MODIS tree cover validation for the circumpolar taiga-tundra transition zone

P.M. Montesano[a], R. Nelson[b], G. Sun[c], H. Margolis[d],
A. Kerber[b], and K.J. Ranson[b]

a. *Science Systems and Applications, Inc., Lanham, MD,20706 USA*
b. Biospheric Sciences Branch, NASA/Goddard Space Flight Center, Code 614.4, Greenbelt, Maryland, 20771, USA *USA*
c. *Department of Geography, University of Maryland, College Park, MD, 20742 USA*
d. *Centre d'études de la forêt, Université Laval, Québec City, Québec, Canada G1K 7P4*

Abstract

A validation of the 2005 500m MODIS vegetation continuous fields (VCF) tree cover product in the circumpolar taiga-tundra ecotone was performed using high resolution Quickbird imagery. Assessing the VCF's performance near the northern limits of the boreal forest can help quantify the accuracy of the product within this vegetation transition area. The circumpolar region was divided into 7 longitudinal zones and validation sites were selected in areas of varying tree cover where Quickbird imagery is available in Google Earth. Each site was linked to the corresponding VCF pixel and overlaid with a regular dot grid within the VCF pixel's boundary to estimate percent tree crown cover in the area. Percent tree crown cover was estimated using Quickbird imagery for 396 sites throughout the circumpolar region and related to the VCF's estimates of canopy cover for 2000-2005. Regression results of VCF inter-annual comparisons (2000-2005) and VCF-Quickbird image-interpreted estimates indicate that: (1) Pixel-level, inter-annual comparisons of VCF estimates of percent canopy cover were linearly related (mean $R^2 = 0.77$) and exhibited an average root mean square error (RMSE) of 10.1% and an average root mean square difference (RMSD) of 7.3%. (2) A comparison of image-interpreted percent tree crown cover estimates based on dot counts on Quickbird color images by two different interpreters were more variable ($R^2 = 0.73$, RMSE = 14.8%,

31 RMSD = 18.7%) than VCF inter-annual comparisons. (3) Across the circumpolar boreal

32 region, 2005 VCF-Quickbird comparisons were linearly related, with an R^2 = 0.57, a

33 RMSE = 13.4% and a RMSD = 21.3%, with a tendency to over-estimate areas of low

34 percent tree cover and anomalous VCF results in Scandinavia. The relationship of the

35 VCF estimates and ground reference indicate to potential users that the VCF's tree cover

36 values for individual pixels, particularly those below 20% tree cover, may not be precise

37 enough to monitor 500m pixel-level tree cover in the taiga-tundra transition zone.

38

39 MODIS; validation; tree cover; ecotone; transition; taiga; tundra; circumpolar

40

41 **Introduction**

42 Over the next century, estimates of global temperature increases range from 1.8°

43 C to 4.0° C, but this warming will not be distributed evenly around the globe (IPCC

44 2007). The Arctic regions are expected to warm approximately twice as fast, 4° C to 7° C

45 by 2100, and much of that warming may occur in the autumn and winter months (Serreze

46 & Francis 2006; Hassol 2004). Large changes in high-latitude forest structure and extent

47 are possible and, in fact, have already been documented. For example, in the 1980s and

48 1990s between 40° N and 70° N, increases in photosynthetic activity have been linked to

49 changing spring land surface temperatures (Goetz et al. 2005; Zhou et al. 2003;

50 Randerson et al. 1999; Myneni et al. 1997). There is modeled, remotely-sensed, and

51 field-measured evidence of changes in fire regime, insect infestations, treeline migration,

52 age class structure and species distribution in northern forests (Soja et al. 2007, Kurz et

53 al. 2007, 2008a,b). Trees may respond to a warming environment by moving northward

54 and upslope, however, southward treeline movements are possible as permafrost melts,

55   soils saturate, and bogs and wetlands replace forests (Crawford et al. 2003; Schuur et al.

56   2008). As the northern boundaries of the circumpolar boreal forests become less certain,

57   there is a need to develop remote sensing techniques to monitor changes in the transition

58   zone between the taiga and tundra, forest extent and treeline location as forests respond to

59   the changing climate and disturbance regimes (Frey et al. 2007). The taiga-tundra

60   transition zone is the largest ecotone on earth, stretching over 13,400 km.

61   The MODIS sensor, part of NASA's Earth Observation System (EOS) onboard

62   the Terra and Aqua satellites, is designed for moderate-resolution global monitoring of

63   the Earth. The Vegetation Continuous Fields (VCF) product is derived from multiple

64   temporal composites and provides a percent canopy cover value for each 500m pixel.

65   This continuous value (0-100%) is related to the amount of skylight obstructed by tree

66   canopies equal to or greater than 5 meters in height. This percent canopy cover term

67   differs from percent tree crown cover (TCC) measured for the Quickbird imagery (see

68   below), where the latter refers to the sum of the canopy cover and the within-crown

69   skylight (Hansen et al. 2003). This continuous value approach to land cover mapping can

70   be used along with more conventional land cover categorical data to more reliably map

71   global land covers. The heterogeneous nature of land cover at ecotone boundaries lends

72   itself to the continuous mapping method of the VCF and there have been other efforts to

73   map tree cover in this manner around the world (Cross et al. 1991; Zhu and Evans 1994;

74   Mayaux and Lambin 1997; Tottrup 2007; Heiskanen and Kivinen 2008).

75   We validated the accuracy of this product within the taiga-tundra transition zone

76   using validation sites distributed throughout the circumpolar boreal and taiga-tundra

77   transition zone. We addressed how VCF pixels compare inter-annually and the

78   differences in TCC interpretations from high-resolution optical Quickbird satellite

79 imagery between different interpreters to assess the variation that is inherent in human

80 estimation of TCC from high-resolution imagery. Finally, we examined comparisons of

81 VCF pixels and visual interpretations from Quickbird imagery.

82

83 **Background**

84   The changing abundance of trees along the taiga-tundra transition zone has been

85 used to monitor shifts in vegetation patterns that may signal a response to human

86 activities and/or climatic changes. Treeline changes occur where temperature changes

87 are most profound and a possible circumpolar trend has been suggested (Esper and

88 Schweingruber 2004). Satellite observations have shown that northern latitudes (Zhou et

89 al. 2003), and specifically the boreal forests (Bunn and Goetz 2006), respond in different

90 ways to changing climate. Because of the variability related to the response of the boreal

91 treeline to climate change, there is a critical need for accurate mapping of land cover and

92 monitoring of changes with consistent monitoring methods (Frey et al. 2007; Esper and

93 Schweingruber 2004). Sun et al. (2004) used a spectral un-mixing method with Landsat

94 data to characterize the transition from taiga to tundra for a forest island in northern

95 Russia. Heiskanen and Kivinen (2008) concluded that the use of multi-angular and –

96 temporal data can increase the accuracy of 1km resolution tree cover estimates of the

97 peak growing season in the taiga-tundra ecotone in northern Scandinavia.

98   With a nominal 500m spatial resolution, the VCF product incorporates multi-

99 spectral and -temporal data, capturing phenological differences, to estimate the

100 proportion of canopy cover in each pixel. The VCF resolution may also be suitable for

101 studying ecotone dynamics (Stow et al. 2004). This technique allows for heterogeneous

102 areas to be better represented than is possible by discrete land cover classes.

4

103   Heterogeneous areas, such as the taiga-tundra transition zone, lend themselves to this

104   type of depiction because groups of percent cover pixels can represent gradients across

105   space. The product is derived from a regression tree that interprets the biophysical

106   relationship between the satellite spectral signals and vegetation cover throughout a

107   season and a linear regression model that helps improve the regression tree's predicted

108   values (Hansen et al. 2002b). The VCF algorithm is kept constant for each year's data,

109   so robust MODIS data processing is a key factor in the year-to-year consistency of the

110   VCF. Furthermore, this method of statistically estimating sub-pixel tree cover at global

111   scales is necessary because individual trees cannot be resolved by the sensors required for

112   regional and global level satellite monitoring (Rees et al. 2002).

113   MODIS data products have been validated at a sub-regional scale by using

114   previously collected field data or higher resolution image data extending across the study

115   area (Pisek and Chen 2007; White et al. 2005; Hansen et al. 2002a; Hansen et al. 2003).

116   White et al. (2005) used 3954 plots from two independent datasets to test the VCF

117   correlation with ground measurements of tree cover in the southwestern United States.

118   They found an overall negative bias, whereby VCF underestimated tree cover (31% and

119   24% overall for the two datasets) and the RMSE generally increased with increasing tree

120   cover. The authors advised using the Version 1 VCF with caution in this part of the

121   world. However, as this version has been superseded with subsequent versions, with

122   consequent changes to the algorithm, the relationship of ground data to VCF values for

123   this area should be reassessed using the most currect product. Hansen et al. (2002a)

124   performed a validation of an early VCF version in Zambia using high resolution

125   IKONOS and Landsat data. After classifying the higher resolution imagery and scaling

126   up to MODIS resolutions, the RMSE was 5.2%. The two different levels of validation

5

127   success revealed by these studies indicate that the VCF's ability to reliably estimate

128   percent canopy cover may vary greatly depending on factors such as ecoregion and

129   latitude. In northern Finland, Heiskanen (2008) compared the VCF with other global

130   scale map products, finding that the VCF overestimates values in low tree cover and

131   underestimates values in high tree cover. The author also noted the difficulty in mapping

132   tree cover in the taiga-tundra transition zone and the difficulty of obtaining field data.

133   There have also been validations of MODIS land products at the pixel level. Liu

134   and Mishchenko (2008) compared measurements of aerosol optical thickness from

135   MODIS and MISR for collocated pixels. Morisette et al. (2003) discussed the use of

136   IKONOS and Landsat ETM+ for exploring variation within a MODIS pixel.

137   Salomonson et al. (2006) compared percent snow cover within individual MODIS 500m

138   pixels to the percentage of Landsat ETM+ derived snow cover for corresponding grid

139   cells. Hall et al. (2008) examined MODIS level 2 land surface temperature (LST) pixels

140   in the sinusoidal grid with LST values derived from Landsat ETM+ and ASTER of the

141   same area. The LST values from MODIS pixels were also compared with automatic

142   weather station data at high latitudes (Greenland). The authors found that single weather

143   station point observations of temperature could not be reliably compared with 1km

144   MODIS, 90m ASTER, or 57m ETM+ pixels because the point measurements don't

145   represent the variation of temperature across the area covered by each pixel. The authors

146   suggested that a local array over areas corresponding to pixels used in the comparison

147   would appropriately characterize the internal heterogeneity of surface temperature within

148   a pixel. The latter two studies used regression to relate the MODIS product values in

149   high latitudes to the ground reference data at the pixel level.

150   In place of on-site validation, some studies have acquired ground reference from

151   high-resolution imagery displayed in Google Earth (GE). Using the Quickbird imagery

152   in GE, measurements have been made of river channels and oxbow lakes (Constantine

153   and Dunne 2008) as well as irrigated areas (Thenkabail et al. 2008). Luedeling and

154   Buerkert (2008) visually interpreted GE's high-resolution data in areas randomly

155   distributed across a study area to validate a Landsat-based classification of desert oases in

156   Oman. In this study, we use GE's archive of Quickbird imagery to validate the MODIS

157   VCF product for northern boreal forests. These Quickbird images in GE cannot be

158   spectrally enhanced nor can band combinations be reconfigured. The imagery is

159   displayed in true-color and the near infrared channel cannot be accessed, which reduces

160   the information content available to image interpreters. There is, however, some basic

161   image information metadata accessible through the "DigitalGlobe Coverage" sub-layer

162   under the pre-packaged "More" layer contained in the GE "Primary Database."

163   Activating this information reveals an icon at the center point of each Quickbird image

164   that links to a web thumbnail version of the image along with the Catalog ID number,

165   acquisition date, latitude and longitude of the image center, off nadir angle, target

166   azimuth, cloud cover, and environmental quality.

167

168   **Methods**

169   *Identifying validation sites in Google Earth*

170   We selected sites throughout the circumpolar boreal forest that included a range

171   of percent TCC for which there was high resolution, true-color Quickbird imagery

172   available in GE. Only images with resolutions that allowed interpreters to identify

173   individual tree crowns were considered in the selection process and the majority of

7

174   images were acquired from May-September.   To ensure that our sites were well

175   distributed throughout the circumpolar northern boreal region, we divided Eurasia above

176   60° N into 4 regions and North America above 55° N into 3 regions and selected

177   approximately 60 validation sites for each of these seven zones (Figure 1).   The

178   validation sites extended southward to 55 ° N in North America to include the taiga-

179   tundra transition zone of eastern Canada, where biophysical characteristics in this region

180   begin to limit tree growth at this latitude.   The southern limits of study sites in both

181   Eurasia and North America were chosen based on information from various

182   interpretations of treeline, ensuring that these limits allowed us to consider tree cover

183   within the transition zone from boreal forest to tundra (Timoney et al. 1992, Callaghan et

184   al. 2002a, Olsen et al. 2001).   The seven circumpolar regions include: Scandinavia,

185   Western Eurasia, Central Eurasia, Eastern Eurasia, Alaska, Central/Western Canada, and

186   Eastern Canada. In each zone, at least 20 test sites were identified in the boreal forest in

187   each of three TCC categories: 0-20%, 20-60%, 60-100%. In a given region, we selected

188   every point based on (1) the availability of high-resolution (2.44m at nadir) Quickbird

189   imagery in GE, (2) our initial, quick-look, visual assessment/estimate of percent TCC

190   category (i.e., 0-20%, etc), and (3) on landscape/forest homogeneity surrounding the

191   location (approximates a 3x3 MODIS 500m pixel window size). This third criterion was

192   introduced to try to mitigate mis-registration errors between the GE imagery and the

193   MODIS VCF pixels. The percent canopy cover values of the corresponding local VCF

194   pixels were not considered during this site selection process.

195   Sites were located at least 15 km apart to spread the VCF evaluation pixels across

196   a given zone and to introduce spatial independence into the site selection process.   We

8

197 selected 431 sites (7 zones x 3 TCC classes x 20 locations + 11 additional unique/unusual

198 sites).

199
200 *MODIS VCF Mosaics*

201   We obtained the latest VCF data, downloaded by MODIS tile, through the

202 University of Maryland's Global Land Cover Facility

203 (ftp://ftp.glcf.umiacs.umd.edu/modis/VCF). The Collection 4 version of the data used for

204 this validation was collected for MODIS tile rows 2 and 3. Tile row 1 (above 70° N) was

205 not available at the time of the analysis. Tile mosaics in the native sinusoidal projection

206 for North America and Eurasia were created using the MODIS Reprojection Tool version

207 4.0 from the Land Processes Distributed Active Archive Center (http://lpdaac.usgs.gov).

208 Horizontal tiles 9-14 were used for North America and tiles 18-26 for Eurasia (Figure 2).

209 While each MODIS pixel is nominally 500m X 500m, they are in fact 463m on either

210 side in the original sinusoidal projection. This detail is important for generating precise

211 TCC sampling grids.

212

213 *Matching validation sites, VCF pixels and Google Earth imagery*

214   The comparison of percent TCC estimates from Quickbird imagery in GE to

215 MODIS VCF pixel values of percent canopy cover requires accurately matching the VCF

216 pixels with their corresponding spatial locations in GE to the extent possible given

217 MODIS geo-location error limits. The VCF pixel whose centroid was closest to each

218 location selected in GE was identified and the boundary of each pixel, geo-located

219 according to the four corners, was overlaid on the GE imagery. We then generated

220 regular 10 X 10 dot grids for each VCF pixel such that each dot represented the center of

221 an area that was one-tenth the length and width of the VCF pixel. The boundaries of the

222   selected pixels for each of the 7 circumpolar zones along with the regular dot grids were

223   then converted to Keyhole Markup Language (KML) files and displayed in GE. Since

224   the MODIS data was in its native sinusoidal (equal area) projection with an accuracy of

225   50m (1σ) (Wolfe et al. 2002), the VCF pixel boundaries displayed in GE and associated

226   dot grids did not maintain the square shape of a gridded cell, but preserved the areal

227   coverage, location and extent that the pixel represented (see Figure 3). We note that the

228   sinusoidal projection does not increase inaccuracy beyond the geo-location accuracy

229   stated above.

230

231   *Estimating percent tree cover*

232   Five interpreters systematically viewed a subset of the validation sites, counting

233   the number of dots that lay atop tree crowns. Prior to the image interpretation,

234   interpreters collectively reviewed test sites composed of a range of TCC densities and

235   types at various latitudes and various regions in order to standardize interpretations of

236   tree cover. The interpreters counted dots, for each assigned VCF pixel, lying atop tree

237   crowns to provide an estimate of percent TCC. The following guidelines were used to

238   determine if a dot lay atop a tree crown:

239   1. The entire dot must intercept a tree crown. As each dot represents an

240   infinitesimally small point on the earth, i.e. it has no area, increasing the

241   scale (zooming in) reveals the placement of the dot in relation to tree

242   crown pixels. Note: Increasing the scale will not alone determine whether

243   a dot intercepts a tree crown, because the context of surrounding pixels is

244   needed to identify features.

245   2. The feature beneath a dot had to be distinguished from shrubs. The two

246   central indications of tree crown presence were:

247   a. the shape of the crown, and

248   b. the presence of a shadow.

249   In cases of continuous dense tree cover, where shadows and shape did not help

250 identify individual trees, the broader spatial context and texture of the forest patch and

251 local landscape patterns helped to determine whether a dot intercepted a tree crown or a

252 shrub.

253   The VCF product measures percent canopy cover, where canopy cover is defined

254 as the component of TCC that obstructs skylight (Hansen et al. 2003). Percent TCC was

255 interpreted, rather than percent canopy cover, because our validation approach relied on

256 Quickbird imagery where canopy cover and within-crown gaps could not be

257 distinguished. From an interpreter's perspective, TCC is the metric that can be visually

258 estimated.

259   Each site was attributed a value based on an interpreter's count/estimate of tree/no

260 tree for each of the 100 dots regularly arranged within a VCF pixel's boundary. Each

261 site's value was then associated with the corresponding values of VCF pixels for years

262 2000-2005. This data set allowed us to examine the consistency of the VCF from one

263 year to the next. Furthermore, 58 of the 396 sites were systematically selected for

264 replicate image-interpretations by two of the five image-interpreters. This re-

265 examination of 15% of the sample allowed us to investigate the variability among the five

266 image interpreters. Interpretations of percent TCC were made for 424 of the 431 sites

267 because seven sites were judged, after examining the imagery, to be too difficult to

268 confidently discern trees from shrubs, i.e., the images were either too dark or too blurry

11

269   to permit accurate estimation of TCC. Quickbird – VCF regressions were based on 396

270   sites since there were no Collection 4 VCF values for 28 sites above 70° N. Sites for

271   which multiple interpretations of tree cover differed by more than 30 percentage points

272   were examined to investigate and resolve the differences in the tree cover estimates. In

273   these instances, the interpreters involved discussed the scene, and interpretations were

274   redone to better standardize image-interpreted estimates. A reduced major axis (RMA)

275   regression was used to compare a later VCF year's data with an earlier year's, and one

276   interpreter's TCC estimates to another's. The RMA regression is appropriate for

277   comparing 2 year's of VCF data and 2 interpreter's estimates of TCC because there are

278   similar amounts of error in data on both axis, i.e., in both X and Y (Curran and Hay

279   1986).

280

281   **Results**

282   *Inter-Annual Comparisons*

283   Figure 4 reports the slopes of the simple linear regression lines relating Quickbird

284   estimates to VCF estimates for each zone for 2000-2005. While there are differences

285   between longitudinal zones, most slopes (64%) are between 0.4 and 0.6 and all slopes for

286   all zones were less than 1 reflecting the varying relationship of Quickbird image-

287   interpreted estimates and VCF values at the low and high ends of the canopy cover range.

288   This relationship is examined in subsequent figures. Results for Scandinavia report

289   regressions with low slope values throughout the study time period. Visual examination

290   of the data in this region reveals what may be some systematic processing bias in that

291   VCF canopy cover values show an abrupt change in northern Scandinavia that are not

292   indicative of natural patterns or land use.

293   We compared the VCF canopy cover values with the Quickbird TCC

294   interpretations at the 396 validation sites. Figure 5 shows the distribution of each

295   Collection 4 VCF dataset from 2000-2005 with the validation site estimates using an

296   amalgam of cloud-free data (2002-2008) from the Quickbird satellite. This provides

297   information on the variability of the VCF from 2000-2005, as the appearance of each

298   frequency distribution plot helps to explain the consistency of the VCF and where there

299   may be systematic differences from ground-truth validation. To generate Figure 5, each

300   of the yearly VCF estimates and the image-interpreted estimates of percent TCC were

301   assigned to a 10% bin (e.g., 0-10%, 11-20%,... 91-100%) and a count made. The

302   majority of VCF pixels at the 396 sites report canopy cover values between 20 and 60%.

303   While the difference between iterations of the VCF is minimal in this range, they are

304   distinctly higher than the image-interpreted estimates of percent TCC. A Kolmogorov-

305   Smirnov Test for two independent samples was performed to compare the distributions of

306   the earliest and latest (2000 and 2005) VCF data and confirmed that the differences in

307   these yearly VCF distributions are in fact not significantly different (p = 0.939). The

308   distributions of the image-interpreted estimates and the 2005 VCF data were tested in the

309   same manner, confirming a significant statistical difference (p < 0.001). The frequency

310   distributions of each yearly VCF estimate suggest that the VCF tends to underestimate

311   canopy cover in areas where Quickbird reports TCC < 20% and overestimates in areas

312   where Quickbird reports TCC > 20%. In addition, this figure shows that there are no

313   VCF values above 80% canopy cover (discussed in a subsequent section).

314   We also examined VCF inter-annual variability with regression in order to

315   characterize the precision of the VCF product. Figure 6 shows the inter-annual variation

316   on a pixel-by-pixel basis for 403 circumpolar ground sites for each combination of VCF

317 years (2000-2005). This set of regressions explains the relationship of all VCF estimates

318 available to date for 403 sites. In interpreting the regressions in Figure 6, we assume that

319 from 2000-2005 actual canopy cover changes little at the each site, though certainly

320 forest disturbances, such as fire, could alter the canopy cover in this time period. Simple

321 t-tests indicate that most slopes are not significantly different from 1.0 and most

322 intercepts are not significantly different from zero (Table 1). At the pixel level, the inter-

323 annual scatter, 2004 to 2005, amounted to RMSE < 10% and RMSD = 7%, though the

324 greatest range discrepancy across all sites was 38% (61% in 2005 with 23% in 2004 and

325 47% in 2005 with 9% in 2004). The inter-annual scatter for all combinations of years

326 ranged from a RMSE = 9.5 to 10.8% and a RMSD = 6.8 to 7.8%.

327

328 *Variability of Quickbird TCC Interpretations*

329   Figure 7 presents the variation inherent among Quickbird image interpreters.

330 Different pairs of interpreters estimated %TCC on 58 of the 396 validation sites to

331 quantify the variability associated with our image-interpreted tree counts. The VCF

332 product (Figure 6) is less variable, i.e., more stable (average RMSE = 10.1%, average

333 RMSD = 7.3%) than the Quickbird image-interpreted estimates (RMSE = 14.8%, RMSD

334 = 18.7%) which serve as our ground reference data product. There were 15 estimates for

335 which at least a 30 percentage point difference in tree cover interpretation occurred. The

336 majority (12) of these involved shrub/tree confusion by one of the interpreters. These

337 estimates were re-done after interpreters re-calibrated their understanding of tree cover

338 appearance for these sites, to further standardize ground reference.

339
340 *Percent Tree Crown Cover by Circumpolar Zone*

14

341   Using the most recent (2005) version of the Collection 4 VCF, we correlated the

342   observed Quickbird percent TCC values with the percent canopy cover estimates of the

343   VCF for each of the 7 zones and for all zones combined (Figure 8). Table 2 summarizes

344   the simple linear regression results for each of the plots in Figure 8. Each zone had a

345   minimum of 45 validation sites. The scatter varies for each zone, with no apparent bias

346   towards any percent TCC interval. Some of the more inconsistent data points were

347   investigated and found in some cases to be at the boundary of an abrupt change in VCF

348   percent canopy cover or in an area where shrub cover had a tone and texture similar to

349   that of nearby stands of trees. Tests confirmed that all slopes were different than 1 and

350   all Y-intercepts were different than 0 (Table 2). The Scandinavian sites had a slope that

351   was significantly different than the slope for all zones combined ($p < 0.001$).

352   We note the following in Table 2 and Figure 8: (1) an upper VCF bound of 80%

353   in all graphs in Figure 8 (also apparent in Figure 6), (2) a consistently positive y-intercept

354   significantly different from zero in all zones, and (3) RMSE values ranging from ~10-

355   15% and RMSD values ranging from ~ 14–23% (with the exception of Scandinavia).

356   The 80% upper VCF limit noted in Figures 6 and 8 in both North America and Eurasia is

357   consistent with other reports on the VCF product (White et al. 2005; Mark Carroll,

358   personal communication). Hansen et al. (2003) explains that this 80% upper VCF limit is

359   due to the relationship between TCC and canopy cover. On average, 80% canopy cover

360   corresponds to complete crown cover. However, we might expect the canopy cover to

361   TCC ratio to vary depending on the different crown cover densities of the dominant tree

362   types across regions.

363   For low percent TCC values, the y-intercepts show a range from about 11% to

364   nearly 25%, indicating the VCF's tendency to overestimate crown cover in areas that are

365 more lightly forested. The consistent, positive y-intercepts may indicate some tendency

366 of the VCF to confuse tall shrubs as trees. Certainly the Quickbird image interpreters had

367 this same problem, and a considerable amount of time was spent by the interpreters trying

368 to discern and define the difference between shrubs and trees in some of the imagery.

369 However, where the 2005 VCF values were less than 20% canopy cover (49 sites), 73%

370 of the Quickbird estimates were also less than 20 (36 sites) and only 7 of the remaining

371 13 fell outside the RMSE range (14%). This indicates that while the VCF may err in

372 areas where trees are scarce and shrubs are abundant, the Quickbird estimates tend to

373 confirm the VCF low percent canopy cover values.

374 The entire circumpolar region had an overall RMSE and RMSD of 13.4% and

375 21.3%, respectively. We view the zonal and circumpolar RMSE and RMSD values

376 reported in Table 2 and Figure 8 as worst-case estimates given that some unknown

377 portion of this error term is due to variation associated with the image-interpretation

378 results. This 13.4% RMSE is lower than the values reported by White et al. (2005) for

379 the southwestern United States (24% and 31%) and higher than the values reported in

380 Hansen et al. (2002a) for Zambia (5.2%).

381

382 **Limitations and Uncertainties**

383

384 It is helpful to understand the spatial and temporal context in which our ground

385 reference was gathered and the VCF produced. Limitations to, and uncertainties in,

386 accurately assessing tree cover at a given location may stem from the method of

387 collecting ground reference, satellite data geo-location/gridding, and the timing of

388 satellite data for compiling both the ground reference and the VCF data product.

389   The variability in Figure 7 indicates the ease with which similar estimates will be

390   repeated and the inherent difficulty of remote satellite and visual estimation of percent

391   canopy cover and TCC. Although standardization protocols were established prior to the

392   validation effort, there were a number of cases in which the five interpreters differed in

393   their interpretation of tree cover. The replicated site interpretations of percent TCC were

394   done by five interpreters, which likely increased the level of disagreement between site

395   interpretations than if sites had been re-interpreted by the same interpreter. Thus, the

396   overall uncertainty in ground reference is not surprising. The similar image tone and

397   texture of trees and shrubs combine with different interpretations of these features by five

398   interpreters to create noisy validation data. The purpose of re-examining those replicated

399   sites that differed by at least 30 points was to minimize differences based on strong

400   interpreter bias where an errant interpretation of tree cover could dramatically effect

401   overall interpreter precision. A lower threshold would have allowed for examination of

402   smaller discrepancies and refinement of a greater number of the replicated estimates,

403   resulting in a higher level of consistency.

404   We chose this conservative threshold so as not to disguise the true noisy nature of

405   the ground reference. The level of variability in the ground reference makes it difficult to

406   pinpoint the extent to which the VCF estimates of low canopy cover, in the transition

407   zone and throughout the boreal forest, vary from true tree cover. However, the overall

408   RMSE and RMSD (13.4% and 21.3% respectively) are evidence that below 20% VCF

409   canopy cover the ability of the product to provide reliable 500m pixel-level delineation of

410   the circumpolar taiga-tundra transition zone is limited.

411   The accuracy of the MODIS VCF product may also introduce discrepancies

412   between VCF percent canopy cover values and interpreted percent TCC. MODIS' sub-

17

413   pixel geo-location accuracy (at nadir) may contribute some positional inaccuracies, as

414   will Quickbird's geo-location accuracy (+/- 23m, CE90%), and we also note that the

415   gridding of some observations could also contribute to disagreement between

416   interpretations of percent TCC and VCF values. Heiskanen and Kivenen (2008) indicate

417   that per-pixel comparison of reference data with gridded MODIS Level 1B data is

418   complicated by the fact that some gridded values represent different areas on the ground

419   than do the original satellite observations from which the gridded values were derived.

420   This is due to the MODIS bow-tie effect (Gomez-Landesa et al. 2004, Tan et al. 2006)

421   whereby observations further from nadir account for a larger land surface area than those

422   observations at nadir. Tan et al. (2006) also explain that the surface area contributing to a

423   MODIS observation is always larger than the grid cell because of the triangular point-

424   spread function. MODIS Level 3 products, such as the VCF, are multi-temporal

425   composites of the best Level 1B data available for a given location within a time interval.

426   Factors such as proximity to nadir and aerosol contamination contribute to the selection

427   of the data that will best describe ground features. Higher latitudes will have increasingly

428   more nadir and near-nadir pixels available for incorporation into a Level 3 data because

429   of the increasing degree of scan line overlap toward the poles, and more opportunities for

430   larger overlap between satellite observations and the grid cells into which they are

431   mapped (Robert Wolfe, personal communication).

432   Figure 9 shows a lake outline and one pixel's corners in northern Alaska (68.9N,

433   151.3W) overlaying both MODIS and Landsat TM from 6/14/2008 and Google Earth

434   basemap data. This example demonstrates that (1) MODIS Level 3 grid cells overlay the

435   correct earth features irrespective of the projection in which they are cast, (2) there is no

436   apparent geo-location degradation at high latitudes, and (3) features maintain their spatial

437   arrangement when displayed in Google Earth. The regular 10x10 dot grids, constrained

438   by VCF grid cells, provide a way of estimating TCC in an area with which the boundaries

439   are presumed to correspond. Each dot is used to represent an area that is 1/10 the area of

440   a MODIS pixel. The 100 regularly arranged dots within the boundary of each site's

441   corresponding pixel provide point-based estimates of areal values of TCC, while the

442   pixels themselves are geo-located approximations of the actual land surface that provided

443   the radiance values recorded by the satellite. This method, recommended by Hall et al.

444   (2008), provides an intuitive and easily replicated means of performing pixel-level

445   validation, and assessing sub-pixel variability, of MODIS Level 3 geophysical maps.

446   While our Quickbird TCC interpretations are not the definitive assessment of tree

447   cover at the 396 sites, they represent our best estimates. We did not examine how the

448   differences in acquisition time between the Quickbird and MODIS VCF imagery may

449   have increased variability between the two estimation methods. However (1) the VCF is

450   a temporal composite designed to minimize sensitivity to a specific time of year, (2) we

451   gave preference to Quickbird imagery obtained from May to September when ground

452   features aren't obscured by long shadows from low sun angles, and (3) the vast majority

453   of sites were dominated by conifers whose crowns are visible regardless of season.

454   Cases where forest disturbance occurs between the acquisitions of each data

455   source could contribute to variability of our ground reference with MODIS estimates.

456   One case involves a forest disturbance at a validation site after the 2005 MODIS data and

457   prior to a Quickbird acquisition used for ground reference. For example, a disturbance

458   after the 2005 VCF dataset but prior to the acquisition of a Quickbird image used to

459   select a site could affect forest cover that the VCF had already mapped. Depending on

460   the nature of the disturbance, the ground reference and MODIS estimates could be based

461   on significantly different forest conditions. Our ground reference sites were chosen in

462   GE with Quickbird imagery where there were relatively homogenous forest conditions

463   and no visual signs of recent disturbances. This site selection method was a screening

464   process, decreasing the likelihood of validating mixed MODIS pixels and changed pixels.

465   A second case may occur if a site selected using a 2002 Quickbird acquisition

466   experienced a subsequent tree cover altering disturbance prior to the 2005 VCF. Such

467   cases will introduce noise into the relationship of the VCF to the ground reference

468   estimates. Regarding the likelihood of a burn at one of our validation sites, Hicke et al.

469   2003 indicate that 1% of the North American boreal forest biome may burn each year.

470   SPOT-VEGETATION estimates of boreal forest fire indicate that about 1% of the

471   circumpolar boreal burned from 2000-2004 (Bartalev et al. 2007). Similarly, Potapov et

472   al. (2008) estimated that approximately 1% of the area of forest cover loss in the boreal

473   biome was due to wildfire from 2001-2005, while forest cover loss as a percent of year

474   2000 forest area was 5.63% for North America.

475

476   **Discussion**

477   This validation helps establish the limits of the VCF for mapping the taiga-tundra

478   ecotone on a global scale. The accuracy and precision of global products such as the

479   VCF must be assessed to determine their utility for detecting subtle changes to the taiga-

480   tundra ecotone.

481   A comparison of VCF-to-VCF variability and interpreter-to-interpreter variability

482   provided an unexpected result. We note the following: (1) each inter-annual VCF percent

483   crown cover RMSE (9.5 – 10.8%) and RMSD value (6.8 – 7.8%) was smaller than the

484   image-interpreters' percent TCC RMSE (14.8%) and RMSD (18.7%) and (2) each inter-

485   annual VCF R^2 value (0.74 – 0.80) was higher than the interpreter R^2 (0.73). In the

486   majority of cases, slopes were not significantly different from 1 and intercepts were not

487   significantly different from zero, however the results, specifically the RMSE and RMSD

488   values, point out the difficulty of working in short-stature, open forests interspersed with

489   shrubs of varying heights, both for image-interpreters and for analysts generating

490   machine-processed standard products.   In the ideal case of perfect, completely

491   reproducible estimates of tree canopy cover from Quickbird, then we would expect the

492   VCF-interpreter comparisons to have higher R^2 values and smaller RMSE and RMSD

493   values. Interpreters were themselves quite variable, and that variability, which has no

494   bearing on the quality of the VCF estimates, is manifested in the comparison.

495   Google Earth saved money and time by providing free high-resolution imagery

496   across tens of thousands of kilometers of inhospitable, inaccessible terrain. However, the

497   use of GE imagery comes at a cost.   Some key factors that affected the interpretation of

498   tree cover were the inability to manipulate or enhance Quickbird imagery in GE, the

499   combination of view and high-latitude sun angle and the similar appearance of trees and

500   shrubs. The Quickbird data's near-infrared channel is unavailable in GE, removing a

501   potentially valuable tool for interpreting vegetation on screen. The GE imagery also

502   lacks the ability to perform contrast enhancements. Additionally, in some instances the

503   combination of the arrangement of forest patches and tree shadows with certain sun and

504   view angles may have obscured tree cover, or led to false identification of tree cover.

505   Regional differences in tree and shrub types and appearance may have, in some cases,

506   affected tree cover estimates.

507   The VCF is a MODIS satellite map of a geophysical variable (tree canopy cover).

508   The pixels of this map have attribute and geo-location information and are intended to be

509   used along with other geographically referenced spatial data. Thus, the pixel locations

21

510   are presumed to represent those earth features with matching geo-location. It is for this

511   reason that pixel-level validation is useful, even with the variety of documented

512   uncertainties from the input satellite data, as it provides users, who will relate these pixels

513   to other spatial data, with a fundamental understanding of the accuracy of the product at

514   its finest resolution.

515

516   *Implications for Future Work: Mapping Tree Cover Changes at the Taiga-Tundra*

517   *Ecotone*

518   Callaghan et al. (2002a) reviewed the global importance of the taiga-tundra

519   boundary. They highlight the difficulties in understanding and representing the location

520   of this transition zone and in how to monitor its fluctuations. Although there are a number

521   of characteristics that aid delineation of this boundary, such as tree growth form, height,

522   age and aggregation, the distribution of trees is probably the most logical way of

523   presenting the regional transition from trees to tundra (Callaghan et al. 2002b). Rees et

524   al. (2002) discussed the importance of spatial resolution on the definition of treeline.

525   Stow et al. (2004) acknowledge the need for multi-scale mapping of vegetation

526   variability in the northern latitudes. Mapping the distribution of trees along the taiga-

527   tundra boundary in the circumpolar Arctic will be mostly at continental to global scales

528   but must also account for finer scale patterns that reveal gradual transitions. Examining

529   the heterogeneity of groups of 500m VCF pixels may help identify forest patches and

530   gaps (i.e., areas of similar tree distribution). Identifying landscape components formed

531   by groups of similar, adjacent pixels at a few scales establishes a foundation on which to

532   apply clear rules and definitions of ecotone location and extent. This process uses the

533   spatial context in which individual pixels exist (the values of surrounding pixels) to

534 reduce the data volume while advancing understanding of landscape patterns and

535 components. The performance of the VCF at the pixel level is important for

536 understanding the fundamental association of the predicted data with observed data. This

537 pixel-level approach assesses the VCF's utility for regional tree cover mapping in the

538 taiga-tundra transition zone. The patterns that these pixels reveal can help form the basis

539 of the dataset that can be observed and analyzed for change through time.

540   McMahon et al. (2004) discussed a robust basis for mapping ecological regions

541 and established a framework for both identifying and mapping ecoregions. This

542 framework involves identifying an extent, a grain/resolution, and a scale for the area

543 being examined and accounts for how the heterogeneity of conditions affects the

544 delineation of a boundary. The forest gaps and patches that form the spatial patterns of

545 the taiga-tundra ecotone represent internal heterogeneity that is difficult to capture on a

546 continental-global scale map. The continuous tree cover mapping provided by the VCF

547 product allows groups of pixels that represent patches and gaps to have attributes

548 representing internal variability that capture the gradual nature of the boundary. The

549 spatial variability, or texture, of the VCF product along with ancillary data may produce

550 maps that replicate forest cover variability in a way that discrete land cover classification

551 cannot, and may facilitate closer monitoring of subtle changes in the ecotone. In fact,

552 according to Hansen et al. (2002b), the VCF was developed because continuous classes

553 can represent the spatial heterogeneity better than fixed land cover classes, particularly in

554 land cover transition zones.

555   Given the inter-annual scatter noted between individual VCF observations, we

556 suggest that the taiga-tundra ecotone may be more reliably delineated and monitored over

557 time by considering groups of VCF observations rather than individual pixels. In this

558 scenario, any number of grouping algorithms – clustering processors, decision-tree

559 classifiers, image segmentations – could be used to identify VCF pixels with similar

560 canopy cover attributes. The mean and variance of the grouped pixels might then be used

561 to define a band of canopy cover conditions that represents the taiga-tundra transition

562 zone. Such an approach, one that groups adjacent pixels based on a defined level of

563 similarity, may mitigate spurious canopy cover error caused by VCF variability at the

564 pixel level.

565   To further improve the VCF product for work in northern regions, we provide the

566 following recommendations for subsequent iterations of the VCF products:

567   1.   The addition of a water mask, which was available in the Collection 3

568   version, would be useful particularly for northern latitudes where low

569   tree cover and water bodies can be difficult to distinguish without

570   water-coded pixels.

571   2.   Completion of the Collection 4 VCF tiles above 70° N would be

572   helpful for tracking changes at the northern edge of the taiga-tundra

573   transition zone, particularly in the lowlands just north of the Siberian

574   Trappes (Taimyr-Central Siberian tundra ecoregion) in Russia and at

575   the northern tip of Scandinavia (Scandinavian Montane Birch forest

576   and grasslands ecoregion) (UNEP/GRID-Arendal 2007).

577   3.   Continued refinement of the VCF algorithm for low percent canopy

578   cover areas.

579

580 **Conclusion**

24

581   Accurate mapping of the taiga-tundra ecotone is important for monitoring subtle

582   changes at the northern limits of the boreal forest. Documenting tree cover in this region

583   using satellite-derived maps provides a means to continually assess changes in the

584   transition between forest and tundra. The results of this study indicate the VCF's

585   potential utility for the circumpolar taiga-tundra region and limitations inherent in its use,

586   highlighted by its tendency to overestimate percent canopy cover in sparsely forested

587   areas of the circumpolar taiga-tundra transition zone. This overestimation demonstrates

588   to potential users of the VCF maps that this product may not be suitable for detailed

589   mapping and monitoring of tree cover, particularly tree cover below 20%, at the

590   product's finest (500m pixel) level of detail.

591

592 **References**
593
594 Bunn, A.G., & Goetz, S.J. (2006). Trends in satellite-observed circumpolar
595 photosynthetic activity from 1982 to 2003: The influence of seasonality, cover type, and
596 vegetation density. *Earth Interactions, 10*, 19.
597
598 Callaghan, T.V., Crawford, R.M.M., Eronen, M., Hofgaard, A., Payette, S., Rees, W.G.,
599 Skre, O., Sveinbjornsson, J., Vlassova, T.K., & Werkman, B.R. (2002a). The dynamics of
600 the tundra-taiga boundary: An overview and suggested coordinated and integrated
601 approach to research. *Ambio*, 3-5.
602
603 Callaghan, T.V., Werkman, B.R., & Crawford, R.M.M. (2002b). The tundra-taiga
604 interface and its dynamics: Concepts and Applications. *Ambio,* 6-14.
605
606 Constantine, J.A., & Dunne, T. (2008). Meander cutoff and the controls on the production
607 of oxbow lakes. *Geology, 36*, 23-26.
608
609 Crawford, R.M.M., Jeffree, C.E. & Rees, W.G. (2003). Paludification and forest retreat
610 in northern oceanic environments. *Annals of Botany, 91*, 223-226.
611 doi:10.1093/aob/mcfl85.
612
613 Cross, A.M., Settle, J.J., Drake, N.A. & Paivinen, R.T.M. (1991). Subpixel measurement
614 of tropical forest cover using AVHRR data. *International Journal of Remote Sensing,*
615 12,1119-1129.
616
617 Curran, P.J. & Hay, A.M. (1986). The importance of measurement error for certain

618 procedures in remote sensing at optical wavelengths. *Photogrammetric Engineering*
619 *and Remote Sensing, 52*, 229–241.
620
621 Esper, J., & Schweingruber, F.H. (2004). Large-scale treeline changes recorded in
622 Siberia. *Geophysical Research Letters, 31*, 1-5.
623
624 Frey, K. E., & Smith, L. C. (2007), How well do we know northern land cover?
625 Comparison of four global vegetation and wetland products with a new ground-truth
626 database for west Siberia. *Global Biogeochemical Cycles, 21*, GB1016.
627 doi:10.1029/2006GB002706.
628
629 Goetz, S.J., Bunn, A.G., Friske, G.J., & Houghton, R.A. (2005). Satellite-observed
630 photosynthetic trends across boreal North America associated with climate and fire
631 disturbance. *Proceedings of the National Academy of Sciences U S A.,* 102,13521-13525.
632
633 Gomez-Landesa, E., Rango, A., & Bleiweiss, M. (2004). An algorithm to address the
634 MODIS bowtie effect. *Canadian Journal of Remote Sensing, 30:*4, 644-650.
635
636 Hall, D.K., Box, J.E., Casey, K.A., Hook, S.J., Shuman, C.A., & Steffen, K. (2008).
637 Comparison of satellite-derived and in-situ observations of ice and snow surface
638 temperatures over Greenland. *Remote Sensing of Environment, 112,* 3739-3749.
639
640 Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Marufu, L., & Sohlberg, R. (2002a).
641 Development of a MODIS tree cover validation data set for Western Province, Zambia.
642 *Remote Sensing of Environment, 83*, 320-335.
643
644 Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Sohlberg, R., Dimiceli, C., & Carroll,
645 M. (2002b). Towards an operational MODIS continuous field of percent tree cover
646 algorithm: examples using AVHRR and MODIS data. *Remote Sensing of Environment,*
647 *83*, 303-319.
648
649 Hansen, M.C., Defries, R., Townshend, J., Carroll, M., Dimiceli, C. & Sohlberg, R.
650 (2003). Global percent tree cover at a spatial resolution of 500 meters: first results of the
651 MODIS vegetation continuous fields algorithm. *Earth Interactions, 7,* 1-15.
652
653 Hassol, S.J. (2004). Arctic Climate Impact Assessment: Impacts of a Warming Arctic.
654 Cambridge University Press, http://www.acia.uaf.edu (la=Aug. 27, 2008). 146 pgs.
655
656 Heiskanen, J. (2008). Evaluation of global land cover data sets over the taiga-tundra
657 transition zone in northernmost Finland. *International Journal of Remote Sensing, 29*:13,
658 3727-3751
659
660 Heiskanen, J., & Kivinen, S. (2008). Assessment of multispectral, -temporal and -angular
661 MODIS data for tree cover mapping in the tundra–taiga transition zone. *Remote Sensing*
662 *of Environment, 112*, 2367-2380.
663

664   IPCC. (2007). Climate Change 2007: The Physical Science Basis. Contribution of
665   Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on
666   Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Avery,
667   M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United
668   Kingdom and New York, NY, USA, 996 pp.
669
670   Kurz, W.A., Stinson, G., & Rampley, G. (2007). Could increased boreal forest ecosystem
671   productivity offset carbon losses from increased disturbances? *Philosophical*
672   *Transactions of the Royal Society B,* (doi:10.1098/rstb.2007.2198).
673
674   Kurz, W.A., Dymond, C.C., Stinson, G, Rampley, G.J., Neilson, E.T., Carroll, A.L.,
675   Ebata, T., & Safranyik, L. (2008a). Mountain pine beetle and forest carbon feedback to
676   climate change. *Nature, 452,* 987-990.
677
678   Kurz, W.A., Stinson, G., Rampley, G.J., Dymond, C.C., & Neilson, E.T. (2008b). Risk of
679   natural disturbances makes future contribution of Canada's forests to the global carbon
680   cycle highly uncertain. *Proceedings of the National Academy of Sciences U.S.A.,105,*
681   1551-1555.
682
683   Liu, L., & Mishchenko, M.I. (2008). Toward unified satellite climatology of aerosol
684   properties: Direct comparisons of advanced level 2 aerosol products. *Journal of*
685   *Quantitative Spectroscopy & Radiative Transfer, 109,* 2376-2385.
686
687   Luedeling, E., & Buerkert, A. (2008). Typology of oases in northern Oman based on
688   Landsat and SRTM imagery and geological survey data. *Remote Sensing of Environment,*
689   *112,* 1181-1195.
690
691   McMahon, G., Wiken, E.B., & Gauthier, D.A. (2004). Toward a scientifically rigorous
692   basis for developing mapped ecological regions. *Environmental Management, 34*, S111-
693   S124.
694
695   Mayaux, P., & Lambin, E.F. (1997). Tropical forest area measured from global land-
696   cover classifications: inverse calibration models based on spatial textures. *Remote*
697   *Sensing of Environment, 59*, 29-43.
698
699   Morisette, J.T., Nickeson, J.E., Davis, P., Wang, Y., Tian, Y., Woodcock, C.E.,
700   Shabanov, N., Hansen, M., Cohen, W.B., Oetter, D.R., & Kennedy, R.E. (2003). High
701   spatial resolution satellite observations for validation of MODIS land products: IKONOS
702   observations acquired under the NASA Scientific Data Purchase. *Remote Sensing of*
703   *Environment, 88,* 100-110.
704
705   Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G. & Nemani, R. R. (1997).
706   Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature, 386,*
707   698–702.
708
709   Pisek, J., & Chen, J.M. (2007). Comparison and validation of MODIS and
710   VEGETATION global LAI products over four BigFoot sites in North America. *Remote*
711   *Sensing of Environment,109,* 81-94.

712

713 Potapov, P., Hansen, M.C., Stehman, S.V., Loveland, T.R., & Pittman, K. (2008).
714 Combining MODIS and Landsat imagery to estimate and map boreal forest cover loss.
715 *Remote Sensing of Environment, 112*, 3708-3719.
716

717 Randerson, J. T., Field, C. B., Fung, I. Y. & Tans, P. P. (1999). Increases in early season
718 ecosystem uptake explain recent changes in the seasonal cycle of atmospheric CO_2 at
719 high northern latitudes. *Geophysical Research Letters, 26*, 2765–2768.
720

721 Rees, G., Brown, I., Mikkola, K., Virtanen, T., & Werkman, B. (2002). How can the
722 dynamics of the tundra-taiga boundary be remotely monitored? *Ambio*, 56-62.
723

724 Salomonson, V.V., & Appel, I. (2006). Development of the Aqua MODIS NDSI
725 fractional snow cover algorithm and validation results. *IEEE Transactions on Geoscience*
726 *and Remote Sensing, 44(7),* 1747-1756.
727

728 Schuur, E.A.G., Bockheim, J., Canadell, J.G., Euskirchen, E., Field, C.B., Goryachkin,
729 S.V., Hagemann, S., Kuhry, P., Lafleur, P.M., Lee, H., Mazhitova, G., Nelson,
730 F.E., Rinke, A., Romanovsky, V.E., Shiklomanov, N., Tarnocai, C., Venevsky, S., Vogel,
731 J.G., & Zimov, S.A. (2008). Vulnerability of permafrost carbon to climate change:
732 implications for the global carbon cycle. *Bioscience, 58*, 701-714.
733

734 Serreze, M.C., & Francis, J.A. (2006). The arctic amplification debate. *Climate Change,*
735 *76*, 241–264.
736

737 Soja, A. J., Tchebakova, N. M., French, N. H. F., Flannigan, M. D., Shugart, H. H.,
738 Stocks, B. J.,Sukhinin, A.I., Parfenova, E.I., Chapin III, F.S., & Stackhouse, Jr., P.W.
739 (2007). Climate-induced boreal forest change: Predictions versus current observations.
740 *Global and Planetary Change, 56*, 274–296.
741

742 Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Avery, K. B., Tignor, M. &
743 Miller, H. L., editors. (2007). Summary for Policymakers. *Climate Change 2007: The*
744 *Physical Science Basis. Contribution of Working Group 1 to the Fourth Assessment*
745 *Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press,
746 Cambridge, United Kingdom.
747

748 Stow, D.A., Hope, A., McGuire, D., Verbyla, D., Gamon, J., Huemmrich, F., Houston,
749 S., Racine, C., Sturm, M., Tape, K., Hinzman, L., Yoshikawa, K., Tweedie, C., Noyle,
750 B., Silapaswan, C., Douglas, D., Griffith, B., Jia, G., Epstein, H., Walker, D., Daeschner,
751 S., Petersen, A., Zhou, L.M., & Myneni, R. (2004). Remote sensing of vegetation and
752 land-cover change in Arctic Tundra Ecosystems. *Remote Sensing of Environment, 89*,
753 281-308.
754

755 Sun, G., Ranson, K.J., Kharuk, V.I., & Kovacs, K. (2004). Changes in the taiga-tundra
756 boundary observed with Landsat. *International Geoscience and Remote Sensing*
757 *Symposium Proceedings. 2004 IEEE International , 2*, 722-724.
758

759 Thenkabail, P.S., Biradar, C.M., Noojipady, P., Cai, X.L., Dheeravath, V., Li, Y.J.,
760 Velpuri, M., Gumma, M., & Pandey, S. (2007). Sub-pixel area calculation methods for
761 estimating irrigated areas. *Sensors, 7*, 2519-2538.
762
763 Timoney, K.P., Laroi, G.H., Zoltai, S.C., & Robinson, A.L. (1992). The high subarctic
764 forest-tundra of northwestern Canada: Position, width, and vegetation gradients in
765 relation to climate. *Arctic, 45*,1–9.
766
767 Tottrup, C., Rasmussen, M. S., Eklundh, L., & Jönsson, P. (2007). Mapping fractional
768 forest cover across the highlands of mainland Southeast Asia using MODIS data and
769 regression tree modeling. *International Journal of Remote Sensing, 28,* 23-46.
770
771 UNEP/GRID-Arendal, 'Ecoregions prioritised for conservation, in the Arctic (WWF
772 Global 200)', *UNEP/GRID-Arendal Maps and Graphics Library*, 2007,
773 <http://maps.grida.no/go/graphic/ecoregions-prioritised-for-conservation-in-the-arctic-
774 wwf-global-200> [Accessed 15 January 2009]
775
776 White, M.A., Shaw, J.D., & Ramsey, R.D. (2005). Accuracy assessment of the vegetation
777 continuous field tree cover product using 3954 ground plots in the south-western USA.
778 *International Journal of Remote Sensing, 26*, 2699-2704.
779
780 Wolfe, R.E., Nishihama, M., Fleig, A.J., Kuyper, J.A., Roy, D.P., Storey, J.C., & Patt,
781 F.S. (2002). Achieving sub-pixel geolocation accuracy in support of MODIS land
782 science. *Remote Sensing of Environment, 83,* 31-4.
783
784 Zhou L., Kaufmann R. K., Tian Y., Myneni R. B., & Tucker C. J. (2003). Relation
785 between interannual variations in satellite measures of northern forest greenness and
786 climate between 1982 and 1999. *Journal of Geophysical Research 108(DI),* 4004,
787 doi:10.1029/2002JD002510.
788
789 Zhu, Z., & Evans, D. L. 1994. U.S. forest types and predicted percent forest cover from
790 AVHRR data. *Photogrammetric Engineering & Remote Sensing, 60*, 525–53.
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806

807
808
809
810
811
812
813
814
815 **Figures & Tables**
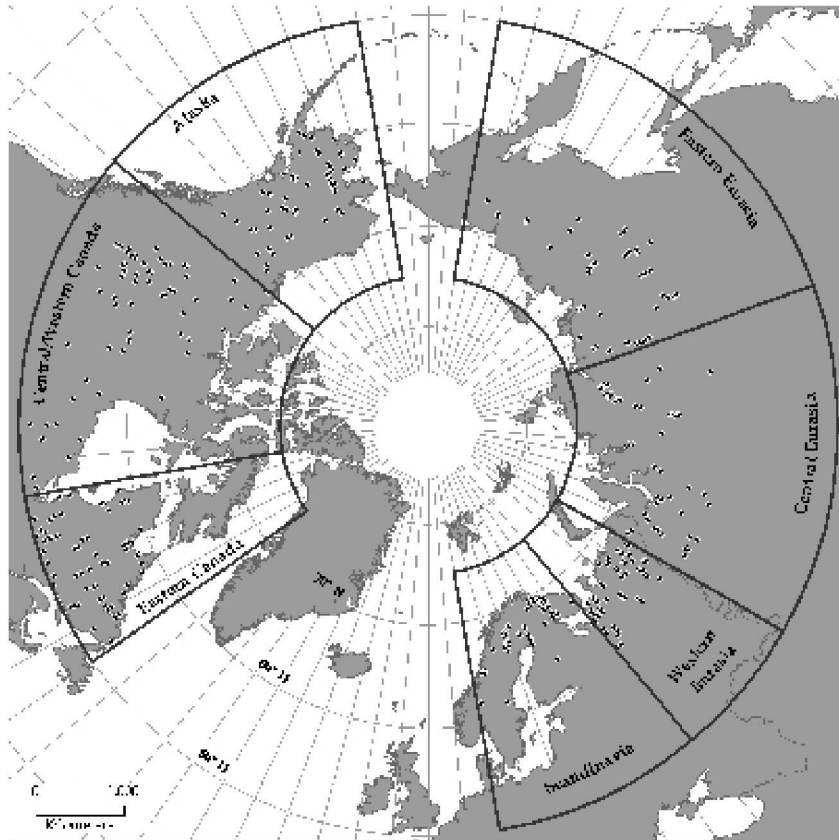
816

817 **Figure 1** Validation sites by zone in Eurasia and North America.
818
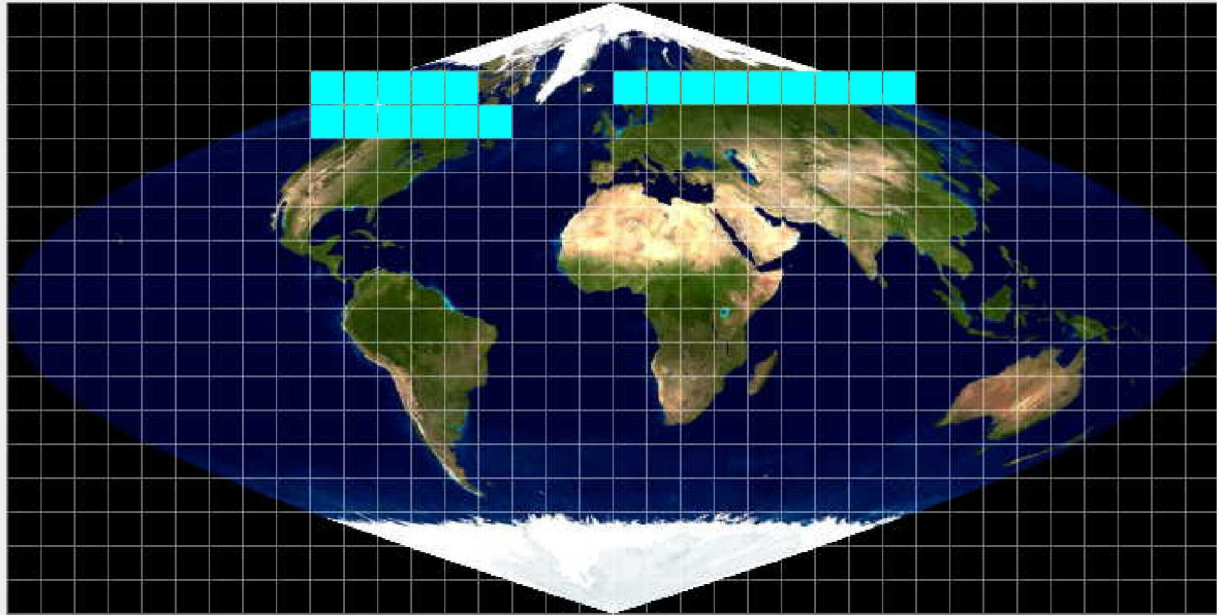
819

820 **Figure 2** Areas (in light blue) of MODIS tiles within which Quickbird imagery validation

821 sites were selected.
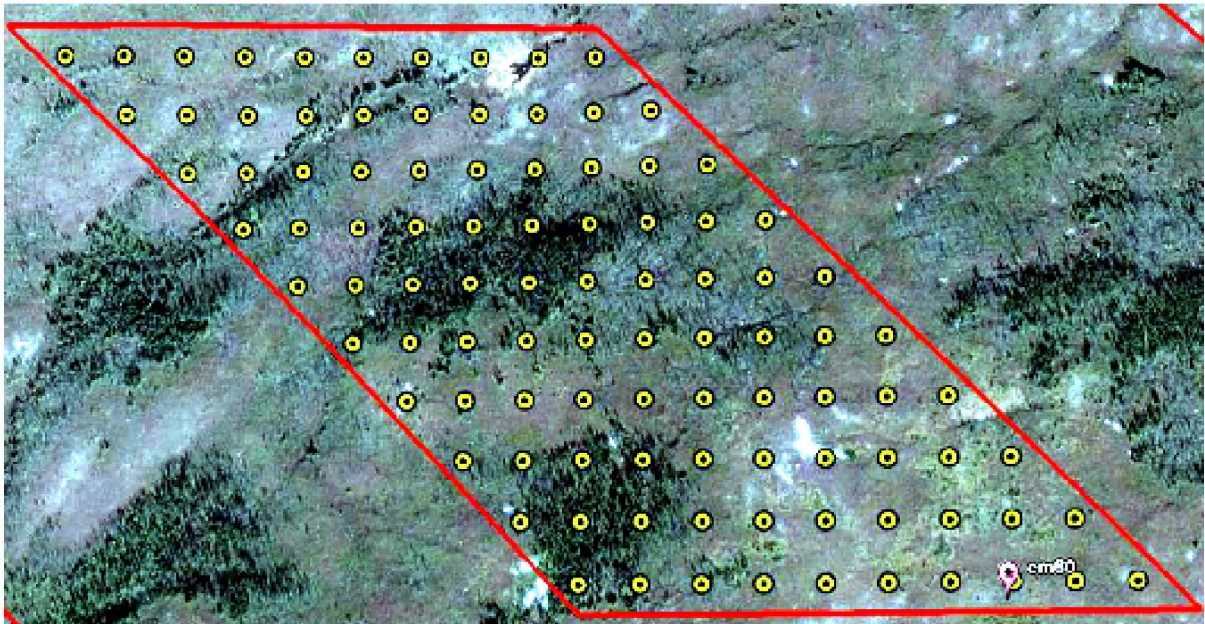
822

823 **Figure 3** An example of a VCF validation site in eastern Canada. The marker in the

824 lower right specifies the exact site selected using true-color Quickbird imagery in GE

825 with fixed enhancement. The red box represents the boundary of the corresponding VCF

826 pixel in the sinusoidal projection. The yellow dots mark the location of the 100 points

827 used to estimate %TCC. For this site, interpreters observed 23% TCC (i.e., 23 of 100

828 yellow dots fall on a tree crown) and the VCF estimated 25% canopy cover. Image

31

829 metadata was accessed through the "DigitalGlobe Coverage" layer in GE to provide

830 background information that in some cases aided interpretation of tree cover.
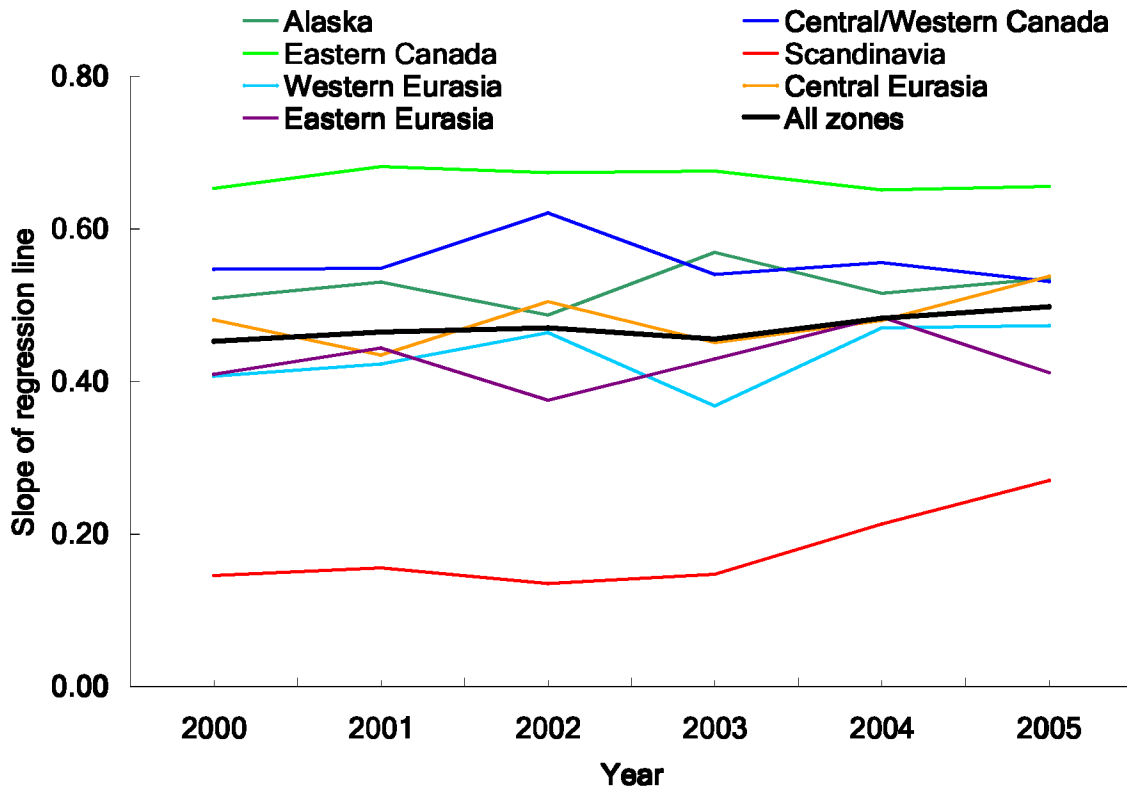
831

832 **Figure 4** Slopes of the regression line between VCF values and Quickbird estimates of

833 percent TCC for each year of the Collection 4 VCF data. The slopes for Scandinavia

834 may indicate a possible systematic bias in VCF processing for that area.

835
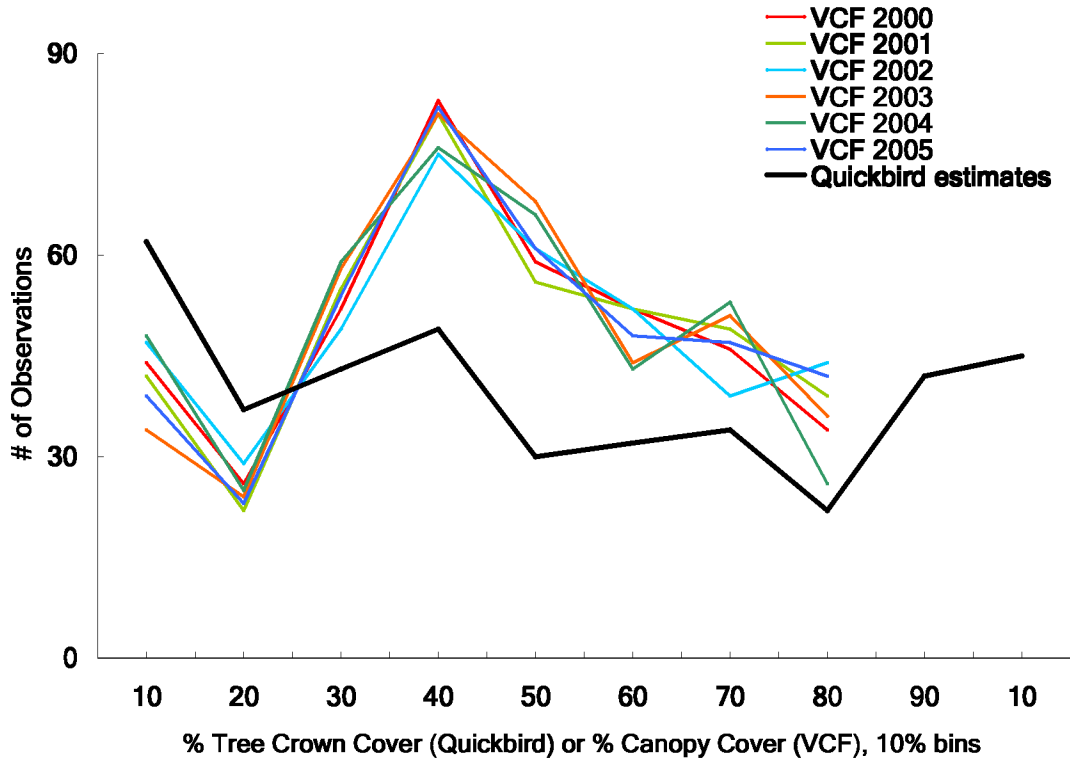
Figure 5 Frequency distribution of percent TCC from Quickbird or percent canopy cover

from the VCF for 396 circumpolar boreal sites. The numbers along the X-axis denote the

bin maximum values. The Quickbird estimates report percent tree crown cover while the

VCF maps percent tree canopy cover.

843

844

845
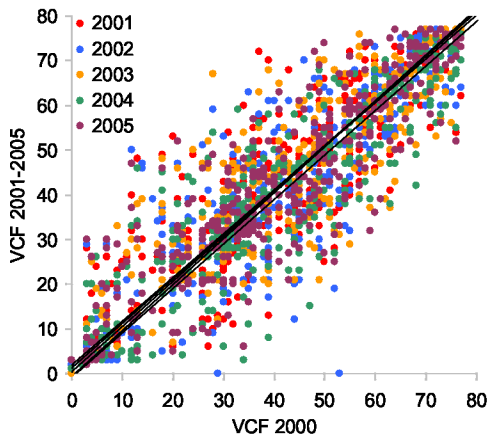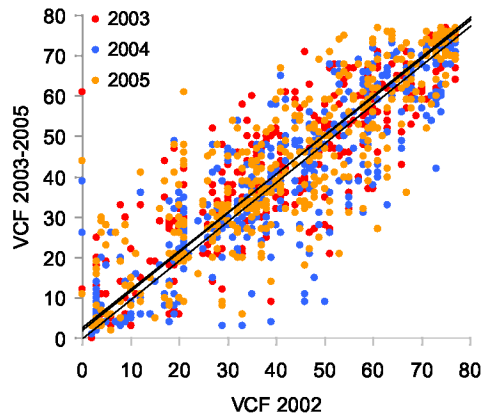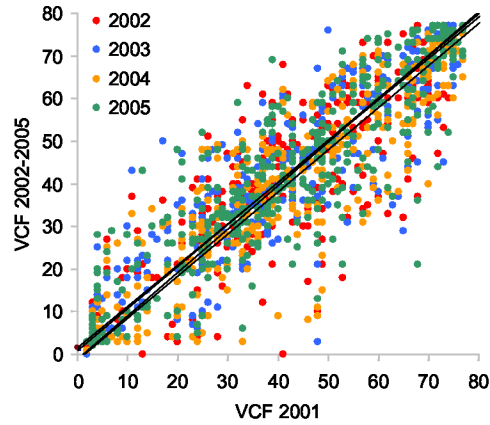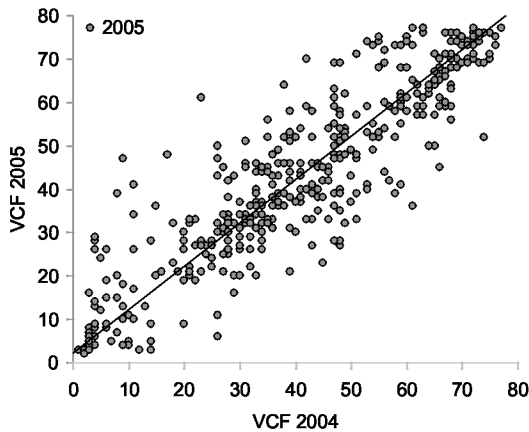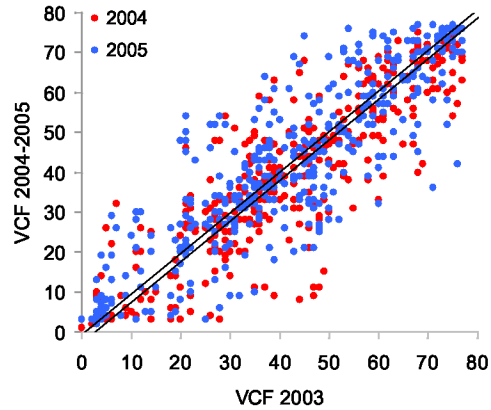
Figure 6 Plots of all VCF inter-annual combinations using RMA regression to demonstrate the variability inherent within the VCF. Each observation represents the value of each year's VCF data at a validation site (n = 403). For each plot, the earlier year is shown on the x-axis, with all subsequent years plotted on the y-axis.

850 **Table 1** Summary of VCF RMA regression results showing Y-intercept (top row), slope

851 (second row), RMSE and RMSD (third row), and R^2 values (bottom row) for all VCF

852 inter-annual combinations (2000-2005) at 403 sites. P-values for tests of Y-intercept $\neq 0$

853 and slope $\neq 1$ are shown in italics adjacent to their corresponding regression coefficients.

854 Significant differences shown in bold (p<0.05).

855

		2001		2002		2003		2004		2005	
2000	Y-intercept	0.42	*0.697*	-0.65	*0.566*	1.89	*0.082*	-0.93	*0.362*	1.27	*0.213*
	slope	1.01	*0.647*	1.03	*0.247*	0.98	*0.408*	1.00	*0.900*	1.00	*0.883*
	RMSE | RMSD	10.0	7.2	10.4	7.5	10.6	7.6	9.5	6.8	9.5	7.0
	R^2	0.78		0.77		0.76		0.80		0.80	
2001		--	--	-1.07	*0.248*	1.49	*0.159*	-1.34	*0.219*	0.86	*0.384*
		--	--	1.02	*0.454*	0.97	*0.181*	0.99	*0.566*	0.99	*0.549*
		--	--	10.2	7.3	9.9	7.0	10.2	7.3	10.2	7.3
				0.78		0.78		0.77		0.77	
2002		--	--	--	--	2.51	***0.015***	-0.30	*0.784*	1.90	*0.090*
		--	--	--	--	0.95	***0.035***	0.97	*0.192*	0.97	*0.200*
		--	--	--	--	10.0	7.1	10.4	7.5	10.8	7.8
						0.78		0.76		0.75	
2003		--	--	--	--	--	--	-2.85	***0.012***	-0.65	*0.587*
		--	--	--	--	--	--	1.02	*0.479*	1.02	*0.514*
		--	--	--	--	--	--	10.1	7.3	10.7	7.8
								0.77		0.74	
2004		--	--	--	--	--	--	--	--	2.20	***0.030***
		--	--	--	--	--	--	--	--	1.00	*0.982*
		--	--	--	--	--	--	--	--	9.6	7.0
		--	--	--	--	--	--	--	--	0.79	

856

857

858

859

The chart contains the following annotations:

- - 1 to 1 line
y = 1.03x - 1.06
$R^2 = 0.73$
n = 58
RMSE = 14.8%
RMSD = 18.7%

X-axis: Quickbird % Tree Crown Cover Estimate 2

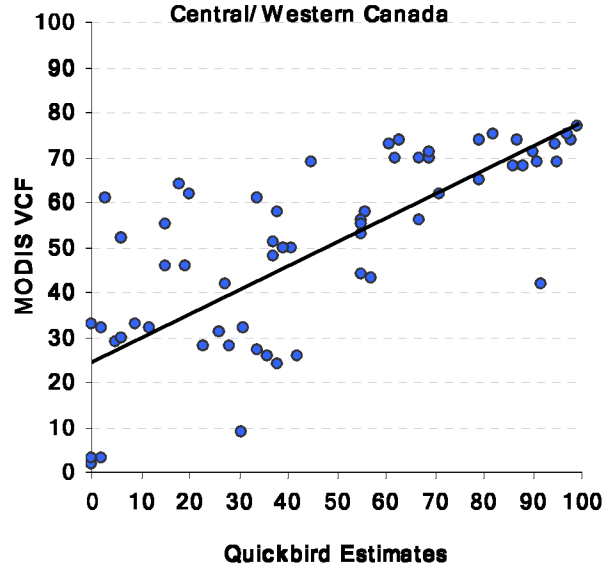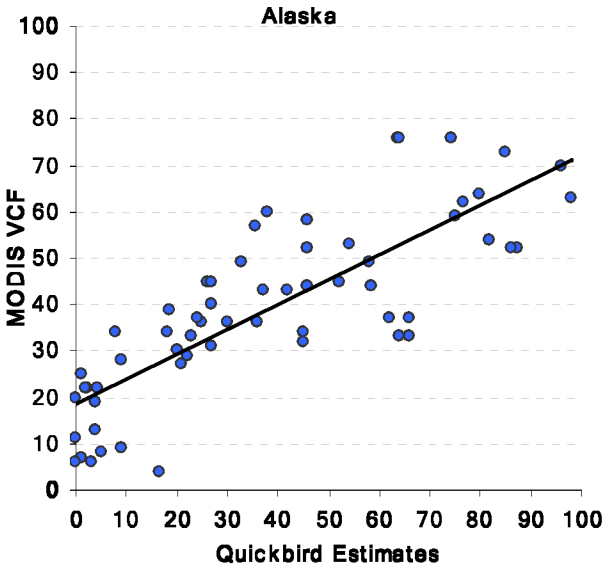Y-axis: Quickbird % Tree Crown Cover Estimate 1
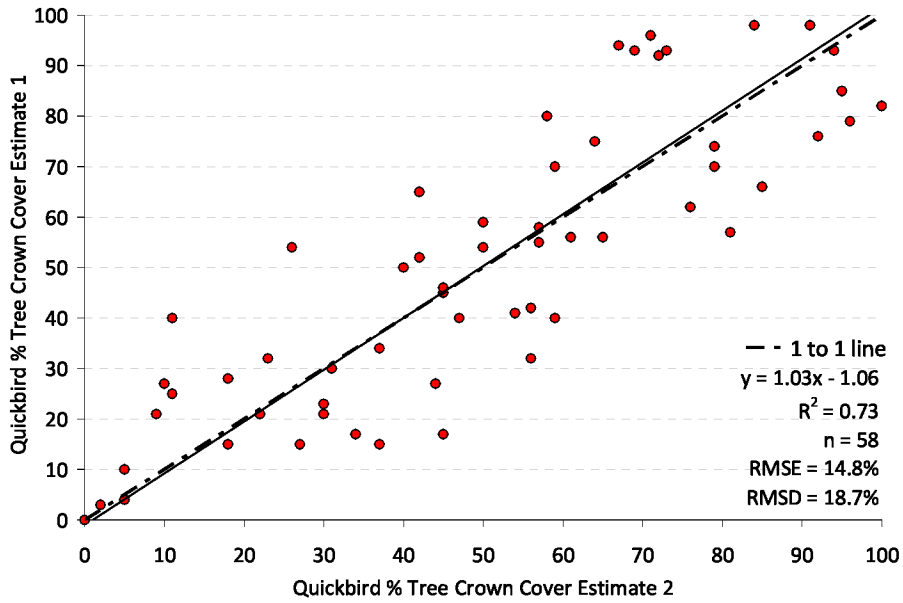
860

861 **Figure 7** RMA regression showing the relationship of replicated TCC interpretations

862 from Quickbird data by different interpreters for selected sites. The difference between

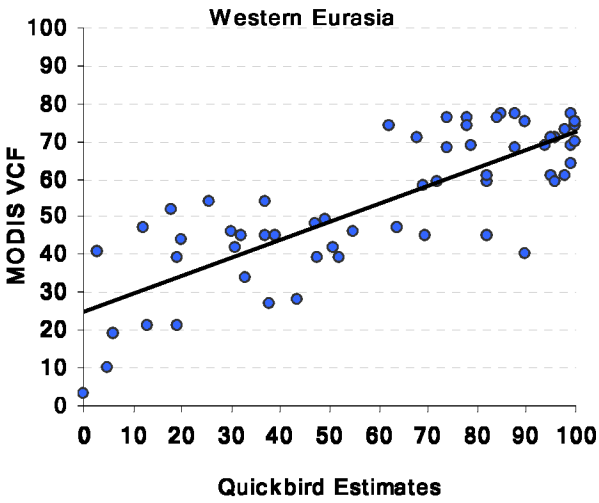863 the Y-intercept and 0, and the slope and 1 are not significant (p = 0.12 and p = 0.09).

864
865
866
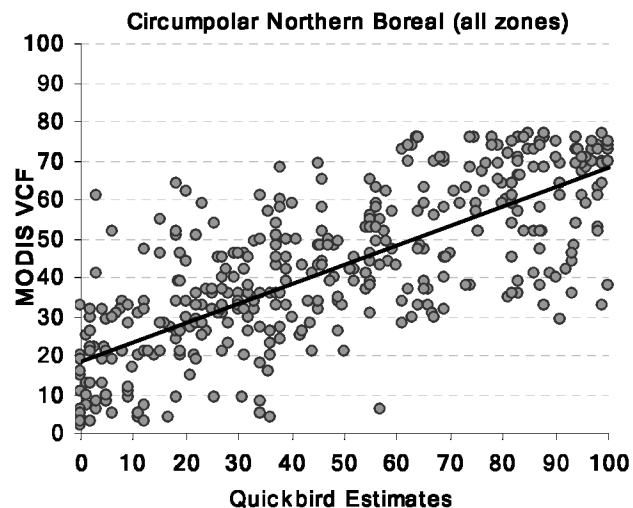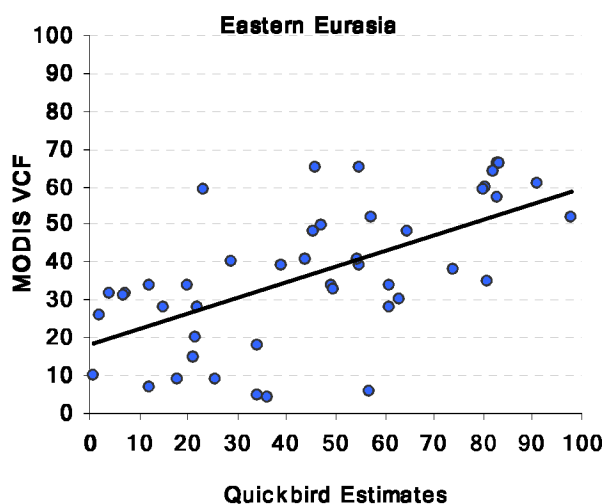
867

868

869

870
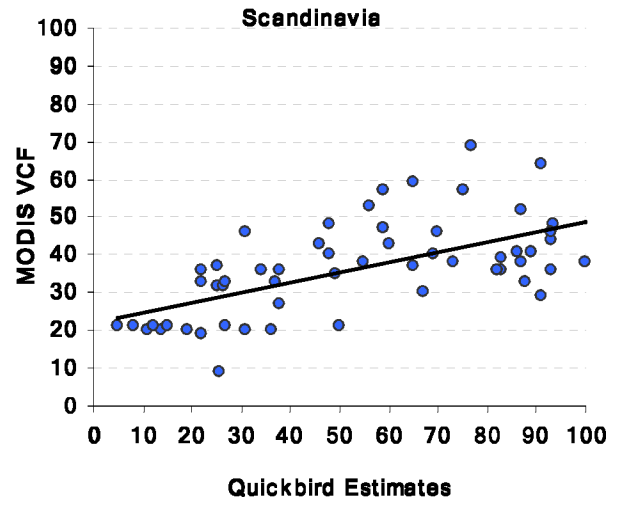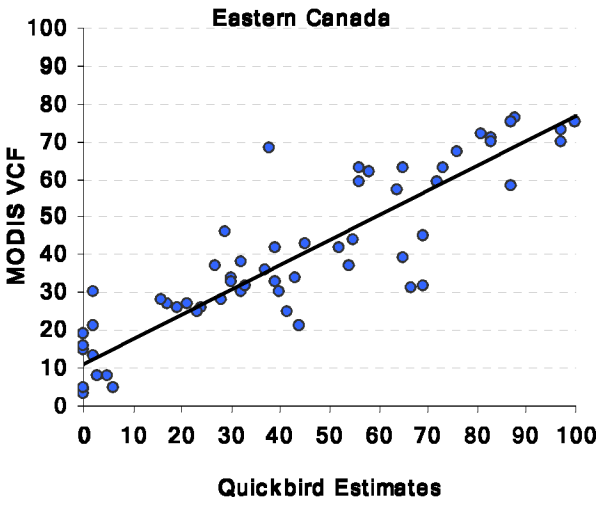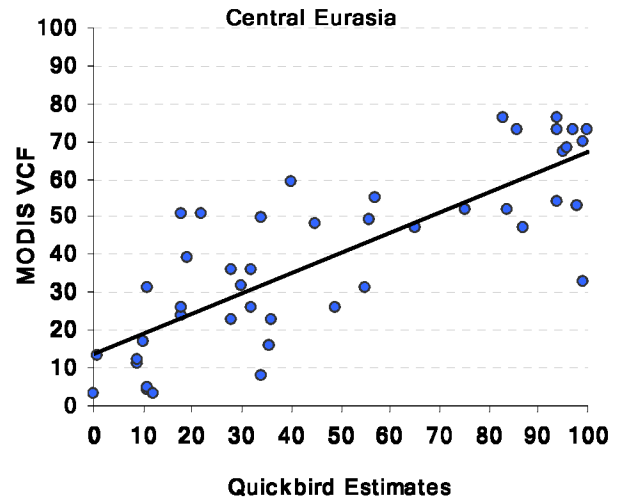
Figure 8 Simple linear regression showing the relationship of MODIS 2005 VCF percent
canopy cover values with estimated percent TCC values from Quickbird imagery.

871
872

873
874 **Table 2** Summary of simple linear regression values representing the relationship of the
875 2005 VCF with the observed % TCC values from Quickbird data for each region and all
876 combined regions.

2005 VCF	R^2	RMSE	RMSD	Slope	P-value for slope ≠ 1	Y-intercept	P-value for Y-intercept ≠ 0	n
Alaska	0.66	10.9	17.2	0.54	< 0.001	18.6	< 0.001	60
Central/Western Canada	0.57	14.5	21.1	0.53	< 0.001	24.58	< 0.001	64
Eastern Canada	0.82	9.5	14.4	0.66	< 0.001	10.96	< 0.001	64
Scandinavia	0.38	9.7	28.6	0.27	< 0.001	21.83	< 0.001	58
Western Eurasia	0.65	10.9	21.0	0.47	< 0.001	24.91	< 0.001	59
Central Eurasia	0.66	13.2	22.4	0.54	< 0.001	13.51	< 0.001	45
Eastern Eurasia	0.38	14.4	22.9	0.41	< 0.001	18.37	< 0.001	46
All zones	0.57	13.4	21.3	0.5	< 0.001	18.36	< 0.001	396

877
878
879
880

38
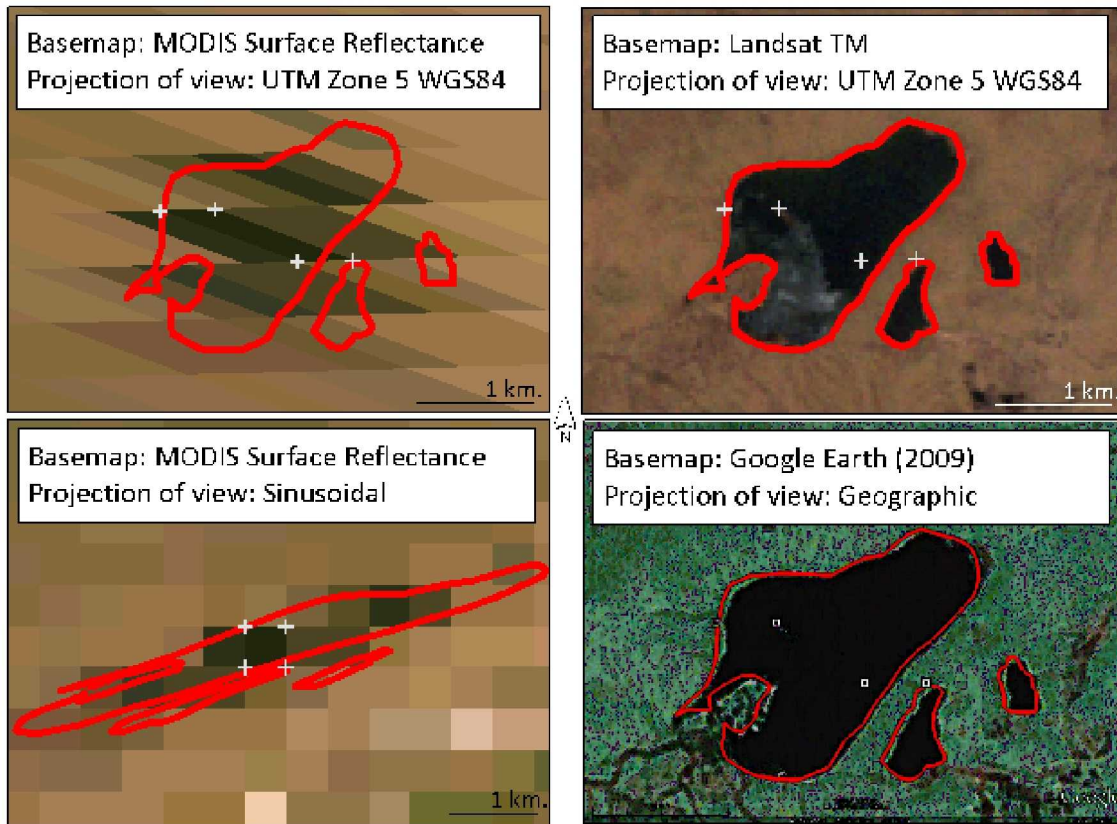
881
882 **Figure 9** An example of a feature (a lake) denoted by the red polygon, shown in MODIS
883 Level 3 data and Landsat TM data for 6/14/2008 and in Google Earth. The MODIS (in
884 the sinusoidal grid) is shown in both Landsat TM's UTM projection and MODIS's
885 sinusoidal projection. The 4 corners of the lake's center pixel are show in each image as
886 reference, to show that the geo-location of the pixel with respect to the lake polygon does
887 not change regardless of view projection.