PRC Docket No. N2006-1

USPS Library Reference N2006-1/17

END Optimization Modeling Report

DRAFT: For Internal Use Only

UNITED STATES POSTAL SERVICE™

Optimization Modeling Report

Evolutionary Network Development

Table of Contents:

1 Introduction **2**

2 Distribution Concept **3**

3 Overall Formulation **4**

3.1 Cost Modeling 4

3.2 Translating Volume Into Workload. 8

3.3 Service Considerations 10

3.4 Capacity Constraints 10

3.5 Modeling Constraints 11

3.6 Data Inputs 13

4 Modeling Steps **17**

4.1 Optimization Decomposition 17

4.2 Optimization Iteration 18

4.3 LogicNet Plus Model Refinements 19

1 Introduction

This report will provide an explanation of how the optimization modeling was developed and implemented for the USPS Evolutionary Network Development (END) program. Although this model focuses only on the optimization and cost modeling components of the END program, these elements represent the first two steps in our overall modeling approach.

Simulation is used to estimate several metrics for a pre-defined process in a "test" environment. It is also used to conduct "what if" scenarios by modifying selected inputs to a specific process in a simulated environment. Unlike simulation, optimization can be used to calculate an optimal solution given certain constraints, inputs, and an objective function to solve. In our approach, we use optimization to determine and develop a network and then we use simulation to test that strategy and understand the effects of certain "what if" scenarios.

USPS manages one of the most complex and disperse distribution and processing networks in the world. Thus, meeting demanding service requirements, while simultaneously shedding excess cost, is a difficult challenge. To approach this challenge, USPS has worked diligently to formulate a model that meets the critical operating and management requirements for strategic analysis, yet is feasible to solve with today's computing technology.

In summary, the following techniques were used to allow us to solve this problem:

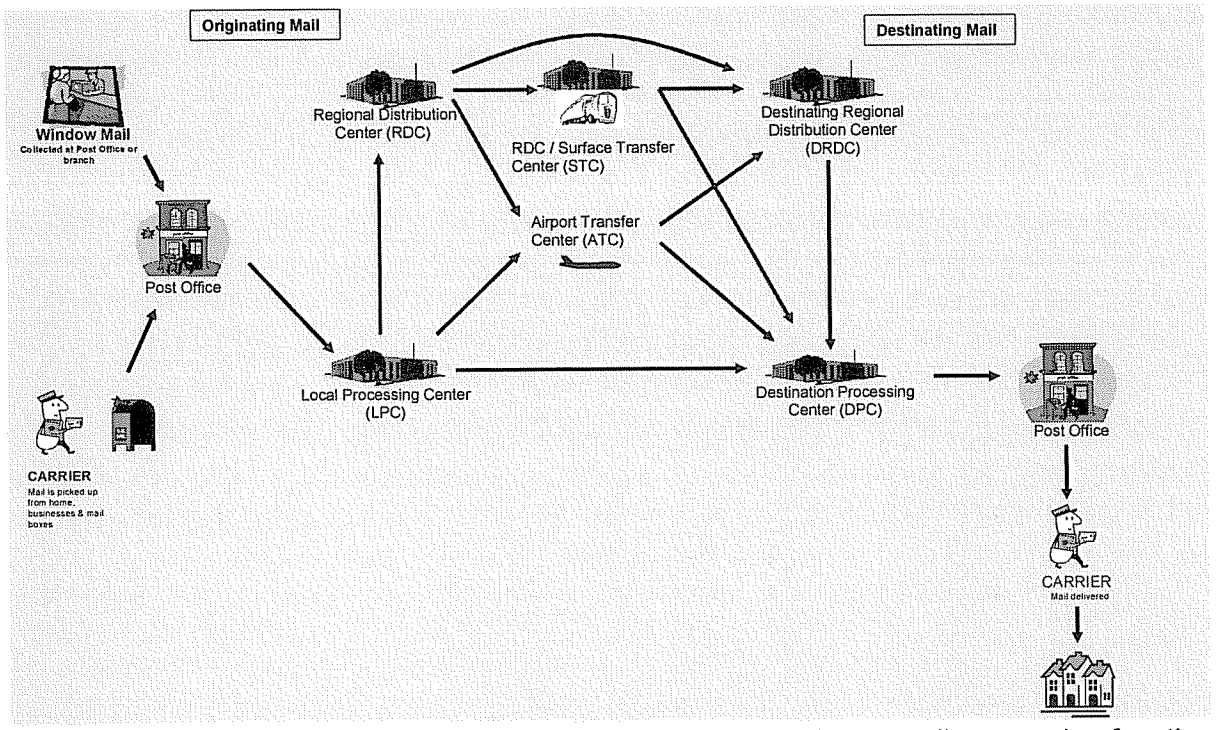
- Rather than requiring the model to determine what operations would look like in different parts of the network, the model assumes a pre-defined, standard distribution concept. This concept specifies the types of facilities that exist and the roles associated with those facilities. This reduces the number of decision variables in the model.
- Aggregate volumes at the 3-Digit level instead of the 5-Digit level. First, reliable mail flow data at the 5-Digit level are not available. Second, aggregating volumes at the 3-Digit level significantly reduces the computational complexity of the model.
- Split the solver into two, almost identical versions of the model: (a) Regional Assignment Model and (b) National Assignment Model. Both models simultaneously determine the optimal locations for consolidation operations while assigning each 3-Digit ZIP Code to an originating/destinating processing facility (LPC). There are six Regional models that analyze a segment of the country and the National model evaluates all of the country at once.

Even though we are "splitting" the problem, we will still answer the fundamental questions that we need to in a structured, unified manner by iterating back and forth between sub-models. Each sub-model will use the solutions of the other sub-models to inform the overall model solution. This interaction will serve to overcome the effects of splitting the models up.

2 Distribution Concept

USPS' goal is to develop a more efficient network designed to handle multiple products with a trend toward more shape-based mail processing streams. The new network design will simplify both mail processing and transportation flows. See Figure 1 below.

Figure 1: RDC Distribution Concept



The illustration above depicts potential mail flows between different mail processing functions in the future network. Note that in the future network different functions depicted in the illustration may end up being performed at the same location.

Essentially the backbone of the network's infrastructure is a Regional Distribution Center (RDC). RDCs will consolidate parcel and bundle distribution to take advantage of shape-based efficiencies. They will serve as mailer entry points and Surface Transfer Centers (STC) to enable shared product transportation. They will act as concentration points for subordinate Local Process Centers (LPCs). LPCs will handle most of the letter and flat process workload for both originating and destinating sorts. Destinating Processing Cetners (DPCs) handle the same shapes as LPCs but only conduct destinating sorts and do not have outgoing processing. Some RDCs are co-located with LPCs, where both roles are supported in one geographic location; we call these COLOCs. Air Transfer Centers (ATCs) facilitate the exchange of mail with the air carriers.

The overall modeling formulation assumes this distribution network concept and finds ways to implement it in the most cost effective way.

3 Overall Formulation

The objective of the Assignment Model is to minimize the complete transportation and processing costs of the system. The assignment model needs to make two decisions:

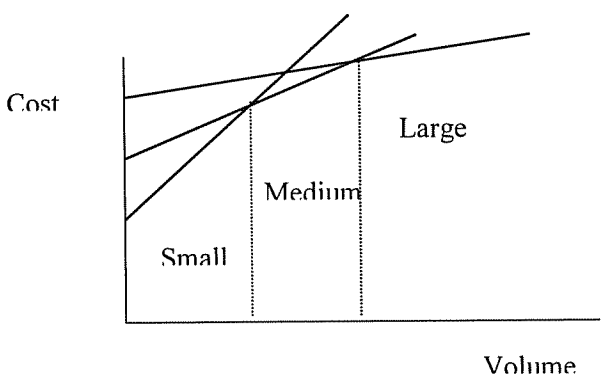
- Every 3-Digit ZIP Code must be assigned to one LPC for all originating processing and one DPC for destinating processing.
- Every LPC must be assigned to at most one RDC for each mail type where consolidation activities occur.

3.1 Cost Modeling

For the model to make the best decision, it must have good cost information. While cost equations can be complex (non-linear), mathematical programming solutions become very difficult to attain when non-linear cost equations are used. Therefore, we assume that all costs can be modeled by a cost function that has both a fixed cost and a variable cost component. In these models, a cost of a facility, transportation link, etc. can be represented with a fixed value for opening the facility along with a per unit cost for each mail item that flows through the facility. These costs can be made up of many sub-costs, but must aggregate as a fixed/variable cost.

For example, facilities that do more of some process have a lower per-unit cost for that process. It seems sufficient to linearize this aspect by having three sizes: small, medium, and large, and allowing a different cost for each of them. This size, however, can vary by product type. So, depending on the actual volumes assigned to a facility, a facility can be small for letters, medium for flats, and large for parcels. Each of small, medium, and large has a fixed cost component, and a variable cost component, leading to a cost curve of the following form:

Figure 2: Facility Size Cost Curve



The model will cover both current P&DC operations and anticipated roles by consolidators and dispersers. The cost model will be used to help the optimization model answer strategic questions. For these types of questions, the costing efforts will focus on replicating the costs of the essential mail flows, determining the costs associated with the work done by consolidators and dispersers and identifying the types of facilities in which these additional roles will be located.

In developing the structure of the cost model, we considered two dimensions, the coverage or "breadth" of the modeling and the mathematical structure or "depth" of the model.

3.1.1 Determining the Breadth of the Model

The breadth of the model is determined by the needs of the optimization model and the requirement to accurately represent mail processing costs. These two criteria lead to explicit modeling, in terms of individual cost functions for all of the main direct sorting and allied operations for letters, flats, and parcels. Cost functions will be developed for the following individual sorting operations:

- OCR
- MPBCS
- DBCS
- Manual Letters
- Manual Flats
- FSMs
- Manual Priority
- Manual Parcels
- SPBS - Priority
- AFCS
- APPS
- AFSM100
- PSM
- NMO

In addition, cost functions will be developed for the major allied operations. Although these operations do not have direct sorting responsibilities, they are a major part of mail processing labor costs and are required to facilitate sorting. Five allied operations were modelled:

- Platform
- Opening Unit Pref
- Opening Unit Bulk
- Pouching
- SPBS Other (mostly Lips/Rapistan)

Together, these operations represent over 70% (as of 2001) of mail processing labor costs at P&DCs. The remaining costs will be modeled more simply, for one of three reasons. In some cases, the cost of the operation can be linked directly to a modeled sorting operation, as in the case of RBCS costs, which will be linked to OCR costs. In other cases, the operation's cost is not part of the optimization analysis. For example, the model does not contemplate "consolidating" facing and canceling operations. These operations will thus stay in the P&DCs. Although these costs must be accounted for in determining which P&DCs will be operating, and how ZIP Codes should be assigned to P&DCs, there is no requirement for detailed cost modeling of their characteristics. Finally, some operations are simply too small to justify detailed modeling and increased complexity in the optimization model. For example, the empty equipment operation generated only $38 million worth of cost in FY 2000. Need to confirm when these cost functions were last updated?

3.1.2 Determining the Depth of the Model

One of the key issues in the optimization model is whether or not the Postal Service can obtain economies of scale and scope through consolidating operations. Thus it is essential that the mail processing cost model allow for such economies. At the same time, the optimization model is highly complex and imposes severe demands on the computing algorithm. Consequently, our model structure is guided by two considerations: 1) our attempt to best represent the actual cost within facilities incorporating both economies (and diseconomies) of scale and scope and (2) the need for the optimization model to be tractable.

Scale and scope economies require unit costs to vary with the level of production, but this typically implies complex functional forms that are computationally difficult. Accurate approximations of complex functions can be obtained by applying simple affine functions and then allowing the parameters of those functions to vary with operational size. Based on marginal accuracy and simplicity, three linear functions were used to approximate the non-linear functions for each operation. The model is structured so that as an operation increases/decreases workload, the linear function used would change as to yield the minimum of all three; therefore closely following the path of the non-linear curve. We have classified these three linear curves as "small", "medium", and "large".

This means that for any direct operation, we model the cost function as follows. For any direct operation j for a particular volume range i:

$$C_{ij} = F_{ij} + \delta_{ij} \; v_{ij} \; ,$$

where: i = volume range (small, medium, or large)
j = direct operation

where the F_{ij} represents the fixed cost associated with that specific operation of size "i" (e.g. small, medium or large), the δ_{ij} represents the variable cost associated with handling a piece in the operation of a given size, and the v_{ij} represents the workload in the operation.

This formulation can be contrasted with allied labor operations in which the fixed cost will not be associated with a specific shape or product family. Having common fixed cost in the allied labor provide a mechanism for incorporating economies of scope. In addition, allied labor operations may be driven by more than one operational volume. We can specify an allied labor operation, k, as:

$$C_{ik} = F_{ik} + \sum_j \sum_i \alpha_{ijk} \; v_j$$

where: i = volume range (small, medium, or large)
j = direct operation
k = allied labor operation

Note that most of the α_{ijk} will be multiplied by zero volumes. That is, a particular facility will only have non-zero volume in one of the volume ranges for each of the direct operations. Finally, we may wish to consider an overall common fixed cost, which would be associated with opening the facility and not related to any specific operation. This also is a mechanism for generating economies of scope.

With these pieces we can build the cost equation for the entire facility. It would include any common fixed cost (F) and the sum of the costs in the direct and allied operations. The cost function in each operation is approximated by an affine function with parameter choice dependent upon operation size. Despite the complexity of choosing the appropriate parameters for any particular facility (Z), the overall cost equation remains linear:

$$C_z = F_z + \sum_j \sum_i (F_{ij} + \delta_{ij}\, v_{ijz}) + \sum_k \left(F_{ik} + \sum_j \sum_i \alpha_{ijk}\, v_{ijz} \right).$$

where: i = volume range (small, medium, or large)
j = direct operation
k = allied labor operation
z = facility

While this function looks complex in its general form, it is quite straightforward in application. The process of using it is as follows:

- Determine the volume of piece handlings in each operation. Based upon that volume, classify the size of the operation (e.g., small, medium or large.)
- Choose the appropriate parameters for each operation.
- Combine the operational cost functions into the overall cost function

To make this a bit more concrete suppose that facility Z has 2 direct operations, BCS letter sorting and manual flat sorting and 1 allied operation, the platform. Now apply the costing process:

- Step 1: Classify the size of the operation (e.g., small, medium or large). Suppose the data show that the BCS operations is "medium," and the manual flat operation is large. The platform operation, based upon the sizes of the direct operations, can be classified as "medium."
- Step 2: Choose the appropriate parameters for each operation. The model allows for varying parameters by operational size. Given there are two direct operations and two parameters per function, this means that the model allows for 12 different direct sorting parameters:

Direct Sorting Parameters				
Operation Size:	BCS Fixed	BCS Variable	Man. Flat Fixed	Man. Flat Variable
Small	$F_{BCS,S}$	$\delta_{BCS,S}$	$F_{MF,S}$	$\delta_{MF,S}$
Medium	**$F_{BCS,M}$**	**$\delta_{BCS,M}$**	$F_{MF,M}$	$\delta_{MF,M}$
Large	$F_{BCS,L}$	$\delta_{BCS,L}$	**$F_{MF,L}$**	**$\delta_{MF,L}$**

In any actual application, only 4 of these parameters will be picked. In our example, they are the parameters for the medium BCS operation and large flat operation, (highlighted in bold and shaded).

A similar exercise is required for the allied operation. The choice of the fixed cost parameter is similar to the choice for the direct operations in the sense it is dependent upon the size of the operation. In other words, the fixed cost for the platform depends upon whether it is, for example, a small, medium, or large platform operation. The choice of the parameters for the variable costs, however, is slightly different. Because no direct measures of workload in the platform operation exist, the platform variable cost parameters will depend upon the size of the operations that the platform supports. The fixed component of an allied operation changes with size to account for varying platform operations; however the variable piece is the same for all sizes.

Finally, the actual unavoidable fixed costs, apart from any operation-specific costs, are added to determine the total cost of operating a facility.

- Step 3: Combine the operational cost functions into the overall cost function. Once the first two steps are completed, the combined cost function is derived. In our example, the facility cost function looks like:

$$C_Z = F_z + F_{m,BCS} + \delta_{m,BCS} v_{BCS,z} + F_{l,MF} + \delta_{l,MF} v_{MF,z}$$

$$+ F_{m,PL} + \alpha_{m,BCS,PL} v_{BCS,z} \; \alpha_{l,MF,PL} v_{MF,z}$$

Despite it simplicity, this function allows for both economies of scale and economies of scope in mail processing costs.

3.2 *Translating Volume Into Workload.*

The costs generated in a plant are dependent upon the work accomplished in that plant. This means that the relevant measure of volume for a plant is determined by the piece handlings in different operations. The ZIP Code Assignment Model, in contrast, focuses on assigning originating and designating workload to facilities. One of the key factors in making that assignment is the cost associated with different sets of assignments. The mail processing cost model must therefore provide the cost of handling mail in different facilities under different ZIP Code assignments. To do so, the model must translate originating and destinating volumes into workload. This is done in three steps:

3.2.1 Step 1: Identity 3-digit ZIP Code volumes by mail class.

The process begins by collecting origin 3-digit to destination 3-digit volumes for each of the following mail classes (Express Mail will not be included):

- First-Class Mail (Presort and Single Piece separately)
- Priority Mail
- Package Services
- Standard Mail
- Periodicals

3.2.2 Step 2: Breakdown by shape.

Once the class-based mail flows have been identified, they must be segregated into separate flows for letters, flats, and parcels (i.e. volumes by 3-digit pair, class, and shape). This must be done because the workload content varies by shape. Ideally, this could be done separately for every set of 3-digit origin/destination pairs. However, data thinness precludes such an analysis. Instead, the shape breakdown will be made by origin ZIP Code and combined with the assumption that the distribution by shape is the same to all destinations from the origin. The shape breakdown for any specific 3-digit destination will be calculated by using the shape breakdowns, weighted by volume, from all origins that send mail to that 3-digit ZIP Code.

3.2.3 Step 3: Measure the piece handlings by operation per piece.

The last step in the volume to workload process is to determine the piece handlings, by operation, for each originating and destinating piece. We will use the current ZIP Code assignments to plants to calculate the ratio of, for example, flat piece handlings in outgoing sorting to the number of originating flats assigned to that facility. The ratios will be calculated by shape and operation and will allow us to assign a workload content to each originating piece in each ZIP Code. A similar method will apply to destinating volume.

Once this process is finished, we will be able to approximate the workload associated with any 3-digit ZIP Code. When alternative ZIP Code assignment schemes are investigated by the optimization model, the alternative costs of the workload can then be calculated and included in identifying the optimal outcome.

3.3 Service Considerations

In consolidating operations, the Postal Service to the greatest extent possible attempts to minimize the impact on service standards, however it is difficult to design a network of this scale and scope without some re-mapping of 3-digit ZIP Code service areas. Simulation modeling was primarily used to test and understand the feasibility of the optimization results from a service and capacity perspective.

Using data from the Texas Transportation Institute, each 3-Digit ZIP Code and facility was given a designation based on the average traffic delays and congestion observed in that location. Based on these designations and the distances between facilities (from PC Miler), we calculated travel times for all possible combinations of ZIP Code to LPC/DPCs and LPC/DPCs to RDCs. We then limited these to feasible options by applying the following rules:

- ZIP Codes could be no more than 2 hours from an LPC/DPC
- LPC could be no more than 3 hours from an RDC

We made exceptions to these rules were geographical barriers (e.g., mountain ranges and bridges) existed. We also stretched distances in some cases to allow for multiple choices at sites that didn't have any alternatives. We also used feasible paths to constrain the model in it's determination of how to use facilities. In particular:

- Specific downtown LPC/DPCs could not act as RDCs
- BMCs and Busse could not be selected as an LPC

3.4 Capacity Constraints

A key aspect of this model is the handling of the capacity constraint. The capacity constraint defines the amount of work that can be done within the confines of a particular building. While the amount of space needed to handle a particular amount of originating or destinating mail is potentially straightforward to calculate, determining the total space usage is difficult due to two types of overlap:

- Overlapping machine usage, where originating and destinating mail use the same machines
- Overlapping time windows, where there are times both originating and destinating mail are being handled.

We are exploiting the distribution concept to more accurately gauge the true space usage. The key is that for much of the system, both originating and destinating mail is handled by the same facility. So we can calculate the overall space usage for the assignment of a 3-Digit ZIP to a particular facility.

For example, mail can be handled at three types of facilities: the LPC A, LPC B, and LPC C. For a particular 3-Digit ZIP, if it is assigned to A, then it is possible to determine the space required for handling both the originating and destinating mail for A as follows: *For each major machine type, determine the number of machine hours needed for originating and destinating mail under this distribution concept to handle the 3-digit ZIP.*

Suppose the required machine hours for machines M1, M2, and M3 are as follows (based on volume at each machine and the machine-specific throughputs).

- M1: 32 hours originating, 12 hours destinating
- M2: 0 hours originating, 16 hours destinating
- M3: 12 hours originating, 9 hours destinating

Take the true operating windows lengths, say 8 hours originating (6PM-2AM) and 6 hours destinating (midnight-6AM) for all three-machine types. Determine the number of machines needed, for originating (8 hour window), destinating (6 hour window) and both (12 hour window):

- M1: originating: 32/8 = 4, destinating: 12/6 = 2, both: 44/12=3.67
- M2: originating: 0/8 = 0, destinating 16/6 = 2.67, both 16/12 = 1.33
- M3: originating 12/8 = 1.5, destinating 9/6 = 1.5, both 21/12 = 1.75

And take the maximum values:

- M1: 4
- M2: 2.67
- M3: 1.75

Multiply by the square footage per machine, and you get the total square footage required.

Similar calculations can be done for the RDCs. The result is, for 3-digit ZIP i, there is a space usage u(i,k) for every role k in the distribution concept. These are the data used to determine space utilization in the model. If x(i,j,k) is 1 if 3-digit ZIP i is assigned to facility j for role k, and if the square footage of facility j is S(j), then we require

$\Sigma i \Sigma k\ u(i,k)\ x(i,j,k) <= S(j)$ for all j

We will use the operating windows as part of a pre-processor that will calculate the square footage needed for each product and ZIP Code combination outside the Assignment Models.

3.5 Modeling Constraints

The optimization model and the constraints it uses helps us with three questions

- Which 3-digit ZIP Codes should be assigned to plant for both origins and destinations?
- What processing roles should be assigned to each facility?
- Which facilities can be absorbed by surrounding facilities?

The following table provides insights into the constraints we model.

Table 1: Modeling Constraints Algorithms

Constraint	Algorithm	Description
1. ZIP Assignment	sum_j w(i,j) = 1 for all i in Z	Every ZIP must be assigned to exactly one LPC
2. ZIP Assignment	w(i,j) ≤ x(j) for all i in Z, j in P	A ZIP can be assigned only to an open LPC
For the RDC, denote the set of roles as R, and let R(r) be the set of RDC's with role r (so R (letter concentrator) would be all k with "letter concentrator" as its role).		

Constraint	Algorithm	Description
3. Processing Role	sum_(k in R(r)) v(j,k) = x(j) for all j in P and r in R	Every open LPC must be assigned to an RDC
4. Processing Role	v(j,k) ≤ y(k) for all j in P and k in C	An LPC can only be assigned to an open RDC.
Costs can be assigned to each variable, so we can model: • *Facility costs to open/close LPC (costs on the x(j))* • *Processing costs at the LPC (costs on the w(i,j))* *To get the processing costs at the RDC, we need an auxiliary variable u(i,j,k): binary variable if ZIP i is assigned LPC j which is assigned to RDC k.* *This is a standard method to force u(i,j,k) to be 1 if and only if both w(i,j) and v(j,k) are 1.* *With these variables, we can put in* *-Transportation costs from ZIP to LPC and from LPC to RDC* *-Processing costs at LPC, DPCs and RDCs.* *-Inter-RDC transportation costs*		
5. Processing Costs	u(i,j,k) ≥ w(i,j)+v(j,k)-1 for all i in Z, j in P, k in C sum_k u(i,j,k) ≤ w(i,j) for all i in Z, j in P sum_i u(i,j,k) ≤ v(j,k) for all j in P, k in C	Linkage between *u*, *v*, and *w* variables
Capacities are handled by assuming the machines used in each role are separate. Associated with each w(i,j) is a space utilization s(i,j) given the space (square feet) needed at j by 3-digit ZIP i. Similarly each u(i,j,k) are space utilizations s'(i,k) for the usage by i at the facility in role k. If facility j has size S(j), then we have a constraint. *Note that the s and s' values offset any overlap between originating and destinating mail.*		
6. Capacity	Σi (s(i,j) w(i,j) + Σj' s'(i,j',j)) <= S(j)	Assignments must recognize capacity limitations at each facility by mail type and operation

We include facility size constraints, using the cost curve introduced earlier by adding auxiliary variables:

- size(j,s,r): binary decision variable if facility j has size s for product r.
- size'(k,s,r): binary decision variable if RDC facility k has size s for product r.
- asg(i,j,s,r): binary decision if 3-digit ZIP i is assigned to LPC j at size s for product r.
- asg'(i,k,s,r): binary decision if 3-digit ZIP i is assigned to RDC k at size s for product r.

Now, we need to link the various variables together through the following constraints:

- sum_s size(j,s,r) = x(j) for all j and r
- sum_s size'(k,s,r) = y(k) for all k and r
- sum_j sum_s asg(i,j,s,r) = 1 for all i and r
- sum_k sum_s asg'(i,k,s,r) = 1 for all i and r
- asg(i,j,s,r) <= w(i,j) + size(j,s,r) – 1 for all i,j,s,r
- asg'(i,k,s,r) <= sum_j u(i,j,k) + size'(k,s,r) – 1 for all i,k,s,r

With these definitions, the costs are straightforward: the fixed, or product core, cost for facility j at size s for product r is the cost of size (j,s,r) (and similarly for k) while the variable processing costs for ZIP code i assigned to j at size s for product r is the cost of asg(i,j,s,r).

For each 3-Digit ZIP code, the amount of mail volume in cubic feet is given as an input. In addition, based on determined truck sizes, the model is given the cost per trip for both ZIP-to-LPC and LPC-to-RDC trips by distance. Using the truck size capacity for ZIP-to-LPC transport and the amount of cubic feet volume, the number of trips needed is calculated. The model then multiplies the number of trucks needed by the per-trip cost corresponding to the distance between that ZIP and LPC. For example, if the model were given a ZIP-to-LPC cost per trip of $5 for one mile, $6 for two miles and $7 for three miles, and the distance between a ZIP and LPC was two miles, then the cost per trip would be $6. Applying this to the calculated number of trucks needed for that ZIP yields the ZIP-to-LPC transportation cost applied for that ZIP-to-LPC decision. LPC-to-RDC transportation cost works essentially the same way except that the volume isn't directly given into the model, but rather is a function of the ZIP codes assigned to that LPC.

Ideally, for each originating 3-Digit ZIP we would like to calculate the weighted average cost of transporting volumes from the assigned originating RDC to the destinating RDCs. Since we do not know the locations of the RDCs in advance of running the model, we will approximate these costs. We do know the distance between each potential RDC and each 3-digit ZIP, and we know the volumes from each origin ZIP to each destination ZIP. Thus, for each feasible assignment of origin ZIP to potential RDC, we can approximate the cost of transportation beyond the RDC as the average of the distances from that potential RDC to the destination ZIPs, weighted by the volumes going from the origin ZIP to those destination ZIPs. This cost will be specific to whichever RDC the ZIP is assigned to, and will allow us to approximate the inter-RDC cost without actually knowing the location of the RDCs[1]. The inter-RDC cost will be added to the other transportation and processing costs associated with assigning a 3-digit ZIP to an LPC – RDC combination in the optimization model.

3.6 Data Inputs

The data requirements are driven by inputs required by the optimization model. The optimization model requires the following inputs:

1) For every 3-digit ZIP Code and P&DC, the cost of assigning the 3-digit ZIP Code to that P&DC. This should include:
    a. Transportation cost to get mail from the 3-digit ZIP Codes to the P&DC
    b. Transportation cost to get mail to the 3-digit ZIP Code from the P&DC
    c. Cost of doing initial separation of mail at the P&DC
    d. Cost of doing any final sorts at the P&DC
2) For every P&DC and processing product, the cost of doing the originating product sorts at the P&DC. For this and all such calculations, a product may be broken down into sub-products (manual, presorted, etc.) with the cost being the sum of the costs for the sub-products.

[1] This is an approximation and is not a precise calculation of the transportation costs beyond the originating RDC. These costs are only used to help the optimization model consider the impact of these costs when locating facilities. These are not the transportation costs used in the final cost calculator.

3) For every product and processing concentrator (including size) for that product, the cost of doing the originating product sorts at that concentrator (and similar for dispersers).
4) For every P&DC and processing concentrator, the transportation cost of getting the product from the P&DC to the concentrator (and similar for dispersers).
5) For every P&DC and transportation concentrator, the transportation cost of getting mail from the P&DC or processing concentrator to the transportation concentrator (and similar for dispersers).
6) For every processing concentrator and transportation concentrator that can handle that mail, the cost of transporting the product from the processing concentrator to the transportation concentrator (and similar for destinating mail).
7) For each pair of 3-digit ZIP Codes, the amount of mail of each product (and sub-product if needed) to be sent from one to the other.
8) For each 3-digit ZIP Code, the amount of destinating entry volume sent by high-volume mailers to its transportation disperser.
9) For every pair of transportation concentrators/dispersers, the per unit cost of sending mail along the leg.
10) Space capacities at all facilities where processing could be done.

The following eight subsections explain the specific data requirements needed to calculate the 10 inputs listed above.

Mail flows and volumes

We need 3-digit to 3-digit piece volumes for each mail class and shape combination. We need these volumes in a file with the following fields:

- Year
- Mail Class (First-Class Presort and Single-Piece, Priority, Standard, Periodicals, Package Services)
- Shape (letters, flats, parcels)
- Origin 3-digit ZIP Code
- Destination 3-digit ZIP Code
- Piece Volume
- Pound Volume
- Cube Volume

Workloads

We need the number of piece handlings by operation for each plant in the current network.

Facility locations and ZIP Code assignments

We need the following information for each function 1 processing facility:

- Finance number
- Plant name
- Plant type
- Address, city, state, and ZIP Code
- Square feet (processing and administration)
- Origin 3-digit ZIP Code to SCF assignments by mail class (and shape where applicable)
- ADC/AADC to Destination 3-digit ZIP Code (i.e., SCF) assignments by mail class (and shape where applicable)
- SCF to Destination 3-digit ZIP Code assignments by mail class (and shape where applicable)

Facility equipment and capacities

We need the following equipment and capacity information for each facility:

- Number of machines by type for each plant
- Footprints for each type of machine
- Throughput per machine hour for each type of machine

Labor costs and productivities by operation/facility

We need the following data for labor costs and productivities for each operation and facility in order to develop the cost functions.

- Hours by operation along with the pieces handled during those hours for each operation
- A fully loaded wage rate

Operating Plan

We need to know the specific operating plans for each facility by product

Transportation Mileage and Times

For every origin 3-digit to destination 3-digit ZIP Code combination, we need to know the following information:

- Actual driving miles between points
- Average time needed to travel by mode: air, highway, rail, water

Transportation costs

We need transportation cost data by mode as described below.

- **Highway Contract Costs.** We need cost per truck mile

- **PVS Costs.** We need cost per truck mile
- **Shared Networks Costs.**

  Cost per cubic foot rates

  Available capacity by leg (origin to destination) per day
- **Commercial Air Costs.**

  Cost per pound mile rates

  Available capacity by leg (origin to destination) per day
- **Amtrak Rates.**

  Cost per rail car by leg

  Available capacity by leg (origin to destination) per day
- **Freight Rail Rates.**

  Cost per trailer by leg

  Available capacity by leg (origin to destination) per day
- **Air Taxi Costs.**

  Cost per pound mile

  Available capacity in pounds by leg (origin to destination) per day
- **Inter-Alaska Costs.**

  Cost per pound mile

  Available capacity in pounds by leg (origin to destination) per day
- **Water Costs.**

  Cost per container per leg

  Available capacity in pounds by leg (origin to destination) per day

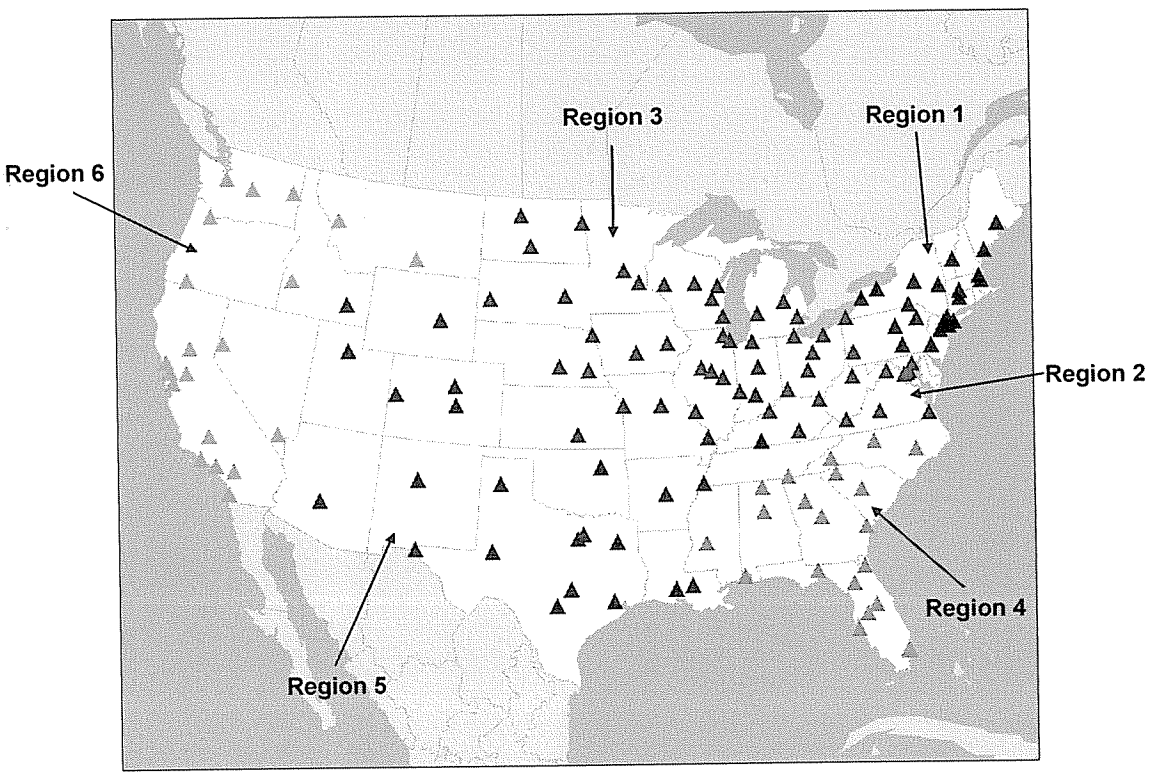
4 Modeling Steps

The ideal approach to most optimization problems would be to develop a single optimization model that represents the entire problem and identifies the single, optimal solution. However, creating such a model has proven to be intractable. Splitting models into sub-models is a well accepted and proven approach for solving models as complex as this one. Each sub-model will use the solution from the other sub-model to inform the overall model solution. This interaction, or iteration, between sub-models will serve to overcome the effects of splitting the models into pieces.

4.1 Optimization Decomposition

The Assignment model was decomposed into two types of solvers: Regional and National Assignment Models. Figure 3 demonstrates the regions we typically use. Depending on the number of constraints in a given model run, we can sometimes combine regions together and have fewer model iterations.

Figure 3: Regional Definitions



Feasible Paths helped us ensure that each ZIP Code was assigned to the same region as its' corresponding LPC. If this was not the case, we transferred the 3-Digit ZIP Codes between regions until this condition was satisfied across all regions. The optimization model had to be subdivided into geographic regions to achieve a solution within a reasonable amount of time. To run the model in regions, we needed to assign each 3-Digit ZIP Code to a single region. Using the geocoded 3-Digit ZIP Codes, we divided the country into regions of approximately the same number of ZIP Codes.

4.2 Optimization Iteration

There are two types of iterations that occur. First we iterate from the Regional to the National Model. We also iterate between model sets varying our assumptions the size of each operation.

The number of iterations completed impacts how closely the final result approaches the optimal solution. To determine an appropriate termination point that represents the best solution within an accepted tolerance, it is important to iterate this loop many times and assess the trend as the results converge toward the optimal solution. Provided enough iterations are completed, we retain the integrity of our results and the ultimate solution will still be very close to optimal.

4.2.1 Iterating between Regional and National Models

The Regional Models is formulated exactly like the National model except it includes only a segment of the national data. By limiting the data it evaluates to a region, it can solve within a practical amount of time. The National Model solves because it uses the results of the Regional Models as a "hot start". All sites and LPC-RDC assignments chosen in the regional optimizations were candidates in the National Model. In addition, cross-region assignments are now considered. If a site or assignment was not chosen in a regional model, and it did not represent a cross-region opportunity, it was not included in the National Model.

4.2.2 Iterating Facility Sizes

To keep the models to a reasonable size, it will initially be run assuming a certain size classification (e.g. LG LTTR, LG FLAT, LG PRCL, LG PRTY) for operations. This size classification will determine the fixed and variable costs used in the optimization model. Therefore we are not initially using size as a decision variable within the optimization.

In addition, we can perform reasonable heuristics to further mitigate this issue. The general approach would be to run this model iteratively, with different size classification assumptions in each successive run. Therefore, we can run a model with a certain size classification, use the solution to that model to re-consider the size classification assumption and then run the model again. This method should ensure that each successive iteration results in a solution as good as or better than the previous solution.

4.2.3 Model Run Sequence

We will run the optimization model and its pre-processors and post-processors, refine the model for iteration, and then iterate through this process. Steps the optimization modeling process follows are in Table 2.

Table 2: Modeling Steps

Step	Description
1. Run Pre-Processor	Prepare data for regional models. Develop capacity requirements for each product and ZIP combinations. Identify all feasible ZIP to LPC and LPC to RDC paths.
2. Regional Models with Large Size	Run all regional models assuming a large size classification for each facility. This will encourage the most consolidation. Produce assignments by region for ZIPs to LPCs and for LPCs to RDCs
3. National	Regional outputs are candidates in the National Model (as the "hotstart").

Step	Description
Models with Large Size	Assume large size classifications for all facilities. Make cross-region assignments were necessary (both ZIP and LPCs to RDC). "Home ZIP" constraint applied.
4. Refine model for Iteration	"Re-size" the facility functions based on assigned volumes.
5. Regional Models with True Sizes	Run all regional models assuming the true size classification for each facility. Produce assignments by region for ZIPs to LPCs and for LPCs to RDCs
6. National Models with True Sizes	Regional outputs are candidates in the National Model (as the "hotstart"). Make cross-region assignments were necessary (both ZIP and LPCs to RDC). "Home ZIP" constraint applied.

Our experience suggests that two iterations, repeating steps 4 to 6 a second time, is an appropriate termination point that presents the best solution with an accepted tolerance.

4.3 LogicNet Plus Model Refinements

In LogicNet Plus software, plants produce product and then ship to warehouses that in turn ship to customers. Below, we will discuss how to map the refined END RDC optimization into the components of LogicNet Plus.

4.3.1 Customers, Products and Demand

The customers will be the 3-Digit ZIP codes. The ZIPs will have demand for two types of product, LPC product and RDC product. The demand for these products will be in terms of square feet required.

4.3.2 Warehouses

The warehouses in this model will represent a combination of an LPC facility and an RDC location. Since customers receive products from warehouses, this will effectively allow us to assign a customer to both an LPC and an RDC simultaneously. For example, suppose the Chicago plant has only Palatine and Carol Stream as potential feasible RDC assignments. At the Chicago location, there would be two warehouses, "Chicago LPC - Palatine RDC" and "Chicago LPC - Carol Stream RDC". Then, when a ZIP code is assigned to one of these warehouses, it is actually assigned to an LPC and an RDC.

4.3.3 Plants and Production Lines

There will be a plant for every location in the network. Within that plant, there may be one, two, or three production lines. For a facility that is only capable of being an LPC and is incapable of being an RDC, it will have a single production line that produces only LPC product. For a facility that may be both an LPC and an RDC, there will be two production lines, one producing RDC product and one producing LPC product. For those plants that can be either an LPC or an RDC but where co-location of LPC and RDC is not allowed, the model will be restricted to only allow one of those two production lines (i.e., either LPC product or RDC product). If it is determined during pre-processing analysis that a given facility may be required to perform both roles, we

will add a third line that is allowed to produce both types of products. The square foot capacity of facilities will be captured as production capacity on the production lines.

4.3.4 Costs

In this model, all transportation and variable costs will be applied on the arc between the customer (ZIP) and the warehouse (LPC - RDC combination). To calculate the transportation and variable cost for a given ZIP, all we need to know is its assignment to an LPC and RDC and the sizes of those facilities for each shape. Because each warehouse represents both an LCP and RDC, we can place all of these costs on the arc between the customer and the warehouse.

The fixed costs for facilities may be placed at the production line or the warehouse, or some combination of both.

4.3.5 Key Constraints

4.3.5.1 ZIPs must be assigned to a single LPC/DPC

This is a straightforward constraint in LogicNet Plus by simply single-sourcing the customers (ZIPs). This says that a customer (ZIP) can only be assigned to a single warehouse (LPC – RDC - DPC combination).

4.3.5.2 Each LPC/DPC can only be assigned to a single RDC

We have a warehouse at each location for each potential RDC assignment (for example, "Chicago LPC/DPC - Palatine RDC" and "Chicago LPC/DPC - Carol Stream RDC"). We group these warehouses at each location and write a constraint that allows at most one to be used. In addition, we will construct transportation lanes in such a way that "Chicago LPC/DPC - Palatine RDC" can only receive LPC product from Chicago and RDC product from Palatine, etc. Together, this effectively eliminates the possibility of an LPC/DPC being assigned to more than one RDC.

4.3.5.3 A location can be an LPC/DPC or an RDC but not both

At each plant, we can group the production lines and write a constraint that allows at most one to be used. Since each production line will produce only LPC/DPC or RDC (with some exceptions for sites that require co-location), allowing at most one line to be used effectively allows that site to perform only one role.