

Modelos simples para estimar rendimiento de cultivos agrícolas a partir de imágenes satelitales: una herramienta para la planificación

Mónica Bocco¹, Silvina Sayago¹, Soraya Violini² y Enrique Willington¹

¹ Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba. Córdoba (Argentina).

² Instituto Mario Gulich - Universidad Nacional de Córdoba – CONAE (Argentina).

Resumen. En agricultura la estimación de los rendimientos de los cultivos es necesaria para la planificación que deben realizar tanto los actores públicos como privados y la producción agropecuaria es el principal uso de la tierra y un importante recurso económico en Argentina. El objetivo de este trabajo fue de desarrollar herramientas que permitan con suficiente anticipación la toma de decisiones en el ámbito de la agricultura, específicamente desarrollamos modelos a partir de imágenes satelitales para estimar rendimiento de cultivos de soja y maíz, antes de la cosecha. Utilizando modelos de regresión lineal que incluyen las bandas de los satélites Landsat y SPOT, que presentan la ventaja práctica de simplicidad y buena precisión, se obtuvieron muy buenos estadísticos de ajuste ($R^2 = 98.1$ para soja y $R^2=92.0$ para maíz).

Palabras clave: Cultivos agrícolas, Modelos predictivos, SPOT, Landsat.

1 Introducción

La estimación de la producción agrícola es necesaria para la planificación que deben realizar tanto los actores públicos como privados; desde un alerta temprana de la seguridad alimentaria y la sustentabilidad ambiental hasta en el plano científico la validación de modelos biofísicos de cultivos [1]. La información confiable, anticipada y oportuna de rendimiento de los cultivos tiene influencia en la gestión de las actividades de cosecha, almacenamiento, importación/exportación, transporte y comercialización. Definir políticas de precios, optimizar la eficiencia y la mayor rentabilidad de las inversiones, así como minimizar riesgos económicos es importante para gobiernos, productores, compañías de seguros y otros actores sociales [2,3,4,5, 6].

La agricultura extensiva es el principal uso productivo de la tierra en la zona central de Argentina y una de las mayores fuentes de ingresos económicos. Los cultivos de soja (*Glycine max* (L.) Merrill) y maíz (*Zea mays* L.) constituyen la principal actividad agrícola en la campaña de verano, alcanzando alrededor del 86% del área total sembrada en Argentina. Dentro de estos cultivos se destaca la soja, con

19.781.812 ha sembradas en la campaña 2013/2014, en particular, en la provincia de Córdoba se sembraron aproximadamente el 26% del total del país [7].

La información provista por sensores remotos es importante para estudios relacionados con el uso, la ocupación y la producción de la superficie del suelo. También permite la obtención de información de campo a un menor costo respecto a otros métodos, la cobertura de grandes áreas y la posibilidad de repetir la obtención de los datos en forma periódica a lo largo del tiempo [8,9]. Existen varios tipos de imágenes que se pueden utilizar para monitorear la superficie terrestre. En áreas agrícolas donde los lotes son pequeños, debido a su resolución espacial, los datos de LANDSAT TM y SPOT son muy adecuados; no obstante la frecuencia temporal de estos satélites y la posibilidad de nubosidad durante sus pasadas son, a veces, una limitante [10].

En épocas pasadas, las estimaciones gruesas del rendimiento de un cultivo se construyeron, en general, a partir del conocimiento experto de los agricultores o bien a partir de procedimientos de muestreo en áreas representativas, pero estos métodos son en general destructivos. En los últimos años información a priori de rendimiento se produce, generalmente, a partir de modelos, ya sea con información sobre el manejo del cultivo, datos del clima y de suelo, entre otros. Varios estudios demuestran el poder de los modelos de crecimiento de cultivos para predecir el rendimiento [11,12], y también fueron aplicados para cultivos en Argentina [13,14].

Los productos de la teledetección han permitido, en la actualidad, anticipar y estimar el rendimiento de un cultivo agrícola con mayor precisión [15,16]. Distintos modelos para estimación del rendimiento de cultivos se realizaron a partir del uso de imágenes satelitales y con técnicas estadísticas, en [17,18] se pueden encontrar estimaciones muy adecuadas del rendimiento de soja y maíz para regiones productoras en EEUU.

El objetivo de este trabajo fue de desarrollar herramientas que permitan a priori la toma de decisiones públicas y privadas en el ámbito de la agricultura. Específicamente se desarrollaron modelos simples que permiten estimar rendimiento de los cultivos de soja y maíz antes de la cosecha, a partir de imágenes satelitales.

2 Materiales y métodos

2.1 Área de aplicación y datos base para los modelos

El área de estudio es la zona central de la provincia de Córdoba (31°-32° S. y 63°-64° O.), Argentina, la cual presenta un relieve de lomas muy extendidas, casi planas. El suelo se clasifica como Haplustol éntico, con textura limosa gruesa, mineralogía mixta; las precipitaciones promedio anuales son de 800 mm, aproximadamente, y concentradas principalmente en verano. El clima corresponde al tipo seco sub-húmedo. En esta área (Figura 1) la producción agrícola es principalmente en secano y los dos cultivos predominantes, en verano, son soja y maíz [9].



Fig. 1. Área de estudio. Imagen SPOT (22/03/2014)

A fin de evaluar los modelos se relevaron 26 lotes en los cuales se sembraron ambos cultivos en la campaña 2013/2014, de los cuales correspondieron 14 lotes con soja y los restantes 12 lotes con maíz.

La verdad de terreno, para la construcción y validación de los modelos, se obtuvo registrando el rendimiento promedio para cada lote, a partir de:

- 1- Datos recolectados e informados por los productores dueños de los establecimientos.
- 2- En cada uno de los lotes y en el momento anterior a la entrada de la cosechadora a campo (Figura 2), se realizaron cortes y recolección de plantas de 3 muestras tomadas al azar con un aro de $0,25 \text{ m}^2$ de superficie. Se secaron las muestras en estufa hasta peso constante y se realizó la trilla manual. Para calcular el rendimiento promedio de cada lote se ajustó el peso de las muestras obtenidas a un 13% de humedad. El resultado final se expresó en quintales por hectárea (qq.ha^{-1}).

Los valores obtenidos de rendimiento para los lotes en estudio, se encontraron en un rango entre 28 qq.ha^{-1} y 49 qq.ha^{-1} para la soja. Para el cultivo de maíz el rendimiento medido varió entre 80 qq.ha^{-1} a 105 qq.ha^{-1} , excepto para dos lotes que reportaron muy bajo rendimiento 45 y 67 qq.ha^{-1} ; de todas maneras se decidió la inclusión de éstos para la construcción y validación de los modelos ya que permitían ver la robustez de los mismos en presencia de una variabilidad importante. Debemos notar que existieron en la muestra varios lotes con igual rendimiento, por ejemplo para soja se registraron varios pares de lotes con igual rendimiento (32 , 35 , 37 y 38 qq.ha^{-1} respectivamente); en el caso de maíz ocurrió igual situación, por ejemplo 2 lotes registraron de 87 qq.ha^{-1} y otros dos 105 qq.ha^{-1} , respectivamente.

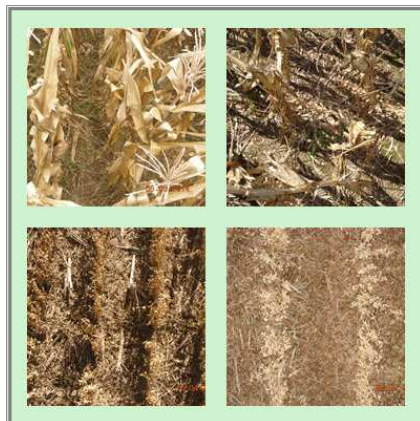


Fig. 2. Fotografías de lotes de maíz y soja previas a la cosecha.

2.2 Datos satelitales

Para el presente estudio se utilizaron dos imágenes satelitales provenientes del satélite Landsat 8, de fechas 28/12/2013 (I1) y 13/01/2014 (I2) y dos imágenes del satélite SPOT 5, correspondientes a 12/03/2014 (I3) y 22/03/2014 (I4). No se contó con imágenes del mes de febrero debido a que los días de pasada de los satélites estuvieron nublados.

En cada imagen se hizo un recorte del área de estudio y los valores de contadores digitales de las bandas fueron transformados a valores de reflectancia en el límite superior de la atmósfera y posteriormente a reflectancias a nivel de superficie mediante el uso del software ENVI 4.6.



Fig. 3. Fotografías de un lote del presente estudio donde se observan las diferencias entre borde e interior del mismo.

Los datos de reflectancia para todas las bandas de ambos satélites y cada lote se extrajeron ubicando sobre cada imagen una región de interés (ROI), las mismas se trazaron en la parte central de cada lote, abarcando la mayor superficie posible sin llegar a considerar los bordes, ya que en éstos se presentan variaciones en el estado de desarrollo de los cultivos y/o píxeles perturbados con cobertura aleadaña (Figura 3). Los tamaños de las ROI variaron entre 498 y 4977 píxeles en las imágenes SPOT y se correspondieron con análogos tamaños para las imágenes Landsat.

2.3 Modelos propuestos

Se propusieron modelos de regresión lineal del tipo:

$$\text{Rendimiento} = \sum_{i=1}^k a_i v_i + b \quad (1)$$

donde las variables regresoras son v_i = banda i del satélite SPOT (XS1 a XS3 y SWIR) y/o Landsat (B1 a B7) y las constantes del modelo a_i y b .

Los modelos fueron ejecutados con el software Infostat [19]. En la Tabla 1 se resume para cada modelo las variables utilizadas. Los mismos se diseñaron teniendo en cuenta la utilización de la máxima cantidad de variables e imágenes o bien contemplando imágenes de un único satélite.

Tabla 1. Modelos de Regresión: variables y satélites utilizados en su construcción.

Modelos	Cultivo	Imágenes satelitales	Variables regresoras (bandas)
M1	Soja	I1 a I4	B1 a B7 y XS1 a SWIR
M2			B1 a B5 y XS1 a XS3
M3		B4, B5, XS2 y XS3	
M4		I1 e I2	B1 a B7
M5		I3 e I4	XS1 a SWIR
M6	Maíz	I1 a I4	B1 a B7 y XS1 a SWIR
M7			B1 a B5 y XS1 a XS3
M8		B4, B5, XS2 y XS3	
M9		I1 e I2	B1 a B7
M10		I3 e I4	XS1 a SWIR

El estadístico usado para la evaluación y validación de los modelos fue el coeficiente de determinación (R^2) calculado a partir de los valores de rendimiento relevado para cada tipo de cultivo en el lote y estimado por el modelo.

3 Resultados y discusión

Se presentan los valores registrados y estimados y el coeficiente de determinación para los modelos¹ realizados utilizando las bandas de los satélites SPOT y Landsat, en la Tabla 2 para el cultivo de soja y en la Tabla 3 para maíz.

Tabla 2. Rendimiento registrado y estimado para soja por los distintos modelos y coeficiente de ajuste de cada modelo.

Rendimiento registrado en lote a campo	Rendimiento estimado por los modelos				
	M1	M2	M3	M4	M5
28	29,8	29,6	31,3	30,1	31,3
30	32,7	33,1	33,1	33,0	32,7
31	32,2	31,5	32,2	36,8	32,1
32	31,7	32,3	32,5	32,3	31,3
35	34,3	33,9	33,8	33,9	34,1
37	37,3	37,4	37,4	36,6	37,5
38	38,1	38,0	37,8	38,4	39,4
47	45,1	44,8	44,2	42,8	42,1
49	45,9	45,7	43,9	41,7	43,3
R²	98,1	97,5	96,4	83,8	92,4

Tabla 3. Rendimiento registrado y estimado para maíz por los distintos modelos y coeficiente de ajuste de cada modelo.

Rendimiento registrado en lote a campo	Rendimiento estimado por los modelos				
	M6	M7	M8	M9	M10
45	48,3	50,6	52,0	53,0	52,1
67	67,5	67,9	68,3	67,5	67,6
80	89,5	88,6	89,0	90,3	92,3
83	88,7	90,3	90,6	90,1	85,7
87	86,8	88,3	90,5	90,8	83,7
95	96,1	96,6	96,8	94,4	98,1
96	88,9	87,8	84,3	90,5	87,0
99	95,9	88,8	86,3	97,3	94,4
100	92,6	90,0	90,2	89,7	94,0
105	101,4	100,6	99,3	99,4	99,2
R²	92,0	86,5	80,6	88,6	88,1

Como puede observarse si se dispone de imágenes de varios satélites se logran los mejores resultados para la estimación (modelos M1 y M6 para soja y maíz respectivamente).

¹ Los coeficientes de las regresiones, para todos los modelos, obtenidos con el software Infostat resultaron significativos ($p < 0,01$)

Ahora bien, si sólo se dispone de imágenes de un solo satélite, puede verse que a través de estos modelos simples el SPOT proporciona un muy buen ajuste para estimar rendimiento de la soja (modelo M5), no obstante el coeficiente R^2 obtenido a partir de Landsat ($R^2= 83,8$) es también un resultado muy bueno. Para el cultivo de maíz, ambos satélites, usados por separado, permiten obtener un ajuste que supera el 88% (M9 y M10).

Los modelos realizados presentan resultados comparables a los obtenidos por otros autores. Por ejemplo para estimar rendimiento de soja utilizando imágenes del sensor MODIS, Li *et al.* [18] obtuvieron para las regresiones realizadas valores de R^2 entre 31,0 y 92,0 y cuando los aplicaron a lotes con maíz este estadístico varió entre 71,0 y 90,0.

Para soja en el estado de Rio Grande do Sul (Brasil) y a partir de rendimiento registrado por el Instituto Brasileiro de Geografia e Estatística (IBGE), Gusso *et al.* [20] obtuvieron un índice de ajuste $R^2 = 82,0$ al estimar para un período de diez años el rendimiento utilizando el índice EVI (Enhanced Vegetation Index) de imágenes MODIS.

Para maíz, Prasad *et al.* [17] en el estado de Iowa (EEUU) utilizando un modelo de regresión lineal a partir de NDVI, temperatura de superficie, precipitación y humedad de suelo, estimaron rendimiento de maíz obteniendo un coeficiente de determinación de 78,0.

Para la misma provincia que este estudio pero a escala departamental, De la Casa y Ovando [21] encontraron que el NDVI (Índice de Vegetación Normalizado), obtenido a partir de imágenes del satélite NOA AVHRR, permite inferir rendimientos con valores de R^2 entre 69,0 y 84,0, según la fecha de siembra y toma de imágenes.

En la Figura 4 se puede observar, para soja, la relación entre los rendimientos reales a campo y los estimados por los modelos construidos a partir de todas las imágenes (M1) y las provenientes de un único satélite (M4 y M5).

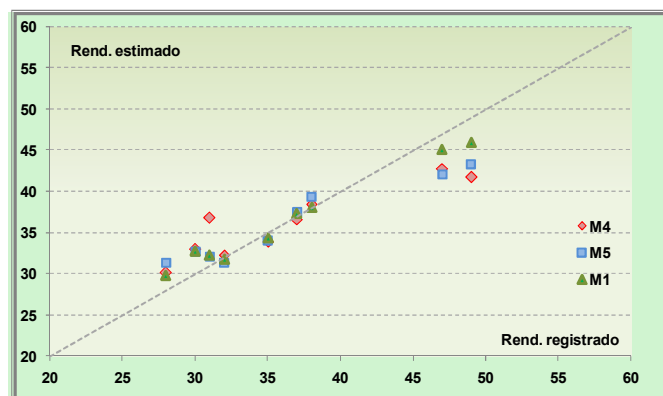


Fig. 4. Relación entre rendimiento medido y estimado para lotes de soja y modelos con mayores estadísticos resultantes.

Como se observa en la Fig. 4 cuando los rendimientos son superiores a 45 qq.ha⁻¹ los modelos tienden a subestimar los valores, mientras que para lotes con bajos rendimientos (menores a 32 qq.ha⁻¹) todos los modelos sobreestiman el resultado.

Para el cultivo de maíz se presenta en la Fig 5 el scatter-plot que da cuenta de la relación entre rendimientos medidos y estimados, para los mejores tres modelos, que coinciden con utilizar imágenes de todos los satélites (M6) o de cada uno separadamente (M9 y M10). En dicha figura se puede observar que, aún para los lotes con rendimiento muy bajo para este cultivo todos los modelos estiman razonablemente el mismo, aunque siempre con un grado de sobreestimación, en cambio para los rendimientos más altos subestiman.

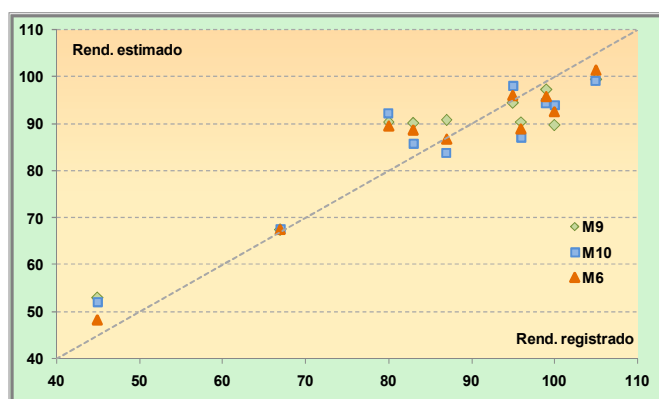


Fig. 5. Relación entre rendimiento medido y estimado para lotes de maíz y modelos con mayores estadísticos resultantes.

Los resultados de los modelos para predecir rendimiento en lotes con maíz, para la zona de trabajo, seguramente podrían mejorarse ya que no se pudieron obtener, para el año 2014, imágenes satelitales en el mes de febrero (todos los días de pasada del satélite coincidieron con nubes o lluvia). Según lo expresan Mkhabela *et al.* [22] el mejor período para estimar el rendimiento es la etapa de floración y llenado de grano, pues dichas etapas comprenden la fase más crítica para la producción.

4 Conclusiones

Modelos matemáticos simples permiten describir adecuadamente el rendimiento a obtener en cultivos de soja y maíz a partir de imágenes satelitales obtenidas dos o tres meses anteriores a la cosecha.

Utilizando modelos lineales que incluyen todas las bandas de los satélites Landsat y SPOT en conjunto o bien con las correspondientes a uno de éstos, que presentan la ventaja práctica de simplicidad y buena precisión, se obtienen muy buenos

estadísticos de ajuste para estimar rendimiento de los dos principales cultivos agrícolas de la provincia de Córdoba.

Este trabajo pretende mostrar, además, que la aplicación de modelos en base a información satelital se constituye en un instrumento efectivo para evaluar alternativas y así determinar las mejores decisiones a asumir tanto por productores como por los organismos que diseñan las políticas y contemplan las necesidades a nivel agropecuario. Es precisamente porque apostamos a esta concepción que cobra relevancia esta aplicación y los resultados obtenidos.

Agradecimientos. Los autores expresan su gratitud a la Secretaría de Ciencia y Técnica – Universidad Nacional de Córdoba - Argentina (Secyt-UNC) por el apoyo financiero que permitió la realización de este trabajo.

Especial agradecimiento a los Ing. Agr. Miguel Nolasco, Juan Pablo Clemente y Daniel Faraoni quienes contribuyeron a la toma de datos a campo y en particular el último, por su relación con los productores, recolectó los rendimientos registrados por éstos. Los alumnos Gianluca Selva y Samir Mohuanna participaron en la cosecha manual de los cultivos, muchas gracias a ambos.

Referencias

1. Lyle, G., Lewis, M., Ostendorf, B.: Testing the Temporal Ability of Landsat Imagery and Precision Agriculture Technology to Provide High Resolution Historical Estimates of Wheat Yield at the Farm Scale. *Remote Sens.* 5, 1549–1567 (2013)
2. Lobell, D.B., Asner, G.P., Ortiz-Monasterio, J.I., Benning, T.L.: Remote sensing of regional crop production in the Yaqui Valley, Mexico: Estimates and uncertainties. *Agr. Ecosyst. Environ.* 94, 205–220 (2003)
3. Liu, L., Wang, J., Bao, Y., Huang, W., Ma, Z., Zhao, C.: Predicting winter wheat condition, grain yield and protein content using multi-temporal EnviSat-ASAR and Landsat TM satellite images. *Int. J. Remote Sens.* 27, 737–753 (2006)
4. Vicente-Serrano, S.M., Cuadrat-Prats, J.M., Romo, A.: Early prediction of crop production using drought indices at different time-scales and remote sensing data: Application in the Ebro Valley (north-east Spain). *Int. J. Remote Sens.* 27, 511–518 (2006)
5. Wall, L., Larocque, D., Leger, P.M.: The early explanatory power of NDVI in crop yield modelling. *Int. J. Remote Sens.* 29, 2211–2225 (2008)
6. Bocco, M., Ovando, G., Sayago, S., Willington, E.: Modelos simples para estimar cobertura de suelo agrícola a partir de Imágenes MODIS: aplicación a la zona central de Córdoba (Argentina). In: 42 JAIHO Congreso Argentino de AgroInformática, CAI 2013, Córdoba, pp. 61–69 (2013)
7. Ministerio de Agricultura Ganadería y Pesca, Estimaciones Agrícolas: Series Históricas, http://www.siaa.gob.ar/sst_pcias/estima/estima.php (2015)
8. Massey, R.E., Myers, D.B., Kitchen, N.R., Sudduth, K.A.: Profitability maps as an input for site-specific management decision making. *Agron. J.* 100, 52–59 (2008)
9. Bocco, M., Sayago, S., Willington, E.: Neural network and crop residue index multiband models for estimating crop residue cover from Landsat TM and ETM+ images. *Int. J. Remote Sens.* 35(10), 3651–3663 (2014)
10. Bocco, M., Heredia, S., Ovando, G., Sayago, S., Willington, E.: Estimación de cobertura vegetal con redes neuronales a partir de imágenes de distintos satélites. In: X Congreso Argentino de Ingeniería Rural y II del MERCOSUR, Rosario, pp. 1250–1256 (2009)

11. Batchelor, W.D., Basso, B., Paz, J.O.: Examples of strategies to analyze spatial and temporal yield variability using crop models. *Eur. J. Agron.* 18, 141–158 (2002)
12. Thorp, K.R., DeJonge, K.C., Kaleita, A.L., Batchelor, W.D., Paz, J.O.: Methodology for the use of DSSAT models for precision agriculture decision support. *Comput. Electron. Agric.* 64, 276–285 (2008)
13. Serra, G., Trumper, E.V.: Influencia de los daños provocados por el barrenador del tallo (*Diatraea saccharalis*) en maíz sobre el rendimiento por planta. Serie: Modelos bioeconómicos para la toma de decisiones de manejo de plagas. Edic. INTA. Nro. 5 (2004) <http://anterior.inta.gov.ar/manfredi/info/documentos/docprodveg/entomo/bioeco5.htm>
14. Penalba, O.C., Bettolli, M.L., Vargas, W.M.: The impact of climate variability on soybean yields in Argentina. *Multivariate regression. Meteorol. Appl.* 14, 3–14 (2007)
15. Fang, H., Liang, S., Hoogenboom, G., Teasdale, J., Cavigelli, M.: Corn-yield estimation through assimilation of remotely sensed data into the CSM-CERES-Maize model. *Int. J. Remote Sens.* 29, 3011–3032 (2008)
16. Wang, J., Li, X., Lu, L., Fang, F.: Estimating near future regional corn yields by integrating multi-source observations into a crop growth model. *Eur. J. Agron.* 49, 126–140 (2013)
17. Prasad, A. K., Chai, L., Singh, R. P., Kafatos, M.: Crop yield estimation model for Iowa using remote sensing and surface parameters. *International Journal of Applied Earth Observation and Geoinformation* 8, 26–33 (2006)
18. Li, A., Liang, S., Wang, A., Qin, J.: Estimating crop yield from Multi-temporal satellite data using multivariate regression and neural network techniques. *Photogrammetric Engineering & Remote Sensing* 73(10), 1149–1157 (2007)
19. Di Rienzo J.A., Casanoves F., Balzarini M.G., Gonzalez L., Tablada M., Robledo C.W.: InfoStat versión 2012. Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina. <http://www.infostat.com.ar>
20. Gusso, A., Ducati, J.R., Veronez, M.R., Arvor, D., Gonzaga da Silveira Jr. L.: Spectral model for soybean yield estimate using MODIS/EVI data. *International Journal of Geosciences* 4, 1233–1241 (2013)
21. de la Casa, A., Ovando, G.: Integración del índice de vegetación de la diferencia normalizada (NDVI) y del ciclo fenológico de maíz para estimar el rendimiento a escala departamental en Córdoba, Argentina. *Agricultura Técnica (Chile)* 67(4), 362–371 (2007)
22. Mkhabela M.S., Bullock P., Raj S., Wang S., Yang Y.: Crop yield forecasting on the Canadian Prairies using MODIS NDVI data. *Agricultural and Forest Meteorology* 151(3), 385–393 (2011)