**Supplementary Information to**

Estimating fractional green vegetation cover of Mongolian grasslands using digital camera images and MODIS satellite vegetation indices

Jaebeom Kim1, Sinkyu Kang1\*, Bumsuk Seo2, Amratuvshin Narantsetseg3 and Youngji Han1

*1 Department of Environment Science, Kangwon National University, Chuncheon 24341, Republic of Korea;*

*2 Institute of Meteorology and Climate Research Atmospheric Environmental Research, Garmisch-Partenkirchen 82467, Germany;*

*3 Institute of General and Experimental Biology, Mongolian Academy of Sciences, Ulaanbaatar 21051, Mongolia*

Corresponding author. Email: kangsk@kangwon.ac.kr

**1. Flow diagram of the study**



Figure S1. Flow diagram of the study.

**2. Color space conversion**  
Color spaces are designed to describe colors in mathematical spaces often in two dimensional (2-D) or three dimensional (3-D). In most of the digital camera images (DCI), color spaces are Red-Green-Blue (RGB). In our study, we converted the original images in sRGB space to Hue-Intensity-Saturation space and CIE L\*a\*b space.

**RGB to Hue-Intensity-Saturation (HIS)**HIS color space constitutes of Hue, Intensity, and Saturation. Hues are colors and ranging from 0˚ to 360˚. Hue of 0˚ = Red, 60˚ = Yellow, 120˚ = Green, 180˚ = Cyan, 240˚ = Blue, and 300˚ = Magenta. Intensity ∈[0,1] represents lightness of a color and Saturation ∈ [0,1] represents purity of a given hue (Chien and Tseng, 2011).

(Eq. S1)  
 (Eq. S2)  
 (Eq. S3)  
 (Eq. S4)

**RGB to L\*a\*b\* (at D65)**L\*a\*b\* color space is a uniform perceptual color space established *by Commission international de l'éclairage* (CIE) in the basis of studies on human’s color perception. It represents color stimulus with respect to an illuminant by L\* (Lightness index), a\* (chromaticness index for green and magenta), b\* (chromaticness index for yellow and blue). Negative value of a\* indicates the color is more greenish and positive more magenta. Likewise, negative b\* is for bluish colors and positive for yellowish (G. Hoffmann., 2000). To convert RGB colors into an L\*a\*b\* color space, first a RGB color is transformed into a XYZ space.

(Eq. S5)

The 3 X 3 matrix b is a tristimulus matrix of a light source. The tristimulus values are determined by the color temperature and the color space of the input (e.g., sRGB, adobe RGB). In our study (Eq. S5), we used the values for D65, which is the standard illuminant for daylight, and sRGB. The L\*, a\*, and b\* values are derived using the following equations.

(Eq. S6.)  
 (Eq. S7)  
 (Eq. S8)  
 (Eq. S9)

, where *X**n, Yn, Zn*represent the normalized white point in the XYZ color space. For D65, *Xn* = 95.047, *Yn* = 100, and *Zn* = 108.883.

**3. Auto white balance**

For taking photo with a uniform quality, digital cameras provide an “auto mode” for the shutter speed, aperture opening and white balance. The shutter speed and the aperture opening are related to the light exposure and the focal depth respectively. The white balance is related to the reflected spectrum from an object under vary illuminants. Generally, human eyes cannot well feel the changing illuminant because of the color constancy except at sunset and dawn time. However digital camera cannot avoid the effect of illuminant’s color temperature because it does not have a color constancy. Auto white balance function can solve it using adjust technique to fit daylight illuminant (i.e., D65). The equation below is a simple gain-control model of chromatic adaptation for understanding the adjust technique based on Von Kries hypothesis1) (Fairchild, 2013; Lam and Fung, 2008).

(Eq. S10)  
 (Eq. S11)  
 (Eq. S12)

, where , and represent the initial cone (spectrum peak) response; , and are the coefficients used to scale the initial cone signals; and , and are the post-adaptation cone signal. The coefficients , and are taken from inversing each initial cone response for the white scene or maximum stimulus as below.

(Eq. S13)  
 (Eq. S14)  
 (Eq. S15)

Post adaptation signals for the second condition can be calculated as follow Eqs. S16-S18, where , and represent the second cone response and , and are the second cone response for the white scene or maximum stimulus. If the second condition is like daylight condition, the initial cone response is calculated close to the second cone response under daylight.

(Eq. S16)

(Eq. S17)

(Eq. S18)

**1)** Von Kries hypothesis explains the concept of human visual systems like color recognition and adaptation. The organs of vision are completely independent and each color receptor (i.e., red cone cell, green cone cell, blue cone cell) can be fatigued or adapted to an illuminant. For example, we can recognize the original color of the object under tungsten-lamp because the color receptor of the eyes (i.e., cone cell) are fatigued or adapted to tungsten-lamp wavelength (i.e., color temperature)

**4. Comparisons of MODIS vegetation indices between at the center pixel and averaged from a 3-by-3 window**

Table S1. Mean bias (MB, %) of MODIS vegetation indices.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **NDVI** | **EVI** | **SAVI** |
| **MB (%)** | +0.14 | -0.68 | -0.84 |

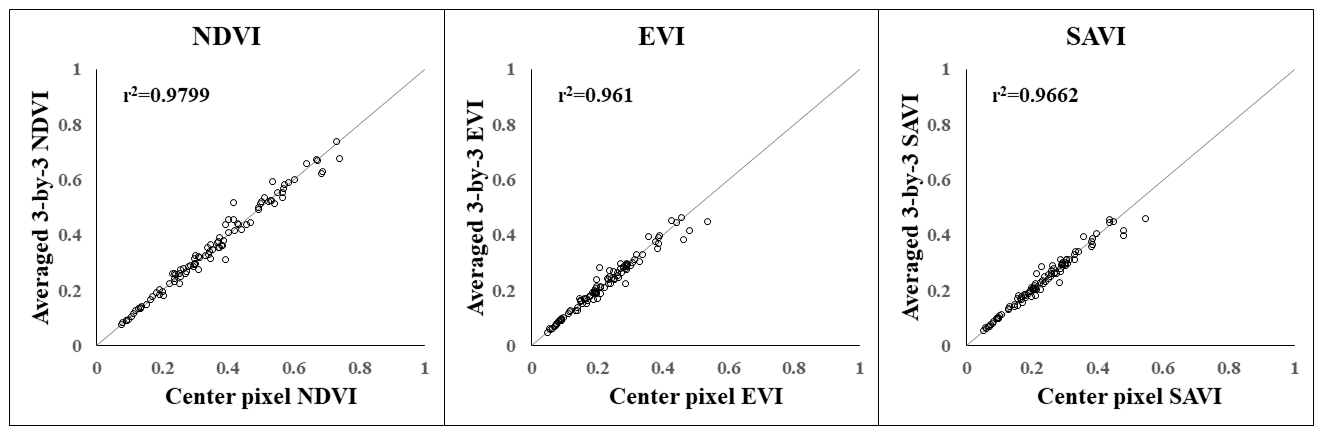


Figure S2. Scatter plot between the center pixel and the averaged 3-by-3 window.

**5. Use of Landsat NDVI to evaluate a matter of scale mismatch between sampling and MODIS data**

Cloud-free Landsat 8 NDVI images for our field survey periods were collected at 34 out of 96 sites from USGS Land Product Characterization System (LPCS, https://lpcsexplorer.cr.usgs.gov/). The NDVI images were produced with Landsat level-1 reflectance data radiometrically calibrated and orthorectified (Roy et al., 2014). For the collected NDVI images, subsets of 7-by-7, 15-by-15 and 25-by-25 pixel windows centered at each sampling site were clipped and then, the coefficient of variation (CV, %) (Figure S3) and the mean NDVI were calculated within and to the subsets, respectively. Pearson’s correlation coefficient and RMSE between the window-mean Landsat and MODIS NDVIs were calculated (Figure S4). Here, the sizes of the 7-by-7 (210 m) and 25-by-25 (750 m) pixel windows correspond to the spatial scales of the field-sampling and MODIS 3-by-3 pixel window, respectively.

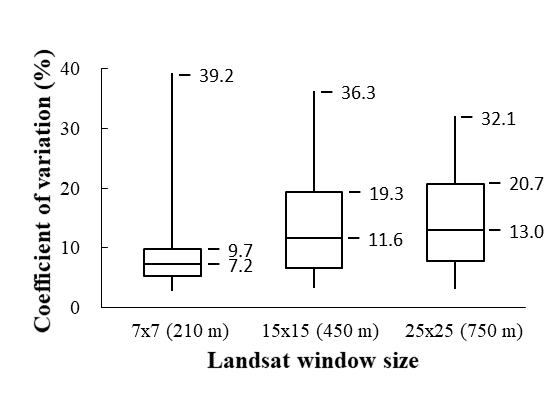


Figure S3. Interquartile boxplot of coefficient of variation (CV, %) of Landsat NDVI subsets for 7-by-7, 15-by-15 and 25-by-25 pixel windows. Values of each boxplot displayed median, 75 percentile, and maximum CVs from bottom to top, respectively.

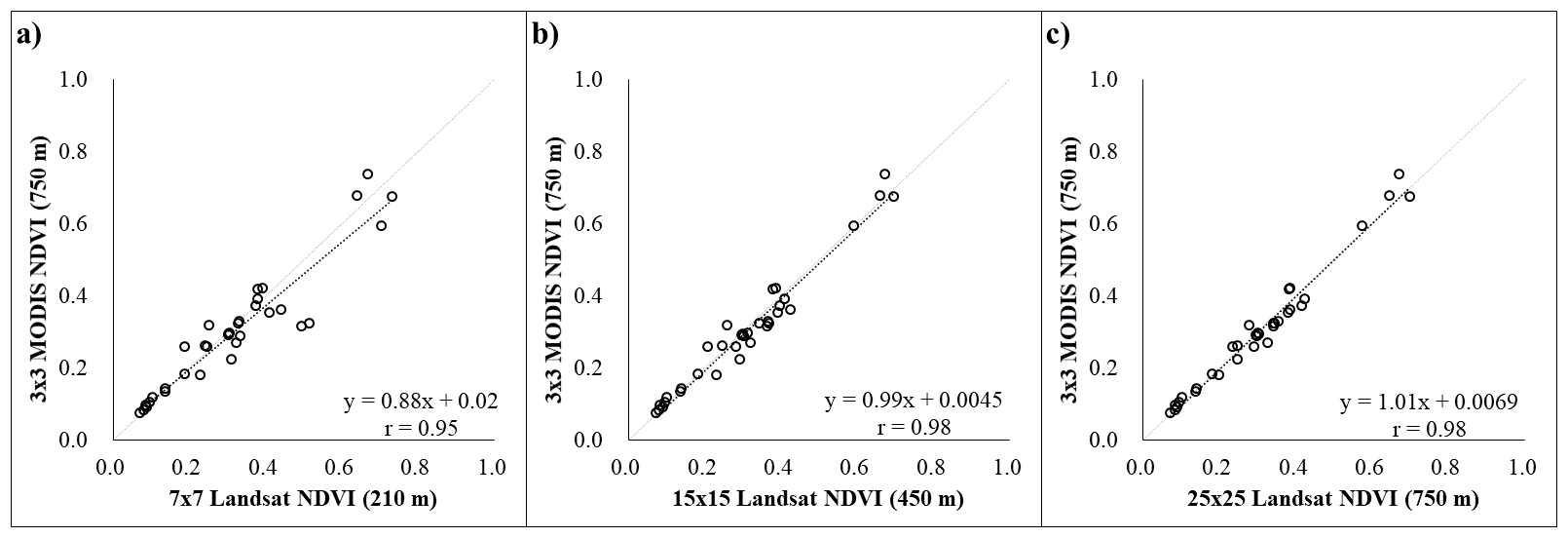


Figure S4. Scatter plots between the window-mean Landsat and MODIS NDVIs. Horizontal axis of each sub-figure means mean Landsat NDVI from (a) 7-by-7 (210 m), (b) 15-by-15 (450 m), and (c) 25-by-25 (750 m) subsets. Vertical axis is the mean MODIS NDVI from 3-by-3 pixel window (750 m). Pearson’s correlation coefficient (r) appeared on sub-figures. RMSEs between Landsat and MODIS NDVIs were evaluated as 19.2, 11.0, and 8.7%, respectively.

**References**

Hoffmann, G. (2003) CIELAB color space. online material http://docs-hoffmann.de/cielab03022003.pdf

Chien, C., D. Tseng. (2011). Color image enhancement with exact HSI color model. Innovative computing, information and control 7: 6691-6710.

Fairchild, M. D. 2013. “Color appearance models” John Wiley & Sons.

Lam, E. Y. and G. S. Fung. 2008. “Automatic white balancing in digital photography.” *Single-sensor imaging (pp. 287-314) CRC Press.*

Roy, D. P., M. A. Wulder, T. R. Loveland, C. E. Woodcock, R. G. Allen, M. C. Anderson, and D. Helder et al. 2014. "Landsat-8: Science and Product Vision for Terrestrial Global Change Research." *Remote Sensing of Environment* 145: 154-172.andour advice. o that the vegetation distribution is oriented in a certain direction so that the sample can be repre