Developing a Routing Priority Map for Geospatial Modeling of CO_2 Pipeline Deployment in the People's Republic of China

Ari A. Zwick
Office of Science, Science Undergraduate Laboratory Internship Program
Correspondence: azwick2@uic.edu

University of Illinois at Chicago

Pacific Northwest National Laboratory
Richland, Washington

August 26, 2009

Prepared by the requirement of the Office of Science, Department of Energy's Science Undergraduate Laboratory Internship under the direction of Casie Davidson in the Energy and Environment Directorate at Pacific Northwest National Laboratory

Developing a Routing Priority Map for Geospatial Modeling of CO_2 Pipeline Deployment in the People's Republic of China

Ari A. Zwick (University of Illinois at Chicago, Chicago, IL), Casie L. Davidson (Pacific Northwest National Laboratory, Richland, WA 99352).

ABSTRACT

Carbon capture and sequestration (CCS) technologies are currently being researched as a potential component of a global portfolio of technologies to help reduce anthropogenic emissions of carbon dioxide (CO_2) to the atmosphere. In China, currently a leading emitter of CO_2 and a potentially critical player in future carbon emissions reduction strategies, it is important to evaluate the economic feasibility of CCS to understand its potential for large-scale deployment. This paper describes the development of a high resolution geospatial model to assist in efforts to estimate the construction costs of pipelines for transport of CO_2 from sources to storage sites. The model assigns relative weights to geographic features throughout mainland China to form a relative prioritization map that may be used to model pipeline routing along paths that are likely to represent the lowest cost paths. The final routing priority map (RPM) differentiates between areas according to their relative cost for routing from sources to sinks. The RPM represents the weighted combination of all overlapping geographic and cultural features included in the model. By using the RPM in conjunction with a routing protocol, grid cells with low priority values (i.e., those for which construction and/or societal costs would be higher) would be avoided in favor of cells with higher priority values, all else equal. This mode of estimating least-cost pipeline routing could represent a significant enhancement to existing methodologies used to estimate CO_2 transport costs for CCS in China.

INTRODUCTION

With growing public and political consciousness of climate change, there is a growing international dialogue about addressing global emissions of carbon dioxide (CO_2), with China playing a key role as the world's largest developing economy and one of its largest emitters of CO_2. With its vast coal reserves accounting for approximately 70 percent of the nation's current energy consumption [1], it is unlikely that China will significantly reduce its dependence on this cheap, indigenous energy source in coming years. Given the projected growth rates for the nation's energy use and infrastructure, CCS may be a key technology for reducing carbon emissions in the coming decades.

Dahowski, et al., [2] estimated the potential CO_2 storage capacity within China to be in excess of 3000 billion metric tons of CO_2 ($GtCO_2$), which is enough to store all of China's coal emissions for hundreds of years at current emissions rates. In that study, sources were matched with sinks, and costs calculated using a straight-line routing algorithm to estimate CO_2 pipeline transport costs. However, the data exist to allow for the modeling of more realistic pipeline routing in estimating pipeline costs for source-sink pairs. This present study is designed to facilitate modifications to the Dahowski methodology by creating a pipeline routing priority map (RPM) that can serve as the basis for a modifiable cost surface to drive pipeline route modeling in China, and for future iterations of the China CCS cost curve work.

This was accomplished within a geographic information system (GIS) using almost exclusively free datasets primarily derived from official government sources. The purpose of this routing priority map is to model pipeline placement under real geographic constraints from point sources such as coal fired power plants, cement, and ethanol plants to sinks such as deep saline aquifers, depleted oil fields, and deep unmineable coal beds. These non-linear routing paths will provide a higher-resolution method for estimating the costs for routing hundreds of potential pipelines all throughout mainland China. This is a first of its kind attempt to create a national scale pipeline prioritization map for CCS in

China, and therefore relies on correlating assumptions from various smaller scale individual projects as proxies for this model. This paper describes the methodology used to create the RPM, including key features and assumptions applied to each dataset.

MATERIALS AND METHODS

Pipeline routing and associated costs are commonly evaluated using a relative weighting system in which multiple data layers are assigned relative weights and combined to form a cost map [3]. The RPM differs from a cost map in that the weighting assumptions used reflect relative routing preferences as opposed to actual cost[1]. Each data layer falls under a data category, which has an assigned multiplication factor (category weight) reflecting its relative importance, compared to other data categories (see Table 1). Within each category, each feature is represented by one or more data layers and is assigned a relative weight in accordance with its estimated cost or routing preference within the category. Each data layer therefore has a relative weight for each feature it contains, and a multiplication factor from its category to yield its total weight. The total weights from each data layer are then summed to form the final RPM.

Table 1: *Categorical and relative weighting hierarchy used for the present study.*

Category	Relative Weight	Category Weight	Category	Relative Weight	Category Weight
Landslide Probability (% per anum)		10	**Slope** (Degrees)		5
0.01−1.5	1000		>70	9	
0.0025−0.01	9		60−70	7	
0.00125−0.0025	8		50−60	5	
0.00025−0.00125	6		40−50	4	
0.0001−0.00025	4		30−40	3	
0−0.0001	2		20−30	2	
0	0		0−20	1	
Population (Persons per square kilometer)		10	**Infrastructure**		5
>10 000	1000		Highway	9	
3500−10 000	9		National Road	7	
400−3500	4		State Road	5	
0−400	0		Railroad	5	
Protected Areas (IUCN Category)		7	**Rivers and Lakes**		5
Level 1a	1000		Primary River	9	
Level 1b	1000		Secondary River	7	
Level 2	9		Tributary River	5	
Level 3	9		Secondary Tributary River	3	
Level 4	5		Water Reservoir/Lake	7	
Level 5	3				
Level 6	3				

Weightings applied to data layers are the keystone that allows for the approximation of real world conditions. Categorical weights modify the relative weighting system, in which each individual layer is weighted on a scale of 0-9. A relative weight of 9 is the highest value where routing is discouraged around these features, while a value of 0 is the most desired weight and would preferentially routed to. Certain features were deemed unsuitable for routing, and in order to block routing expect where no other option existed, these features have been assigned values of 1000 in the RPM. The categorical weight is then applied to assign relative importance of all the features included in the category to emphasize certain categories over others in the total cell value. All category and relative weights were inferred from Yildirim [3] and Saha [4] and modified to accommodate the RPM. Category weights of 10 are assigned for population and landslide probability to emphasize their

[1] It should be noted however that relative prioritization is expected to correlate with cost, though a true cost map would imply a cost multiplier effect based on the weighted value of a given cell relative to another (as calculated in the model). A true cost map would require data currently unavailable for China regarding the actual cost impact functions of each parameter. Instead, the RPM method presented here represents a method of pipeline routing via ordinal ranking of all cells within the modeled grid for subsequent lowest-value path analysis.

importance over features such as roads and rivers, each of which have been assigned a category weight of 5. Protected areas are assigned a category weight of 7 to emphasize conservation in these areas as having a higher impact on costs and feasibility than crossing rivers and roads, but of lower importance than population and landslide probability**.** This again emphasizes the importance of the routing nature of the map as opposed to cost.

Each data layer must be standardized in order for multiple layers to be combined to form the finished RPM. All data layers were converted to a grid with an edge length of 250 meters, and a total area of 62 500 m^2 per cell. This cell size was chosen as an intermediate value between three lower resolution (1 km^2) grids and higher resolution features such as roads and rivers, which were obtained at resolutions of around 100 meters (10 000 m^2). The 250 m cell size was chosen to take advantage of the higher resolution of the roads and rivers data, because they are present across so much of mainland China. This higher resolution allows for better routing decisions without losing much detail. Once the RPM has been created, a shortest path function can be run between source-sink pairs to determine the best route. The shortest path is determined by adding cell values traveling from source to sink, where the best path reflects the lowest possible additive value for all cells traversed, therefore avoiding the least attractive paths, those with the highest value as defined by the combination of weighted features, where possible.

The RPM was constructed using ESRI's ArcGIS Model Builder, which provides an excellent framework for expansion and revision of the model at any time. This model was built under stringent time constraints, and therefore does not include all features that might have been included given more time (see Discussion and Conclusions for a discussion of suggested future improvements). The assumptions used to create the RPM are described in detail in the following sub-sections. These assumptions will likely be refined as the model undergoes sensitivity analysis and new information becomes available. The RPM however does include enough high quality data for the most heavily weighted (i.e., the most critical) parameters to provide a good approximation of likely pipeline routing criteria, even without including all possible data layers.

Landslide Probability

Landslide probability is one of the most significant considerations for pipeline routing and construction [4], and thus, for the purposes of this analysis, was weighted most heavily. Landslides create acute hazards capable of damaging or even destroying sections of pipelines. A landslide frequency map was obtained from the United Nations Environment Programme's Global Risk Data Platform [5]. The layer is a 1 km^2 grid of landslide hazards with a z-value denoting annual percent probability of slide occurrence. This grid was resampled to the standard 250 meter cell size used in this analysis. The methodology presented by Nadim and Kjekstad [6] was employed to estimate high risk landslide areas from the UNEP data. Per Nadim and Kjekstad, landslide probability values exceeding 0.01% were classified as "Very High Risk" and this class was assigned a value of 1000 to restrict any unnecessary routing in these areas. Probabilities between 0.00025% and 0.01% are considered moderate risk and were assigned high values of 6-9 to avoid routing through these areas. Lee [7] indicates that pipelines can be buried and slopes can be modified to avoid minor landslides in higher risk areas; however construction in these areas is likely to be more costly, justifying their classification as moderate risk areas. Low risk areas were assigned values of 2-4 to deter pipeline routing, but still make it feasible to route through these areas. All cells with a 0% probability were assigned a value of zero, to denote that pipeline routing is unconstrained by landslide hazards in these areas.

Population

Population is also a crucial component for routing pipelines. Populated centers provide unique construction considerations in which available land is scarce and expensive to purchase [4]. Also, heavily populated areas typically present a high density of obstructions within the built environment that must be bypassed for pipeline construction. It is also often difficult to obtain right-of-way passage rights and to obtain approval by land owners [7], which could increase the required pipeline length and cost. Thus, it was necessary to adjust for population gradients that previous routing studies have not accounted for [3], [4]. Weightings were applied to differentiate between high- and low-density populated areas in order to differentiate between the likely cost impacts at various population densities. Population is assigned the category weight of 10, the highest given in this analysis. The Oak Ridge National Laboratory's Landscan dataset [9] was used as the population layer. In addition to available population data, this dataset estimates population density based on such factors as, land use, slope, and satellite imagery of nighttime illumination, among other factors. It must be emphasized that this grid only estimates population density; it is not a direct measure of where people are. Given this caveat, it is currently the highest resolution population data available free of cost. This is a 1 km^2 grid with a unique population density for each cell and was resampled to a 250 m cell size. Concentrated urban centers are assigned a weight of 1000 where the population density is greater than 10 000 people per km^2, which corresponds to the suburban limit described by Pozzi and Small [10]. Semi-urban areas are defined by population densities between 3500 and 10 000 people per km^2, corresponding to the average urban population density for the whole world [11]. This value approximates concentrated urban areas not located in the center of cities, but which are still very difficult to navigate pipelines through, and these areas are given a relative weight of 9. Suburban areas are defined as those having population densities between 400 and 3500 people per km^2. Demographia [11] states that the traditional world definition for an urban area starts at 400 people / km^2, and thus defines the lower bound for suburban areas. Routing in these areas is still somewhat discouraged with a weight of 4, but still making them more suitable for pipelines than urban or semi-urban areas. The final category is rural populations, which are either agricultural or undeveloped land. Rural areas are defined by population densities between 0 and 400 people per km^2. These areas are assigned a value of 0 to connote the lack of population-related constraints on pipeline routing due to less existing infrastructure and cheaper land values in these areas.

Protected Areas

China is home to hundreds of protected areas, with numbers expected to increase over the next 20 years [12]. On a national scale, it may be unavoidable to route through certain protected areas in China due to their large areas being in close proximity to some CO_2 point sources. The World Database of Protected Areas (WDPA) [13] was used to model these protected areas. A dataset from 2003 was used instead of the more current 2008 database because it offered higher spatial resolution. This database includes polygons reflecting the actual shape and location of many protected areas, and also points indicating the location and area of other protected areas not represented by polygons, though these point locations lacked any representation of shape. The point locations were extrapolated to areas using the buffering tool in GIS, which creates a circle around each point sized as a function of the areal extent of the protected area. These buffered areas are intended to approximate the extent of the protected areas with point representations, but do not reflect the actual geographic shape of the areas. Some buffered areas overlapped, which would cause misrepresentation via added weights in the RPM. Buffered area overlaps were excluded from the layer to avoid this problem. This is an acceptable omission because the buffered areas do not represent actual geographic coverage, and can be excluded when causing problems. The resulting layer is a combination of polygons and buffered protected areas with weightings based on their International Union for Conservation of Nature (IUCN) classifications.

Protected areas are given a category weight of 7 to emphasize conservation of internationally recognized areas. The protected areas were assigned relative weights based on their relative importance to society inferred from IUCN Guidelines [14]. All protected areas have been designated based on their value to humanity, and ideally all would be free of pipelines, but for this study it was necessary to rank and weight the protected areas according to their relative degree of protection. Category I protected areas are intended to be free from infrastructure and as much human influence as possible and are therefore assigned values of 1 000 to incorporate these constraints into the pipeline routing model and thus effectively marking these areas as extremely unattractive for routing except for very short distances or as a very last resort. In contrast, areas VI and V are defined by daily human interaction and alteration within the protected area such as food gathering and firewood collection, but are still discouraged with a relative weight of 3, but can be routed through if necessary. Categories II and III represent important protected areas, and are assigned values of 9 to deter nearly all routing, but if necessary these areas can be routed through. Category IV protected areas aim to protect individual species, and are given a weight of 5.

Slope

The slope layer was obtained from the United States Geological Survey [15] at a 1 km grid cell size. This grid was resampled to a 250 meter cell size to standardize with the other layers. Slope is an important parameter for pipeline routing, and is given a category weighting factor of 5. The methodology for slope weighting is used from Yildirim [3]. This methodology is used for oil and gas pipelines, and can be used as an appropriate proxy for CO_2 pipelines for the purpose of the RPM.

Infrastructure

Infrastructure modeled on a national scale creates challenges for accurately determining pipeline placement. Infrastructure in the RPM includes roads and railroads. Road and railroad layers were obtained from the 2000 China Census data provided by the China Data Center at the University of Michigan at Ann Arbor [16]. It is important to include every possible road and railroad to obtain the most accurate RPM possible. This was the best available data, and included highways, national roads, state roads and railroads. This study excludes local roads, which will likely never be included in the national RPM due to problems with data availability and resolution. Weightings were adapted from Yildirim [3] based on the available data. Highways were given the highest value of 9 and National and State roads were given lesser values to account for less area and greater ease of routing pipelines. The category weight for all types of infrastructure is a 5.

Rivers and Lakes

Rivers and lakes were modeled using two datasets. The river dataset was obtained from Harvard's China Earthquake Geospatial Research Portal [17]. This dataset includes 4 orders of streams, which were given relative values based on the 6-order stream classification used in Saha [4]. The lake dataset was taken from the 2000 Census dataset used for infrastructure. This dataset also includes large rivers that are modeled by the rivers dataset as well. Lakes are given a value of 7 to deter routing directly through the center of the lake, due to a cumulative effect occurring when routing the shortest path for pipelines. Rivers and lakes are assigned a category weight of 5.

Excluded Layers

Due to the scope of this project, it was not possible to obtain reliable land use data for inclusion in this iteration of the model. Land use typically represents one of the most highly weighted layers in a traditional cost map. Land use includes naturally occurring land formations such as permafrost and forests, as well as human altered landscapes such as cities and farms. This layer was implied in the population weighting, because land use typically correlates to population density [18].

A reliable geology layer with bedrock type was also not able to be obtained for the RPM. Hard bedrock is more difficult to drill through and is more costly to anchor pipelines, compared to softer bedrock such as sandstone [4]. Because of this impact of bedrock type on construction costs, this parameter is always often considered in pipeline models. However, it is typically given a low relative weight, so excluding it from this analysis is not likely to have major impacts on the final RPM.

RESULTS

The synthesis of the all of layers described above was accomplished via the model shown in Figure 1, and resulting in the RPM shown in Figure 2. Blue ovals indicate input layers, while green ovals are intermediate layers calculated to obtain the final weighted layer. The yellow boxes are all operations performed on data layers in order to standardize each layer with the relative weighting scale, and limit them to the extent of China's borders. These operations include resampling, reclassifying features to relative weights, converting features like roads and rivers to grid cells, and clipping layers to China's borders.

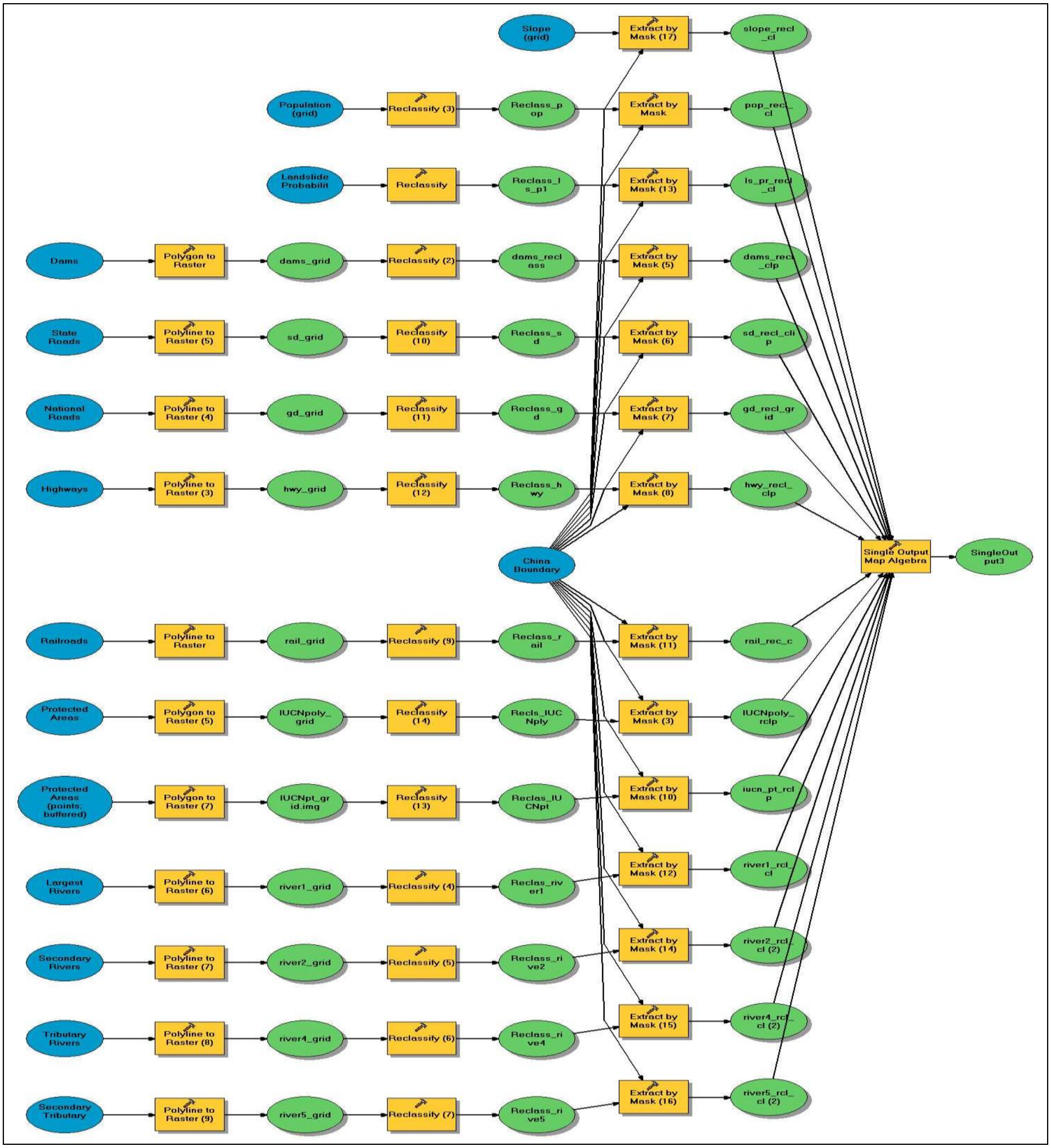

Figure 1: *Workflow diagram showing data processing steps in creation of the final routing priority map.*

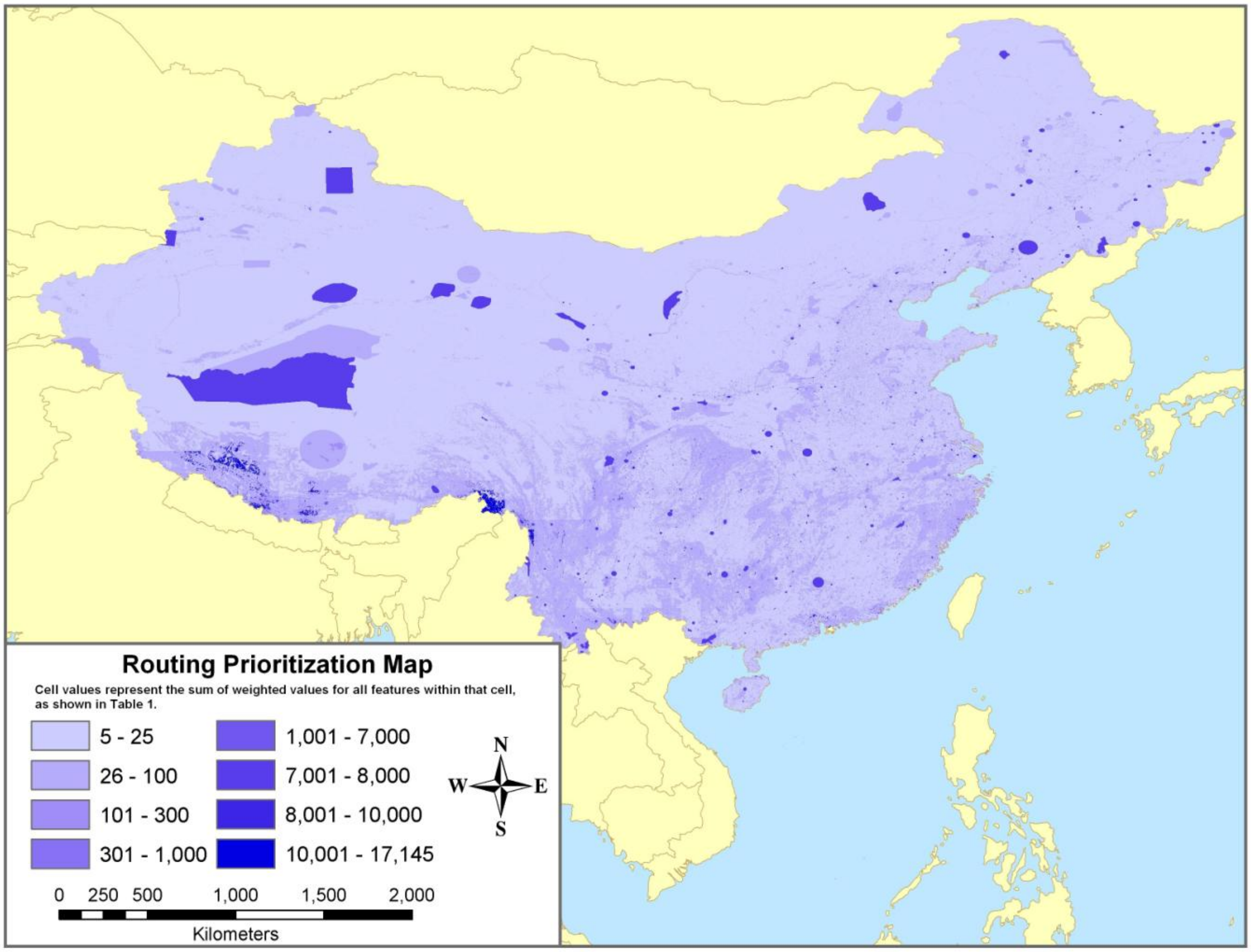

Figure 2: *Final routing prioritization map showing differentiated suitability for each 250-meter grid cell*

The differences in color on the map (Figure 2) denote differences in routing priority, with the lighter cells representing more attractive areas and the darker cells representing those areas that are less desirable to route through. Level 1 protected areas and urban centers, where their respective weights are 1000, display as large polygons with darker shades of blue, and are the easiest features to identify on the map. The protected area buffers display as circles with varying shades depending on their IUCN classification. Semi-urban and suburban areas are most highly concentrated in southeast China and contrast slightly darker compared to background values. High landslide probabilities, which display as the darkest shade, are present in Himalayan Mountain Range in southwest China. Infrastructure and rivers cannot be clearly identified on a national scale, but display accurately when zoomed in.

This study did not allow time to test a shortest path function on the RPM, however the map is ready to be used to analyze the complex set of point sources and sinks in China using a program to pair multiple sources and sinks with the least-cost distance provided by the RPM.

DISCUSSION AND CONCLUSIONS

This relative routing priority map will be applied to help model realistic pipeline routes for potential CO_2 pipeline deployment in China. The RPM is the first attempt to model pipeline routing for CCS throughout all of mainland China. The RPM is intended to inform potential pipeline routing by applying the shortest path function. Using these modeled pipeline routes, costs can be estimated by adapting existing cost methodologies. This future cost estimate will eventually be used to inform discussion of the economic feasibility of using commercial-scale CCS for reducing CO_2 emissions in China.

Due to time constraints, the first iteration of the model does not include all features desired. The excluded layers will be included in future iterations. There are also noted overlap issues occurring between the rivers and lakes layers that need to be corrected on a case by case basis. The next iteration will also include a feature class descriptor to convey the number of cells for each feature traversed by each pipeline. Peer review and sensitivity analysis will likely lead to future modifications of category and relative weights. During this process, the cumulative effect occurring in the shortest path function will be examined to determine the optimal weighting for the large geographic coverage layers such as protected areas and lakes. Future refinements will also include an attempt to adjust relative weightings to reflect true relative costs in order to develop a cost map with realistic cost assumptions as opposed to the RPM presented here. The RPM currently has not been tested or applied in a final cost model, though progress is underway.

ACKNOWLEDGEMENTS

This project was supported by the U.S. Department of Energy at the Pacific Northwest National Laboratory and the Office of Science's Summer Undergraduate Laboratory Intern (SULI) program. Additional funding was provided by the Global Energy Technology Strategy Program. A special thanks to my mentor Casie Davidson who provided me with the flexibility, support and encouragement I needed to make this project possible. The model would also not have come to maturity without greatly appreciated help from Sarah Finne and Robert Dahowski.

REFERENCES

[1] Energy Information Administration, "China Coal," July, 2009, http://www.eia.doe.gov/emeu/cabs/China/Coal.html.

[2] R.T. Dahowski, X. Li, C.L. Davidson, N. Wei, J.J. Dooley, and R. Gentile, "A Preliminary Cost Curve Assessment of Carbon Dioxide Capture and Storage Potential in China," *Energy Procedia*, vol. 1, no. 1, pp. 2849-2856, February 2009.

[3] V. Yildirim, R. Nisanci, T. Yomralioglu and B. Uzun, "TRANSPORTATION - Raster-based GIS data guide economic pipeline construction," *The oil and gas* journal, vol. 106, no. 13, pp. 62-68, April, 2008.

[4] A.K. Saha, M.K. Arora, R.P. Gupta and E. Csaplovics, "GIS-basedroute planning in landlisde-prone areas," *International Journal of Geographical Information Science*, vol. 19, no. 10, pp. 1149-1175, November, 2005.

[5] United Nations Environment Programme, "Global Risk Data Platform," May, 2009, http://preview.grid.unep.ch/index.php?preview=home&lang=eng.

[6] F. Nadim and O. Kjekstad, "Assessment of Global High-Risk Landslide Disaster Hotspots," in *Landslides - Disaster Risk Reduction*, 1st ed. K. Sassa and P Canuti Eds. New York, Ny: Springer Verlag, 2008, ch. 11, pp. 213-221.

[7] E.M. Lee, J.M.E. Audibert, J.V. Hengesh and D.J. Nyman, "Landslide-related ruptures of the

Camisea pipeline system, Peru," *Quarterly Journal of Engineering Geology and* Hydrogeology, vol. 42, pp. 251-259, May, 2009.

[8] E.R.A.N. Smith, K. Michaud, J. Carlisle, "Public Opinion about Energy Development: Nimbyism vs. Environmentalism," Prepared for delivery at the Annual Meeting of the American Association of Public Opinion Research, Phoenix, Arizona, May 13-16, 2004.

[9] Oak Ridge National Laboratory, U.S. Department of Energy, "LandScan™ Global Population Database," http://www.ornl.gov/sci/landscan/index.html.

[10] F. Pozzi and C. Small, "Exploratory analysis of suburban land cover and population density in the U.S.A.," *Remote Sensing and Data Fusion over Urban Areas, IEEE/ISPRS Joint Workshop 2001*, pp. 250-254, November, 2001.

[11] Demographia, "Demographia World Urban Areas & Population Projections: 5th Comprehensive Edition," April, 2009, http://www.demographia.com/db-worldua.pdf.

[12] J. Lopez-Pujol, F. Zhang, and G. Song, "Plant biodiversity in China: richly varied, endangered, and in need of conservation," *Biodiversity and Conservation* vol. 15, no. 12, pp. 3983-4026, November, 2006.

[13] The China Archive, Texas A&M University, "Protected Areas in China Data Set," http://chinaarchive.tamu.edu/portal/site/chinaarchive/menuitem.5d929e440604b297140eadb4f00011ca.

[14] N. Dudley, "Guidelines for Applying Protected Area Management Categories," International Union for Conservation of Nature, pp. 1-86, 2008.

[15] Earth Resources Observation and Science, United States Geological Survey, "HYDRO1k Elevation Derivative Database," January 5, 2009, http://eros.usgs.gov/products/elevation/gtopo30/hydro.

[16] China Data Center, University of Michigan, "2000 China County Population Census Data with GIS maps," July 23, 2009, http://chinadatacenter.org.

[17] Harvard University, "China Earthquake Geospatial Research Portal," July, 2009, http://cegrp.cga.harvard.edu.

[18] K. Kok, "The role of population in understanding Honduran land use patterns," *Journal of Environmental Management*, vol. 72, no. 1-2, pp. 73-89, August, 2004.