DATA AVAILABILITY AND NEEDS
Dave Weekly

Modeler's Task: The modeler needs more than the PNG's10 steps, he needs to understand the engineering and economic relationships and how they are integrated. The modeler needs to show decision makers what key drivers are in analysis; and often must do this without much time, money or data.

DATA: There are two main databases; the Waterborne Commerce Statistics Center (WCSC) and the Lock Performance Monitoring System (LPMS). There are also vessel operating costs developed by IWR

WCSC data:

- This database strives to be a 100% sample of commerce on the inland waterway. All operators are required to report their point-to-point (a.k.a. dock to dock) movements. These are waterway movements only, and do not include the landside legs of the original origin or final destination.
- The operator is the towing company. They are required to report all types of commodities, including passengers & containers. They must also report route (if more than one route, we need to know how it got there.)
- There are 146 commodity types in the database.
- In 2001 582.6 M tons where reported traversing 4000 unique origin destination pairs. There are 51,000 origin-destination-commodity triplicates.
- Users of this data are data challenged because of the internal changes in the database. It is difficult for a programmer to go back past 1990 because the format has changed.

LPMS Data

- Flotilla Information:
    - The flotilla record was added in recent years. It includes the set of barges going through the lock.
    - Vessel types is provided for powered vessel and barges (not detail- just category reporting- type and general stats)
    - This data set has much less specificity on commodity type and tonnage compared to the WCSC data (LPMS records on tonnage are maybe 5% lower than WCSC in this area).
- Issues in reporting:
    - Problems in accuracy because sometimes one chamber is working okay while the other is not and outages are sometimes not reported if no barges are present.
    - How long does it take to get through the lock? Many will report only 1 number for processing time. There are many different types of times that could be reported. Some examples: "start of lockage", "approach time" (time to transit the lock approach), "bow over sill" (when the vessel enters the lock, "end of lockage". This is important because speeding the approach and

exit of lock may produce many efficiencies. It is important that the actual times be recorded.

Comparing the two data sources:

- Barge Data: WCSC data is best for detailed barge information because it is more detailed and comprehensive. Use LPMS data for counts on number of loaded barges and empty. The quantity of loading (tonnages) is more accurate in the WCSC data, while the LPMS is best for powered vessel (that is associated with set of barges) and flotilla data.
- Lock operators provide the LPMS data, while vessel operators report the WCSC.
- Availability of data (timing of availability): Waterborne 2001 data just became available in 9/02. Meanwhile, there is only a 2 month lag in the LPMS data.

Operating Costs:

- IWR provides vessel operating costs for Deep and Shallow Draft vessels, provided on the HQ guidance website with the current interest (discount) rate. The data on operating costs is very detailed by type of ship and size of vessel.
- Operating cost are estimated using the flotilla data and applying the IWR costs
- Other types of costs: Outside the VOC, there is not much detailed cost information available.
    - Value of commodity on vessel and inventory-holding costs can be important to know. How to factor that into the analysis? In LRD bulk commodities don't have much impact on the delay costs (1-2% of costs in LRD) which is trivial.

Other Data

- OMBIL data is not available to the general public. It contains information from the specific site in layered format. It has data on lock performance, output and costs (from Corps of Engineers Financial Management System (CEFMS), on monthly data). It has information on Corps hydropower plants including all outages. This information is useful because it makes comparisons of spending levels in different areas are possible. It is secure for Corps employee to use, and it is on an internal web. There is no method for contractors to get access unless through a Corps employee. Training for OMBIL is available through the PROSPECT (Corps training) program.
- Port Series Manual- not all ports but the bigger ports- comprehensive record including all of the terminals for most of the nation's larger ports (not all ports are included).
- Characteristics of all locks in system.
- Dredging costs are also available.
- Price data: WCSC data does not include price. Without the ability to relate the available information to price it has limited value. With some commodities, grain and coal for example, there exit time series on spot rates back to 1990.

RATE DATA

- Over time various projects have collected rate data. Is there a method to access the rates? Rate information can be proprietary, so it cannot be given out except in aggregated format.
- LRD does an index of movements on its rates every year. TVA has a long history of rate analysis (since the 1970s) but the information is better since the 1990s. Corps wide there is no systemic collection of rate data. No public storage of rate data- the data has been shared with some but they have to go through an extensive process to show that the information will be used only for COE related projects. Some shippers take the information very seriously- rates can define their competitive edge- if the information is going to be given away they are less likely to provide it.
- To do the analysis- original origin to final destination- need to access the total cost of movements. We have high quality data for the water portion of the route, but this must be supplemented to get the entire trip cost. Additional movements include the land side cost, but this information is less clear because of the difficulty in getting at the rates on the landside.
- Index updates are used to update the landside cost. We use an 85% sample of the basin for updates in LRD.

FORECASTS

- WCSC gives a good history, we know who the shippers are, but the data doesn't say who the drivers are; why they are operating there, what will they do in the future? Different futures- what things will effect the future? How does congestion affect the WTP? Are there alternative sources for the commodity and how might that play into things for the future?
- COE does forecasts of the unconstrained condition. Movements are forecasted without regard to the capacity of the locks.
- Need to be aware of the relevant range of forecasts.
- COE has operating rules in some areas, such as 3 cuts set aside for commercial, then the 4th for recreational if any are waiting. June through August can have a really high percentage of the locks eaten up by recreational users. What is experienced in terms of commercial delays as a result of this restriction? The traffic is not distributed evenly-there are different commodities at different times- not same.
- When forecasting future conditions, how do management measures (improved efficiency measures) fit into the analysis? Are they factored into the without project condition? Why not rearrange the priorities to allow the commercial traffic to go first? Political and legal constraints may prevent such an analysis, or implementation of such an alternative. If considered in the analysis, such alternatives would be non-structural alternatives in the with-project condition. P&G states that everything legal to be considered should be rolled into the without project condition, whereas something beyond existing legal bounds should be considered outside as part of the with project alternative.
- Relating forecasts to delay: Annual tonnage forecasted through each lock will define the lock specific delay. The delay time and quantity is then used to compute the additional travel cost.

- Hazard functions are used to represent structural reliability. They are directly related to the forecasts. The forecasted use of the structure implies the degree of structural wear and tear. This complicates the analysis because performance of the locks drives the maintenance costs. There are various maintenance requirements that could require closures which would influence system performance (closures may equal delays). If we have many components needing repair it may require down time, and this will affect the way a particular lock is operated.

MODELING

Waterway networks:

- A typical waterway diagram has links and nodes representing different rivers and tributaries. The Huntington model network typically has 200 ports. There are 90 miles of tributaries in the OH river model. It is important to have specific barge information (an 8 barge tow can be followed by a 5 barges tow) and to model the different segments for costing.
- The navigation models need to be integrated with environmental models. The number of trips and vessel characteristics are important to understand how the environment will be impacted (e.g. wake erosion of riparian habitat). We also need the seasonality of vessel movements as well because there may exist different types of species at different times of year.
- The navigation model is on an annual basis, but what about seasonality?
    - We have looked at the Ohio mainstem and other places. No seasonality or time of day affects where found. However, environmental impacts are seasonal and need to be married to the annual navigation forecasts. The environmental information (now seasonal) can be converted into an annual basis or the annual economic data can be converted into seasonal movements.
- Equilibrium simulation models have been done with one lock and have shown no major system effects. However, when we tried to do system analysis with a large system it fell apart. This gets at the issue of what we are doing, should be doing, and what we are trying to do differently.
    - Items that really make a difference to the answer should be done
    - Also an issue is which of the important items in the system can possibly be reproduced?
    - Factor of transparency, many system models are not transparent. The more items that are added to the model the harder it is to make the model transparent.
- The project will have some impact on recreational use. There is extreme peaking of recreation traffic in certain months. However, it doesn't have a major impact on the commercial traffic but is significant enough to drive the NED plan.
- The goal of modeling is to capture the changes in a waterway system due to project improvements. We want it to be realistic and consistent with theory so as to satisfy the comments made by NAS.

- Where we are with modeling:

    - Green: (Have a good understanding of these parameters and they are included in analyses.) System analysis, modeling lock congestion, rates, structural reliability
    - Yellow: (Are making progress in incorporating these aspects into the models). Assessments of nonstructural alternatives; better integration of engineering, economics and environmental components
    - Red: (Work is needed to understand and model these aspects.) Forecasts and uncertainty; sensitivity of barge traffic to rates (WTP)
- Transportation rates fall into the green category for what we are currently modeling. But there is a data gap because we don't have rates for all of the landside legs (original origin to the water and water to the final destination). The data is not currently available uniformly. However, we do have some info on some of the different areas, but the rates when linked to the WTP item fall in the "red" category.
- The person on the margin is the one that really matters in the analysis. We need to examine the movements and identify those that have suspicious willingness to accept delays. Those suspect movements are the ones that need detailed data collection.
- Forecasts and uncertainty. We use scenario analysis now. In the past we did sensitivity analysis by adjusting the forecast up or down X%. With scenario analysis we say "hey, what if…" Making an effort to move this into the green grouping.
- In forecasting we are looking at what the key drivers will be on the different scenarios. For example, we looked at the 1990's to see what drove prices and demand for coal in order to use that information in making better projections on future trends

DISCUSSION:

- In the presentation you stated that you have to limit the analysis because of time and budget constraints. Who imposes time constraints and why don't you have enough?
    - These are complex analyses. There are issues such as how far do you go with the analysis? Regardless of your results there is something left undone and more being asked for. There is much that has to be done in the study, and we will always be faced with time constraint.
- Recommendation for dealing with data.
    - Start with the grossest information available first. The summary statistics provided on the executive summary cards is a good place to begin, if that isn't sufficient, then go to the public domain info, not the dock-to-dock information. If that level of specificity is needed then go to the master file; this information is proprietary. The NDC CD contains the detailed data and is GIS linked, which helps in visualizing what is going on. If specific information for the different locations is needed then call NDC and have them access the master file.
- Are arrivals random at the lock? Wouldn't it seem logical that once they pass through one lock the next arrival is predictable?
    - Studies have shown that if the locks are within 5 miles of each other then the first lock does act as meter for next. If the distance is greater than 5 miles the inter-arrival time at the next lock looks random with a Poisson distribution.

This is explained by the vagaries of the system (tows leaving the pool, tows entering the pool and differences in tow performance).
    - Can't schedule at one lock because of the interactions.
- If construction of new capacity on the waterway gives relief to rail and road congestions; is this a potential source of benefits?
    - In the analyses we typically assume the overland transportation cost is constant. Don't measure impacts on other modes. If it is a large component, this could be significant to the analysis. (Have potential for congested waterway to dump traffic on the other modes- such as an already congested road system. Conversely if a project relieved the waterway congestion and diverted traffic from the road, it is an additional benefit because of the reduce road congestion.)
- The flip side of that is that there may be some economies of density if shipments are being diverted to another source, such as rail, that may actually result in negative benefits to the project because the other mode is better off with the density benefits
- How are DOT costs of maintaining the roadways incorporated into the analysis?
    - Transport projects impact more than just the rates. The analysis needs to be multi-dimensional with rates; working in cubic space with rate differentials and delay differentials. The analysis needs to look at rising costs in the alternative modes as it may reduce costs and transits
    - Similar issue with deep draft. If you cause diversions from Houston which is high emissions area, there are additional savings. Shippers are considering these impacts when selecting destinations.
    - Need to be careful when developing scenarios. Waterway movements are likely to have a road movement associated with them. An increase in waterway traffic may actually end up increasing the road traffic as well. Therefore, we need to consider all of this when doing scenario analysis in order to actually get the traffic shifting in the right way.