Land-Based Layers Workflow - 2011-2012 run

Path on Neptune:

Run 2011-2012: `/var/cache/halpern-et-al/mnt/storage/marine_threats/impact_layers_2013_redo/impact_layers/work/land_based/201112/` Script only location on Neptune:
`/var/cache/halpern-et-al/mnt/storage/marine_threats/impact_layers_2013_redo/impact_layers/work/land_based/201112_script_only/`

Note:

- The unix steps were ran on Neptune (Ubuntu 12.04.5 LTS)
- On Neprtune, a local install of Python was used; Here is the info to install one if necessary: https://help.nceas.ucsb.edu/local_install_python_on_a_server
- The Windows/ArcGIS ran on Bumblebee in NCEAS Vizlab (Win 7 with ArcGIS 10.2)
- Begin by copying the *201112_script_only* directory and go step by step

Step 1: FAO

Download and format FAO pesticide/fertilizer data.

Get raw FAO fertilizer data

- go to: ~~http://faostat.fao.org~~ New website: http://faostat3.fao.org/download/R/RF/E
- select data options:

```
- country:              Click *Regions* tab and select: *World > (List)*
- year:                 select all (2002-2012)
- item:                 select all 3 "total nutrients" items
- element:              [Consumption in nutrients]
- selected parameters:  click show
```

- click download (csv format)

Get raw FAO pesticide data

- go to: ~~http://faostat.fao.org~~ New website: http://faostat3.fao.org/download/R/RP/E
- select data options:

```
- country:              Click *Regions* tab and select: *World > (List)*
- year:                 select all (1990-2013)
- item:                 select all the "+ (Total)" items
- element:              [Use]
- selected parameters:  click show
```

- click download (csv format)

=> CHANGES in data with new FAO website: They have remove the country id from the dataset and automatically removed the countries with no data

Get FAO country codes table

- go to: ~~http://www.fao.org/countryprofiles/modulemaker/en/~~
- New website: http://www.fao.org/countryprofiles/geoinfo/modulemaker/index.html to get the xml file with counrty codes: Self-governing territories.xml
- go to:http://faostat.fao.org/site/371/DesktopDefault.aspx?PageID=371
- download "the area list", xls format to get the country codes
- Open in in Excel and save it as *country_codes_fao.csv*

Get World Bank AG % of GDP table data

- go to: http://data.worldbank.org/indicator
- click on "Agriculture, value added (% of GDP)"
- this should take you to: http://data.worldbank.org/indicator/NV.AGR.TOTL.ZS/countries

- click 'Download Data->XML'

Get World Bank GDP table data

- go to: http://data.worldbank.org/indicator
- click on "GDP (current US$)"
- this should take you to: http://data.worldbank.org/indicator/NY.GDP.MKTP.CD/countries
- click 'Download Data->XML'

Run scripts

- change worldbank xml files to csv file for next step:

```
./worldbank_gdp_parse.py
```

- create table of fertilizer/pesticide consumption in tonnes per country code (run twice with different year ranges; e.g. 2007-2010 and 2011-2010):

```
R --vanilla <fao_update.R
```

- add the country code to the FAO data:

```
R --vanilla <add_country_codes.R
```

- compare/plot results from previous step (run with correct input files from previous step):

```
R --vanilla <fao_compare_temporal.R
```

- the `fao_compare_2008.R` script was used to compare the new methodology to the one used in the original Halpern 2008 paper

Notes

- Details behind FAO fertilizer methodology can be found here (note that there was a change in the methodology after 2002): http://www.fao.org/fileadmin/templates/ess/ess_test_folder/Publications/Agrienvironmental/Methodologocal_Notes_FAOSTAT2011.pdf
- Also note the distinction between consumption by nutrients and consumption by weight for the fertilizer data. We have used consumption by nutrients for the final analysis, but also tried consumption by weight (see `fert_by_weight` folder for plots).

Step 2: Landcover/Landuse (LCLU)

Download and format LCLU data.

Dataset background

- desired dataset: MCD12Q1 V051: land cover type 1 (IGBP)
- relevant links:
  - http://www.bu.edu/lcsc/data-documentation/
  - http://earthobservatory.nasa.gov/Newsroom/view.php?id=22585
  - https://lpdaac.usgs.gov/products/modis_products_table/mcd12q1

Download dataset

- go to: http://reverb.echo.nasa.gov/reverb/
- follow these steps:

```
- Search Terms: mcd12q1
- Temporal Search:
  start:  2012-01-01 00:00:00
  end:    2012-12-31 23:59:59
- [press enter]
- Select Datasets: MCD12Q1 version 51
- [Press: Search for Granules]
- [Press: Accept]
- [Press: All Shopping Cart button]
- [Press: View Items in Cart]
- [Press: Download]
- Select URLs to Download:  Data
- Select Download Option:  Text File
- Press:  Save
- [Press: Download]
- Select URLs to Download:  Metadata
- Select Download Option:  Text File
- Press:  Save
```

Run scripts

- download hdf files:

```
./go_download_files.sh
```

- merge hdfs into single tif with igbp landcover classes, for both years:

```
./hdf_to_tif.sh
```

Notes

- We also tried to use the UMD classification (see `global_landcover/umd`), but used the IGBP classification for the final analysis.
- Some classification statistics (number of 500 m^2 pixels):

```
umd classification:
  landclass 12 (cropland),  2005:              76716687
  landclass 12 (cropland),  2009:              76463693
  landclass 13 (urban),     2005:              3063063
  landclass 13 (urban),     2009:              3062706

igbp classification:
  landclass 12 (cropland),  2005:              56632531
  landclass 12 (cropland),  2009:              56090829
  landclass 14 (cropland/natural veg),  2005:  42801682
  landclass 14 (cropland/natural veg),  2009:  41760296
  landclass 13 (urban),     2005:              3058152
  landclass 13 (urban),     2009:              3056223
```

Step 3: Dasym Prepare

Prepare data for dasymetric mapping. This step splits global landcover rasters by continents. This step needs to be run on a Windows machine with ArcGIS installed

- run continent/countries raster prep script:

```
C:\Python27\ArcGIS10.1\python.exe
  landcover_countries_continents_split.py
```

- (*Note*: when re-running above script, use fresh/original shp file for countries_units.shp in case .prj disappears during last run)

Step 4: Dasym

Run dasymetric mapping.

Run scripts

- use `run_dasym_loop_*.sh` to run `dasym_map.py` multiple times for each continent, year, pest/fert combination

Notes

- run the sripts at the same time as this step is time consuming
- relevant links:
  http://portal.nceas.ucsb.edu/Members/perry/methods/dasymetric-mapping/
  http://portal.nceas.ucsb.edu/Members/perry/methods/dasymetric_flowchart2.png/image_view_fullscreen
  https://github.com/scw/global-threats-model/tree/master/Dasymetric

---

Step 5: Zonal Stats

Calculate zonal statistics for each continent -- sums of fertilizer/pesticide values for each basin.

- run zonal stats script:

```
C:\Python27\ArcGIS10.1\python.exe terra_stats_fert.py
C:\Python27\ArcGIS10.1\python.exe terra_stats_pest.py
```

- ~~(*Note*: you'll need to create blank af, as, etc dirs first in [5]_zonal_stats\output)~~ Added this to the script

---

Step 6: Global Plume Prepare

Prepare global plume data. This step updates global plume shapefile fertilizer/pesticide sum values to the new basin zonal statistics.

- first, create blank `global_plume*.shp` files by resetting the DBF fields to zero via `zero_dbf.py` (resulting shapefiles can be found in `output/og_blank`)
- run `update_global_plume.py` twice, with different years
- run `clean_global_plume.py` twice for both time periods to remove zero fert/pest values
  (note: this script will not update FIDs)
- run `splitter.sh` to slice the pour points shapefile into 4 parts ******

Step 7: Plume Model

Run global plume model.

- run `plume_distributions.R` script (will need to install maptools package in R: install.packages('maptools')). Use the "0.05%" values in log output to manually update `plume_buffer.py` limits values for pesticides and fertilizer.

```
$ R --vanilla < plume_distributions.R
```

Note:

This next step relies on *grass*, you can find the grass setup I used here: `/var/cache/halpern-et-al/mnt/storage/marine_threats/grass_jb`

To define the location extent, I used the extent of the ocean_mask file.

The mapset projection was set using the projection of global*plume**.shp

- launch grass64
- select a the PERMANENT mapset
- add the ocean mask to the PERMANENT mapset:
  `r.in.gdal -o input="impact_layers_redo/land_based/201112/step0/ocean_mask/ocean_mask.tif" output=ocean`

  Note: This need only to be done once

- run: `./run_2011_2012.sh`

  Note: you need to uncomment the part you would like to run and change the year period accordingly to your global_plume**.shp* ***tip****: run 8 mapsets simultaneously during plume*buffer.py for efficiency

- When all the plume runs are done combine plumes into single global raster:

```
./gdal_add.py -o global_plumes_pest_2011_2012_raw.tif -ot Float32 plume_pest*.tif
./gdal_add.py -o global_plumes_fert_2011_2012_raw.tif -ot Float32 plume_fert*.tif
# note: may need to add plumes in stages
```

- repeat above steps for other time period

Step 8: Plume Finalize

Produce final products.

- run `difference_rasters.py` to create change detection rasters
- run `log_normalize_rasters.py` to log-normalize yearly and difference plume rasters
- run `plume_plots.m` in matlab to produce histogram figures of final rasters

Adapted from John Potapenko (john@scigeo.org)'s version
(`/var/cache/halpern-et-al/mnt/storage/marine_threats/impact_layers_2013_redo/impact_layers/work/land_based/documentation/land_based_layers_wo`
by Julien Brun (brun@nceas.ucsb.edu)