

# Supplementary Material to: Dynamics of the bacterial community associated with *Phaeodactylum tricornutum* cultures

Fiona Wanjiku Moejes<sup>1</sup>, Ovidiu Popa<sup>2,3,†</sup>, Antonella Succurro<sup>3,4,†\*</sup>, Julie Maguire<sup>1</sup> and Oliver Ebenhöh<sup>2,3\*</sup>

<sup>1</sup> Bantry Marine Research Station, Gearhies, Bantry, Co. Cork, Ireland; fmoejes@bmrs.ie (F.W.M.); jmaguire@bmrs.ie (J.M.)

<sup>2</sup> Institute of Quantitative and Theoretical Biology, Heinrich-Heine University, Universitätsstraße 1, 40225 Düsseldorf, Germany; ovidiu.popa@hhu.de

<sup>3</sup> CEPLAS (Cluster of Excellence on Plant Sciences), Heinrich-Heine University, Universitätsstraße 1, 40225 Düsseldorf, Germany

<sup>4</sup> Botanical Institute, University of Cologne, Zùlpicher Straße 47b, 50674 Cologne, Germany

Correspondence: a.succurro@uni-koeln.de (A.S.); oliver.ebenhoeh@hhu.de (O.E.)

<sup>†</sup> These authors contributed equally to this work

## 2 Mathematical model and parameter fitting

This document describes additional checks performed on parameter fitting. Abbreviations used in the following: CM (complete media); MM (minimal media); D (diatom); A (alteromonas); P (pseudomonas); PA (pseudoalteromonas); F (flavobacterium);  $\mathcal{P}_{CM}$  (parameter set for CM conditions as in manuscript);  $\mathcal{P}_{MM}$  (parameter set for MM conditions as in manuscript);  $\mathcal{P}(X)_Y$  (subset of the parameter set fitted in condition  $Y$  and relative to organism  $X \in \{D, A, P, PA, F\}$ );  $\mathcal{S}_D$  (fitness score relative to D as defined in Eq. B.29);  $\mathcal{S}_B$  (fitness score relative to the four bacteria relative abundances as defined in Eq. B.30).

### 2.1 Parameter profiling

We performed parameter profiling to test the sensitivity of the fitness scores to a variation of single parameters in a range  $\pm 90\%$  of the fitted value. We identified 12 parameters whose variation caused little to no variation to the fitness scores. These parameters are grouped at the bottom of Table S3. We verified that fixing those parameters to arbitrary values (chosen as values in-between the fitted values in  $\mathcal{P}_{CM}$  and  $\mathcal{P}_{MM}$ ) did not affect the simulation results.

- Table S3 reports the effect on fitness scores of varying single parameters in a range  $\pm 90\%$  of the fitted value;
- Table S4 reports the effect on the MM simulation results from fixing, one after the other, those 12 parameters identified through the profiling;
- Table S5 reports the effect on the CM simulation results from fixing, one after the other, those 12 parameters identified through the profiling;
- Table S6 reports the original parameter sets for CM and MM, and the same sets with the 12 parameters fixed.

### 2.2 De-novo fit of 43 parameters

After deciding to fix 12 parameters of the 55 free parameters of the model, we performed de-novo fitting of the model parameters with the genetic algorithm described in Appendix B of the manuscript. We first generated a new parameter set to serve as initial values for both CM and MM fits. While the 12 parameters have the defined fixed values, all the other 43 parameters are taken from random uniform distribution in range  $[0, 2]$ , except for  $\epsilon_{DOCA}$  generated in  $[0, 1]$ , and  $\epsilon_{DOM}$  is set to  $1 - \epsilon_{DOCA}$ . Five initial random parameter sets (same for CM and MM) are generated, and the genetic algorithm is run ten times on each initial  $\mathcal{P}$  for both conditions separately, for a total of 100 fitted parameter sets.

Table S7 shows some of the fits with the best scores. The goodness of the fits is however to be critically evaluated not only based on the fitness scores. For example CM-01-09 shows the symptoms of overfitting, with a non-smooth diatom growth curve, bacterial biomasses showing multiple rises and decays and noisy relative abundances. CM-02-05 on the other hand has a perfect and smooth fit to the diatom growth data, but missed completely the two low-abundance bacteria. CM-04-06 is comparable to CM-01-09 even though the biomasses are slightly smoother. Overall, out of the 100 runs the CM-05-02 showed the best characteristics to be chosen for re-iteration of the fit: smooth growth curve of the diatom, well defined rise and decay phases in bacterial biomasses and good, non-noisy fit of relative abundances.

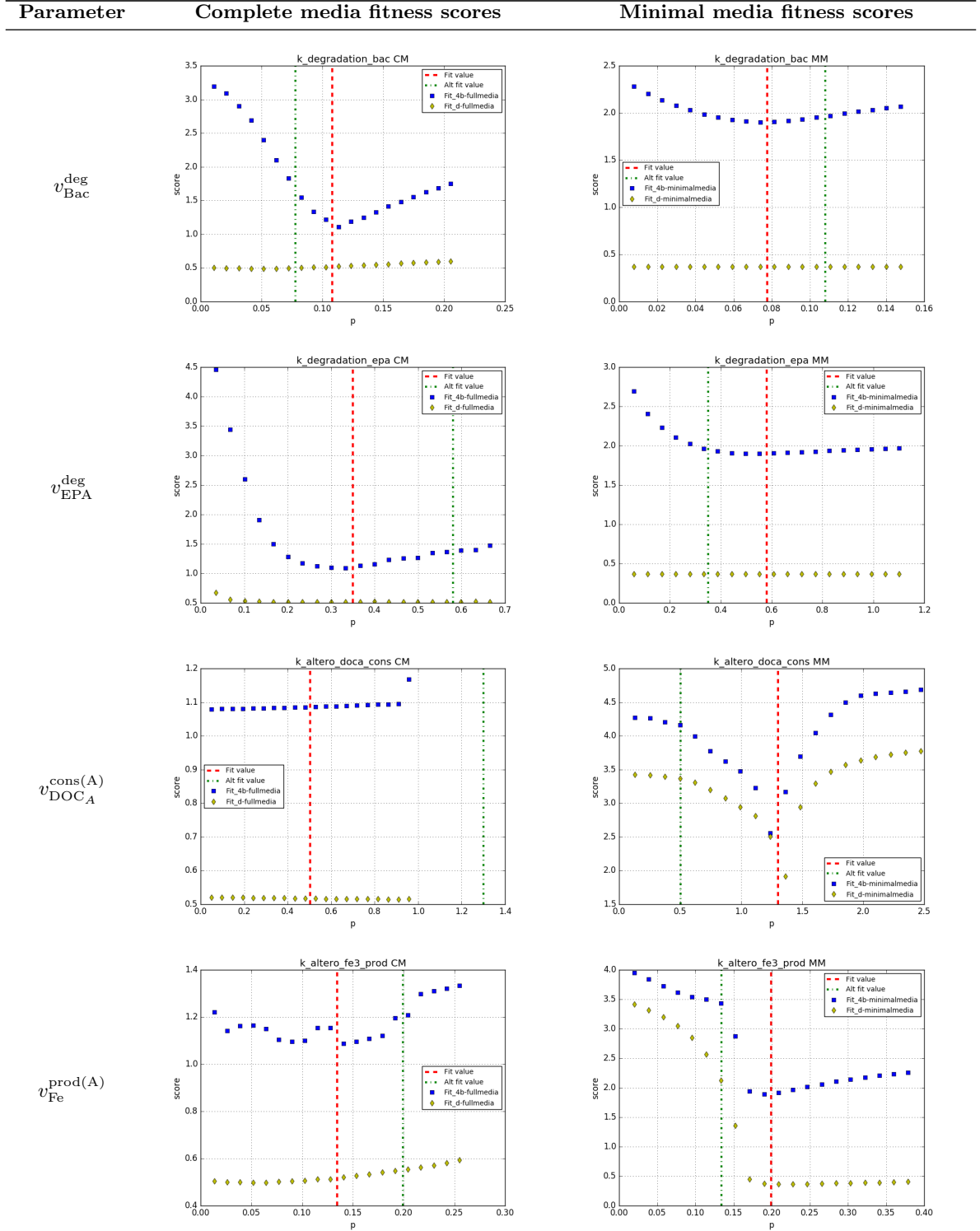
Refits of  $\mathcal{P}_{CM-05-02}$  were performed for CM and MM allowing each of the 43 parameters to vary within  $\pm 90\%$  of the fitted value. The two best fits are shown in Table S8, and they were selected to a further run of genetic algorithm resulting in the best fit for CM of Table S9.

$\mathcal{P}_{refit2-CM-09}$  did not work at all for MM conditions, therefore further rounds of fit were performed in MM conditions only starting from this parameter set. First, only  $\mathcal{P}(D)$  and  $\mathcal{P}(A)$  of the 43 parameters were allowed to vary within  $\pm 90\%$ , all the other parameters were kept fixed. The best results are shown in Table S10.

None of them could capture the bacterial dynamics in a satisfactory way, therefore a further set of genetic algorithm runs was performed. This time for each of the six best re-refits the parameter set was varied:  $\pm 50\%$  for  $D, A$ ,  $\pm 5\%$  for  $PA, F, P$ , and fixed for the rest. This time neither it was possible to capture the bacterial dynamics as well as with the original parameter set  $\mathcal{P}_{MM}$ . The best results are shown in Table S11.

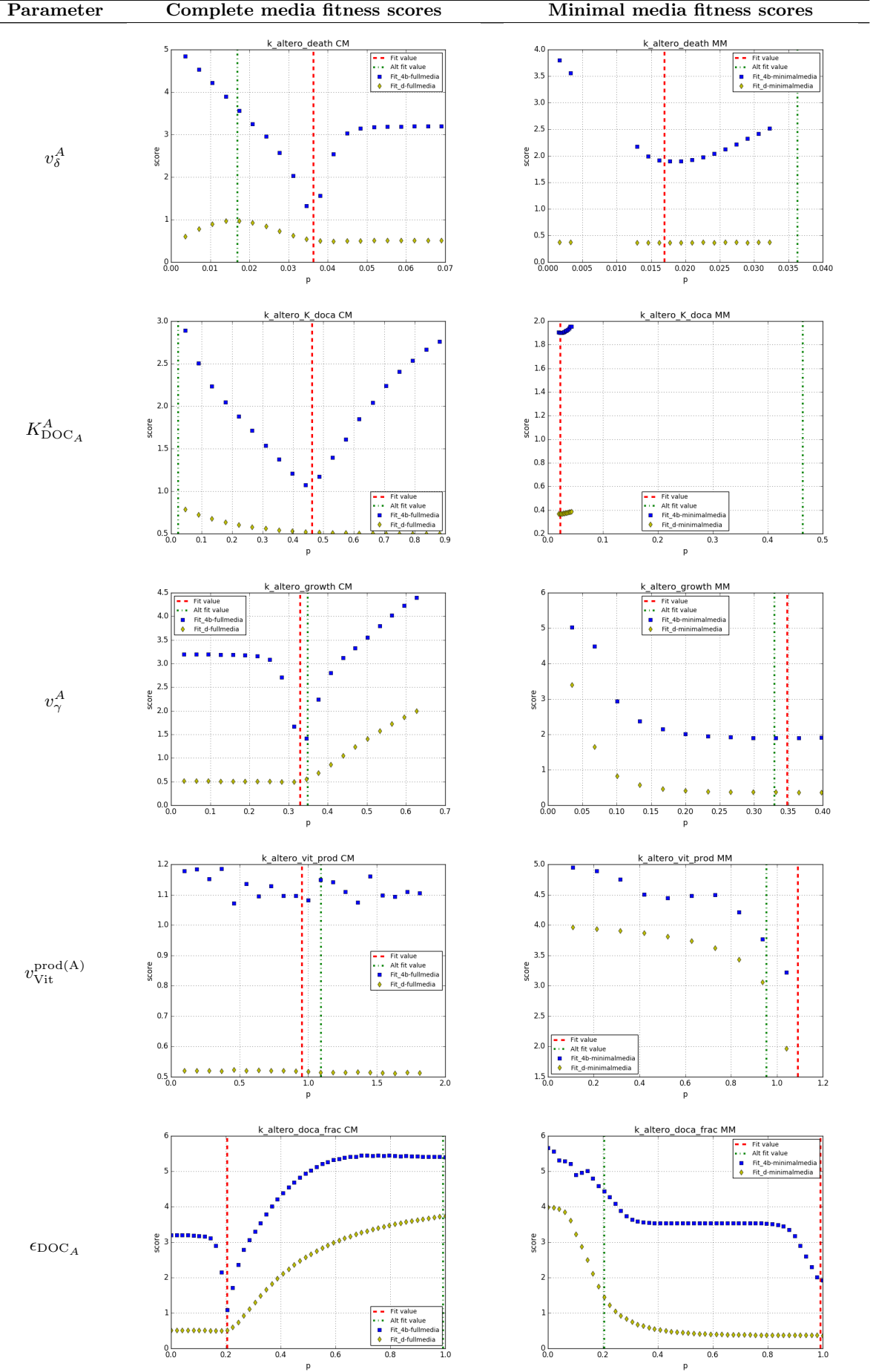
## 2.3 Tables and Figures

Table S3: Parameter profiling. Plots report the diatom (blue squares) and bacteria (yellow diamonds) fitness scores for simulations in complete and minimal media conditions run with the respective parameter sets. Also reported as red dashed line is the fitted value of the parameter being varied in a range  $\pm 90\%$ , and as a green dot-dash line is the value of the parameter in the other parameter set. Highlighted at the bottom of the Table are the 12 parameters chosen to be fixed to an intermediate value.



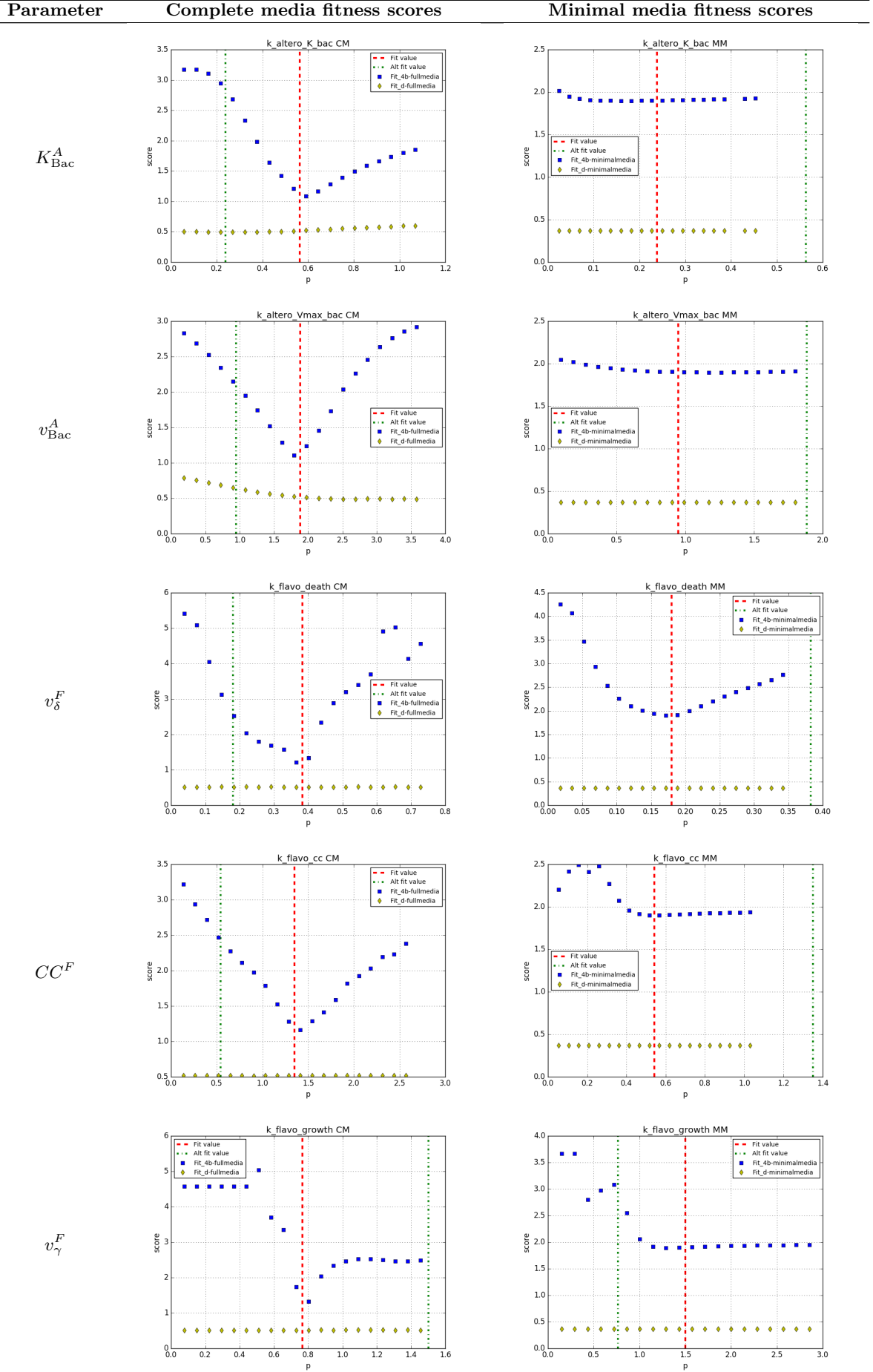
Continued on next page

Table S3 – Continued from previous page



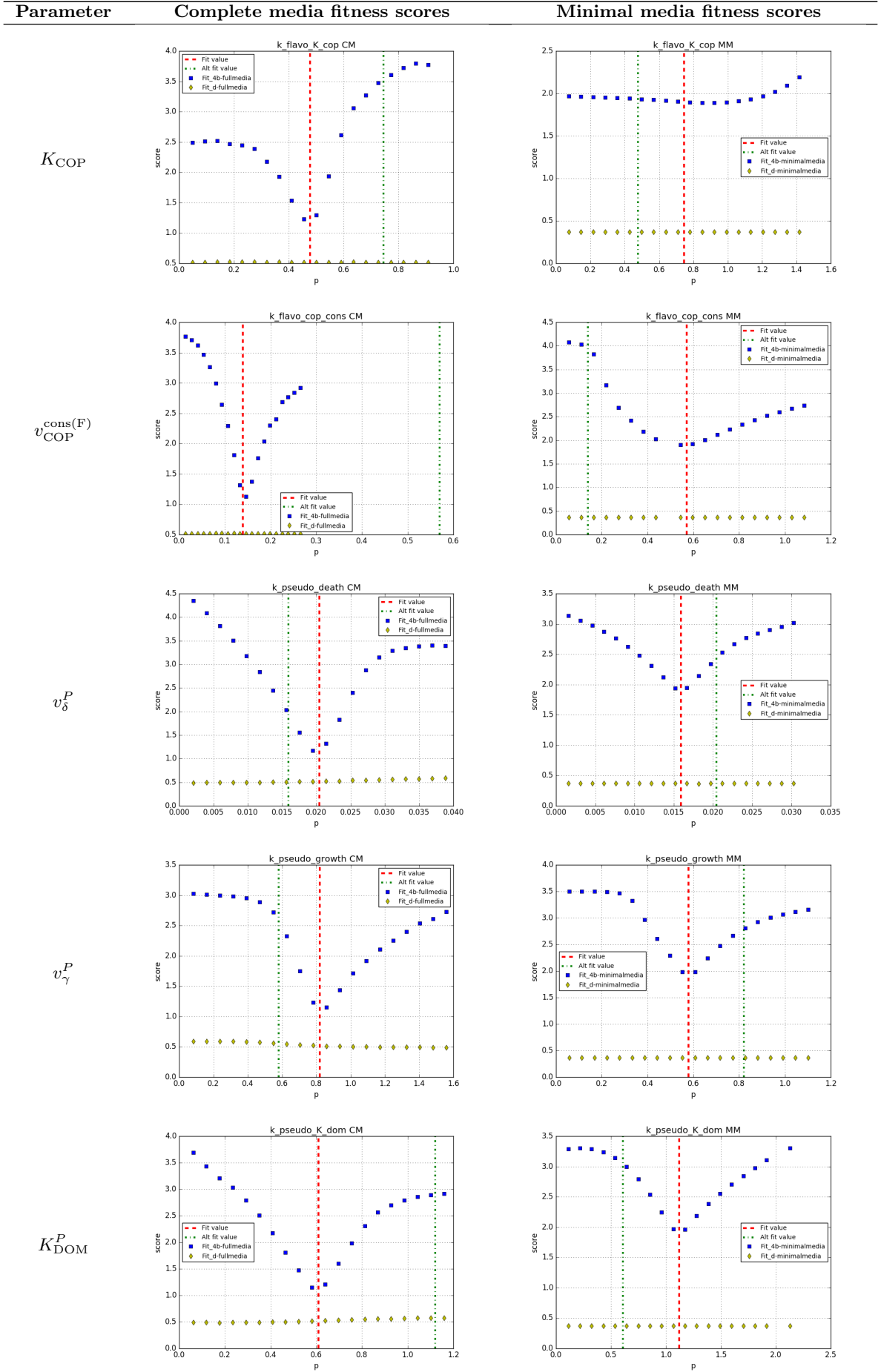
Continued on next page

Table S3 – Continued from previous page



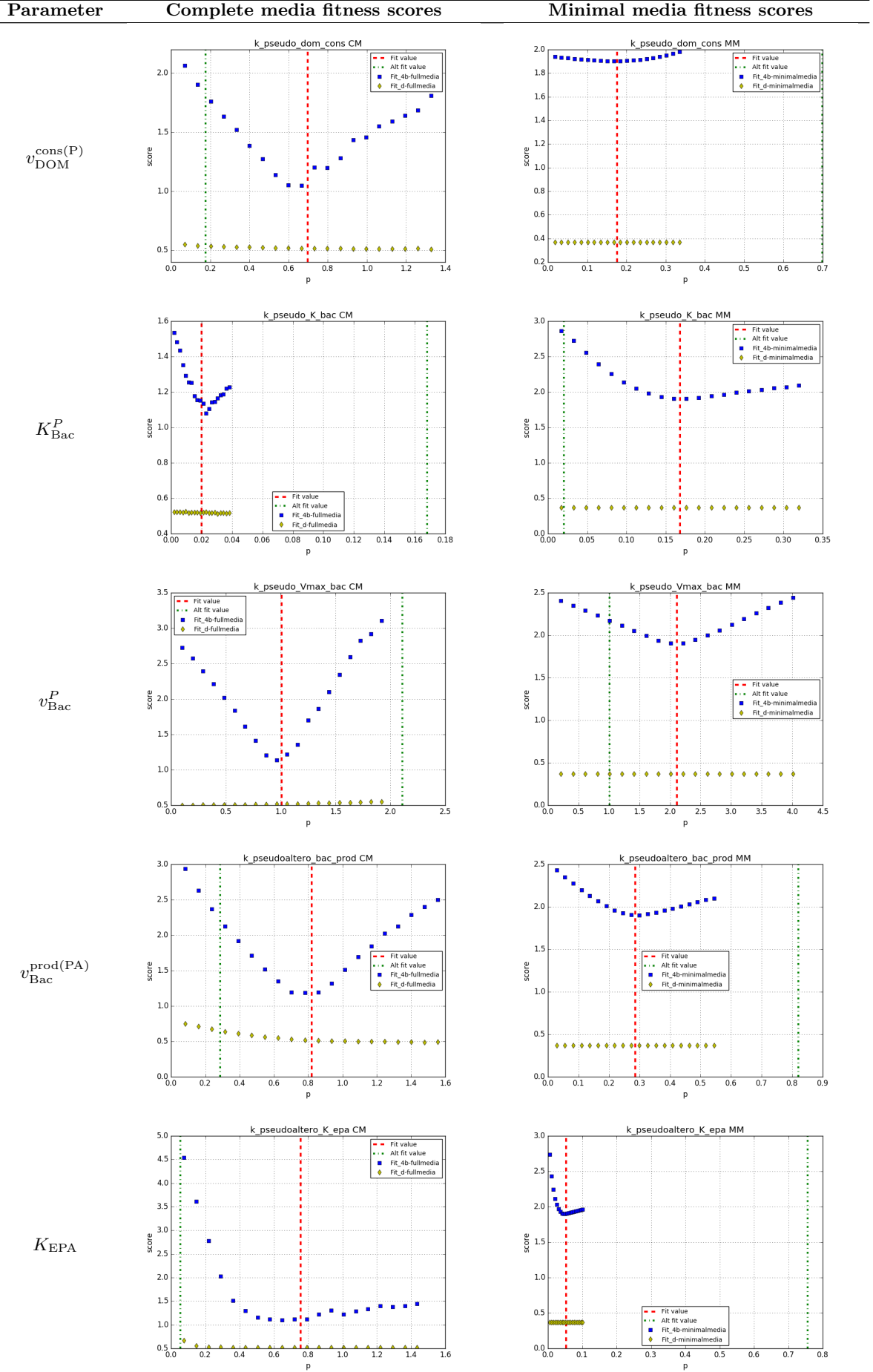
Continued on next page

Table S3 – Continued from previous page



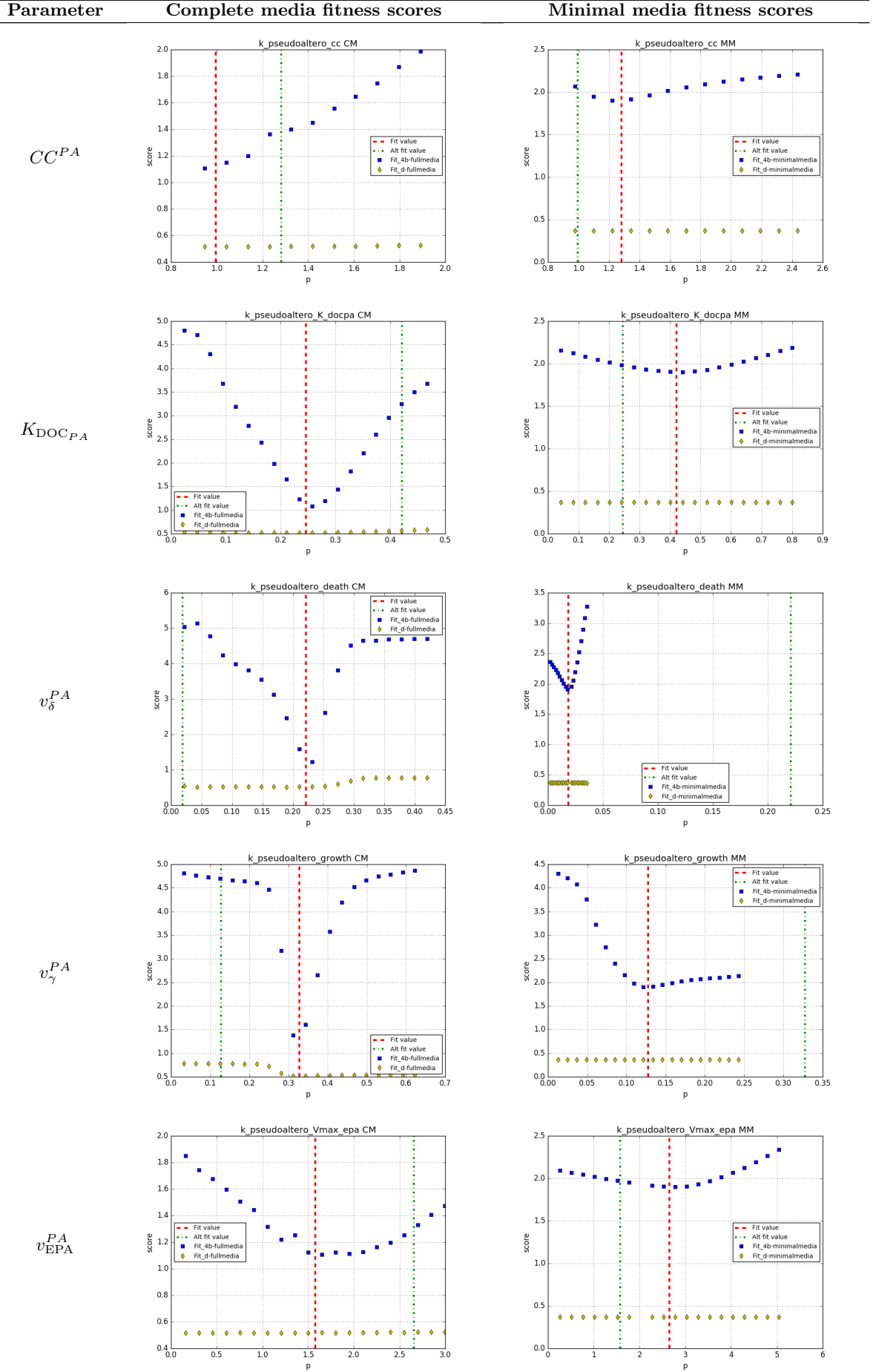
Continued on next page

Table S3 – Continued from previous page



Continued on next page

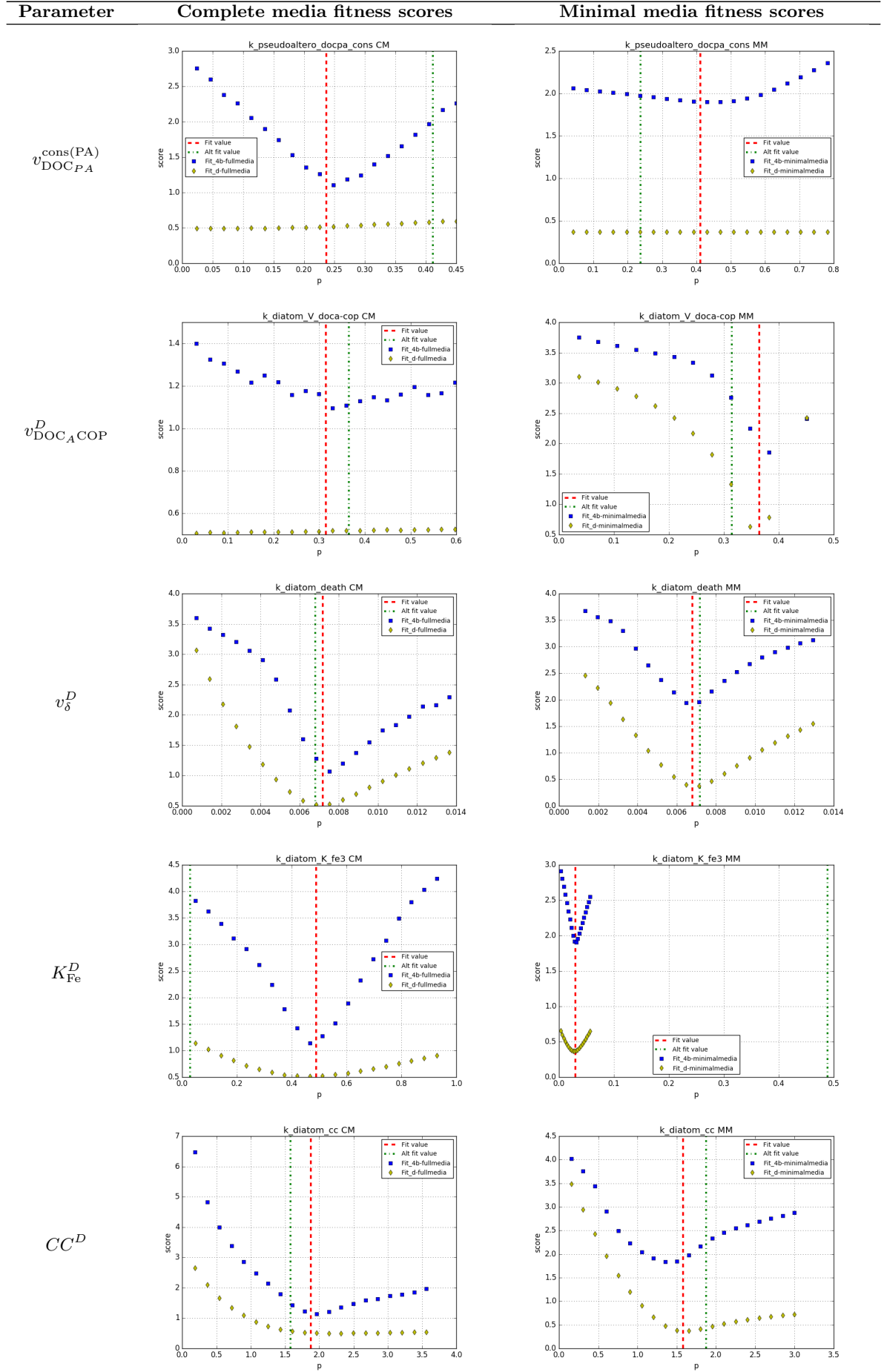
Table S3 – Continued from previous page



Continued on next page

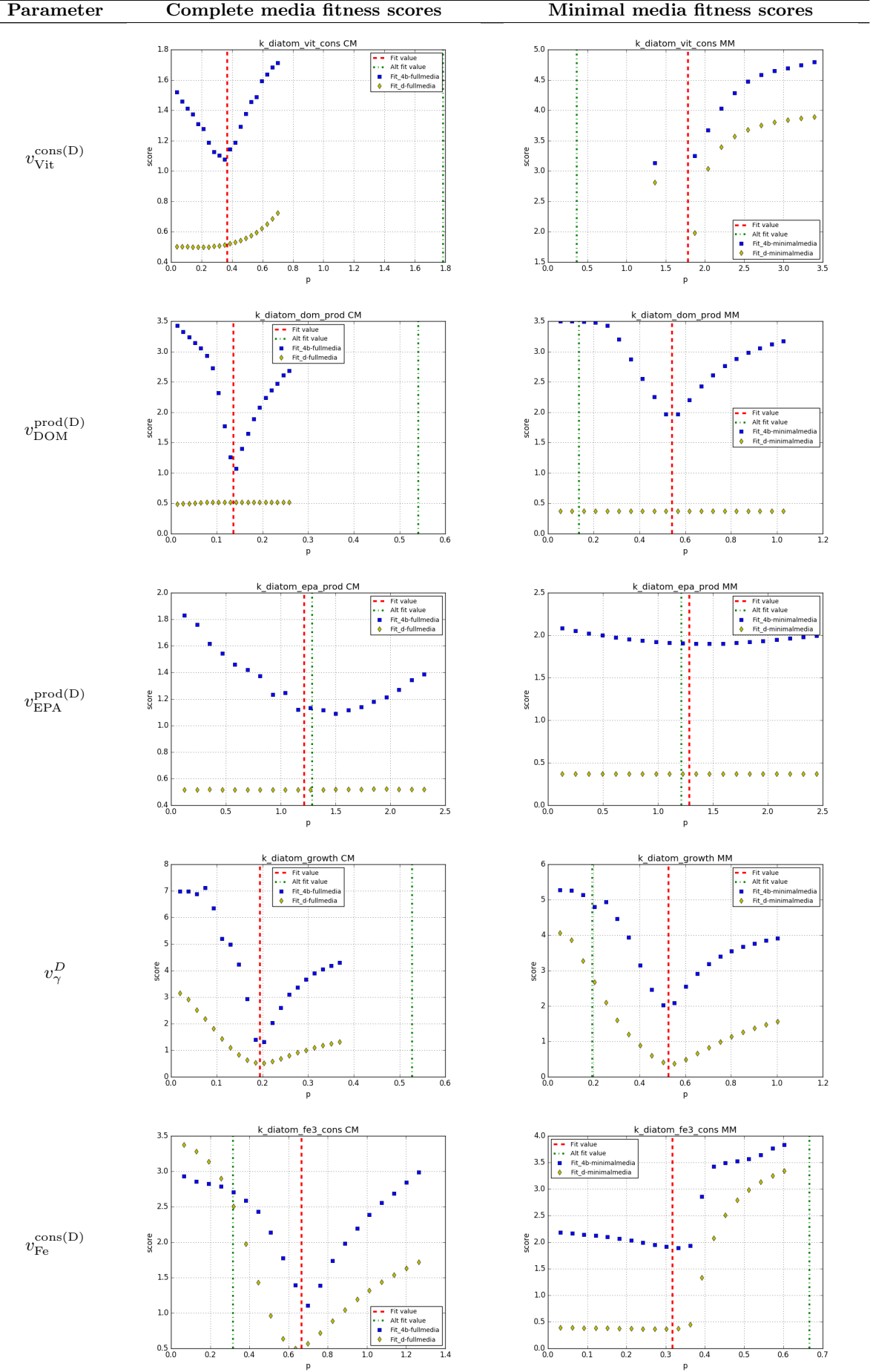


Table S3 – Continued from previous page



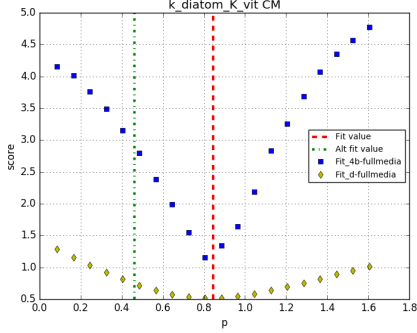
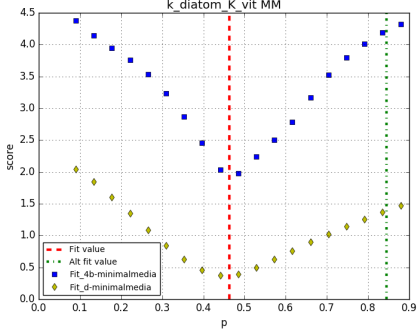
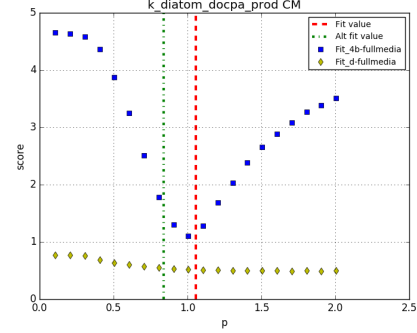
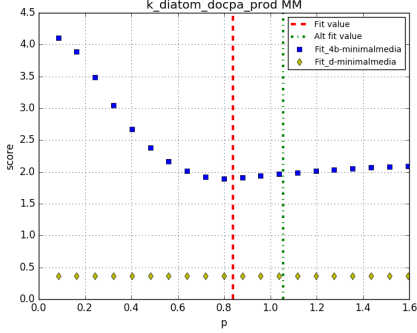
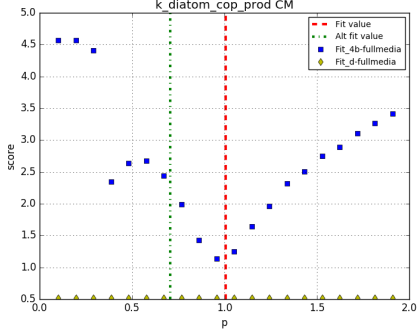
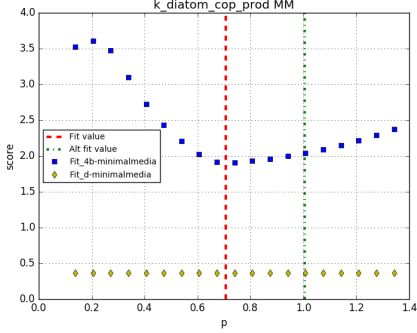
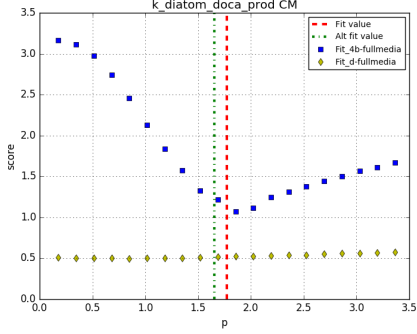
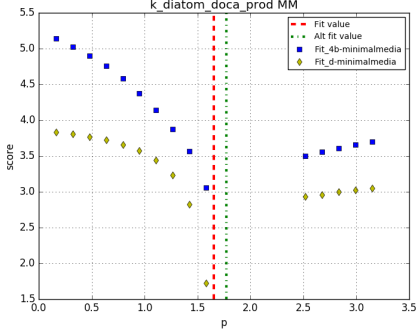
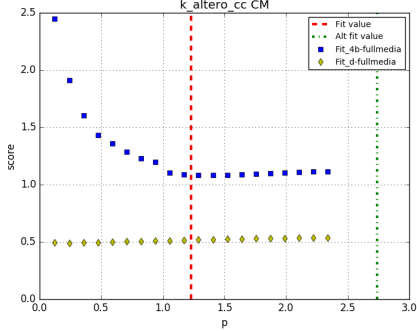
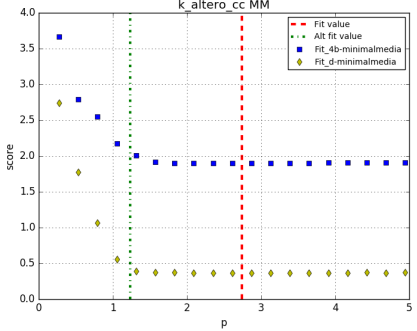
Continued on next page

Table S3 – Continued from previous page



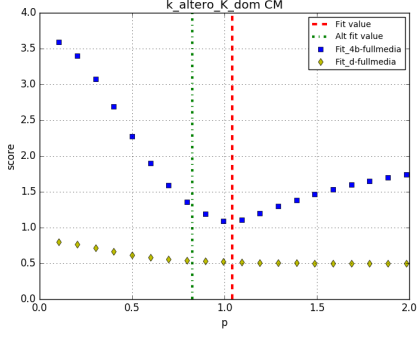
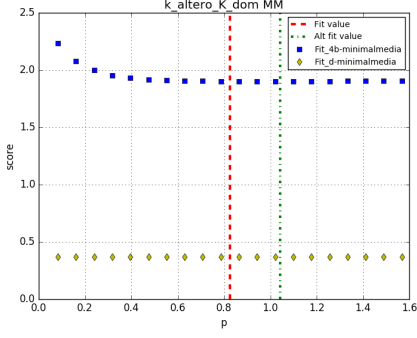
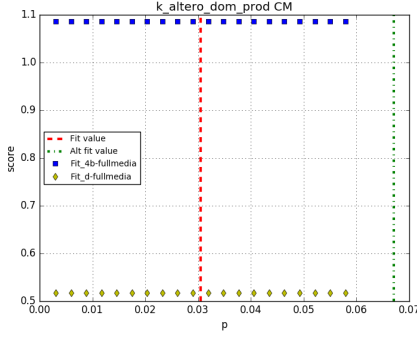
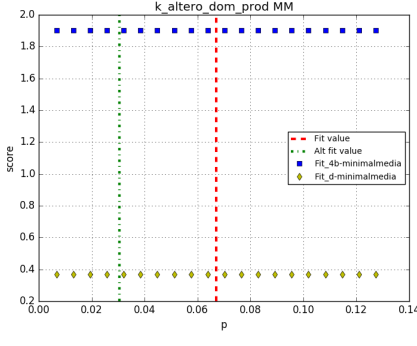
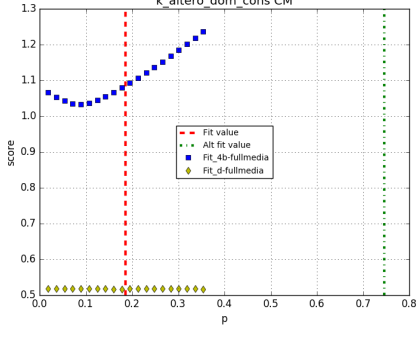
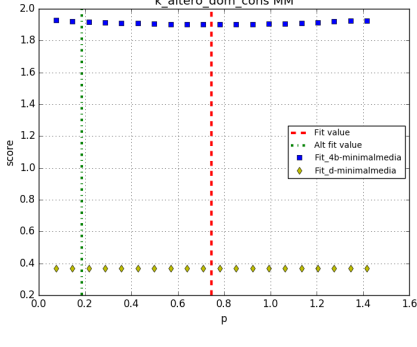
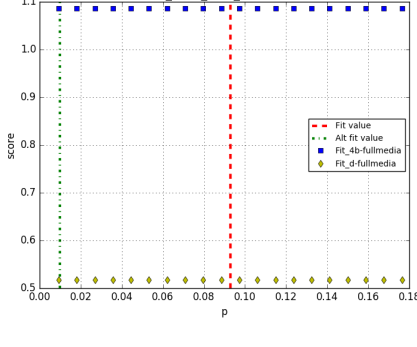
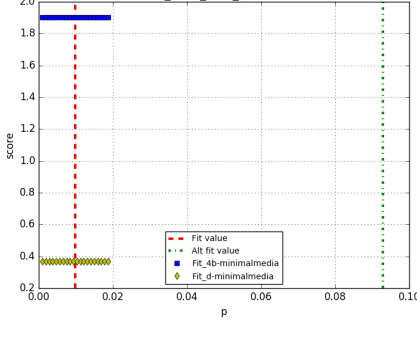
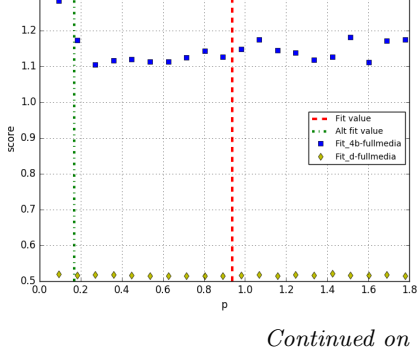
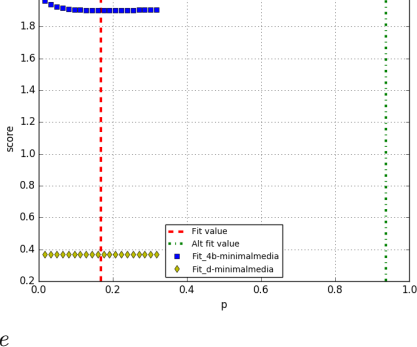
Continued on next page

Table S3 – Continued from previous page

Parameter	Complete media fitness scores	Minimal media fitness scores
$K_{vit}^D$		
$v_{DOC_{PA}}^{prod(D)}$		
$v_{COP}^{prod(D)}$		
$v_{DOC_A}^{prod(D)}$		
$CC^A$		

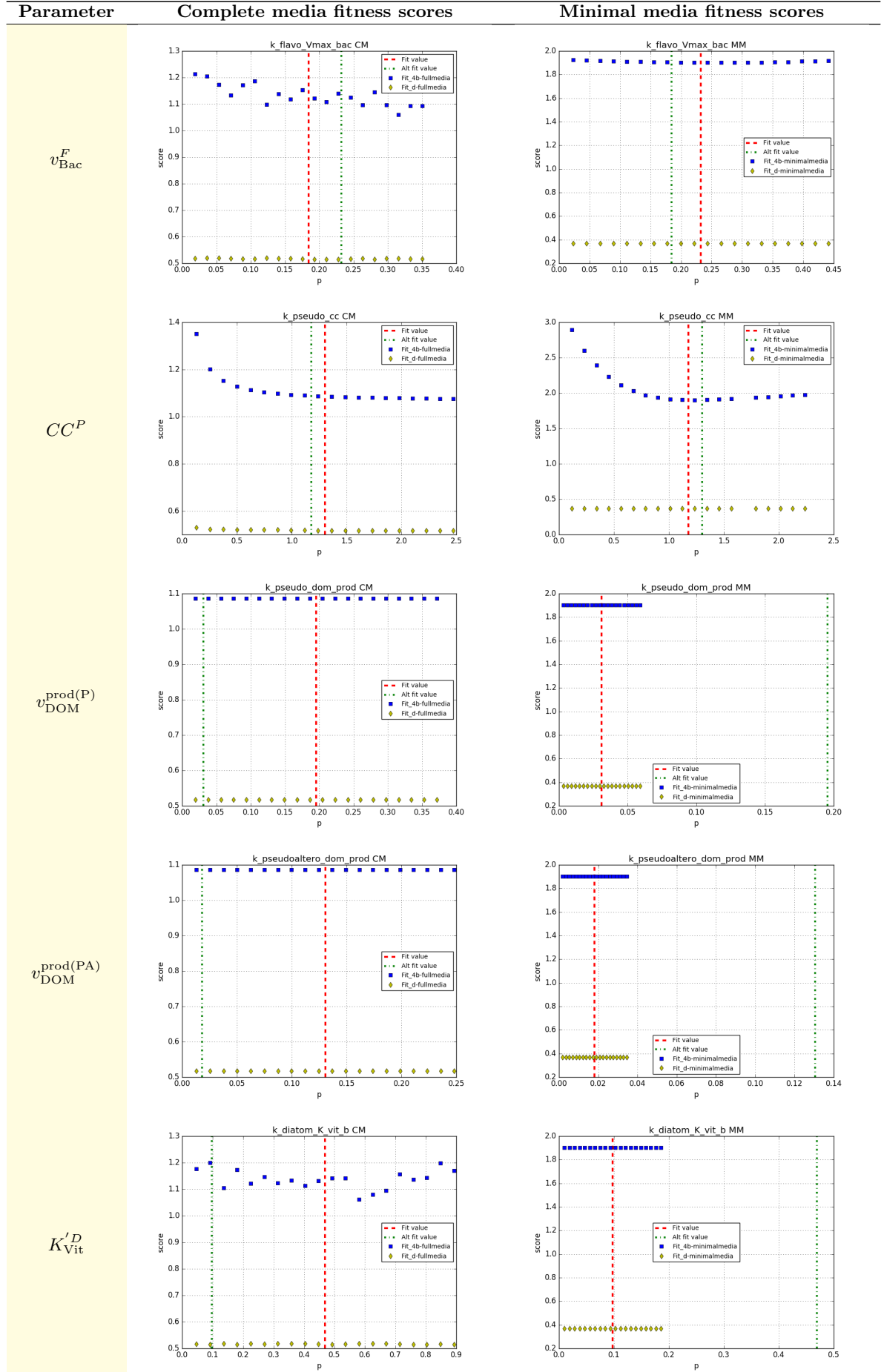
Continued on next page

Table S3 – Continued from previous page

Parameter	Complete media fitness scores	Minimal media fitness scores
$K_{\text{DOM}}^A$		
		
$v_{\text{DOM}}^{\text{cons}(A)}$		
$v_{\text{DOM}}^{\text{prod}(F)}$		
$K_{\text{Bac}}^F$		

Continued on next page

Table S3 – Continued from previous page



Continued on next page

Table S3 – Continued from previous page

Parameter	Complete media fitness scores	Minimal media fitness scores
$K'_{\text{Fe}}^D$		

Table S4: Check of the effect of fixing parameters values one after the other in the  $\mathcal{P}_{MM}$  set. Reported are the original value  $p$  of the parameter in  $\mathcal{P}_{MM}$ , the new value  $p'$ , and the updated fitness scores for diatom and bacteria fits ( $\mathcal{S}_D$  and  $\mathcal{S}_B$ ). The plots show the data (squares) and simulation results: lines refer to simulations run with  $p$ , stars to simulations run with  $p'$ .

			$\mathcal{S}_D$	$\mathcal{S}_B$	plot
					Baseline parameters:
	original $\mathcal{P}_{MM}$		0.36804	1.90259	
parameter	$p$	$p'$	$\mathcal{S}_D$	$\mathcal{S}_B$	plot
					Baseline parameters:
$v_{\text{DOM}}^{\text{prod(A)}}$	0.06702	0.03000	0.36804	1.90259	*
					Baseline parameters:
$v_{\text{DOM}}^{\text{prod(P)}}$	0.03107	0.05000	0.36804	1.90259	*
					Baseline parameters:
$v_{\text{DOM}}^{\text{prod(F)}}$	0.00984	0.01000	0.36804	1.90259	*

Continued on next page

Table S4 – Continued from previous page

parameter	$p$	$p'$	$\mathcal{S}_D$	$\mathcal{S}_B$	plot
$v_{\text{DOM}}^{\text{prod(PA)}}$	0.01816	0.03000	0.36804	1.90259	<p>Baseline parameters:</p>
$K_{\text{Vit}}'^D$	0.09782	0.50000	0.36945	1.90213	<p>Baseline parameters:</p>
$CC^A$	2.74047	1.60000	0.37309	1.91152	<p>Baseline parameters:</p>
$v_{\text{DOM}}^{\text{cons(A)}}$	0.74598	0.10000	0.37257	1.91542	<p>Baseline parameters:</p>

Continued on next page



Table S4 – *Continued from previous page*

parameter	$p$	$p'$	$\mathcal{S}_D$	$\mathcal{S}_B$	plot
$v_{\text{Bac}}^F$	0.23234	0.20000	0.37390	1.91976	<p>Baseline parameters:</p>
$K_{\text{DOM}}^A$	0.82552	1.05000	0.37306	1.91726	<p>Baseline parameters:</p>
$CC^P$	1.17750	1.25000	0.37336	1.92086	<p>Baseline parameters:</p>
$K_{\text{Bac}}^F$	0.16731	0.25000	0.37315	1.92422	<p>Baseline parameters:</p>

*Continued on next page*

Table S4 – *Continued from previous page*

parameter	$p$	$p'$	$\mathcal{S}_D$	$\mathcal{S}_B$	plot
$K_{\text{Fe}}'^D$	0.33321	1.00000	0.37279	1.92504	<p>Baseline parameters:</p> <p>★</p>

Table S5: Check of the effect of fixing parameters values one after the other in the  $\mathcal{P}_{CM}$  set. Reported are the original value  $p$  of the parameter in  $\mathcal{P}_{CM}$ , the new value  $p'$ , and the updated fitness scores for diatom and bacteria fits ( $\mathcal{S}_D$  and  $\mathcal{S}_B$ ). The plots show the data (squares) and simulation results: lines refer to simulations run with  $p$ , stars to simulations run with  $p'$ .

			$\mathcal{S}_D$	$\mathcal{S}_B$	plot
					Baseline parameters:
original $\mathcal{P}_{CM}$			0.51698	1.08592	
parameter	$p$	$p'$	$\mathcal{S}_D$	$\mathcal{S}_B$	plot
					Baseline parameters:
$v_{DOM}^{prod(A)}$	0.03047	0.03000	0.51698	1.08592	*
					Baseline parameters:
$v_{DOM}^{prod(P)}$	0.19545	0.05000	0.51698	1.08592	*

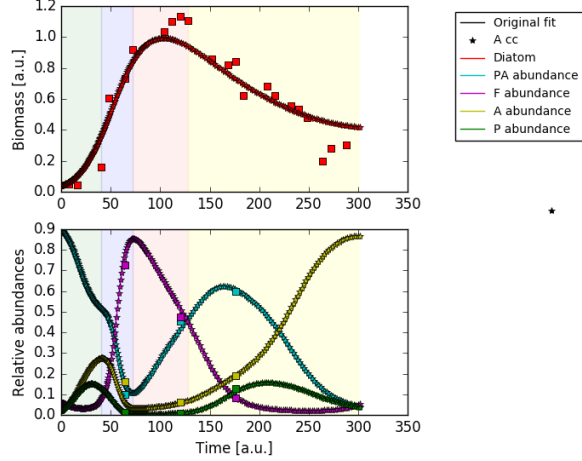
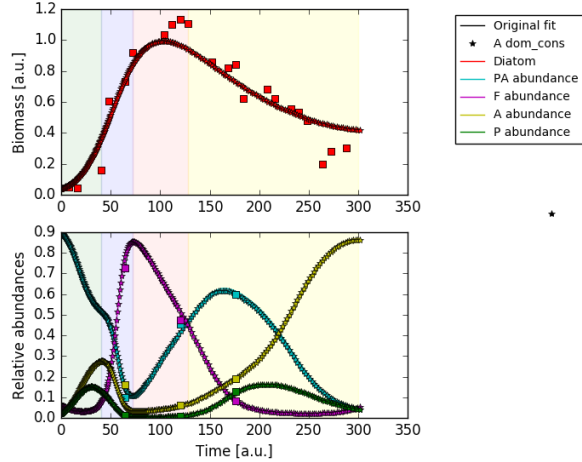
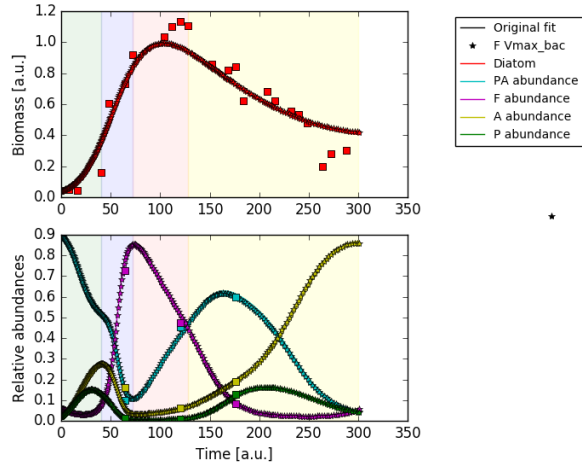
Continued on next page

Table S5 – Continued from previous page

parameter	$p$	$p'$	$\mathcal{S}_D$	$\mathcal{S}_B$	plot
Baseline parameters:					
$v_{\text{DOM}}^{\text{prod}(F)}$	0.09286	0.01000	0.51698	1.08592	
$v_{\text{DOM}}^{\text{prod}(PA)}$	0.13062	0.03000	0.51698	1.08592	
$K_{\text{Vit}}'^D$	0.46960	0.50000	0.51547	1.11314	

Continued on next page

Table S5 – Continued from previous page

parameter	$p$	$p'$	$\mathcal{S}_D$	$\mathcal{S}_B$	plot
Baseline parameters:					
$CC^A$	1.23092	1.60000	0.52307	1.11081	
$v_{\text{DOM}}^{\text{cons}(A)}$	0.18634	0.10000	0.52472	1.05070	
$v_{\text{Bac}}^F$	0.18436	0.20000	0.52400	1.06690	

Continued on next page

Table S5 – Continued from previous page

parameter	$p$	$p'$	$\mathcal{S}_D$	$\mathcal{S}_B$	plot
Baseline parameters:					
$K_{\text{DOM}}^A$	1.04349	1.05000	0.52295	1.04093	
$CC^P$	1.30132	1.25000	0.52362	1.05926	
$K_{\text{Bac}}^F$	0.93642	0.25000	0.52667	1.05065	

Continued on next page

Table S5 – *Continued from previous page*

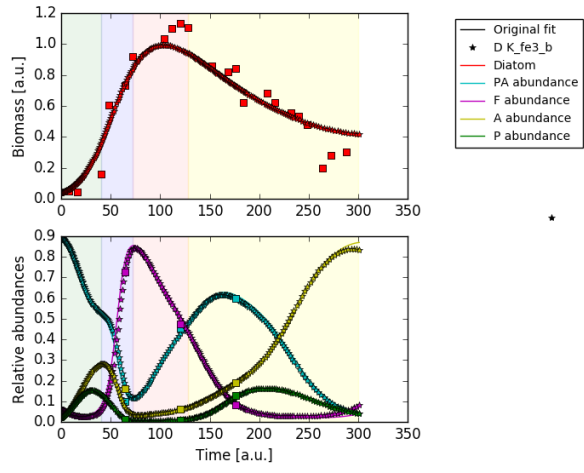
parameter	$p$	$p'$	$\mathcal{S}_D$	$\mathcal{S}_B$	plot
Baseline parameters:					
$K_{\text{Fe}}'^D$	1.19973	1.00000	0.52571	1.04838	

Table S6: Parameter sets fitted in CM and MM ( $\mathcal{P}_{CM}$  and  $\mathcal{P}_{MM}$  respectively) and the respective parameter sets with 12 values fixed after profiling check ( $\mathcal{P}'$ ).

	$\mathcal{P}_{CM}$	$\mathcal{P}'_{CM}$	$\mathcal{P}_{MM}$	$\mathcal{P}'_{MM}$
$v_{Bac}^{deg}$	0.10811	0.10811	0.07769	0.07769
$v_{EPA}^{deg}$	0.35005	0.35005	0.57995	0.57995
$v_{DOC}^{cons(A)}$	0.50422	0.50422	1.30073	1.30073
$v_{Fe}^{prod(A)}$	0.13429	0.13429	0.19948	0.19948
$v_{\delta}^A$	0.03631	0.03631	0.01697	0.01697
$\epsilon_{DOM}$	0.79553	0.79553	0.00743	0.00743
$K_{DOC_A}^A$	0.46369	0.46369	0.02253	0.02253
$v_{\gamma}^A$	0.32952	0.32952	0.34841	0.34841
$v_{Vit}^{prod(A)}$	0.95424	0.95424	1.09226	1.09226
$\epsilon_{DOC_A}$	0.20447	0.20447	0.99257	0.99257
$K_{Bac}^A$	0.56278	0.56278	0.23821	0.23821
$v_{Bac}^A$	1.88469	1.88469	0.94702	0.94702
$v_{\delta}^F$	0.38282	0.38282	0.18005	0.18005
$CC^F$	1.35105	1.35105	0.54187	0.54187
$v_{\gamma}^F$	0.76545	0.76545	1.50156	1.50156
$K_{COP}$	0.47770	0.47770	0.74525	0.74525
$v_{COP}^{cons(F)}$	0.13932	0.13932	0.57005	0.57005
$v_{\delta}^P$	0.02044	0.02044	0.01591	0.01591
$v_{\gamma}^P$	0.82072	0.82072	0.57938	0.57938
$K_{DOM}^P$	0.60980	0.60980	1.12080	1.12080
$v_{DOM}^{cons(P)}$	0.69833	0.69833	0.17625	0.17625
$K_{Bac}^P$	0.02010	0.02010	0.16823	0.16823
$v_{Bac}^P$	1.00974	1.00974	2.11081	2.11081
$v_{Bac}^{prod(PA)}$	0.81958	0.81958	0.28618	0.28618
$K_{EPA}$	0.75557	0.75557	0.05329	0.05329
$CC^{PA}$	0.99513	0.99513	1.28138	1.28138
$K_{DOC_{PA}}$	0.24572	0.24572	0.42128	0.42128
$v_{\delta}^{PA}$	0.22104	0.22104	0.01861	0.01861
$v_{\gamma}^{PA}$	0.32743	0.32743	0.12769	0.12769
$v_{EPA}^{PA}$	1.57705	1.57705	2.65508	2.65508
$v_{DOC_{PA}}^{cons(PA)}$	0.23682	0.23682	0.41130	0.41130
$v_{DOC_{PA}}^D$	0.31433	0.31433	0.36480	0.36480
$v_{\delta}^D$	0.00718	0.00718	0.00681	0.00681
$K_{Fe}^D$	0.48868	0.48868	0.02979	0.02979
$CC^D$	1.87520	1.87520	1.57897	1.57897
$v_{Vit}^{cons(D)}$	0.36788	0.36788	1.78450	1.78450
$v_{DOM}^{prod(D)}$	0.13598	0.13598	0.54133	0.54133
$v_{EPA}^{prod(D)}$	1.21435	1.21435	1.28659	1.28659
$v_{\gamma}^D$	0.19431	0.19431	0.52737	0.52737
$v_{Fe}^{cons(D)}$	0.66574	0.66574	0.31684	0.31684
$K_{Vit}^D$	0.84490	0.84490	0.46274	0.46274
$v_{DOC_{PA}}^{prod(D)}$	1.05527	1.05527	0.83897	0.83897
$v_{COP}^{prod(D)}$	1.00511	1.00511	0.70666	0.70666
$v_{DOC_A}^{prod(D)}$	1.77074	1.77074	1.65657	1.65657
$CC^A$	1.23092	1.60000	2.74047	1.60000
$K_{DOM}^A$	1.04349	1.05000	0.82552	1.05000
$v_{DOM}^{prod(A)}$	0.03047	0.03000	0.06702	0.03000
$v_{DOM}^{cons(A)}$	0.18634	0.10000	0.74598	0.10000
$v_{DOM}^{prod(F)}$	0.09286	0.01000	0.00984	0.01000
$K_{Bac}^F$	0.93642	0.25000	0.16731	0.25000
$v_{Bac}^F$	0.18436	0.20000	0.23234	0.20000
$CC^P$	1.30132	1.25000	1.17750	1.25000
$v_{DOM}^{prod(P)}$	0.19545	0.05000	0.03107	0.05000
$v_{DOM}^{prod(PA)}$	0.13062	0.03000	0.01816	0.03000
$K_{Vit}^D$	0.46960	0.50000	0.09782	0.50000
$K_{Fe}^D$	1.19973	1.00000	0.33321	1.00000



Table S7: Examples of genetic algorithm fit results.

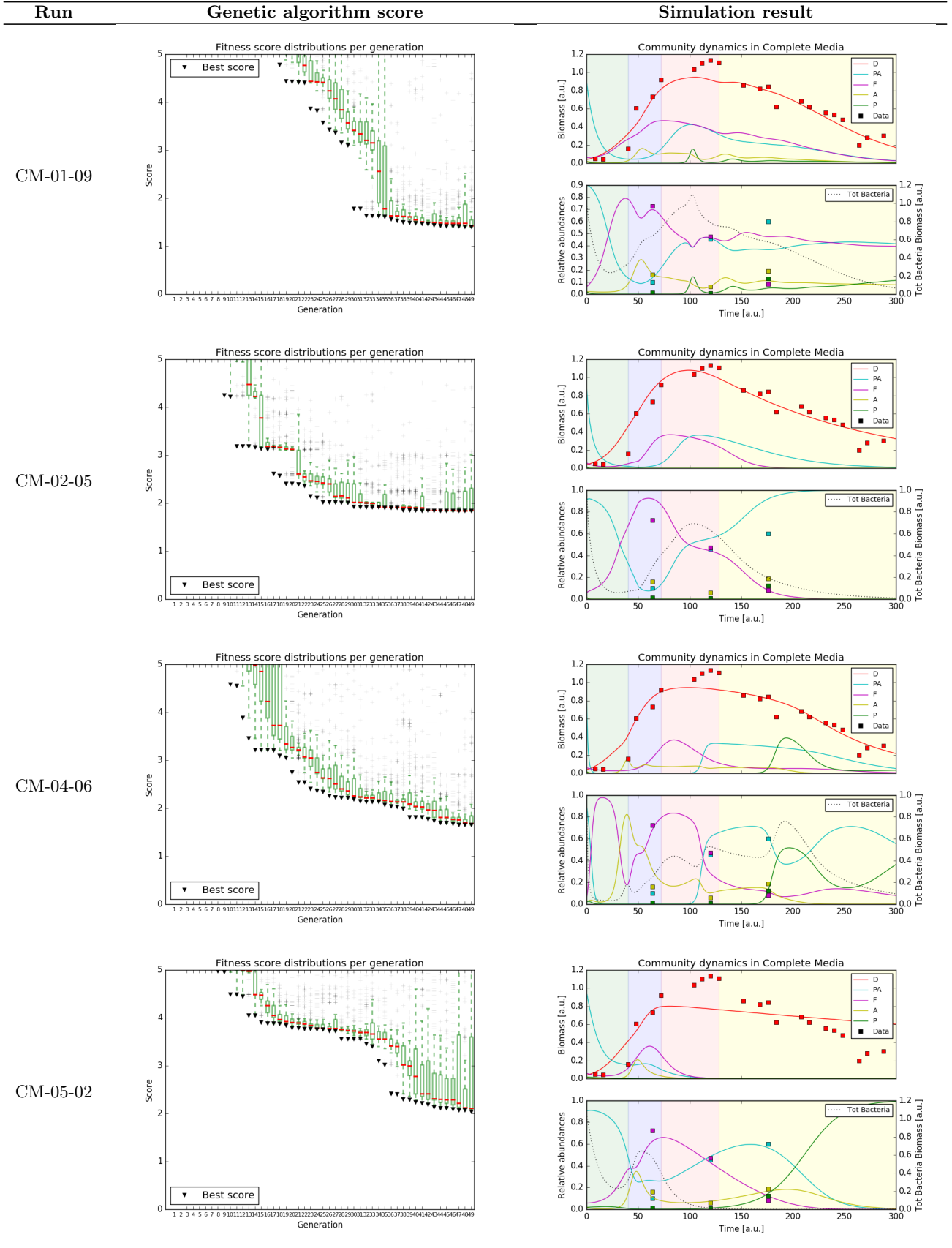


Table S8: Best re-fits of CM-05-02.

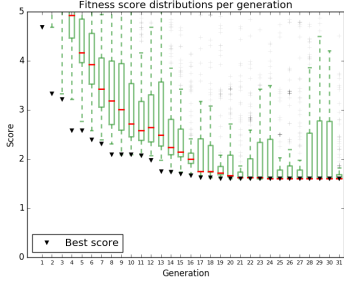
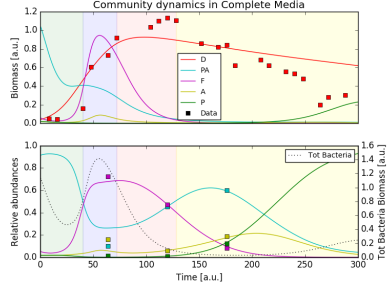
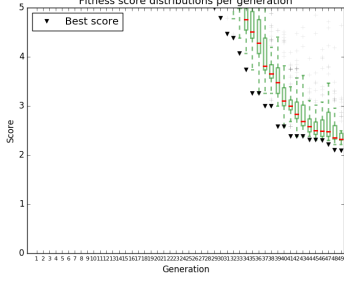
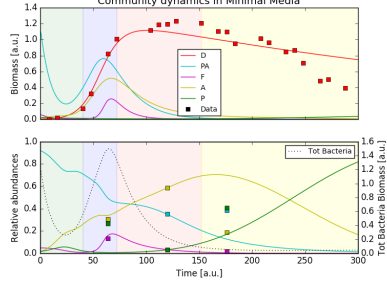
Run	Genetic algorithm score	Simulation result
refit-CM-04		
refit-MM-04		

Table S9: Best re-fit of `refit-MM-04`.

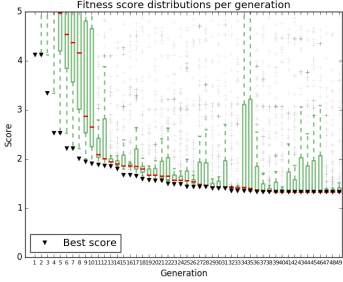
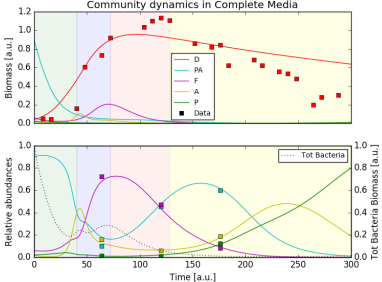
Run	Genetic algorithm score	Simulation result
refit2-CM-09		

Table S10: Best re-fits of refit2-CM-09.

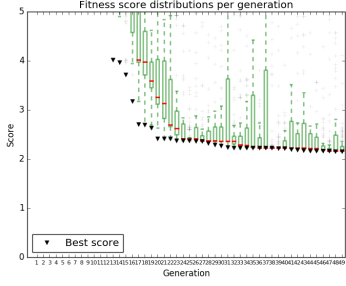
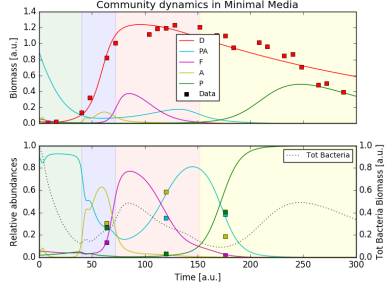
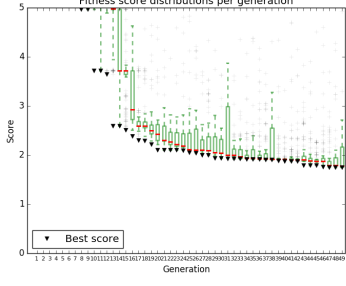
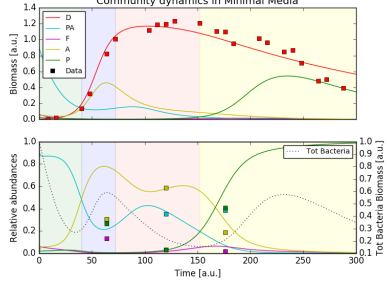
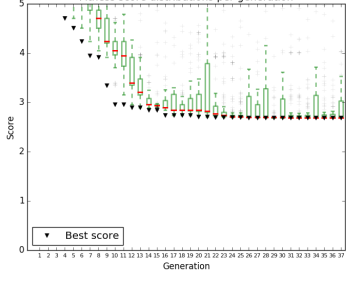
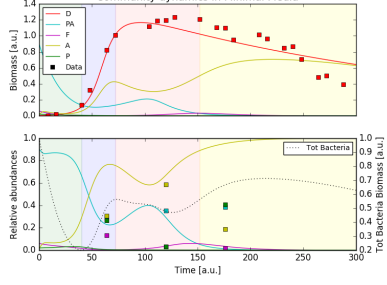
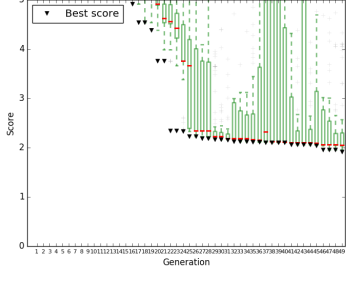
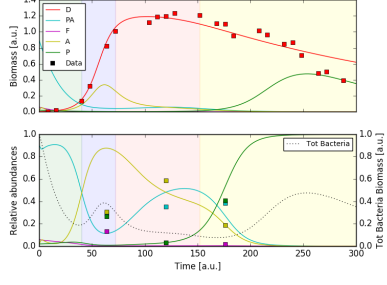
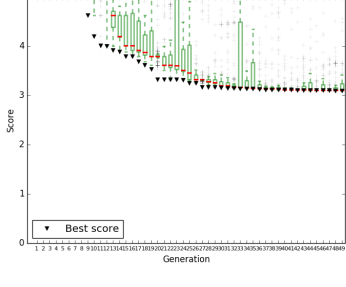
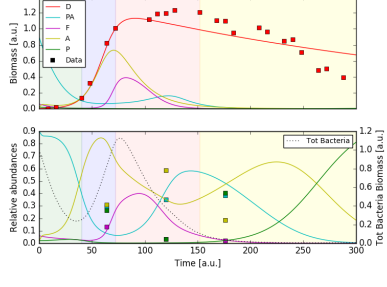
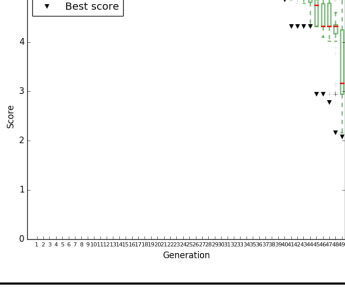
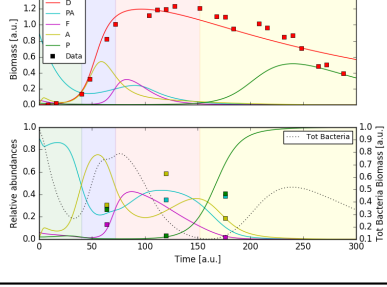
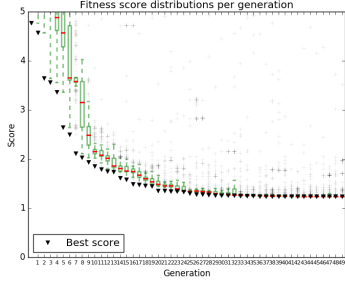
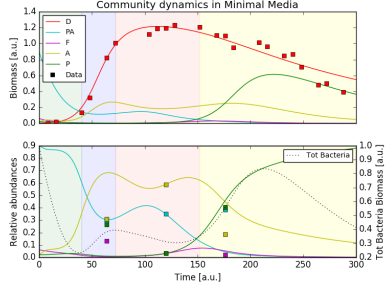
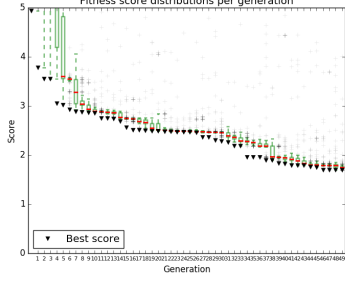
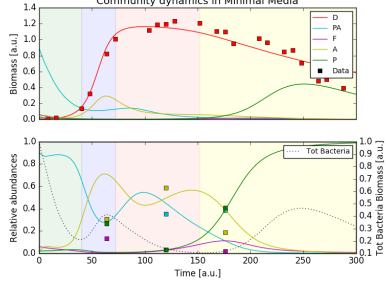
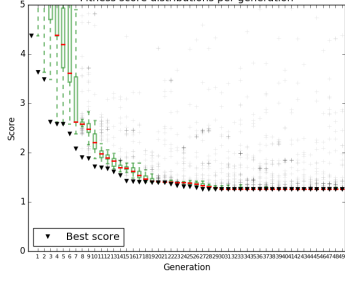
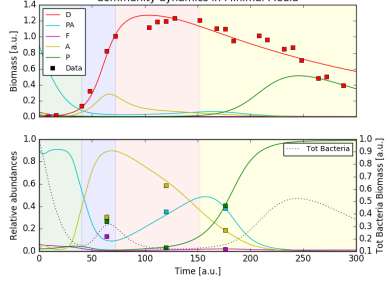
Run	Genetic algorithm score	Simulation result
re-refit-MM-01		
re-refit-MM-05		
re-refit-MM-06		
re-refit-MM-12		
re-refit-MM-13		
re-refit-MM-19		

Table S11: Best re-fits of re-refit-MM-nn.

Run	Genetic algorithm score	Simulation result
re-refit-MM-05-01		
re-refit-MM-06-07		
re-refit-MM-12-01		
re-refit-MM-19-05	