

## Supplementary material

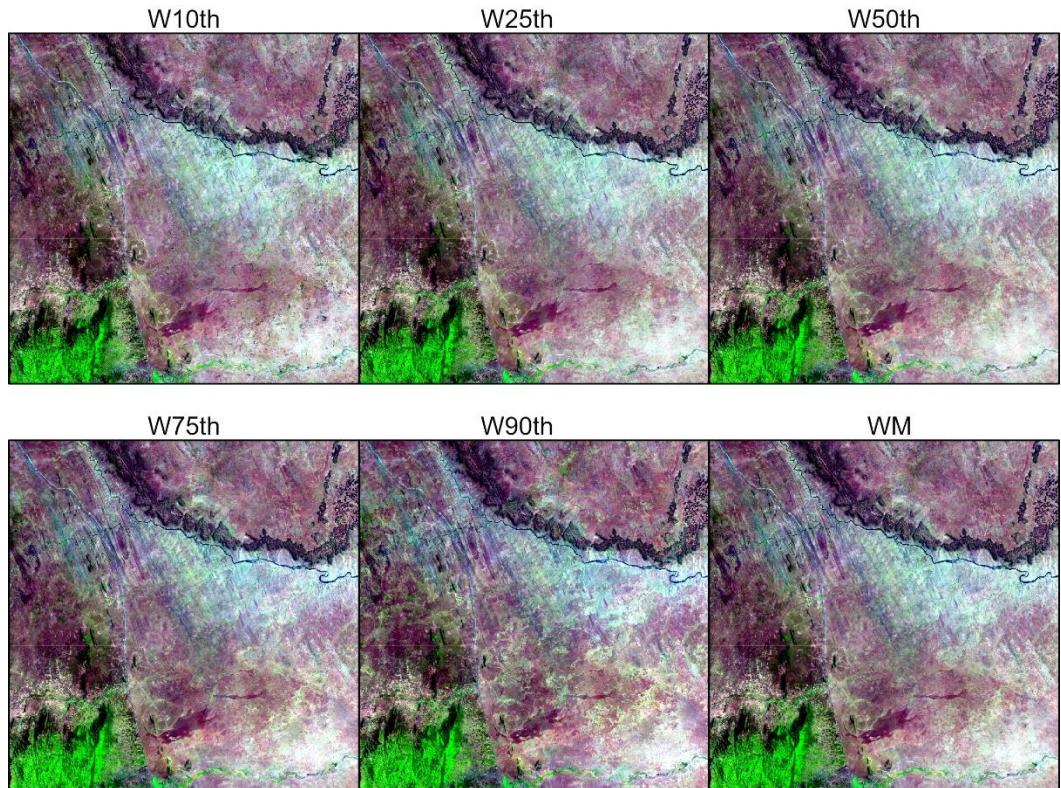


Figure S1. The weighted spectral-temporal metrics (STMs) for one target image in Site 1 acquired on March 9, 2017, displayed in false color of R: SWIR1, G: NIR, and B: red. W10th denotes weighted 10<sup>th</sup>, W25th denotes weighted 25<sup>th</sup>, W50th denotes weighted 50<sup>th</sup>, W75th denotes weighted 75<sup>th</sup>, W90th denotes weighted 90<sup>th</sup> percentiles and WM denotes weighted mean.

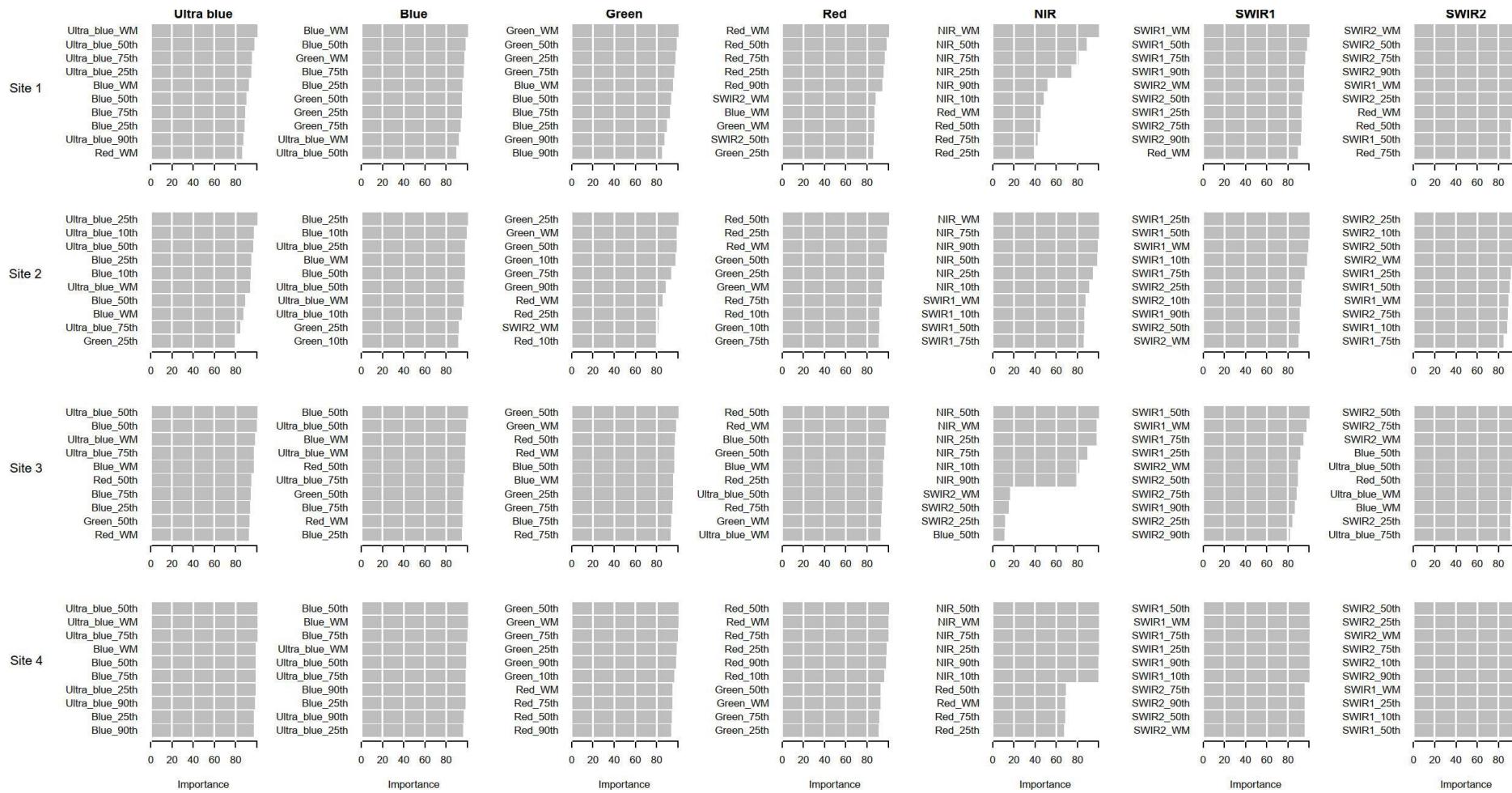


Figure S2. Ten most important variables in the  $k$ -NN regression model. There were 10,000 training and testing pixels from target images. The horizontal axis denotes the importance scores, and the vertical axis denotes the ten most important variables containing weighted mean (WM), and weighted 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup>, and 90<sup>th</sup> percentiles.

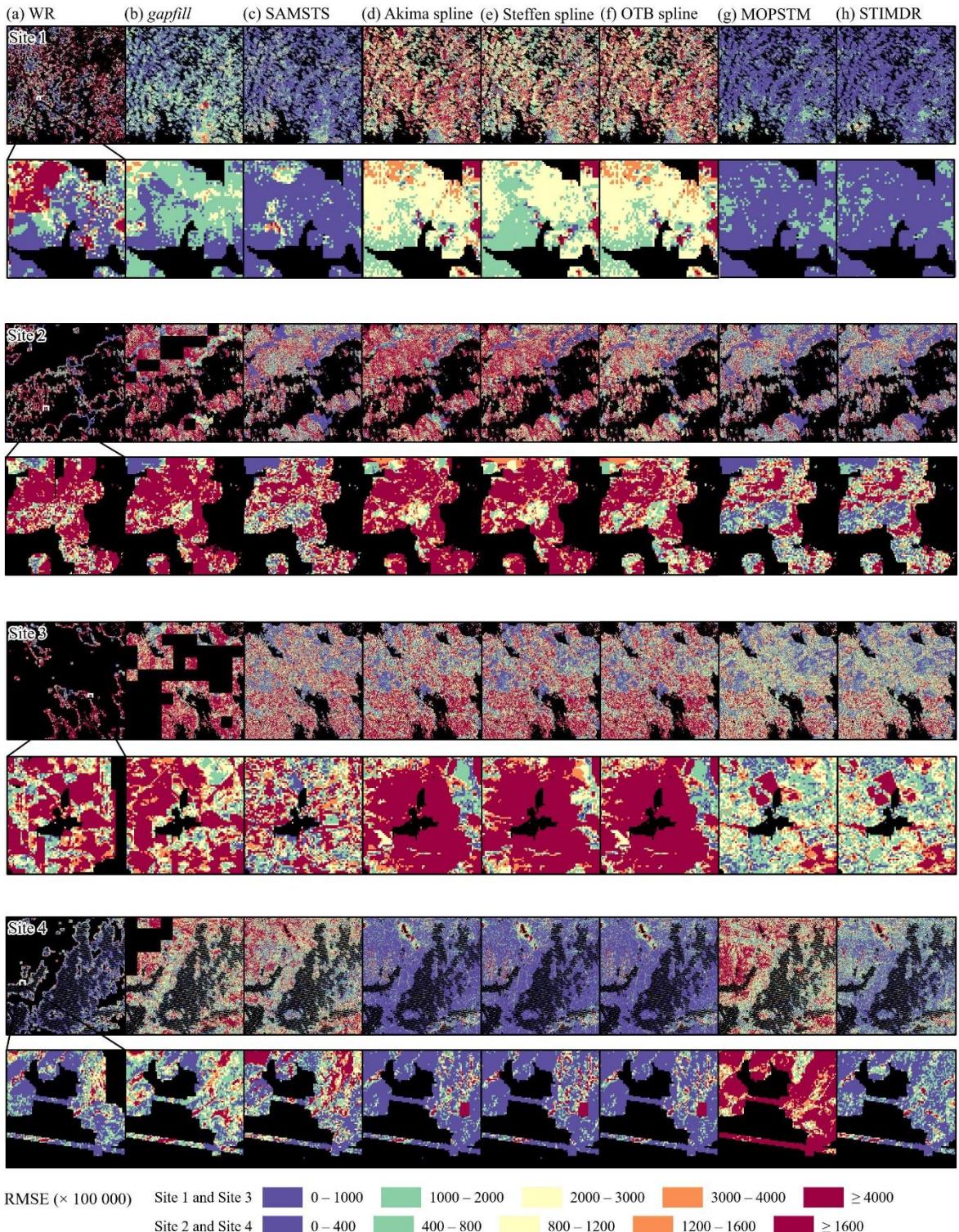


Figure S3. Pixel-based RMSE (multiplied by 100,000) of gap-filled images using different methods (corresponding to Figure 5 in the four sites, displayed in false color of R: SWIR1, G: NIR and B: red surface reflectance. (a) window regression (WR), (b) *gapfill*, (c) SAMSTS, (d) Akima spline, (e) Steffen spline, (f) OTB spline, (g) MOPSTM, and (h) STIMDR. Each image size is  $2000 \times 2000$  30 m pixels, resulting in  $60 \times 60$  km. The valid pixel percentages were 37.3% (Site 1), 39.3% (Site 2), 15.3% (Site 3), and 26.4% (Site 4). Only gap pixels that have reconstructed are displayed, and the gap pixels that failed to be constructed are displayed in the black regions.



Figure S4. Density scatter plot of window regression (WR), *gapfill*, SAMSTS, Akima spline, Steffen spline, OTB spline, MOPSTM, and STIMDR predicted values (y-axis) and validation values (x-axis) for seven spectral bands in Site 2. The black dashed line is the 1:1 line and the solid blue lines show the linear regression fits. Darker color shading: regions with a large density of points; lighter color shading: regions with a small density of points. RMSE was multiplied by 100,000.

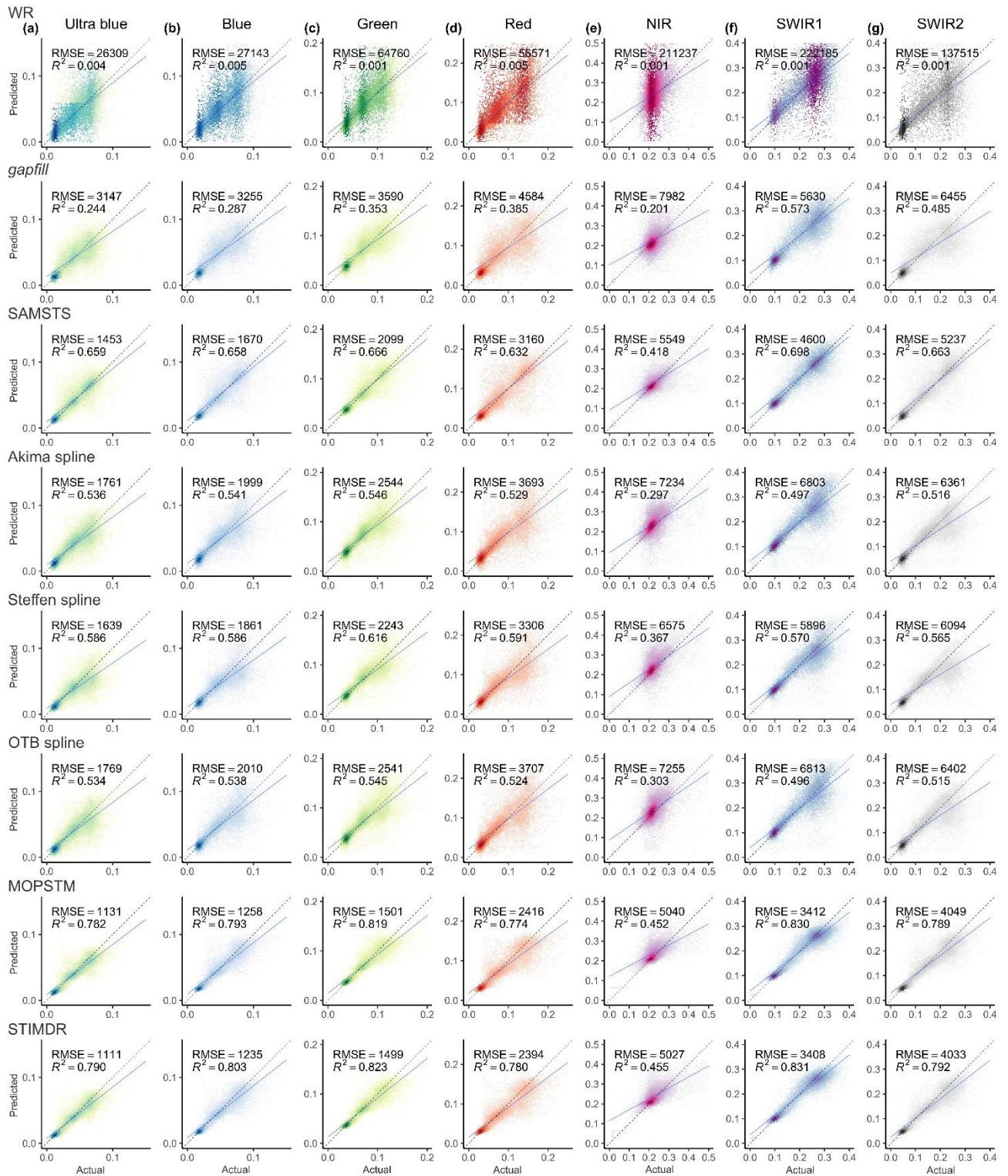


Figure S5. Density scatter plot of window regression (WR), *gapfill*, SAMSTS, Akima spline, Steffen spline, OTB spline, MOPSTM, and STIMDR predicted values (y-axis) and validation values (x-axis) for seven spectral bands in Site 3. RMSE was multiplied by 100,000.



Figure S6. Density scatter plot of window regression (WR), *gapfill*, SAMSTS, Akima spline, Steffen spline, OTB spline, MOPSTM, and STIMDR predicted values (y-axis) and validation values (x-axis) for seven spectral bands in Site 4. RMSE was multiplied by 100,000.

Table S1. The sensitivity of threshold  $M$  with respect to the root-mean-square error (RMSE), multiplied by 100,000.

M	15	16	17	18	19	20	21	22	23	24	25
Site 1	748	748	748	748	748	748	748	748	748	748	748
M	12	13	14	15	16	17	18	19	20	21	22
Site 2	1277	1276	1276	1276	1276	1276	1276	1275	1275	1275	1276
M	7	8	9	10	11	12	13	14	15	16	17
Site 3	2656	2655	2655	2655	2655	2655	2655	2656	2656	2656	2658
M	3	4	5	6	7	8	9	10	11	12	13
Site 4	1222	1204	1203	1204	1206	1207	1208	1210	1212	1213	1215

Table S2. Pairwise comparisons for Site 1 using Wilcoxon rank sum test where p-value larger than 0.05 are labeled in bold.

Site 1	WR	gapfill	SAMSTS	Akima spline	Steffen spline	OTB spline	MOPSTM	STIMDR
WR	-	-	-	-	-	-	-	-
gapfill		< 2e-16	-	-	-	-	-	-
SAMSTS		< 2e-16	< 2e-16	-	-	-	-	-
Akima spline		< 2e-16	< 2e-16	< 2e-16	-	-	-	-
Steffen spline		< 2e-16	< 2e-16	< 2e-16	< 2e-16	-	-	-
OTB spline		< 2e-16	< 2e-16	< 2e-16	<b>0.29</b>	< 2e-16	-	-
MOPSTM		< 2e-16	< 2e-16	< 2e-16	< 2e-16	< 2e-16	< 2e-16	-
STIMDR		< 2e-16	< 2e-16	4.8e-12	< 2e-16	< 2e-16	< 2e-16	-

Table S3. Pairwise comparisons for Site 2 using Wilcoxon rank sum test where p-value larger than 0.05 are labeled in bold.

Table S4. Pairwise comparisons for Site 3 using Wilcoxon rank sum test where p-value larger than 0.05 are labeled in bold.

Table S5. Pairwise comparisons for Site 4 using Wilcoxon rank sum test where p-value larger than 0.05 are labeled in bold.