

# Notes on using the World Database on Protected Areas (WDPA) in the context of the REDD-PAC project

Rebecca Mant<sup>1</sup>, Simon Blyth, Corinna Ravilious, Naomi Kingston and Val Kapos

## Contents

1	Introduction .....	1
1.1	World Database on Protected Areas .....	1
1.2	WDPA in the REDD-PAC project.....	2
2	Important WDPA data considerations .....	3
2.1	Composition of protected areas information .....	3
2.2	WDPA accuracy .....	3
2.3	Database size .....	4
3	Important Analysis considerations .....	4
3.1	Raster or Vector analysis.....	4
3.2	Map Projection for Analysis .....	4
3.3	Different types of protected areas: designated, proposed and designation types.....	5
3.4	Country analysis .....	6
3.5	IUCN Management Categories .....	7
3.6	Overlapping PA's in the WDPA. ....	7
3.7	Point data .....	8
3.8	Status year .....	8
3.9	Lifespan of data from the WDPA .....	9
Annex A	.....	10

## 1 Introduction

### 1.1 World Database on Protected Areas

The World Database on Protected Areas (WDPA) is the most comprehensive global dataset on terrestrial and marine protected areas. It is a joint project between the United Nations Environment Programme (UNEP) and the International Union for Conservation of Nature (IUCN), maintained at the UNEP World Conservation Monitoring Centre and available online at [www.protectedplanet.net](http://www.protectedplanet.net).

The compilation of global protected areas data began over 40 years ago as the United Nations List of Protected Areas, which has been produced under a mandate from the UN General Assembly since 1962. This list developed into a GIS database called the WDPA in 1981, containing mapping more than 190,000 terrestrial and marine protected areas. In accordance with the principles of Conservation Commons and Open Geospatial Consortium Standards, the WDPA is now available in the public domain, through a fully searchable and downloadable GIS database on Protected Planet

---

<sup>1</sup> Contact details: [Rebecca.mant@unep-wcmc.org](mailto:Rebecca.mant@unep-wcmc.org)

Please contact us if you have any questions regarding the guidance

available at [www.protectedplanet.net](http://www.protectedplanet.net). The content of the WDPA conforms to and implements the IUCN Definition, Categories and Governance Types of protected areas as described in the IUCN Guidelines on Applying Protected Areas Management Categories<sup>2</sup>. Together the IUCN Guidelines and WDPA Data Standard<sup>3</sup> specify an international protected area standard to help ensure all protected areas are categorised and recorded in the WDPA with a minimum set of spatial and attribute information.

The IUCN definition of a protected area is ‘a clearly defined geographical space, recognised, dedicated and managed, through legal or other effective means, to achieve the long-term conservation of nature with associated ecosystem services and cultural values’ (Dudley 2008).

The primary sources of the national protected areas data in the WDPA to date continue to be the government agencies responsible for the management and governance of protected areas. However, where there are known gaps in official national data, these gaps are filled by a variety of experts working individually or through affiliations with non-governmental organisations, institutions and more.

The WDPA is a global mosaic comprised of GIS and attribute data produced through a variety of methods. This variation ensures that the WDPA as a whole is ideal only for global analyses, though some parts of it may be well suited to regional or national analyses. This inconsistency must be acknowledged when using the WDPA for analysis as they can impact upon results and should be considered in the interpretation phase of a project. Metadata information found in the WDPA ‘Source’ table help the analyst to assess the suitable scales at which to conduct analysis on the data of interest to their project within the WDPA.

The WDPA is continuously updated and published through monthly releases. It is therefore important to note which version of the WDPA is being used in any analysis, including referencing it in any resulting reports, publications or metadata.

## 1.2 WDPA in the REDD-PAC project

The REDD-PAC project is using the GLOBIOM model to model land use change under different policy scenarios. Protected areas are included within the model to restrict the allowed land uses and land use changes. The WDPA will be used within the global analysis to identify the location and IUCN category of protected areas. The IUCN protected area management categories are a global framework for categorizing the variety of management types applied in protected areas, ranging from strictly protected sites (IUCN categories I-IV) to sites where limited sustainable resource extraction occurs (IUCN categories V-VI). The effectiveness of protected areas in achieving their stated objectives also varies enormously, for a range of reasons, including the pressures on them, the investment in them and prevailing standards of governance. The rules for how to include protected areas in the model are set out in a separate document; including what land uses are allowed in principle in different protected areas and the effectiveness of protected areas for preventing land use change.

---

<sup>2</sup> Dudley, N. (Ed.) (2008) *Guidelines for Applying Protected Area Management Categories*. Gland, Switzerland: IUCN (<http://data.iucn.org/dbtw-wpd/edocs/paps-016.pdf>)

<sup>3</sup> UNEP-WCMC. (2010). *Data Standards for the World Database on Protected Areas*, UNEP-WCMC. (<http://www.wdpa.org/PDF/WDPA%20Data%20Standard.pdf>)

The following document covers considerations on how to use the WDPa to generate the input information for the model on the spatial distribution of protected areas and their attributes. Please read carefully the important WDPa data considerations below in creating the input layer.

## 2 Important WDPa data considerations

### 2.1 Composition of protected areas information

Protected area spatial data is collected in the form of boundary (polygon) GIS data wherever it is available, or as a single latitude/longitude coordinate (point) data where boundary data are not available. Thus protected areas data for a country is likely to have a mixture of polygons and points. The two data types are held in separate files in the WDPa, as points and polygons cannot be stored in the same files, and together they make up all of the protected areas in a country. So it is important that both data sets are used in the analysis for REDE-PAC. How to deal with point data is discussed in section 3.6.

### 2.2 WDPa accuracy

Only those sites with a known geographic boundary (polygons) or a precise geographic location and a known area (points) are included in the WDPa. Sites reported to the WDPa team with no known location are not included in the WDPa.

#### **Spatial Accuracy**

The data within the WDPa is from a wide variety of sources, resulting in a great variation in accuracy and resolution. Due to variations in sources and resolution, if the data is viewed with, for example, Google Earth as a backdrop, slight discrepancies in boundaries, such as coastlines, may be present. Protected areas data is often matched to data created at a national scale which may not 'fit' precisely with Google Earth's web Mercator projection.

#### **Attribute Accuracy**

Within the WDPa information on the attributes of a protected area are given in a set of standard data fields. Several of the data fields are always populated with information e.g. 'Name' (NAME) and 'Country' (COUNTRY), as they are minimum requirements. However, there are other attributes which are not easily completed and therefore gaps in some of the WDPa attributes do exist. 'Reported Area' (REP\_AREA), 'Designation' (DESIG) or 'Status Year' (STATUS\_YR) and 'IUCN Category' (IUCN\_CAT) are examples of fields that are incomplete in the WDPa. If there is a need to use a field for selecting out only a subset of information it is important to check how well populated it is.

Where gaps are present in the reported area, area figures have been generated based on the geometry (GIS\_AREA). All area figures shown within the WDPa are in square km and calculated using the Mollweide projection. In some instances the reported area and the GIS area may show discrepancies. Big differences may be due to incorrect reporting, for example data reported as hectares, or may reflect a real difference in the area reported legally versus what was demarcated on the ground although these cases are rare. Any serious unexplained discrepancies should be reported to UNEP-WCMC. **Within the REDD-PAC analysis the field 'GIS\_AREA' should be used for calculating the input data for the model.**

### 2.3 Database size

Previously, the large size of the WDPA database hindered analysis using the whole WDPA. Therefore, unnecessarily complicated polygons have been simplified. The simplification process was carried out within ArcGIS 10, using the Douglas Peucker algorithm. The simplification means that the sheer size and complexity of the data no longer have the same influence over the way analysis can be conducted. However, the ability to conduct GIS analysis will vary depending on the power of the hardware available to the analyst.

## 3 Important Analysis considerations

### 3.1 Raster or Vector analysis

Decisions on whether to use vector or raster analysis techniques will depend on the nature of the data to be overlain with the WDPA data. If most or all of the other datasets are in raster format then a raster analysis might be the preferred option. In such cases, care must be taken to consider the cell size used, when converting the WDPA data to raster data. A degree of generalisation must be expected in such a conversion. Conversion in which a cell is given the value of the majority of what falls within it will cause small sites to be eliminated. Conversely, if a cell is given a value of being protected if even the smallest part of a site falls within it, an overestimation of the amount of protection will result. Conversion which indicates the percentage of each cell which is protected can be achieved using some software (including ArcGIS), is the conversion used in GLOBIOM for land cover and will give a more accurate analysis, but this is a more complicated calculation that will not be expanded upon here.

**If the analyst is unsure as to whether to use vector or raster analysis techniques for the WDPA data, or how to carry out more complex conversions, they should contact UNEP-WCMC for advice.**

### 3.2 Map Projection for Analysis

In carrying out analysis of the WDPA with other datasets, decisions need to be made as to which map projections should be used. The decision will partly depend on the nature of the analysis to be undertaken and partly on what other datasets are also to be used.

The WDPA is supplied in geographic co-ordinate system WGS84. As a general rule, global analyses in which areas / percentage of protection are being calculated is best conducted within an equal area projection such as Mollweide or Sinusoidal. This will enable accurate area figures to be obtained. Mollweide is used to calculate the 'GIS Area' (GIS\_AREA) field in the WDPA attribute table. For regional analyses a Lambert Azimuthal projection can be used with the centre of the projection put at the centre of the area of interest. This will also probably produce better mapped outputs for a region. For analysis of distance (proximity analysis) an equidistant projection may be preferable. Note that any distance calculations of irregular features (e.g. Coastline) will be strongly influenced by the scale at which the data being used was created (digitised).

If the WDPA is to be analysed with a raster dataset it is important to recognise that the conversion of detailed raster datasets between projections is complex and produces a higher degree of inaccuracy than converting vectors. Each time a raster dataset is transformed into a different projection or coordinate system, the cells are resampled. Minimising conversions will maintain the best data quality. It can therefore be advisable to use the original projection of the detailed raster for analysis.

When conducting national scale analysis, original national data may be more detailed and more appropriate to use instead of the WDPA. If using the WDPA, along with national data layers, it is likely to be appropriate to use only a subset of the WDPA relevant to the particular country and to project it into the countries using standard national level projection, if one exists. It is important to check whether the standard national projection is an equal area projection.

### 3.3 Different types of protected areas: designated, proposed and designation types

The WDPA includes a variety of different types of protected areas. The type of protected area is characterised in three main ways:

- i) Designation Type: Describing sites as either nationally or internationally designated protected areas, where 'national' sites are designated through legal statute or other effective means and where 'international' sites are designated through an international convention such as the UNESCO World Heritage Convention, Man and Biosphere Program or Ramsar Convention. The distinction between national and international sites is recorded in the 'Designation Type (DESIG\_TYPE ) field in the WDPA;
- ii) Designation: The legal/official designation given at the time a protected area is established. It comes in numerous forms in the WDPA including 'National Park', 'Wildlife Reserve', 'World Heritage Site' and so on. The objective and management of designations are not consistent across the world and do not provide a means to compare similar protected areas. Designation is stored in either in the local language through the 'Designation' (DESIG) field or translated into English through the 'Designation English' (DESIG\_ENG) field. In the majority of records both 'Designation' and 'Designation English' are present in the WDPA;
- iii) Marine/Terrestrial: A marine protected area is classified as marine (MARINE = '1') in the WDPA if some or part of its area falls within the marine environment. Protected areas classified as marine within the WDPA can therefore be either completely marine – in a sea, ocean or equivalent – or partially marine – on the coast and encompassing part of the sea/ocean;
- iv) Status: Describing whether a site has been designated or is currently only proposed through legal or other effective means (recorded in the 'STATUS' field);

Almost all protected areas designated through international conventions overlap national designated protected areas in whole or in part. Although, it should also be noted that presently the data for UNESCO Man and Biosphere Reserves often includes core, buffer and transition areas, while only the core zones are considered protected areas under the IUCN definition, and so where buffer and transition areas exist in the data set, they should be excluded from any analysis. UNEP-WCMC is currently working with the UNESCO MAB Program to determine the appropriate level of integration of this data set in the WDPA.

In order to identify the specific type of protected area the 'DESIG\_ENG' field can be used. **The 'DESIG\_ENG' field should be used, rather than the 'DESIG' field, to identify and select consistently**

named categories for both National and International sites, as the 'DESIG' field may have designations spelt in the local language and will not be globally consistent.

**The analysis within the REDD-PAC project should include all sites except Man and Biosphere reserves, as currently (January 2013) the data on these areas is insufficient. However, it is important to note that future updates to the WDPA are likely to improve the Man and Biosphere reserve data and so the issue should be reconsidered in any future analysis.**

Protected areas which have a marine element are shown with an attribute '1' for marine. However the marine element can vary from being very small to the majority or all of the site. Therefore, the marine category should not be used for excluding marine sites from analysis of terrestrial protection as this would exclude sites that have both marine and terrestrial components. Instead a marine base layer, such as the World Vector Shoreline, or the national coastal boundaries, should be used to clip protected areas that fall within the marine and coastal waters.

The WDPA stores information on national sites that have been proposed but have not been formally designated. Many of these sites will already have management in place and are simply awaiting the, formal legal designation process to be completed. Sites whose boundaries have been demarcated and have polygon boundaries within the WDPA database are likely to be further along the designation process. Within the status field there are also categories of 'Adopted' and 'Inscribed' which refer to internationally recognised protected areas. The term used depends on terminology within the convention through which the site has been designated. Both 'Adopted' and 'Inscribed' can be considered equivalent to the 'designated' category used for national sites. **For the analysis in the REDD-PAC project the analysis should include protected areas that are designated and those that are proposed and have polygon outlines. Sites that are proposed and only have point data associated with them should be excluded.**

### 3.4 Country analysis

In carrying out analysis which includes country level work it is important to decide how to assign protected areas to different countries. It is possible to identify the country of protected areas either by the country code within the WDPA or by the country borders within a political boundary data set, such as the ESRI Digital Chart of the World. It is possible to have discrepancies between what is shown in the WDPA as falling within a country and where the political boundaries lie in another data set. If the WDPA country code is used, there are likely to be slight overlaps (slivers) of protected area boundaries on the borders between countries due to different political boundaries being used as the basis for digitising the protected area boundaries. These overlaps need to be eliminated and although this can be done with little adverse consequence at a global scale, at country and particularly at site scale level, more significant discrepancies can be introduced. Alternatively, if the country code is used from the other dataset being overlaid, some parts of the protected areas may be excluded. The necessity to use a large number of "buffered" points within the analyses is likely to increase the level of this type of error in any analyses (see Point data section 3.6 below).

Possible mismatches in country codes can also occur due to disputed country borders and dependent territories. If data are being provided by a country which has a disputed boundary, the protected areas they provide will match the country's interpretation of where the boundary is, not necessarily the UN's boundary. Dependent territories can be defined variously by their own ISO3

codes (as in the WDPA) or other datasets may define them in terms of their mother country e.g. GLP (Guadeloupe) in the WDPA, a French territory, might be FRA (France) in other datasets.

### **Transboundary sites**

Another important consideration when using the WDPA country codes are sites which cover multiple countries through transboundary agreements to establish a co-managed protected area. Transboundary protected areas have their ISO3 codes defined with a comma in between (e.g. USA,CAN), although as this is a new development at the time of this guidance, there may be some transboundary sites which are not yet identified in this way. To ensure that all sites in a country are incorporated in any analyses you must select using a query such as ISO3 = '%CAN%'. However, it is important to be aware that the area of the site will be defined as the complete area of the transboundary site, rather than the area relating to the site in one or other single country. In order to count only the portion of a site which falls within a specific country, a clip or intersection analysis should be done using a political boundary layer.

## **3.5 IUCN Management Categories**

The WDPA stores the IUCN categories of protected areas (recorded in field IUCN\_CAT). However, there are many sites (both polygons and points) for which IUCN category has not been reported. These sites are defined as IUCN\_CAT = 'Not Reported'. Assigning IUCN categories to designated areas is not a mandatory requirement for countries, although recommended through the IUCN guidelines and the Convention on Biological Diversity. The IUCN category is always provided to the WDPA as part of the data set; even where there are gaps in the data reported, categories are not applied by the administrators of the WDPA at UNEP-WCMC. Also, the IUCN category does not relate to the given designation (e.g. National Park), but to its management objectives, and therefore categories cannot be inferred by the designation listed. As it is not a requirement that IUCN category information is reported to the WDPA by protected areas agencies, it is possible that any particular IUCN category information may tend to be under- represented in any analyses. **Therefore the 'Not Reported' category should be treated as a category in its own right rather than excluding these protected areas completely.**

IUCN categories can each be selected out and analysed separately, but as there can be overlaps between (and exceptionally within) categories, it is important not to simply use the sum of all protection as the sole basis for calculating the amount of the country that is protected (as it would give an overestimated value). See section 3.6 for how to deal with overlaps.

## **3.6 Overlapping PA's in the WDPA.**

There are genuine overlapping protected areas (often with different IUCN categories) in the WDPA. Additionally, duplicate sites can occasionally be found as a result of the variety of data sets that have been reported to the WDPA over time and by different sources, and where validation processes did not identify the duplication on account of the sites being reported with difference in attributes and geography.

When undertaking certain spatial analysis with the WDPA, including the REDD-PAC project, **it is important to create a 'flat' layer which contains no duplicates to ensure that there is no double counting of total area protected.** An analysis may want to just look at protection vs. no protection.

This can be simply undertaken by assigning all protected areas with a single value and applying a simple dissolve function which removes duplicates and all boundaries between them.

In the REDD-PAC project, identification of IUCN category will also be required. Subsets of the dataset can be created by selecting sites from the global protected area layer for each IUCN Management category, or groups of categories. A dissolve function should be run on each of the subsets to produce layers that do not contain any overlapping polygons. Where overlaps exist between the subsets, for the REDD-PAC analysis, these should be resolved by assigning **the highest management category**. For example, where a site with IUCN Management Category Ia overlaps with a portion of a site of category III, that portion is assigned the Ia value. For the purposes of this project, the order of highest to lowest is: Ia, Ib, II, III, IV, V, VI, Not Reported. If required, categories can be combined. **In the REDD-PAC project areas should be grouped into eight subsets: each of the IUCN categories Ia, Ib, II, III, IV, V and VI plus those with no management category assigned (i.e. Not Reported)**. The above dissolve function should be used in a similar way on these combined subsets. The analysis would result in a dataset which retains the categories but eliminates overlaps according to the priority of the IUCN category.

### 3.7 Point data

Wherever possible within the WDPA point data are being replaced with boundary (polygon) data. However, there are still a significant number of point features in the WDPA, and these must be included within the analysis. A point buffering process enables point data to be converted into circles that are proportional to the area reported. Data can then be combined with the polygon data to allow analysis of a single protected areas dataset. **If the area of a point feature has not been reported, it should be excluded from the REDD-PAC analysis.** Thus points with 'REP AREA' <> 0 are removed. The remaining points are buffered by calculating the radius of a circle proportional to the area of the site by:

Add field called 'radius'

Calculate field to be equal to

$\text{Sqr} ( [\text{REP\_AREA}] * 1000000 / 3.1416 )$

Then simple buffering can be done using the 'radius' field. It makes no difference at all whether you project your geographical points to Mollweide and buffer these or if you add your WGS84 point data to a Mollweide Projection data frame and buffer them there. The result is the same.

It should be noted that the resulting buffers produce acceptable equal-area results for analysis but due to the projection system, will not be perfectly circular everywhere on the globe. This would be the case no matter what projection system was used. If the goal was to have circular buffers everywhere on the globe, you would buffer the points using the geodesic distance. **However for the REDD-PAC analysis, the method described above should be used.**

### 3.8 Status year

The 'Status Year' (STATUS\_YR) field can be used to identify the year of establishment for protected areas and so carry out time series analysis. However, it is important to note that not all protected areas within the WDPA have a year defined. **In the REDD-PAC analysis which looks at protection in different time periods (for example for year 2000 and 2010) all sites without a 'year' value defined**



**in the WDPAs should be considered to have been present within the first year of analysis (e.g. the year 2000).** The status year provides the year that the protected area receives its current status which provides some potential limitations to the analysis (discussed in Annex A).

### **3.9 Lifespan of data from the WDPAs**

The WDPAs are constantly updated as new sites are protected. However, there will be a lag between a site becoming designated and it appearing in the WDPAs. For this reason, it is not considered likely that statistics for number or area of sites calculated for an analysis in any one year will remain static. It is therefore important to note which version of the WDPAs is being used in any analysis, including referencing it in any resulting reports, publications or metadata.

## **Annex A – Considerations in using the status year in the REDD-PAC analysis.**

The status year field in the WDPa records the year in which the protected area received its current status (i.e., proposed or designated). A geographic area which is under protection, a protected area, could have changed its status several times through its history. A protected area may change from a proposed to a designated site after it has been through an official process (or it may never become designated in which case it would be removed from the WDPa). Additionally, if a protected area comes under stricter regulations or a new national designation, the status year for that geographical area may change in the WDPa. Therefore, when using the Status Year it is important to be aware that it does not necessarily represent the first year in which there was any protection within an area. This is often an assumption made when using the status year to select protected areas present in 2010 by selecting all protected areas that have status year less than or equal to 2010 (or not reported). If a site had been included in the WDPa as proposed (with polygons outlined) in 1999 and was updated in the WDPa as designated after 2000, it would now be in the WDPa as designated with the status year as the year of designation (e.g. 2005) and so it would not be included in the REDD-PAC analysis for 2000, underestimating the level of protection in 2000. Figures can alternatively be slightly overestimated, if a protected area is listed in the WDPa as proposed and never becomes designated but has not yet been removed.

As the WDPa is released in monthly releases, an alternative method for analysing what protected areas were present in a given year, is selecting the release from that year. However, it takes time for additions or changes to national protected area systems to be recorded in the WDPa and therefore is only a best estimate of the global protected area network at any point in time. Hence, **in the REDD-PAC project, using the 'Status Year' attribute in the WDPa is still the best option for selecting the protected areas present in 2000 and 2010.**