Chapter 6 Background

6.1 Geographic units of analysis

This bulletin adopts the Global Administrative Units Layer (GAUL) boundaries for all national and sub-national units, except for China, where official Chinese data have been used. GAUL is available from FAO GeoNetwork (38)

6.1.1 Crop Production System Zones

For a detailed overview of the Crop Production System Zones (CPSZ), see Annex A. The main basis for the delineation of the CPSZs is the global ecological map prepared in the ambit of the FAO Forest Resources Assessment (39), further subdivided when necessary or otherwise modified based mainly on Köppen climate zones (the digital maps from Grieser et al. (40), VASClimO 1976-2000 data), "the most suitable cereal" grids available from the Global Agroecological Zones project (41). Other sources include USDA (42), Ramkutty's Global distribution of cultivable lands (43) and Monfreda et al. (44).

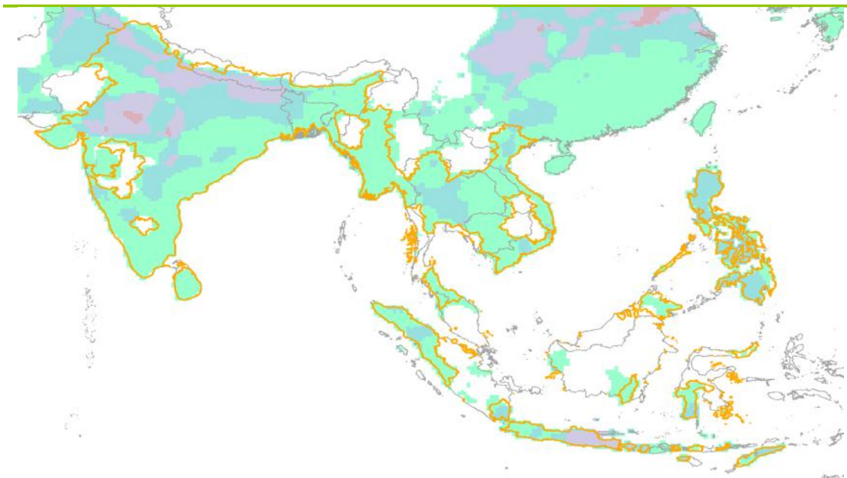
A special mention is needed about China, where the zones were not derived using the same approach, but cover standard Chinese agroecological zones (published in Chinese by Sun He (45); an English language description can be found in Hu Zizhi and Zhang Degang (46).

6.1.2 Major Production Zones

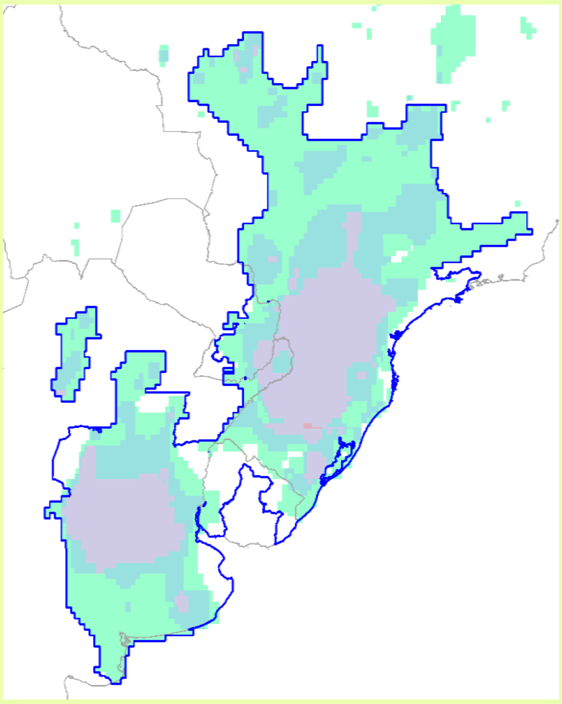
The Major Production Zones (MPZs) were selected based essentially on a combined maize, rice, soybean and wheat distribution raster map based on JRC crop masks (figure 6.1). For the areas of interest, each MPZ includes the area where at least one of the four crops is cultivated, bounded either by the area where none of the four crops is cultivated or by national or sub-national political boundaries. In one instance (central Europe and W. Russia), the northern limit was taken to coincide with CPSZ 59 ("Ukraine to Kazakhstan") and the eastern limit is given by the Ural Mountains.

The CropWatch Bulletin may cover different "major production zones" in subsequent issues.

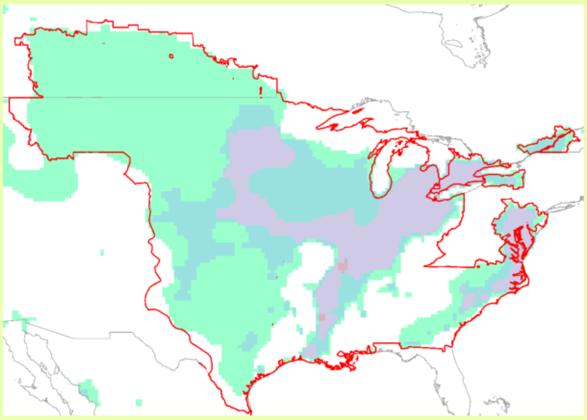
Figure 6.1 Map of the major crop producing zones showing political boundaries and the number of cultivated crops.



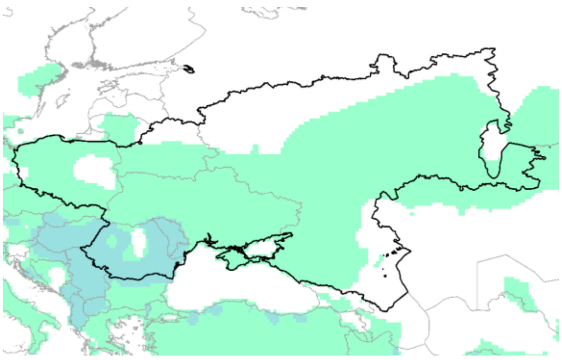
South-East Asia



South America



United States and Canada



Poland and Romania to Ural mountains in Russia.

Note: Figures shows only maize, rice, soybean and wheat. White background: none of the four crops is grown; light green: only one crop is grown; dark blue-green: two crops are grown; purple: three crops; dark purple-red: all four crops. Very few areas cultivate all four crops (mostly in India). 6.1.3 Thirty countries

6.1.3 Country selection

The selection of countries was done based on statistics published by FAO on production and trade, to include 80 percent of both. Several countries near the "end" of the list were included based on other considerations, e.g., their location in Asia (such as Uzbekistan) or in Africa (Ethiopia). Some generic information about the countries and their agriculture is given in Annex F.

6.1.4 CPSZ, MPZ and country boundaries

Environmental indices were computed for several spatial units: CPSZs, MPZs as well as the selection of 173 countries and territories. The polygons constitute a subset of GAUL_0 (the national level) after exclusion of all "small" polygons defined as those which are smaller than a 25 km x 25 km pixel, empirically measured by polygon perimeter length.

6.1.5 Sub-national units

For some of the largest countries (Australia, Brazil, Canada, China, India, Kazakhstan, Russia, and the United States), the first level administrative units were included in the analyses.

6.1.6 Chinese provinces

For China, 24 province-level subdivisions are covered in the report, divided into six official geographic regions, namely North China, Northeast China, East China, South and Central China, Southwest China and Northwest China (47).

6.1.7 Crop masks and arable land masks

Global crop masks were provided by JRC. The original crop masks were created for global water satisfaction indices and cover 11 crops. In our analysis, maize, rice, soybean and wheat masks at 0.25 degree resolution were used. Other sources include major crop areas by USDA (42) and the suitability map for rain-fed plus irrigated crops by FAO/IIASA (48) (41).

The arable land mask was created by joining the arable land from MODIS-derived land use and land cover products for 2010 and 2011 (49), Version 2 International Geosphere Biosphere Programme (IGBP) global land cover dataset (IGBP-DISCover) (50), and GlobCover 2009 (51) (52). The arable land maps for China in 2000, 2005, and 2010 were extracted from ChinaCover 2000, 2005, and 2010 provided by the Institute of Remote Sensing and Digital Earth (RADI) (53).

6.2 Data

6.2.1 NDVI

NDVI data for this bulletin is mostly MODIS NDVI provided by NASA, selected mainly because of its high spatial and temporal consistency. The NDVI data covers January 2002 to the end of September 2013. Only MODIS Terra Land Level 3/Level 4 16-Day Tiled Products (found on the LAADS Website (54)) with one kilometer resolution were used. After downloading the data covering target regions, crop masks were applied to remove the non-arable land from the downloaded data.

In addition, long-term average NDVI over the years from 1999 to 2012 with a resolution of 0.1875 degree based on SPOT-VEGETATION (provided by VITO (55)) was used to produce growing season masks.

6.2.2 Temperature

The air temperature used in this report is a global gridded (0.25x0.25 degree) monthly product from January 2000 to September 2013 generated based on the Global Surface Summary of the Day (GSOD) dataset, available from the Global Change Master Directory (56).

The GSOD dataset is derived from the Integrated Surface Hourly (ISH) dataset, DSI-3505 (C00532), and is produced by the National Climatic Data Center (NCDC). Its online data dates back to 1929 and the latest daily summary data are normally available 1-2 days after the date-time of the observations used in the daily summaries. Over 9000 stations' data (including air temperature, dew point, sea level pressure, wind speed, precipitation, snow depth) are typically available in this dataset.

The mean daily air temperatures from 2000 to 2013 in the GSOD dataset were extracted to calculate the monthly air temperature for each station. Then kriging interpolation was applied combined with STRM_DEM data (57) considering temperature elevation correction to generate the 0.25x0.25 degree global monthly product.

6.2.3 PAR

Photosynthetically active radiation (PAR), which covers radiation in the 400 to 700 nm range, is an important biological variable. It can be derived from shortwave irradiance in the 300 to 2000 nm range (58). For the years between 2001 and 2012, global monthly mean PAR data was calculated from the NASA CERES Energy Balanced and Filled (EBAF)-Surface Ed2.7 data product (59). For the year 2013 global monthly mean PAR data were obtained from original hourly geoland-2 project DSSF products and the FY-2D surface radiation product (60) (61) (62). All three products were converted to the WGS84 with a resolution of 0.25 x 0.25 degree.

6.2.4 Rainfall

CropWatch has assembled composite rainfall grids for the period from 2000.10 to 2013.9 covering the whole land surface at 0.25 degree spatial resolution. Two rainfall products were merged: (i) version-7 TRMM rainfall (63) at 0.25 degree resolution extending from 50 degrees south (50S) to 50 degree northern latitude (50N) and (ii) GPCC rainfall (64) rescaled from 0.5 or 1 degree to 0.25 degree. For the period between 2000.10 and 2010.12, TRMM 3B43 monthly rainfall was used between 50S and 50N and GPCC monthly data in other regions. From 2011.1 to 2013.6, GPCC Monitoring Analysis product (65) replaced GPCC in other regions. From 2013.7 to 2013.9, monthly TRMM rainfall products were combined with 3 hours TRMM 3B42 real time products, and the merged rasters were used between 50S and 50N, GPCC first guess product (66) replacing GPCC monitoring product in other regions. The comparison of satellite-based rainfall with rain gauges shows good consistence, with some “lines” existing at the boundary regions due to the different data sources.

6.2.5 VHI

The Vegetation Health Index (VHI) is an effective indication of the crop growth condition. In this bulletin, the VHI was calculated (67) (68) (69) by weighting the Vegetation Condition Index (VCI) and Temperature Condition Index (TCI), which are downloaded separately from the NOAA Star Center for Satellite Applications and Research GVI-x VH dataset (70). The equation is as follows:

$$\mathrm{VHI} = \mathrm{a} * \mathrm{TCI} + (1 - \mathrm{a}) * \mathrm{VCI}$$

The weighting factor (a) is an empirical coefficient for which the constant value of 0.5 was used.

6.2.6 Biomass

In this report, the net primary production (NPP) according to the Lieth’s “Miami model” (71) (72) is used to describe the global biomass situation. The Miami model expresses NPP as a function of two environmental factors, temperature and precipitation using following equations:

$$\mathrm{NPP} = \min(\mathrm{NPP_T}, \mathrm{NPP_P})$$

$$\mathrm{NPP_T} = \frac{3000}{1 + e^{1.315 - 0.119\mathrm{T}_C}}$$

$$\mathrm{NPP_P} = 3000(1 - e^{-0.000664\mathrm{Prec}})$$

Where T_C is average annual temperature in $^\circ$C and Prec is annual precipitation in mm.

6.3 Methodology

6.3.1 Environmental indices (EI)

To compare the agricultural impact of environmental variables across years and geographic areas (e.g., countries), it is necessary to use an index that gives a high weight to agricultural areas and focuses on the period covered by the cropping cycle. The intention is to derive "one number per year," so that it becomes comparable with other variables that are available at the same spatial and temporal scale, i.e., agricultural statistics and a number of socio-economic indicators and variables.

Such an index has been defined earlier for rainfall (73) and applied for synthetic studies (e.g., (74) (75)). This bulletin expands the notion of Rainfall Index to Environmental Index by applying basically the same approach to PAR and temperature.

The EIs are defined for one variable V (rainfall, PAR, temperature) and one polygon, which may be a CPSZ or a country or any other spatial unit. They are based on spatial grids at a resolution of 25 km and are computed as the spatial average over an arable land mask of the values of V accumulated over a user defined period, using the net primary production potential (1976-2000 VASClimO based NPPP) (72) as the weighting factor: the most productive pixels receive a higher weight than unproductive ones. For this study, PAR and rainfall are simply accumulated, while temperature is taken above a threshold of 5°C. For most maps, the accumulation period is the entire period from October 2012 to September 2013, with some exceptions that are indicated in the text, e.g., January to May to cover winter crops in the northern hemisphere.

6.3.2 Production, Area and Yield

Production

The crop phenology of each individual country was considered and the average NDVI of the current growing season over a specific crop mask combined with the production of two previous years was used to calculate the production of the current year for the four crops (maize, wheat, soybean and rice) using the following equation:

$$\mathrm{Production_i} = \mathrm{a} + \mathrm{b} * \mathrm{NDVI_i} + \mathrm{c} * \mathrm{Production_{i-1}} + \mathrm{d} * \mathrm{Production_{i-2}}$$

where i is the current year; a, b, c and d are four coefficients estimated by stepwise regression with the average NDVI from 2002 to 2012 and the production from 2000 to 2011 from FAOSTAT.

For China, the production is estimated by using the area and yield estimated for each crop with following equation:

$$\mathrm{Production} = \mathrm{Yield} * \mathrm{Area}$$

Area

For China, CropWatch combines remote-sensing based estimate of the crop planting proportion (cropped area to arable land) with a crop type proportion (specific type area to total cropped area) to estimate crop area (76) (77). The planting proportion is estimated based on an unsupervised classification of high resolution satellite images from HJ-1 CCD images. The crop-type proportion for China was estimated by GVG instrument from field transects (see also 6.3.6). The area of a specific crop was computed by multiplying farmland area, planting proportion and crop-type proportion of the crop.

To estimate crop area for wheat, soybean, maize and rice outside China, we employed two methods. One relies on the profile features derived from time series MODIS NDVI data (78) (79) (80) (81), the other one relies on regression of crop area to that of the two previous years using the following formula:

$$\mathrm{Area_i} = \mathrm{a} + \mathrm{b} * \mathrm{Area_{i-1}} + \mathrm{c} * \mathrm{Area_{i-2}}$$

where a, b, and c are the coefficients generated by stepwise regression with area from FAOSTAT or national sources.

Yield

For 30 countries, yield for each crop is estimated using production and area:

$$\mathrm{Yield} = \mathrm{Production/Area}$$

For China, two models, namely an agrometeorological model and a remote sensing index model were applied (82) (83). The values pooled from the two models are combined and averaged to predict the crop yield per unit area.

6.3.3 Cropping intensity (CI)

The Cropping Intensity index describes the extent to which arable is used over a growing season. It is the ratio of total crop area of all planting seasons in a year to the total area of arable land (84). It also can be used to describe the food-producing potential of an area (85). This report adopts the method proposed by Fan and Wu (86) based on an NDVI time-series derived from MODIS Terra and reconstructed by the S-G filter method (87). The following descriptors were derived from the data: number of peaks, width of each peak and peak values at pixel level.

The calculation of Cropping Intensity involves counting the number of peaks in the NDVI profile; we use the difference method to extract the maximum value of the discrete points. Assuming a pixel is in a discrete points series, the values of pixels before and after the point constitute a point series S1 which include N-1 pixels,

S1=DIFF (S)

where DIFF is the function used to calculate the difference among these points. Then, determine the sign of each pixel in S1, if the value of a pixel less than zero, assign the element value to minus one, otherwise, the element value is assigned to one; these values constitute a point series S2,

S2=SIGN (S1)

Where SIGN is the function used to determine whether the values are positive or negative. As a next step, calculate the difference of pixels before and after a point in series S2, which yields the points series S3

S3= DIFF (S2),

Finally, count the number of pixels with value minus two, which is the number of peaks in the NDVI profile. Based on the above equations and with other limitation, the model can generate a Cropping Intensity of one, two, and three per pixel, to illustrate areas with a single, two, or three crop seasons respectively.

6.3.4 Uncropped arable land ratio (UAL)

Uncropped arable land ratio was introduced to demonstrate the proportion of uncropped cultivated land to the total cultivated land over a certain geographic area (countries or sub-national units). Previous studies have shown a high correlation between NDVI and the photosynthetic biomass of cropped fields (88) (89). MODIS NDVI time series were used to identify whether an agricultural pixel is cropped or uncropped. For each cultivated pixel, 23 MODIS NDVI values were extracted from time series NDVI images and smoothed using Savitzky-Golay (S-G) smoothing filter (87). Maximum NDVI peak value was acquired from smoothed NDVI profiles from 2002 to 2013. Average and standard deviation of annual NDVI peaks (NDVIm and NDVIstd) were calculated based on annual maximum NDVI peak values from 2002 to 2013. A NDVI threshold method (90) (91) together with a decision tree were used to identify whether an agricultural pixel was cropped or uncropped over the last whole year (i.e., from October 2012 to September 2013). The difference between average and standard deviation of annual NDVI peaks were incorporated as a threshold to eliminate the inter-annual variability of biomass, crop phenology, and crop rotation. The decision tree can be described as follows:

$$\text{Status} = \begin{cases} \text{uncropped, if } \mathrm{Peak}_{2013} \le 0.15 \\ \text{uncropped, if } 0.15 < \mathrm{Peak}_{2013} < 0.45 \; and \; \mathrm{Peak}_{2013} < (\mathrm{NDVI_m} - \mathrm{NDVI_{std}}) \\ \text{cropped, if } 0.15 < \mathrm{Peak}_{2013} < 0.45 \; and \; \mathrm{Peak}_{2013} \ge (\mathrm{NDVI_m} - \mathrm{NDVI_{std}}) \\ \text{cropped, if } \mathrm{Peak}_{2013} \ge 0.45 \end{cases}$$

6.3.5 Potential biomass ratio

Potential Biomass Ratio (PBR) is a new index proposed in this report to describe the current crop biomass potential compared with the maximum crop biomass potential. Based on the good relationship between NDVI and plant productivity and biomass (92) (93) (94) (95), NDVI is used here as a proxy of crop biomass and the time-series NDVI (2002 to 2013) are applied to calculate crop PBR with following equation:

$$\mathrm{PBR} = \frac{NDVI_{\max\_c} - \mathrm{NDVI}_{\min\_h}}{NDVI_{\max\_h} - \mathrm{NDVI}_{\min\_h}}$$

where NDVI_{max_c} is the maximum NDVI of current year, NDVI_{max_h} is the historical maximum NDVI and NDVI_{min_h} is the historical minimum NDVI. Although NDVI_{max_c} and NDVI_{max_h} can be determined by assigning the maximum NDVI values of current year and the whole time range respectively, this cannot apply to NDVI_{min_h}, because the low values of non-vegetation NDVI may contaminate the crop minimum NDVI. In this report, the empirical minimum vegetation NDVI value (0.15) is introduced to calculate NDVI_{min_h} with following equation:

$$NDVI_{\min\_\mathrm{h}} = \max(0.15, \mathrm{NDVI}_{\min\_h0})$$

where NDVI_{min_h0} is the original minimum NDVI of the whole time series.

Based on above equations, PBR can be used as a relative measure of current biomass condition with value from 0 to 1. The higher the BVR value is, the larger biomass potential it indicates for that year.

6.3.6 Cropping structure

Cropping structure is an additional variable only applied to China. It precisely illustrates the proportion of area under a given crop type to the total sown area (i.e., crop type proportion) for geographic areas (i.e., provinces). The crop type proportion was estimated by combining GPS, video, and GIS data (collectively referred to as GVG) from field transects (76). The specifically developed GVG instrument collects thousands of field photos that are used to estimate the proportion of different crop types with accuracy above 98 percent (96) (97) (98) (85).

6.3.7 Time profile clustering

Based on a time series of pixel-based (raster) images, time profile clustering is a method that automatically or semi-automatically compares the time profiles of all pixels and distributes them among a limited number of "typical" behaviors (classes) that can be mapped. The method has the advantage of very synthetically describing the spatial distribution of typical time profiles (99) (100). In the CropWatch bulletin, we have used the SPIRITS software developed for JRC/MARS by VITO (101). NDVI and VHI profiles have been clustered, especially the difference between the current season and the average of the last five years taken as reference.