VEHICLE ROUTING WITH SPACE- AND TIME-CORRELATED STOCHASTIC SPEEDS

STEIN W. WALLACE, NHH NORWEGIAN SCHOOL OF ECONOMICS, BERGEN
ZHAOXIA (LEVIN) GUO, SICHUAN UNIVERSITY, CHENGDU, CHINA
MICHAL KAUT, SINTEF, NORWAY
Vehicle routing

• Basic logistics problem. Thousands of papers and multiple applications.
• A set of customers with demands
• A set of vehicles with capacities
• Travel times between customers
• Time windows (when you can visit customers)
• Service times (how much time will the delivery take?)
• … and many more variants
Important in city logistics

• Major tool for planning the distribution of goods from depots to “customers”
• Could be from major warehouses to shops
• Could be from depots or warehouses to customers, including “last mile”.

• Which truck / bicycle takes care of which customers, and what routes should the trucks / bicycles follow?
Uncertainty / stochastics

• Obviously (agree?), when you plan a distribution like this, many things are not known to you, depending on when you make the plan, as well as the specific settings
• Speeds / travel times (traffic conditions)
• Demand (especially when you are collecting or the plan is made beforehand)
• Service times
• Sometimes you do not even know the customers (since they appear while you are driving)
Stochastic vehicle routing

• Some literature over the years
• Mostly on stochastic demand, then travel time (speed), service time and random customers.
• Reviews:
Random variables

• So people have started to look at randomness, but almost exclusively independent (or uncorrelated) randomness – sometimes “Who-cares” dependence.
• The reason is not that we believe that the phenomena are independent, but that we do not know how to handle dependence in large-scale optimization models.
• The demand for beer depends on the weather across all our customers
• Service times may also depend systematically on the weather
• And travel times on streets in a city are random but obviously positively correlated to different degrees.
Correlated stochastic speeds

• We cover the case where speeds are correlated in time and space:
  - Speeds on the same link in neighboring time periods are correlated; If the speed is low in one period it is probably low also in the next period for a given link.
  - Speeds on links that meet in an intersection are correlated – if there is trouble on one link into an intersection, there is probably also trouble on other links into the same intersection.
• This is new. Old papers that mention “time dependence” simply mean that speeds depend on the time of day. No correlations.
• A map of Beijing with 142 nodes and 418 links (roads).
  - 3 time periods: $418 \times 3 = 1254$ correlated random variables
  - 60 time periods: $418 \times 60 = 25,080$ correlated random variables
Two-stage model - first example

• Stage 1: Pick routes (sequence of customers) for K capacitated vehicles visiting N customers (living in nodes on the Beijing map) with given demands

• Second-stage: Pick routes between the customers based on realized speeds (using some app based on present traffic conditions)

• Goal: Minimize expected excess travel time (over-time pay for drivers).

• The Value of the Stochastic Solution (VSS) is about 10% in this setup.
Modeling correlated speeds

• Normally we use customer-to-customer speeds (travel times) in deterministic models.

• But with stochastic speeds, that does not work.
  - Measures are most likely on road links – especially if they are collected by the system; toll roads, traffic management systems …
  - We might have collected our own data for customer-to-customer speeds, but dependence will be “impossible” to work out as pairs share links – but not always the same ones
  - We would need an enormous number of data points – and for many customer-pair / time-period combinations we would have no data points at all.
  - *If we assume that customer-to-customer speeds are uncorrelated, we are assuming that no customer pairs share any links.*

• We need to operate on the map

• Ehmke, Campbell, Thomas (2016) Vehicle routing to minimize time-dependent emissions in urban areas, *EJOR* 251(2) 478-494.

Data or not?

- **yes**: Calculate correlations
- **no**: Guess or model correlations

**Calculate correlations**

**Guess or model correlations**

**Positive semi-definite?**

- **yes**: Generate scenarios
- **no**: Fix
So without data, we can just assume a distribution – right?

• All we need is a 25,000-dimensional dependent random variable (correlation matrix)
• At least for testing (and academic) settings, just invent one.
• But should we require that what we use makes sense – that the correlations make sense?

• Claim: It is almost impossible (and truly impossible as a general approach) to guess a meaningful distribution of this type.
With data

• I have not investigated this, so off you go …
• In this kind of dimensions, though a correlation matrix can be calculated, how many measures do we need for ending up with a meaningful one?
• That is, how many measures do we need for the empirical distribution to be useful or meaningful?
Some additional issues of size

- If we have 25,000 dependent random variables, we have over **310 million** distinct correlations to handle (and store ...). *And this problem isn’t even particularly big.*
- So even if we had proper data, and could calculate correlations directly, we would be in trouble ...
- And fixing a matrix that is not positive semi-definite to make it a correlation matrix is numerically challenging in such dimensions
- Putting the small ones to zero will most likely lead to a matrix that is not positive semi-definite.

- We need scenarios, but what can we do in such dimensions? Gendreau, Jabali and Rei (2016) said that the necessary number of scenarios will “likely be quite large”. And they were not thinking of such dimensions.
What is a scenario here?

• A scenario has one outcome for each link in each period.

• So this is still two-stage modeling, it is not truly dynamic.
Scenario generation somewhat generally

- Sampling is always possible if you have something to sample from – which is far from obvious.
- It is actually hard to sample from a 25,000 dimensional (dependent) distribution even if you have one.
- You could sample from the empirical distribution (if you have one), but will it make sense?

- And there is the issue of how many samples you need in these dimensions. I do not know … but I expect it to be huge.
Other than sampling

• I would claim that no available methods would be able to handle, numerically, this kind of dimensions.
• The calculations would kill us ... feel free to disagree.

• Remember that we are talking about dependent (correlated) cases.
Our approach to scenarios

- If we cannot define a distribution, not even a meaningful correlation matrix, and in any case could not properly handle it, what can we do?
- We use a scenario generation method from Kaut (2014), based on copulas, that takes as input
  - Marginal distributions
  - Any subset of correlations, and we use
    • The same link in neighboring periods
    • Neighboring links in same period
- This is very well suited for VRPs
  - A route uses exactly such pairs
  - The expected length of any route will be correct
- In the largest example, we specify about 264,000 correlations.

… but

• The method does not control the other correlations. They will often be OK (close to zero) but not always.

• What problems can that cause?
  - The distribution of route lengths will be off if two links with a spurious correlation sit on the same route.
  - And that could affect the optimal routes as methods will be “smart”. (We believe using heuristics actually helps here).
Solution approach

• We assume that the problem is to be solved by a heuristic
  - A part that finds the sequences of customers
  - A part that evaluates any solution
• We shall mostly be concerned about the evaluation part.
• So, in principle, any deterministic heuristic can be applied to the stochastic case.
Evaluating on one scenario

- Starting point: Sequence of customer nodes
- If two customers are neighbors, travel directly
- Otherwise, use present travel speeds to find the shortest path to the next customer (Dijkstra).
- On this path, calculate properly the travel time using speeds in time and space on the scenario.
The number of scenarios

• Even if we have a distribution, there is a question:
  - How many scenarios do we need? (… likely many …)
  - Too few: Just noise
  - Too many: Can’t handle and solve the problem
• And we don’t even have a proper distribution …

• We develop a (new) way to check how many we need
  – also considering the size and the hardness of the optimization problem
Stability test

• Take a few instances of the problem (in our tests always on the Beijing map).
• For each instance, generate a set of feasible solutions (one coming from the deterministic problem)
• For different $S$ (#scenarios), generate 9 scenario trees of size $S-4$ up to $S+4$ – Kaut (2014) is “almost” deterministic.
• Find the objective function value for each solution and each scenario tree.
• Find the max relative difference for each instance and $S$
• Pick the smallest $S$ with acceptable relative difference.
<table>
<thead>
<tr>
<th>Number of scenarios</th>
<th>10</th>
<th>30</th>
<th>50</th>
<th>100</th>
<th>500</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOV</td>
<td>0.8640</td>
<td>0.8615</td>
<td>0.8619</td>
<td>0.8616</td>
<td>0.8648</td>
<td>0.8646</td>
</tr>
<tr>
<td>RD</td>
<td>4.99%</td>
<td>2.85%</td>
<td><strong>3.79%</strong></td>
<td><strong>1.65%</strong></td>
<td>0.65%</td>
<td>0.33%</td>
</tr>
<tr>
<td>Solution 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOV</td>
<td>1.3138</td>
<td>1.3080</td>
<td>1.3080</td>
<td>1.3107</td>
<td>1.3106</td>
<td>1.3105</td>
</tr>
<tr>
<td>RD</td>
<td>4.56%</td>
<td>1.92%</td>
<td>0.91%</td>
<td>0.87%</td>
<td>0.36%</td>
<td>0.14%</td>
</tr>
<tr>
<td>Solution 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOV</td>
<td>1.7906</td>
<td>1.7811</td>
<td>1.7882</td>
<td>1.7863</td>
<td>1.7864</td>
<td>1.7858</td>
</tr>
<tr>
<td>RD</td>
<td>3.63%</td>
<td>1.33%</td>
<td>1.49%</td>
<td>0.53%</td>
<td>0.32%</td>
<td>0.19%</td>
</tr>
<tr>
<td>Solution 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD</td>
<td>1.58%</td>
<td>1.17%</td>
<td>0.96%</td>
<td>0.57%</td>
<td>0.27%</td>
<td>0.23%</td>
</tr>
<tr>
<td>Solution 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOV</td>
<td>4.7591</td>
<td>4.7572</td>
<td>4.7522</td>
<td>4.7538</td>
<td>4.7552</td>
<td>4.7563</td>
</tr>
<tr>
<td>RD</td>
<td>2.73%</td>
<td>2.59%</td>
<td>1.51%</td>
<td>0.73%</td>
<td>0.58%</td>
<td>0.30%</td>
</tr>
<tr>
<td>Solution 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOV</td>
<td>4.7591</td>
<td>4.7572</td>
<td>4.7520</td>
<td>4.7535</td>
<td>4.7548</td>
<td>4.7559</td>
</tr>
<tr>
<td>RD</td>
<td>2.73%</td>
<td>2.59%</td>
<td>1.51%</td>
<td>0.73%</td>
<td>0.59%</td>
<td>0.30%</td>
</tr>
<tr>
<td>Solution 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOV</td>
<td>3.2539</td>
<td>3.2447</td>
<td>3.2362</td>
<td>3.2350</td>
<td>3.2375</td>
<td>3.2402</td>
</tr>
<tr>
<td>RD</td>
<td>1.62%</td>
<td>1.79%</td>
<td>1.22%</td>
<td>1.19%</td>
<td>0.37%</td>
<td>0.33%</td>
</tr>
<tr>
<td>Solution 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD</td>
<td>1.53%</td>
<td>0.87%</td>
<td>1.35%</td>
<td>1.17%</td>
<td>0.28%</td>
<td>0.27%</td>
</tr>
<tr>
<td>Solution 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOV</td>
<td>2.8224</td>
<td>2.7840</td>
<td>2.7793</td>
<td>2.7780</td>
<td>2.7845</td>
<td>2.7863</td>
</tr>
<tr>
<td>RD</td>
<td>7.66%</td>
<td>1.85%</td>
<td>2.57%</td>
<td>1.32%</td>
<td>0.55%</td>
<td>0.44%</td>
</tr>
<tr>
<td>Solution 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOV</td>
<td>4.2755</td>
<td>4.2517</td>
<td>4.2485</td>
<td>4.2382</td>
<td>4.2469</td>
<td>4.2483</td>
</tr>
<tr>
<td>RD</td>
<td>2.80%</td>
<td>1.85%</td>
<td>1.44%</td>
<td>1.04%</td>
<td>0.35%</td>
<td>0.33%</td>
</tr>
</tbody>
</table>
### Number of scenarios used and corresponding relative differences

<table>
<thead>
<tr>
<th>$(N, K)$</th>
<th>$P = 3$</th>
<th></th>
<th></th>
<th>$P = 5$</th>
<th></th>
<th></th>
<th>$P = 15$</th>
<th></th>
<th></th>
<th>$P = 30$</th>
<th></th>
<th></th>
<th>$P = 60$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S$</td>
<td>$RD$</td>
<td></td>
<td>$S$</td>
<td>$RD$</td>
<td></td>
<td>$S$</td>
<td>$RD$</td>
<td></td>
<td>$S$</td>
<td>$RD$</td>
<td></td>
<td>$S$</td>
<td>$RD$</td>
</tr>
<tr>
<td>(18, 3)</td>
<td>100</td>
<td>1.7%</td>
<td></td>
<td>60</td>
<td>0.7%</td>
<td></td>
<td>65</td>
<td>1.0%</td>
<td></td>
<td>45</td>
<td>0.9%</td>
<td></td>
<td>35</td>
<td>1.0%</td>
</tr>
<tr>
<td>(32, 5)</td>
<td>55</td>
<td>0.9%</td>
<td></td>
<td>20</td>
<td>0.9%</td>
<td></td>
<td>25</td>
<td>0.8%</td>
<td></td>
<td>22</td>
<td>0.9%</td>
<td></td>
<td>32</td>
<td>0.8%</td>
</tr>
<tr>
<td>(48, 7)</td>
<td>25</td>
<td>0.8%</td>
<td></td>
<td>15</td>
<td>1.0%</td>
<td></td>
<td>15</td>
<td>0.6%</td>
<td></td>
<td>15</td>
<td>0.9%</td>
<td></td>
<td>15</td>
<td>0.9%</td>
</tr>
<tr>
<td>(64, 9)</td>
<td>15</td>
<td>1.0%</td>
<td></td>
<td>15</td>
<td>0.6%</td>
<td></td>
<td>15</td>
<td>0.6%</td>
<td></td>
<td>15</td>
<td>0.8%</td>
<td></td>
<td>15</td>
<td>1.0%</td>
</tr>
</tbody>
</table>
What is going on?

• In dimension 25,080 we need only 15 scenarios to obtain a relative difference of less than 1%.

• The larger the problem instances (the map is always the same) the fewer scenarios. True both with respect to time periods and number of routes.

• *Errors caused by strange correlations are part of these error bounds.*

• We need fewer scenarios for larger cases because the errors from incorrect correlations matter less – they are averaged out.

• The errors do not go to zero as $S$ increases due to the correlations – but they are small enough.
Stability on $m$ – the number of trees is $2m+1$

<table>
<thead>
<tr>
<th>$(N,K,P)$</th>
<th>$S$</th>
<th>$m=4$</th>
<th>$m=6$</th>
<th>$m=8$</th>
<th>$m=10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(18,3,3)</td>
<td>100</td>
<td>1.7%</td>
<td>1.8%</td>
<td>1.8%</td>
<td>2.0%</td>
</tr>
<tr>
<td>(18,3,5)</td>
<td>60</td>
<td>0.7%</td>
<td>0.8%</td>
<td>0.8%</td>
<td>0.8%</td>
</tr>
<tr>
<td>(18,3,15)</td>
<td>65</td>
<td>1.0%</td>
<td>1.8%</td>
<td>1.9%</td>
<td>1.9%</td>
</tr>
<tr>
<td>(18,3,30)</td>
<td>45</td>
<td>0.9%</td>
<td>1.5%</td>
<td>1.5%</td>
<td>1.5%</td>
</tr>
<tr>
<td>(18,3,60)</td>
<td>35</td>
<td>1.0%</td>
<td>1.0%</td>
<td>1.5%</td>
<td>1.5%</td>
</tr>
<tr>
<td>(32,5,3)</td>
<td>55</td>
<td>0.9%</td>
<td>1.1%</td>
<td>1.2%</td>
<td>1.2%</td>
</tr>
<tr>
<td>(32,5,5)</td>
<td>20</td>
<td>0.9%</td>
<td>0.9%</td>
<td>1.2%</td>
<td>1.2%</td>
</tr>
<tr>
<td>(32,5,15)</td>
<td>25</td>
<td>0.8%</td>
<td>0.8%</td>
<td>0.8%</td>
<td>0.9%</td>
</tr>
<tr>
<td>(32,5,30)</td>
<td>22</td>
<td>0.9%</td>
<td>1.2%</td>
<td>1.2%</td>
<td>1.2%</td>
</tr>
<tr>
<td>(32,5,60)</td>
<td>32</td>
<td>0.8%</td>
<td>1.1%</td>
<td>1.1%</td>
<td>1.1%</td>
</tr>
<tr>
<td>(48,7,3)</td>
<td>25</td>
<td>0.8%</td>
<td>0.8%</td>
<td>1.1%</td>
<td>1.3%</td>
</tr>
<tr>
<td>(48,7,5)</td>
<td>15</td>
<td>1.0%</td>
<td>1.1%</td>
<td>1.1%</td>
<td>1.1%</td>
</tr>
<tr>
<td>(48,7,15)</td>
<td>15</td>
<td>0.6%</td>
<td>0.6%</td>
<td>0.9%</td>
<td>0.9%</td>
</tr>
<tr>
<td>(48,7,30)</td>
<td>15</td>
<td>0.9%</td>
<td>1.1%</td>
<td>1.1%</td>
<td>1.2%</td>
</tr>
<tr>
<td>(48,7,60)</td>
<td>15</td>
<td>0.9%</td>
<td>1.4%</td>
<td>1.4%</td>
<td>1.4%</td>
</tr>
<tr>
<td>(64,9,3)</td>
<td>15</td>
<td>1.0%</td>
<td>1.0%</td>
<td>1.1%</td>
<td>1.3%</td>
</tr>
<tr>
<td>(64,9,5)</td>
<td>15</td>
<td>0.6%</td>
<td>0.6%</td>
<td>0.7%</td>
<td>0.9%</td>
</tr>
<tr>
<td>(64,9,15)</td>
<td>15</td>
<td>0.6%</td>
<td>0.6%</td>
<td>0.6%</td>
<td>0.6%</td>
</tr>
<tr>
<td>(64,9,30)</td>
<td>15</td>
<td>0.8%</td>
<td>0.9%</td>
<td>0.9%</td>
<td>1.2%</td>
</tr>
<tr>
<td>(64,9,60)</td>
<td>15</td>
<td>1.0%</td>
<td>1.0%</td>
<td>1.1%</td>
<td>1.1%</td>
</tr>
</tbody>
</table>

Even up to $m=10$ we get stability levels under 2%. So $m=4$ was OK.
Classical in-sample test for $m=4$ and $S$ from the earlier tests. All is well.

<table>
<thead>
<tr>
<th>$(N, K)$</th>
<th>$P = 3$</th>
<th>$P = 5$</th>
<th>$P = 15$</th>
<th>$P = 30$</th>
<th>$P = 60$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S$</td>
<td>RD</td>
<td>$S$</td>
<td>RD</td>
<td>$S$</td>
</tr>
<tr>
<td>(18, 3)</td>
<td>100</td>
<td>1.7%</td>
<td>60</td>
<td>0.9%</td>
<td>65</td>
</tr>
<tr>
<td>(32, 5)</td>
<td>55</td>
<td>0.9%</td>
<td>20</td>
<td>0.9%</td>
<td>25</td>
</tr>
<tr>
<td>(48, 7)</td>
<td>25</td>
<td>0.4%</td>
<td>15</td>
<td>1.1%</td>
<td>15</td>
</tr>
<tr>
<td>(64, 9)</td>
<td>15</td>
<td>0.4%</td>
<td>15</td>
<td>0.4%</td>
<td>15</td>
</tr>
</tbody>
</table>

Weak out-of-sample test

So all the stabilities are in place

CPU comments

• For the case with P=30 we needed 37 minutes for the 15 scenarios (but they have 180,600 elements) and they handle about 130,000 correlations.
• You need time to create the scenarios – fixed cost
• The number of scenarios shows how much longer it takes per iteration relative to the deterministic case.
Extended tests

• We have tested alternative VRPs with time windows (hard and soft), fuel consumption (rather than distance), fixed costs for vehicles …

• The results are consistent – about 15 scenarios in the largest cases for stable function evaluations.

• We are also testing Stochastic Shortest Path problems with real data.
Conclusion

• We have a method than can evaluate the objective function for correlated speeds in very high dimensions with very high accuracy. Major issues:
  - An argument that these problems cannot be formulated on a customer-node network, but need the underlying map
  - Proper stability testing to settle the number of scenarios
  - A scenario generation method well suited to VRPs

• The number of scenarios is very modest, implying that solving the stochastic case is not too time consuming.
But can this be extended?

• Major issue: We know which correlations are most important since we know the shape of a solution.
• I think there is hope in some other model classes, but not all.
  - Demand – close by customers,
  - Facility location – demand and speed
  - NOT portfolio management
• So the issue is wide open.
Thank you!