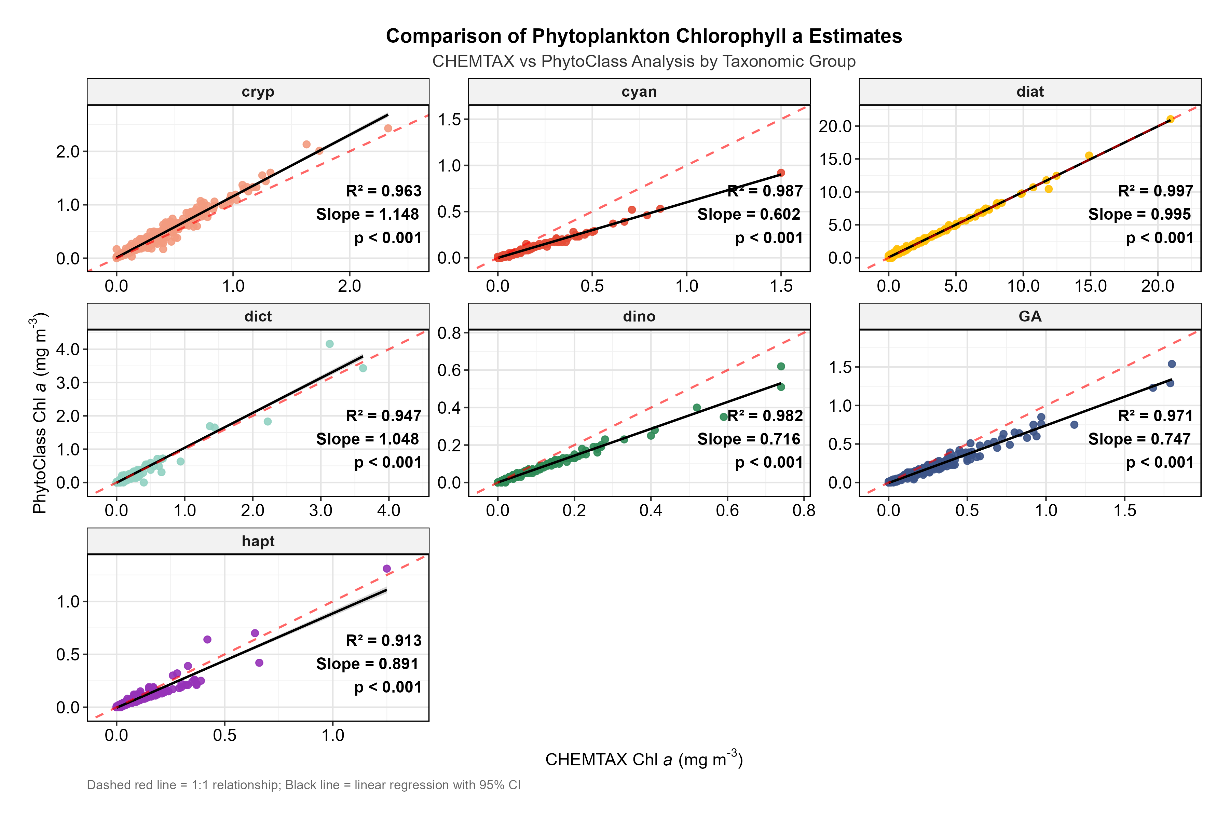
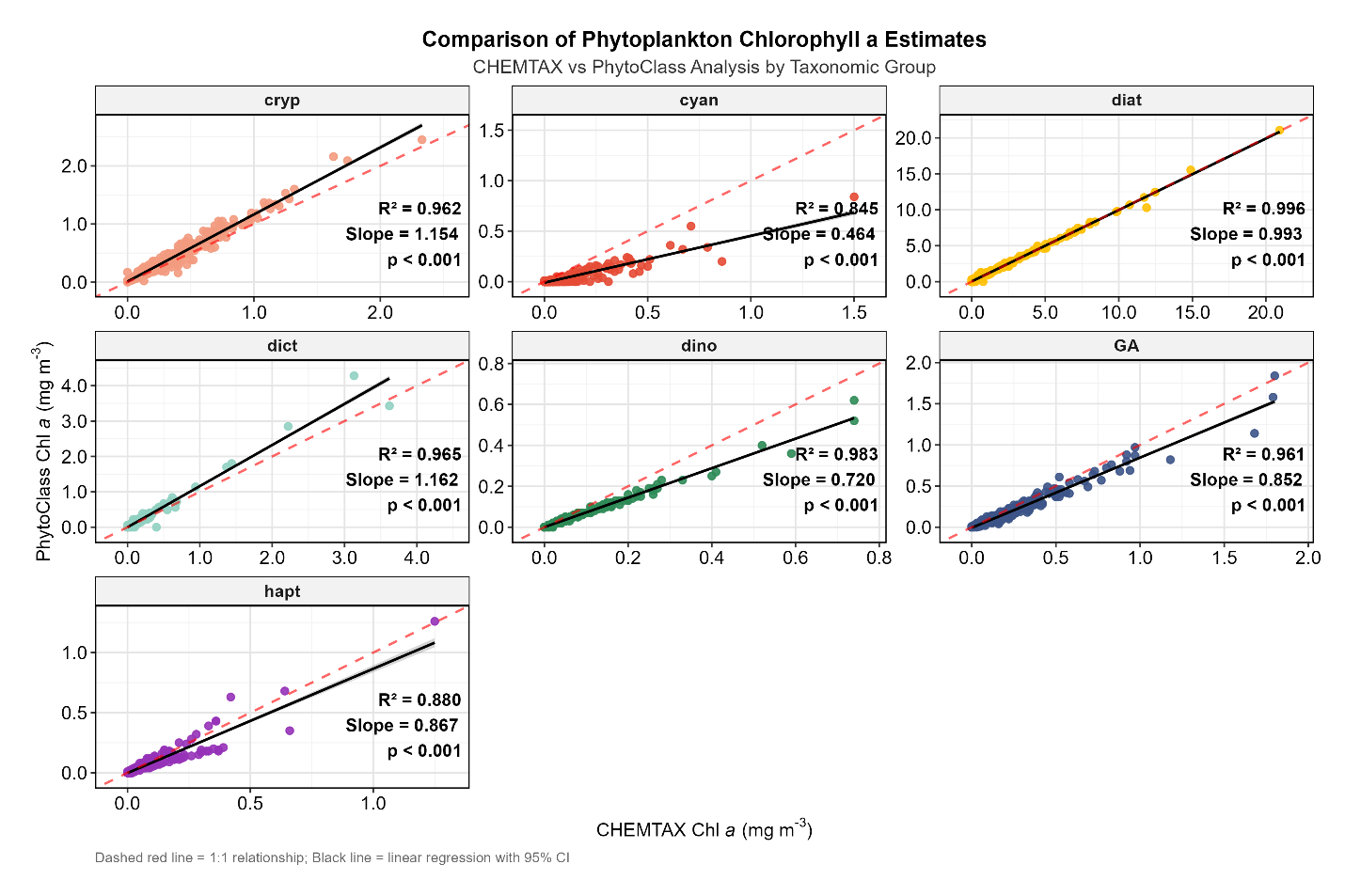
Phytoclass Trial log

1. Min\_chem\_100\_noc1c2 (ran on July 22rd, 2025)
   * I meant to do 100% difference from my CHEMTAX manuscript values, but actually did CHEMTAX ratio/0.5 (min) and CR\*2 (max).
     + Think did to ensure min ratios did not = 0.
   * Deleted chlc1c2 as was causing confusion and large misclassifications.
   * Created my best comparison yet with high R2 for each group   
     (> 0.91) and slight phytoclass underestimates of:
     + Cyanobacteria
     + Dinoflagellates
     + Green Algae
   * Pro
     + Can nearly replicate CHEMTAX with easier method to run
   * Con
     + Defeats the purpose of Phytoclass - meant to not rely on specific input ratios with convergence on optimized ratios from wide min\_max range.
     + Optimized ratios meant to provide better results.
   * Trials
     + Iteratively introduce literature min and max for group to see how it changes.
   * Todo
     + Export converged ratios and analysis stats for comparison to CHEMTAX output ratios.
     + Try to set min\_max based on my field data rather than lit?
     + Should I try to get Phytoclass to exactly replicate CHEMTAX?
     + Run on clusters and see how changes.



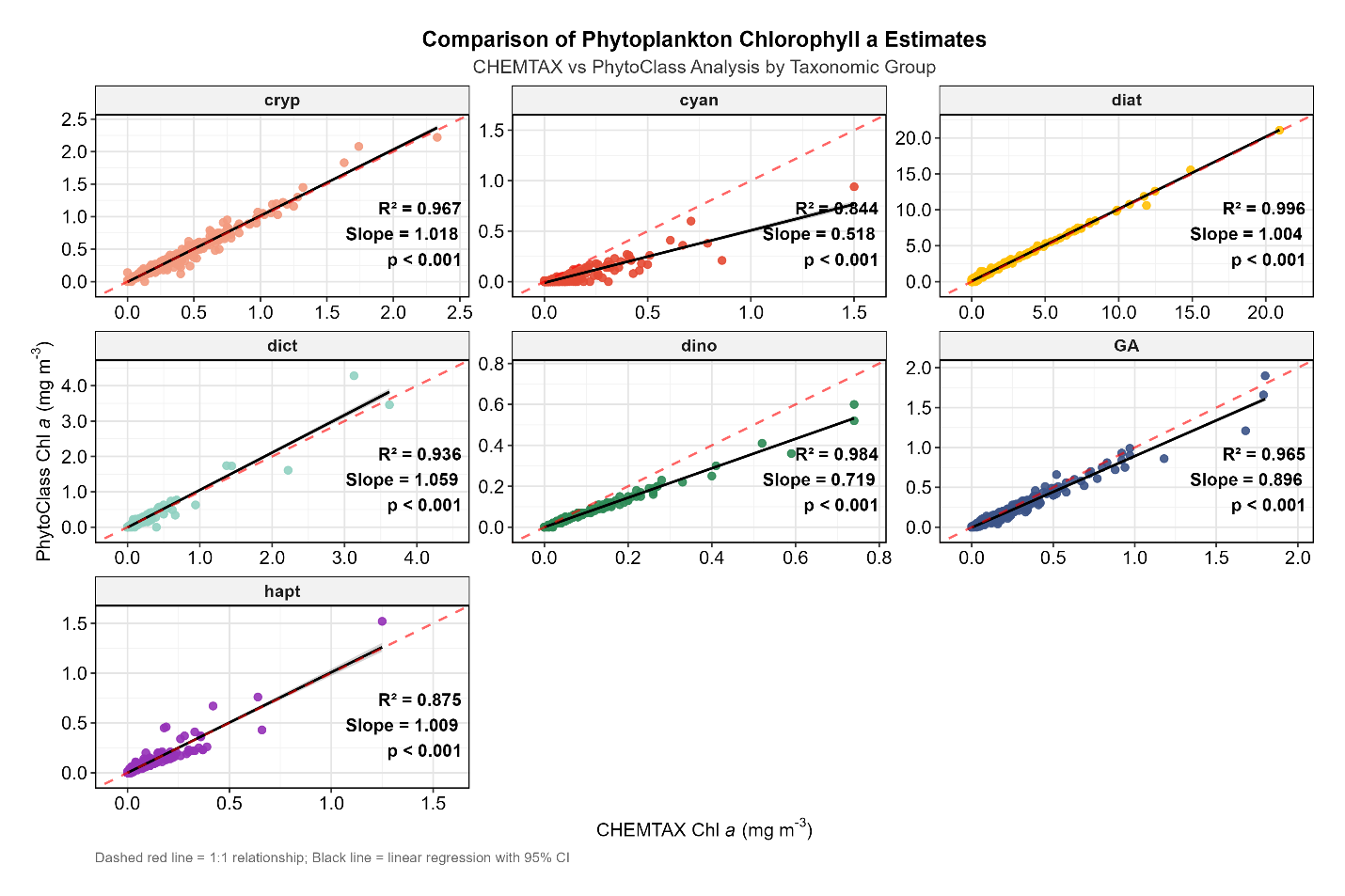
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| dataset | RMSE | condition\_number | type | sample\_size |
| autumn | 0.025509 | 1010.261 | Cluster | 102 |
| summer | 0.025884 | 666.2668 | Cluster | 120 |
| spring | 0.029665 | 727.2185 | Cluster | 113 |
| winter | 0.040668 | 1042.183 | Cluster | 76 |

1. Min\_chem\_100\_noc1c2\_pras\_lit (ran on July 23rd, 2025)
   * Took the prasinophytes-3 literature min and max values.
   * Changed
     + Cyan – much poorer correlation and shallower slope
     + GA – similar correlation shallower slope
     + More divergence in haptos
       - Wonder what this is – two different types?
   * Are these differences just because CHEMTAX has wider ranges? How do I assess which is better?



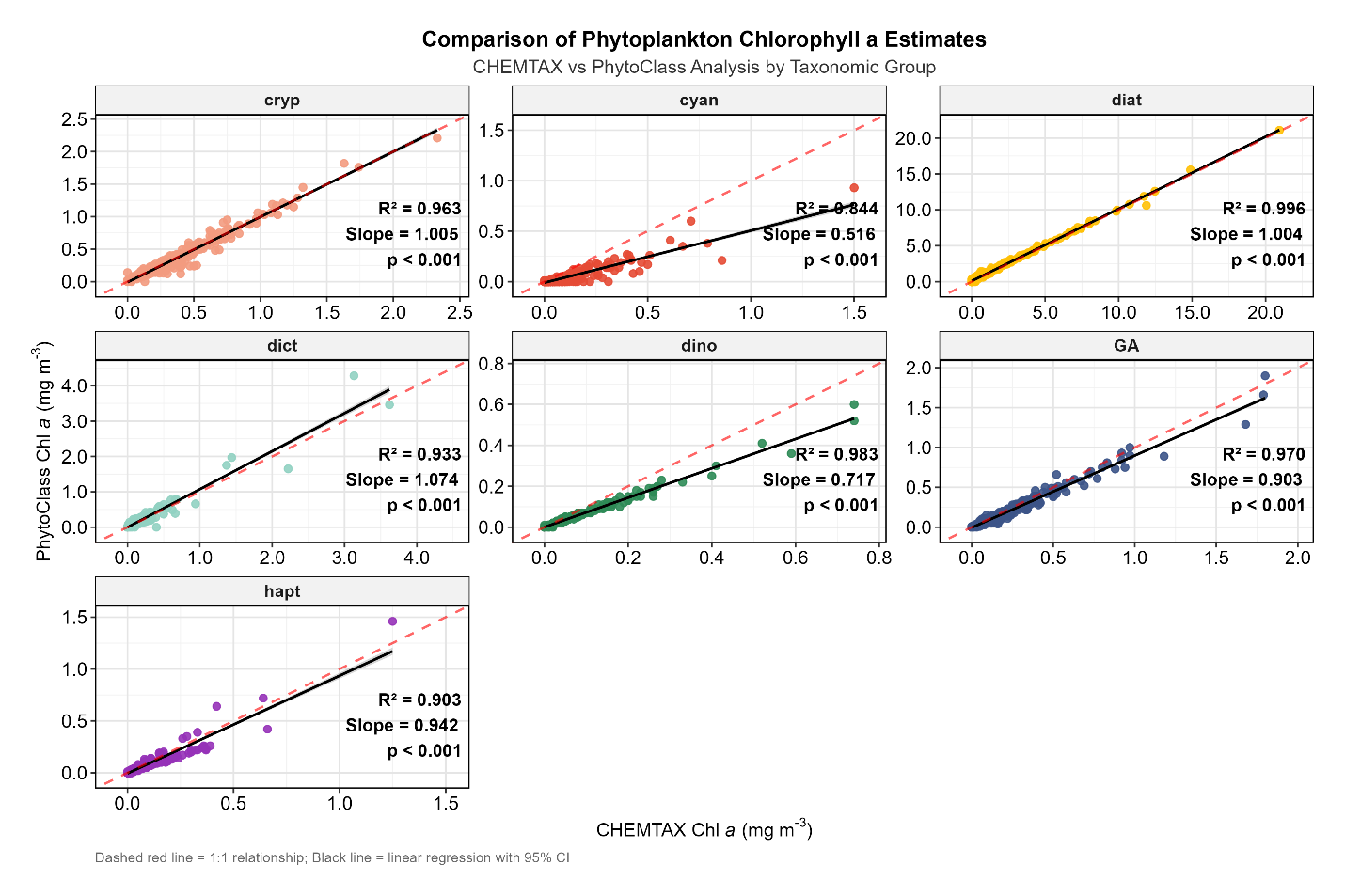
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| dataset | RMSE | condition\_number | type | sample\_size |
| autumn | 0.024956538 | 890.7726 | Cluster | 102 |
| summer | 0.025819509 | 700.1539 | Cluster | 120 |
| spring | 0.030029735 | 1773.824 | Cluster | 113 |
| winter | 0.040801521 | 1128.626 | Cluster | 76 |

1. Min\_pras\_lit\_chem\_field (ran on July 24rd, 2025)
   * Using field data to constrain cryptophyte limits – leaving GA as lit ratios.
     + Our ratios for cryptophytes are considerably lower than those in the literature. For my CHEMTAX paper, I developed my own input ratios and have the ratio limits be 500%.
     + Here, I am going to try to take the min and max from the 10-year time-series.
       - Don’t like how I am using min\_max from entire time-series including periods when cryptophytes are not present – this could skew my min values low.
         * For Alloxanthin – filtered so that the concentration > 0 prior to calculating ratios. Modest change in min from 0.001 to 0.002.
     + Should also try using the literature values.
   * Field values created huge overestimations – nearly completely replaced diatoms!
     + Potentially a result of allowing such low ratios?
     + Allo converged at:
       - 0.002 for W,Sp,Su – confirms hypothesis
       - 0.16 for prior runs where mine 50% lower and max 100% higher than CHEMTAX (I think 0.12).
       - Try literature values – although lower range > our mean
2. Min\_pras\_lit\_cryp\_lit (ran on July 24rd, 2025)
   * Literature min\_max for cryptophytes.
   * Do seem to work better!
   * Convergence at:
     + 0.17 (W)
     + 0.23 (Sp)
     + 0.18 (Su)
     + 0.19 (Au)
       - Some greater variability than the initial tests, which were static – this is good.
   * Outputs pretty comparable to CHEMTAX – best to use literature ratios.
   * Stats pretty comparable – slightly lower conditions numbers.



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| dataset | RMSE | condition\_number | type | sample\_size |
| autumn | 0.024448 | 881.507 | Cluster | 102 |
| summer | 0.02552 | 653.7504 | Cluster | 120 |
| spring | 0.028094 | 440.0372 | Cluster | 113 |
| winter | 0.040292 | 844.0409 | Cluster | 76 |

1. Min\_pras\_lit\_cryp\_lit\_d\_lit (ran on July 24rd, 2025)
   * Trying to use diatom literature values
   * For now, should us the min and max for D1+D2 as both present.
   * Fuco convergence from 0.39-0.49 (low)
   * Not very different than prior runs, but should also compare to CHEMTAX.



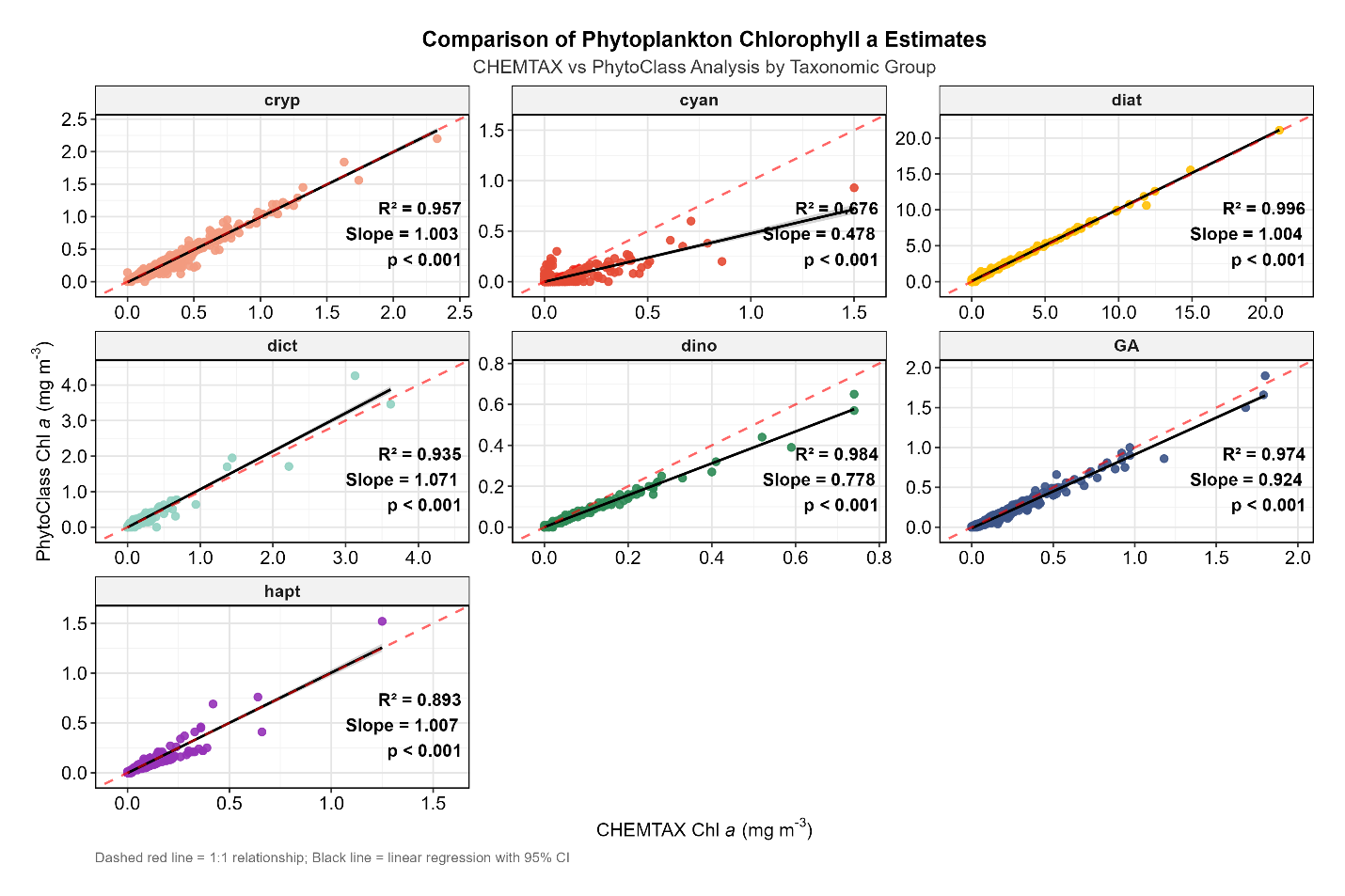
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| dataset | RMSE | condition\_number | type | sample\_size |
| autumn | 0.024607 | 914.2549 | Cluster | 102 |
| summer | 0.025512 | 645.4107 | Cluster | 120 |
| spring | 0.028095 | 476.2789 | Cluster | 113 |
| winter | 0.039708 | 1012.665 | Cluster | 76 |

|  |  |  |
| --- | --- | --- |
| season | pigment | Zea |
| spring | Cyanobacteria | 1.279994 |
| summer | Cyanobacteria | 0.938914 |
| autumn | Cyanobacteria | 1.085118 |
| winter | Cyanobacteria | 0.32 |

1. Min\_lit\_to\_dino (ran on July 24rd, 2025)

* Subbing in literature peridinin
* Could be interesting as one of the groups showing biggest differences from chemtax
* Literature summary values are actually narrower than mine – suspect won’t make a difference
  + Could try cultured values that could be wider
* Difference with CHEMTAX could be that I am still using Chlc1c2, which would attribute more to Dinos during diatom bloom events – possibly.
* Output ratios the same across seasons (1.028) and very similar to prior runs (1.12)
  + Really just a minor group, so maybe why ratios so constrained?
* Try removal of c1c2 from CHEMTAX.

1. Min\_lit\_to\_cyano
   * Skipping haptos and dictyos and moving to cyano to see if it changes the above underestimation.
   * Interesting to see how much it changed the cyano output -an even lower slope, more scatter and a separate low biomass grouping.
   * Everything else seems relatively comparable.
   * Slightly increased condition numbers.



|  |  |  |
| --- | --- | --- |
| season | pigment | Zea |
| spring | Cyanobacteria | 1.716736 |
| summer | Cyanobacteria | 0.952768 |
| autumn | Cyanobacteria | 1.460501 |
| winter | Cyanobacteria | 0.076 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| dataset | RMSE | condition\_number | type | sample\_size |
| autumn | 0.024709 | 885.5163 | Cluster | 102 |
| summer | 0.02564 | 656.658 | Cluster | 120 |
| spring | 0.028117 | 530.4062 | Cluster | 113 |
| winter | 0.038189 | 1300.181 | Cluster | 76 |