Fish, Cows, Fields of Fuel and Optimization

Michael C. Ferris

Joint work with Nathan Pelc (USAF), Adam Christensen, Steven Wangen, Jeff Dischler, Tom Neeson (Oregon), Austin Milt, Pete McIntyre, Allison Moody, Jesse O'Hanley (Kent)

University of Wisconsin, Madison

Applied Math Seminar
University of Wisconsin
September 30, 2016

How to enhance the impact of optimization in applications?

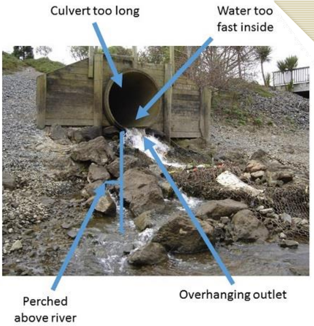
- Engage people with breadth of, and complementary expertise - theory, algorithms, computation, applications

How to enhance the impact of optimization in applications?

- Engage people with breadth of, and complementary expertise - theory, algorithms, computation, applications
- Key impact area: decision making in (environmentally) resource constrained problems
- Feature: shared resource that interacts with complex multi-user systems
- Enhance understanding of decision space, facilitate policy design and operational improvement
- Build appropriate models, fast enough solution for expert interaction, visualize results

Overview

- Anadromous fish migrate from the sea upstream into freshwater to spawn.
- Natural & man-made barriers break stream connectivity and prevent fish from penetrating deep into inland lakes and rivers



- There are over 235,000 identified barriers to migration in the Great Lakes Basin
  - Lake Michigan: >83% of tributaries inaccessible
  - Lake Huron: >86% of tributaries inaccessible
  - Lake Erie: >50% reduction of population size

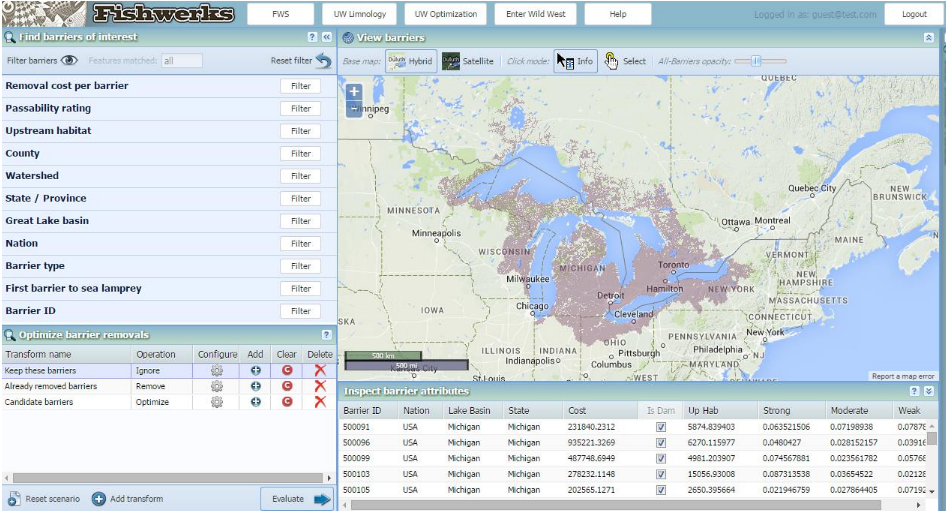
Cont'd

- Barriers can be mitigated to allow for fish passage:
  - Removal of dams, improved road crossings, fish passageways
- However, they are very expensive – Average costs for fixes:
  - Dams: $100,000 - $650,000 each
  - Others: $30,000 - $150,000 per project
- Limited funds necessitate ideal selection of projects
  - Difficult to assess where funds should be used
  - Country/State/County lines make appropriation difficult
- Increasing passability increases risk for the spread of invasive aquatic species (e.g. Sea Lamprey)

Objective

- Provide an interactive tool to consolidate big-data sets across multiple departments (DNR, Fish Wildlife Service, National Fish Passage Program, etc) and visually display in a meaningful way.
- Utilize optimization to maximize efficiency in policy decisions and funds appropriations.
- Allow any user to dynamically solve a large range of models and scenarios without requiring background knowledge of optimization.
- Provide means for certified users to update/validate data sets.

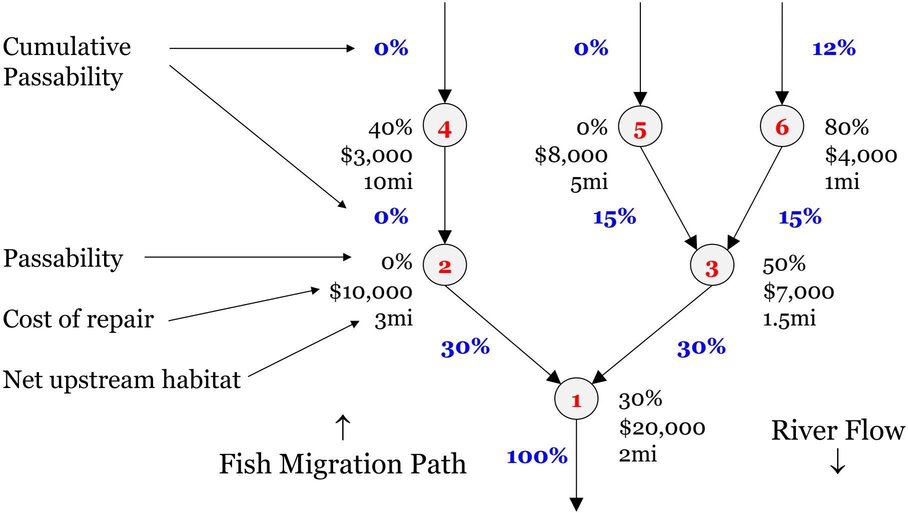
Data Visualization: http://www.greatlakesconnectivity.org/



The data

- For every Barrier $[J]$: 236,264
  - Barrier ID - A unique string identifier
  - Geographical Info - Nation, State, County, Lake Basin, Watershed
  - Barrier Type - Dam or Road Passage
  - Cost - Estimated cost to mitigate the barrier
  - Root - If the barrier is the first in the stream (no downstream barriers)
  - Downstream ID - Identifier of the downstream barrier
- For every Fish Species $[S]$: 36
  - Invasive - If it is an invasive species or not
- For every $[J \times S]$: 8,505,504
  - Passability Rating - % Chance species can pass this barrier
  - Upstream Habitat - Amount of usable habitat upstream of barrier

Probability Chain



The Model

$$\max \sum_{j \in J} \sum_{s \in S \setminus \text{Inv}} v_{js} z_{js}$$

$$\text{s.t.} \sum_{j \in J} c_j x_j \le B$$

$$z_{js} = (\bar{p}_{js} + \pi_{js} x_j) \times z_{d(j)s}, \qquad (j,s) \in J \times S$$

$$x_j \in \{0, 1\}$$

- v_{js} = upstream habitat, $\bar{p}_{js}$ passability rating, π_{js} probability increase if mitigated
- c_j cost of mitigation, B total available budget
- z_{js} cumulative passability rating
- $d(j)$ unique node downstream of j (none at "root" nodes $j \in R$)
- x_j remove barrier j (yes/no)

Linearization of products with binaries

$$z_{js} = (\bar{p}_{js} + \pi_{js} x_j) \times z_{d(j)s}, \quad (j, s) \in J \times S$$

Root nodes have no "downstream" nodes

$$z_{rs} = \bar{p}_{rs} + \pi_{rs} x_r$$

Introduce new variable: $y_{js} = x_j \times z_{d(j)s}$, $(j, s) \in (J \setminus R) \times S$

$$z_{js} = \bar{p}_{js} z_{d(j)s} + \pi_{js} y_{js}$$

Add additional constraints:

$$0 \leq y_{js} \leq x_j$$
$$y_{js} \leq z_{d(j)s}$$

Thus $y_{js} \leq x_j \times z_{d(j)s}$ but equality holds due to objective maximization

Basic Mixed Integer Programming (MILP) Model

$$\max \sum_{j \in J} \sum_{s \in S \setminus \text{Inv}} v_{js} z_{js}$$

$$\text{s.t.} \sum_{j \in J} c_j x_j \leq B$$

$$z_{rs} = \bar{p}_{rs} + \pi_{rs} x_r, \qquad (r, s) \in R \times S$$

$$z_{js} = \bar{p}_{js} z_{d(j)s} + \pi_{js} y_{js}, \qquad (j, s) \in (J \setminus R) \times S$$

$$0 \leq y_{js} \leq x_j, \qquad (j, s) \in (J \setminus R) \times S$$

$$y_{js} \leq z_{d(j)s}, \qquad (j, s) \in (J \setminus R) \times S$$

$$x_j \in \{0, 1\}$$

Interactive Modelling

Allow user to:

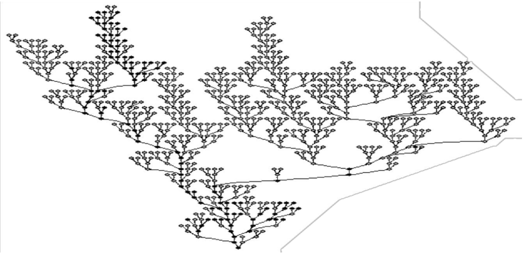
- Select their range of influence (i.e. State, County, etc)
- Select mitigatable barriers using a broad range of criteria
- Manipulate Constraints
- Visualize Results

Let's check it out! ($B = 10^7$)

- Minnesota : 3,458 – 6s.
- Wisconsin: 19,854 – Timed Out!?

The underlying issue

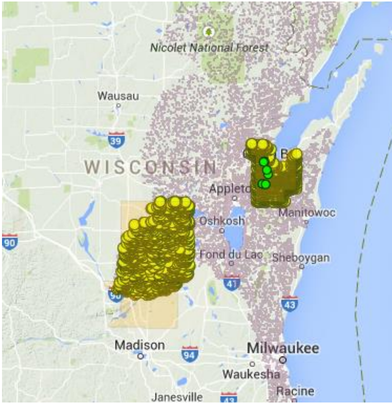
- MILP is NP-Hard
  - Solution time quickly becomes impractical as problem size grows!
  - Web tool requires fast processing to inform user
- Need to find methods to speed up solution time!
- Can we take advantage of the unique structure of our data?



Pre-Processing

Disjoint Counties: Data Compression

- May desire collaboration between counties
- Downstream barriers effected by upstream decisions
  - Barriers in-between are irrelevant
  - Can be removed by smartly incorporating their data into other nodes!



Fish guilds: variable reduction

- 36 species can be divided into 3 different guilds (fast, medium, slow swimmers)
- passability data (but not habitat data) only given by guild

$$\max \sum_{j \in J} \sum_{s \in S \setminus \text{Inv}} v_{js} z_{jg(s)}$$

$$\text{s.t.} \sum_{j \in J} c_j x_j \leq B$$

$$z_{rg} = \bar{p}_{rg} + \pi_{rg} x_r, \qquad (r,g) \in R \times G$$

$$z_{jg} = \bar{p}_{jg} z_{d(j)g} + \pi_{jg} y_{jg}, \qquad (j,g) \in (J \setminus R) \times G$$

$$0 \leq y_{jg} \leq x_j, \qquad (j,g) \in (J \setminus R) \times G$$

$$y_{jg} \leq z_{d(j)g}, \qquad (j,g) \in (J \setminus R) \times G$$

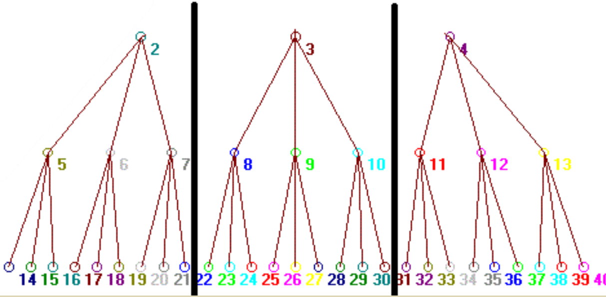
$$x_j \in \{0,1\}$$

Independent Streams?

Each root node corresponds to a completely independent tree!

Can solve separate, smaller MIP on each tree.

- However, budget constraint is global!
- *How do determine budget in each tree?*



Revised solution requirements

- Global constraints (budget, per species performance) not separable over watersheds
- Relaxing $x_j \in [0, 1]$ means these variables model the probability of doing the removal, not the action of removal
- rMIP = LP much faster to solve
- Use rMIP to determine budget allocation
- The combined reformulations provide adequate solution times, but...

Additional constraints

- Supplement base model with additional constraints
- Ensure that available habitat for all (non-invasive) species increases from v_{0s} by specific amount U_s:

$$\sum_{j \in J} (v_{js} z_{jg(s)}) \geq v_{0s} U_s$$

- Prevent over-proliferation of invasive species:

$$\sum_{j \in J} (V_{ji} z_{jg(i)}) \leq v_{0i} U_i, \quad i \in \text{Inv}$$

$$y_{ji} \geq z_{d(j)g(i)} + (x_j - 1), \quad (j, i) \in J \times \text{Inv}$$

- Note that last constraint is necessary to enforce equality in definition of y_{ji} - otherwise model will want to set y_{ji} to 0

Test Data Set: Lake Huron Basin

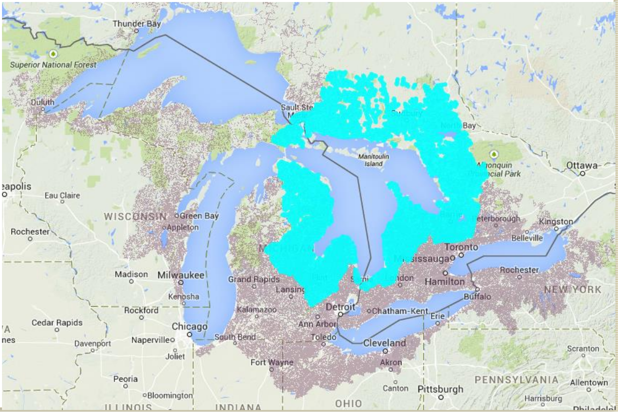
51,149 Barriers

36 Species

- 2 Invasive Species

Model Size:

- 1,934,421 rows
- 1,274,454 columns
  - 753 discrete-columns
- 4,896,386 non-zeroes



Computational issue

- Our data set is extremely large
- Solution times grow exponentially with budget [CPLEX, WID Clusters]:
  - $B = 10^6$: 8211 secs (Gap $= 0\%$)
  - $B = 10^7$: 2132 secs (Gap $= 0\%$)
  - $B = 10^8$:> 4 days (Gap $= 1\%$)
  - $B = 5 \times 10^8$:> 4 days (Gap $= 10\%$)
- Application desires ROI Curve generation, requiring data points over the entire range of budgets and different scenarios!
- Solution time is impractical for dynamic web modelling!

Decomposition approaches

- Imbalanced watersheds (several have most of the barriers) - use better load balancing
- Decompose large watersheds by "precomputing" decisions at nodes near lake
- Minimum increase constraints inactive (remove)
- Force binary decisions on "Left Out" barriers used for load balancing
- Allocate monetary and invasives budgets in each subnetwork based on rMIP solution
- Can split budget and solve "watershed" problems $W_i(B_i)$ independently
- Combining two solutions reduces to solving

$$\min c_1(W_1(B_1)) + c_2(W_2(B_2)) \text{ s.t. } B_1 + B_2 \leq B$$

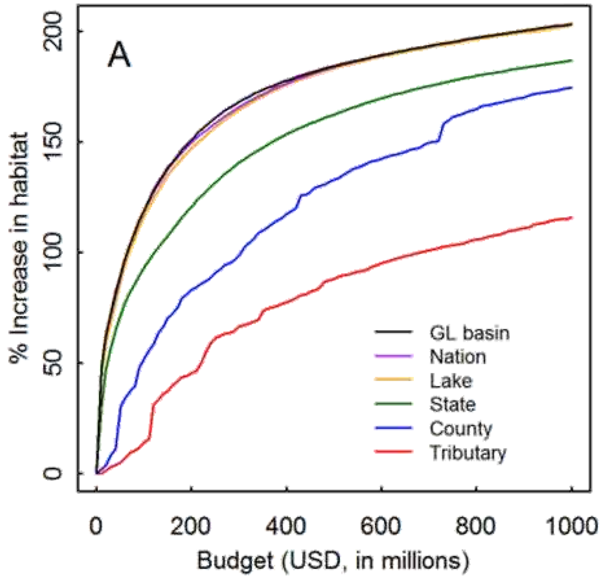
  e.g. using bisection search. More difficult problem when more than two watersheds.

Performance:

Budget ($)	Sol Time (s)	Gap (%)	Sol Time for Best (s)	% Speedup
10^6	573	0.53	8211	1,333 %
10^7	668	0.88	2132	219 %
10^8	2431	1.31	> 4 days	14,116 %

As we can see, we are able to obtain reasonable solutions for most budgets in less than 10 minutes!

Size Matters!



Field of Fuels

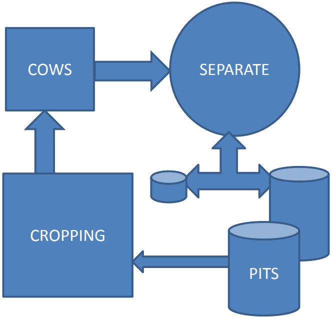
- Multiple agents interacting independently, along with shared resourc e
- Farmers (planting and management, leeching, CO2)
- Economy (supply, demand, money), Environment (bug index), Energy
- Use in schools, undergraduate classes and group of Ag/Econ experts

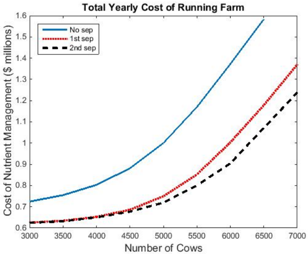
- Repeated game
- Single player not interesting - introduce bots
- Implement bots using GAMS
  - Information in: same as a human player
  - Key step: approximate other players actions/response function
  - Different objectives
  - Approximately solve a mixed integer nonlinear program - fast
  - Information out: planting and management decisions

- Point your google chrome browser at: fieldsoffuel.org

Fields Of Fuel
fieldsoffuel.org/game/play
Round Stage
Year 7
Done Planting
Done Managing
Harvest
Moderator Controls
Your Farm
Scoreboard
Switchgrass
1590 Soil Fert (N)
61% Beneficial Bugs
560 tons CO2 eq
98% Water Quality
Switchgrass
1070 Soil Fert (N)
61% Beneficial Bugs
569 tons CO2 eq
98% Water Quality
Corn
575 Soil Fert (N)
53% Beneficial Bugs
3041 tons CO2 eq
88% Water Quality
Corn
2146 Soil Fert (N)
43% Beneficial Bugs
3227 tons CO2 eq
85% Water Quality
Field Changes Over Time
Economy
Energy
Environment
< Previous Year
6
Next Year >
Your Farm's Yearly Totals
Help
Bug
Quit
Sustainability
42% Avg Score
5 Rank
Overall Score
Year
environment
economy
energy
Economics
28% Avg Score
5 Rank
Energy
35% Avg Score
4 Rank
Environment
61% Avg Score
3 Rank
Overall
Water
Emissions
Score
Year
soil health
water quality
beneficial bugs
emissions

Biomass Research and Development Initiative (BRDI)

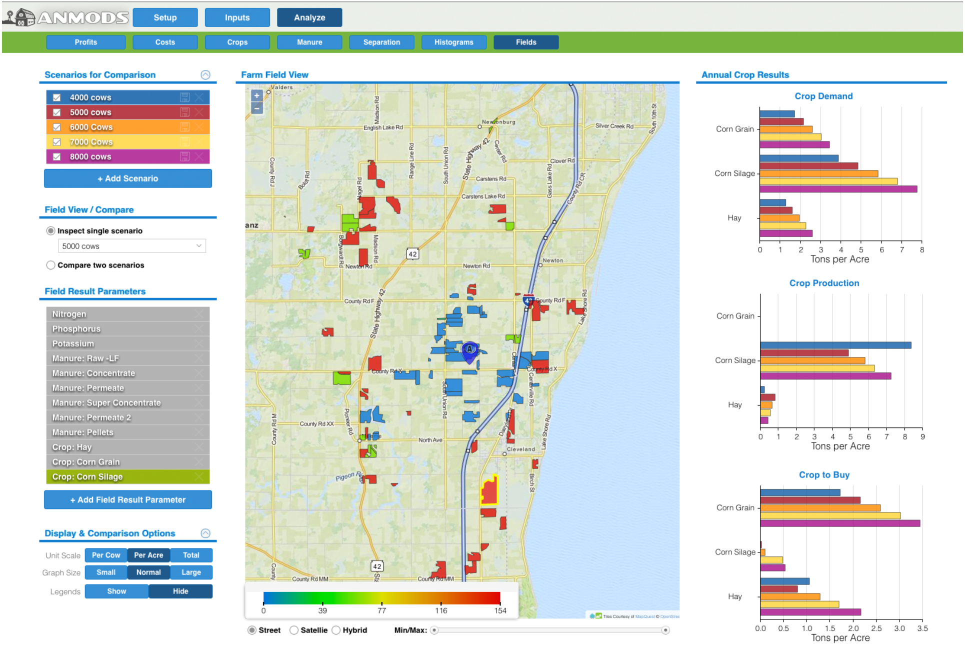




- Whole farm (complex interacting) mathematical model
- Long term sustainable (environment and financial)
- Economic/Logistic Optimization, taking into account phosphorus runoff, other environmental restrictions
- Incorporates data analytics (e.g. SNAP+)
- New insights to operate system efficiently, how to enforce much stricter environmental constraints using blend of rotations, NMP and separations
- Large (mixed integer) optimization

Mathematical details

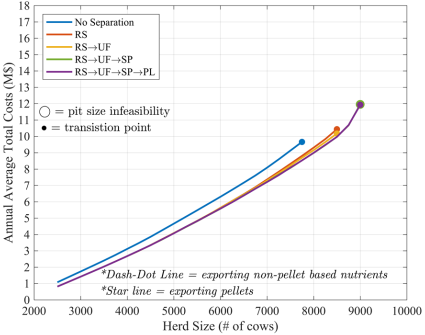
- rotation = (ofas, hay, hay, hay, corn silage, corn silage, grain, grain)
- sustainable: must be able to repeat indefinitely = same state at end of 8 years as at start
- fields (order 150), rotations (order 30 * 8), separation types (order 8), manure applications
- Key variables: use(rot,year,field), apply(src,year,field), frac(cows,src), amount(src,year,field)
- Large MIP, relaxation bound hard to improve - symmetry in start years
- Model is largely indifferent to start year of rotation, solve for rotations, then balance to enforce contraints


ANMODS
Setup
Inputs
Analyze
Profits
Costs
Crops
Manure
Separation
Histograms
Fields
Scenarios for Comparison
4000 cows
5000 cows
6000 Cows
7000 Cows
8000 cows
+ Add Scenario
Field View / Compare
Inspect single scenario
5000 cows
Compare two scenarios
Field Result Parameters
Nitrogen
Phosphorus
Potassium
Manure: Raw -LF
Manure: Concentrate
Manure: Permeate
Manure: Super Concentrate
Manure: Permeate 2
Manure: Pellets
Crop: Hay
Crop: Corn Grain
Crop: Corn Silage
+ Add Field Result Parameter
Display & Comparison Options
Unit Scale
Per Cow
Per Acre
Total
Graph Size
Small
Normal
Large
Legends
Show
Hide
Farm Field View
0
39
77
116
154
Street
Satellie
Hybrid
Min/Max:
Annual Crop Results
Crop Demand
Corn Grain
Corn Silage
Hay
0 1 2 3 4 5 6 7 8
Tons per Acre
Crop Production
Corn Grain
Corn Silage
Hay
0 1 2 3 4 5 6 7 8 9
Tons per Acre
Crop to Buy
Corn Grain
Corn Silage
Hay
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5
Tons per Acre

Nutrient management = water quality problem

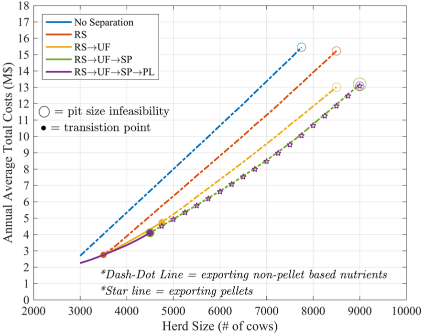
- Nitrogen, as nitrate (NH3), leaches through soil into underground wells/aquifers
- Nitrate poisoning is the leading cause of blue baby syndrome (Methaemoglobinaemia – decreased ability of blood to carry vital oxygen around the body)
- Phosphorus tends to pollute surface waters (rivers, lakes, streams, etc) through runoff (rather than leaching)
- Nitrogen and phosphorus from agricultural sources contribute to eutrophication, have long been known to cause harmful algal blooms, and ultimately result in widespread fish kills as a result of so called "dead" (hypoxic) zones

No organic N constraint



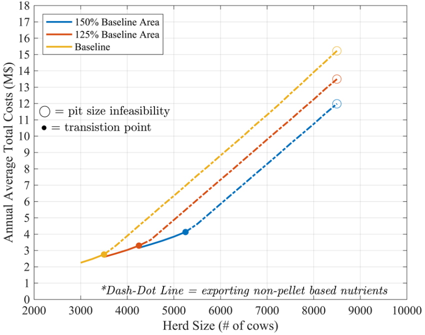
- May compromise sustainability due to nitrogen leaching
- No current regulation on application of organic N – perverse incentive to "dispose of N through over application"

Manure Separation with N constraint

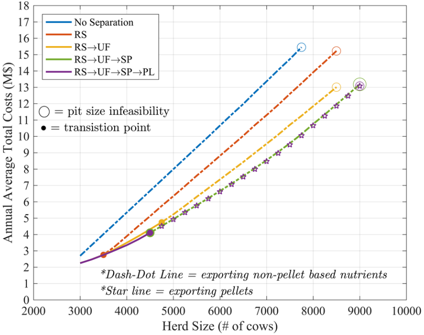


- Separation can enable sustainable nutrient management
- Pellets are valuable in other markets

Rent Land



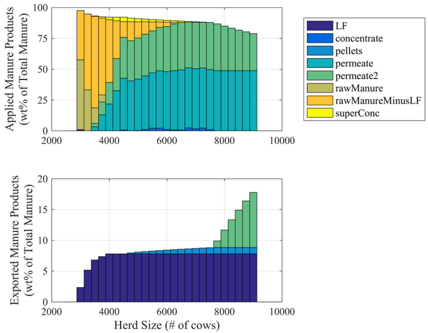
Manure Separation



- Renting land would require more organic N to be applied, which may not be available on farm from manure

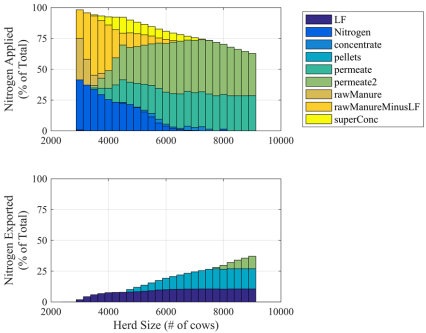
- Advanced separation can enable larger herd sizes per acre of land
- Advanced separation can lead to large economic savings vs. no/single/double stage separation

Separation Products (% weight of total manure)



- Model captures significant detail about products created and applied or exported off farm

Nitrogen application (fields/export) (% weight of total N)



- External nitrogen fertilizer purchases essentially go to zero at large herd sizes

Conclusions

- Optimization guides the development of complex interaction processes within application domains
- Combination of models provides effective decision tool at multiple scales
- Policy implications addressable using optimization
- Problems solved by combination of domain expertise, modeling prowess, good theory/algorithms and efficient implementations all facets needed
- Many new settings available for deployment; need for more theoretic and algorithmic enhancements