On the Stochastic Vehicle Routing Problem with time windows, correlated travel times, and time dependency

Federica Bomboi · Christoph Buchheim · Jonas Pruente

Abstract Most state-of-the-art algorithms for the Vehicle Routing Problem, such as Branch-and-Price algorithms or meta heuristics, rely on a fast feasibility test for a given route. We devise the first approach to approximately check feasibility in the Stochastic Vehicle Routing Problem with time windows, where travel times are correlated and depend on the time of the day. Assuming jointly normally distributed travel times, we use a chance constraint approach to model feasibility, where two different application scenarios are considered, depending on whether missing a customer makes the rest of the route infeasible or not. In addition, we present an adaptive sampling algorithm that is tailored for our setting and is much faster than standard sampling techniques. We use a case study for both scenarios, based on realistic instances, to illustrate that taking correlations and time dependencies into account significantly improves the quality of the solutions, i.e., the precision of the feasibility decision. In particular, the nonconsideration of correlations often leads to solutions containing infeasible routes.

1 Introduction

Vehicle Routing Problems (VRP) were and still are subject of countless studies in the literature of operational research. These problems consist in finding a set of vehicle routes serving certain customer requests with the minimum total travel cost. The term VRP in general refers to the deterministic problem, in which all the data are known and not subject to uncertainty. However, in practice, due to a variety of influences such as traffic jams, weather, customer availability etc., the real situation is often unpredictable, so that a solution of the deterministic problem in most cases produces solutions that are far from optimal in reality. To deal with this problem, specific approaches for addressing these uncertainties are needed.

The aim of this paper is to devise a new approach to check the feasibility of routes in the VRP with time windows (VRPTW) assuming stochastic travel times which can be used in branch-

F. Bomboi
Department of Mathematics and Computer Science, University of Cagliari, Italy
E-mail: fbomboi@unica.it

C. Buchheim
Fakultät für Mathematik, TU Dortmund, Germany
E-mail: christoph.buchheim@math.tu-dortmund.de

J. Pruente
Fakultät für Mathematik, TU Dortmund, Germany
E-mail: jonas.pruente@math.tu-dortmund.de

and-price algorithms as well as many heuristic frameworks. Uncertainty is taken into account in several ways: firstly, the cost of a route depends on its expected travel times and is hence a random variable itself. Secondly, if a driver arrives too early at a costumer's location, the resulting waiting time has to be taken into account. Our model allows to calculate different costs for waiting and driving. Thirdly, and most importantly, a route may turn out to be infeasible under certain realizations of the travel times, due to the risk of missing some time window.

Unlike many other approaches presented in the literature, we explicitly deal with dependent travel times in our approach, assuming a joint normal distribution. In practice, neighboring streets often have highly correlated travel times, so that taking into account the dependency of travel times is important in order to obtain feasible solutions. This is also confirmed by our computational study: considering such dependencies improves the precision of the solutions significantly. The importance of correlations between travel times has also been underlined in [23].

Our approach can be used in any context where single routes are considered separately. This is the case, e.g., in the well-known set partitioning model, where the variables correspond to potential routes that are either enumerated in a first phase or are generated on the fly in a column generation approach. Also many heuristic approaches produce potential routes that have to be checked individually for feasibility. In particular, this allows in principle to use an arbitrary rule for deciding feasibility of a route and to determine its cost.

The solution method proposed in this paper is based on the chance constraint approach, where a route is accepted if the risk of a failure stays below a given threshold. We distinguish between two application scenarios. In the first scenario, the route is considered infeasible if for any of the customers on this route, the probability of missing her time window is too high. We present a (single) chance constraint formulation for this scenario. In the second scenario, missing a customer's time window may render the entire route infeasible. The latter approach applies to variants of the VRP where pickups and deliveries are performed by the same vehicles, so that missing a customer might imply that some following customers cannot be served any more, e.g., due to lack of space in the vehicle. In this situation, we are interested in the (joint) probability that at least one of the time windows is missed. For this setting, we present a joint chance constraint formulation. Moreover, for obtaining a more accurate stochastic model, we propose to consider truncated probability distributions in this case.

In both cases, we use and extend an idea of Ehmke et al. [7]. It addresses the fact that, due to possible waiting times, the arrival times at subsequent customers are not normally distributed any more. It is proposed in [7] to approximate these travel times, which can be defined as maxima between two normally distributed variables, again by normally distributed travel times, with the same expected values and variances. In this context, we also refer the reader to [4], [18] and [24]. We extend this approximation to the correlated case by also considering and updating the covariances between travel times. Additionally, we consider a situation where travel times change over the day, meaning that the time needed to travel an arc depends on when the arc is traversed.

It turns out that these extensions, though being computationally more expensive, lead to much more precise assessments of the feasibility of a route in realisitic situations. In fact, our experiments on a set of realistic instances show that ignoring covariances leads to the creation of routes that are actually infeasible, when validated by sampling. The number of infeasible routes is decreased significantly when including covariance information in our algorithm. When travel times vary over the day, the difference between our approach taking this into account and an approach based on average travel times, is even larger, with many infeasible routes produced in the latter approach and, at the same time, objective values being better for our new algorithm.

Literature Review Other papers which use the chance constraint approach for the time window constraints are [7] with the assumption of independent travel times and [16] with independent travel and service times, others set restrictions on the probability that a vehicle's capacity is exceeded [6] and on the length of the travel time [19]. Some reviews of SVRP literature are [9], the

detailed [22] and [21], and [1] for the case of uncertain, independent and identically distributed customer demands. Articles dealing with uncertain travel times are [7], [19], [25], [26], [27], and [29]. Among these, [19] and [26] study the time dependent case, in which travel times are stochastic and vary during the day. Uncertain demands are considered in [6], [10], [12], and [20]. An SVRP with simultaneous pickup and delivery under uncertain demands and travel times is dealt with by [14]. For works based on stochastic travel and service times, see [16], [17], and [30]. For a VRP with stochastic service times we refer to [8]. Several papers assume soft time windows, namely [12], [25], [26], [27], and [30]. Concerning the assumption of independence and dependence of the uncertain variables, [7] assume independent travel times, [16] and [17] assume independent travel and service times, [6] consider correlated demands and [10] assume that demands follow a Poisson distribution and show extensions with the Binomial, Negative-Binomial and Gamma distributions for solving the case of independent demand and assume multivariate normally distributed demands in case of correlation.

With respect to solution methods, the majority of publications opt for heuristics and meta-heuristics, [12] use a genetic based algorithm, [29] an Ant Colony Optimization heuristic, [6] adapt and extend the Clarke and Wright's heuristic [5] to obtain primal solutions to the chance-constrained VRP and investigate a Dantzig-Wolfe formulation, [10] and [15] also use a heuristic procedure based on [5]. For the computational experiments, [7] embed the feasibility check and the estimation of arrival and start-service times into a tabu search algorithm. Other papers which use the tabu search are [16], [25], and [30]. Recourse methods are used by [8], [16], [20] and [30].

Our Contribution We present a new approach for checking the feasibility of routes of a wide class of variants of the VRPTW subject to stochastic travel times. The most important contributions of this paper are

- an investigation of the importance of considering correlations between travel times by solving realistic instances;
- an improved sampling algorithm for situations in which we only require a decision whether a given threshold is exceeded or not;
- an approach for checking the feasibility of routes for the single chance constrained routing problem with correlations, not based on sampling;
- an investigation of the importance of considering travel times varying over the day by solving instances created with real data;
- an approach for checking the feasibility of routes for the single chance constrained routing problems with time dependencies, not based on sampling;
- an approach for checking the feasibility of routes for the joint chance constrained routing problem with and without correlations, not based on sampling;
- an estimation of the waiting times of the vehicles at the customer locations, that can be used in penalty based approaches.

The main contribution of the paper is the description of an algorithm for checking the feasibility of routes considering correlations and time dependencies at the same time for single chance constrained routing problems. We show that in this setting sampling is not an option and therefore to the best of our knowledge our approach is the only one that can be used.

Outline of the Paper Section 2 presents some notation for the problem we consider. In Section 3, we describe our idea of adaptive sampling. Section 4 deals with the single chance constrained version of the problem. In Section 5, an approximation of the joint chance constrained problem is discussed. The paper terminates with a summary and a discussion of possible extensions in Section 6. In Appendix A, some stochastic formulas used in the proposed algorithms are listed.

2 Preliminaries and Notation

We aim at solving Vehicle Routing Problems with Time Windows subject to uncertain travel times. Extending the terminology of [28] for the deterministic version of the problem, we will refer to this class of stochastic problems as SVRPTW. Since characteristics like capacities and demands are not in the scope of this work, vehicles with infinite capacity are assumed. Every customer is visited exactly once by exactly one vehicle and all vehicle routes start and end at a single depot.

More formally, we assume that a finite set C of nodes is given, where $0 \in C$ corresponds to the *depot* and the remaining nodes in $C \setminus \{0\}$ correspond to *customers* to be served. Moreover, we assume that each pair of nodes $i, j \in C$ is connected by a directed arc $(i, j) \in E$, we thus deal with a complete graph $G = (C, E)$ throughout the paper. A *route* r in G is given by an ordered list of distinct costumers $c_{r_1}, \ldots, c_{r_k}$, the set of its arcs is denoted by $E_r := \{(0, c_{r_1}), (c_{r_1}, c_{r_2}), \ldots, (c_{r_{k-1}}, c_{r_k}), (c_{r_k}, 0)\}$. In our problem, each costumer is assigned a deterministic *time window*. It will be convenient to index the time windows by arcs, thus for an arc $e = (i, j)$ we will denote by a_e (b_e) the *earliest (latest) arrival time* at costumer j, so that the time window is defined by $[a_e, b_e]$.

All travel times are uncertain and thus modelled as random variables. More precisely, we make the assumption that the vector $X \in \mathbb{R}^E$, which defines the travel time X_e for each arc e, is jointly normally distributed with means $\mu \in \mathbb{R}_+^E$ and covariance matrix $\Sigma \in \mathbb{R}^{E \times E}$, i.e., $X \sim \mathcal{N}(\mu, \Sigma)$; we will denote the entries of Σ by $\sigma_{e,f}$ for $e, f \in E$. For simplicity, we assume that Σ is positive definite.

In particular, the travel time of each single arc e is again normally distributed with $X_e \sim \mathcal{N}(\mu_e, \sigma_e^2)$, where we set $\sigma_e = \sqrt{\sigma_{e,e}}$. This also implies that the travel time of a route, given as the sum of travel times of the contained arcs, is normally distributed – however, this is only true as long as no time windows are considered. Finally, we will use $\epsilon \in (0, 1)$ throughout the paper to denote the threshold for the risk of a failure, i.e., we will accept a route if the probability that it is infeasible, with respect to the uncertain travel times, is at most ϵ. For the reader's convenience, we summarize this notation in Table 1.

Table 1: *Basic notation defining our instances*

0	depot node
C	set of customer nodes including depot
E	set of all arcs
E_r	set of arcs contained in a route r
a_e	earliest arrival time at end node of arc e
b_e	latest arrival time at end node of arc e
X_e	random variable describing the travel time of arc e
μ_e	expected travel time of arc e
$\sigma_{e,f}$	covariance between travel times of e and f
σ_e	standard deviation of travel time of arc e
ϵ	threshold for feasibility of a route

In this paper, we restrict ourselves to discussing the following two questions with respect to a *fixed* route r:

(a) is route r feasible with a high enough probability?
(b) if so, what is the expected cost of route r?

This is motivated by the fact that many exact as well as heuristic approaches for solving vehicle routing problems can be reduced to the above tasks, in particular when the decision of (a) is a difficult problem in itself. An important example for this class of approaches is the set partitioning approach to the VRP, which is often combined with column generation. Assuming that the full

set of feasible routes R is known, together with the corresponding cost c_r for each $r \in R$, we can model the final optimization problem as follows:

$$
\begin{aligned}
\min \quad & \sum_{r \in R} c_r x_r \\
\text{s.t.} \quad & \sum_{\substack{r \in R \\ i \in r}} x_r = 1 \quad \forall i \in C \setminus \{0\} \\
& x_r \in \{0,1\} \quad \forall r \in R \, .
\end{aligned}
$$

Here we choose a cheapest subset of all routes such that each costumer is visited exactly once by these routes.

In Sections 4 and 5, we will concentrate on the above questions (a) and (b) in two different application scenarios. In the first one, a route becomes infeasible if we miss the time window of any of the customers with probability larger than $1 - \epsilon$; see Section 4. In the second scenario, a route is infeasible if the probability of missing at least one costumer exceeds $1 - \epsilon$; see Section 5. Even if the difference between the two scenarios seems very subtle, the second approach is much more challenging from a mathematical (and complexity-theoretic) point of view, as it requires to deal with joint chance constraints. In fact, we can only deal with the latter case in an approximate way. In both cases, the expected costs have to take possible waiting times into account.

3 Adaptive Sampling

A standard approach for testing feasibility of a given route r would be to sample the distribution of the involved arc lengths and to check whether the given chance constraints hold with high enough probability. Clearly, the crucial question here is how many samples to compute. If the number is chosen too high, it takes too much time to compute them. If it is chosen too small, the result will be unreliable. However, in our context, we are not interested in the exact probabilities that a route is feasible or not. Instead, we only have to decide whether this probability is greater than the threshold $1 - \epsilon$. This can be exploited in order to adaptively choose the number of samples, using the following classical result due to Hoeffding [13].

Theorem 1 *After sampling independent (not necessarily identically distributed) Bernoulli random variables at most $\frac{1}{2\gamma^2} \ln(\frac{2}{\delta})$ times, the deviation of their mean from the expected value differs by at most γ with probability at least $1 - \delta$, for every $\gamma > 0$ and $\delta \in (0,1)$.*

Here, a Bernoulli distribution is a probability distribution of an experiment with binary outcome. E.g., to be sure with probability 99% that the mean does not deviate more than 1% from the real expected value, at most 26492 samples are needed. Algorithm 1 exploits this result in order to reduce the number of samples, in situations where feasibility can be decided early with a given probability $1 - \delta$.

Algorithm 1 Adaptive Sampling

Input: route r; constant $\delta \in (0,1)$
Output: decision if r is feasible, with confidence level at least $1 - \delta$
1: $s_{\text{tot}} = 0$, $s_{\text{inf}} = 0$ ▷ total number of (infeasible) samples computed
2: **repeat**
3:     sample the travel times of the arcs in r and set $s_{\text{tot}} \mathrel{+}= 1$
4:     **if** route r is infeasible with the resulting travel times **then**
5:         $s_{\text{inf}} \mathrel{+}= 1$
6:     **end if**
7:     $p = \frac{s_{\text{inf}}}{s_{\text{tot}}}$ ▷ percentage of infeasible outcomes
8:     $\gamma = |p - \epsilon|$, $\hat{s} = \frac{1}{2\gamma^2} \ln(\frac{2}{\delta})$ ▷ number of samples needed by Theorem 1
9: **until** $\hat{s} \leq s_{\text{tot}}$ ▷ stop when enough samples
10: **return** whether $p \leq \epsilon$

Clearly, Algorithm 1 can be expected to terminate quickly when the probability of r being infeasible is far from ϵ (in either direction). Otherwise, it might be necessary to bound the total number of samples by some limit $s_{\max}$, in which case the formula in Theorem 1 can be used to compute the reliability $\delta := 2/\exp(2\gamma^2 s_{\max})$ of the resulting decision. Note that Algorithm 1 only deals with the feasibility decision, i.e., with task (a) mentioned in the previous section. For task (b), i.e., for computing the objective value of route r, we always use $s_{\max}$ samples in our experiments. This is acceptable, since we only have to do this for routes classified as feasible. Moreover, Algorithm 1 in general needs more samples anyway to determine that a route is feasible than that it is infeasible, if ϵ is small.

In the following sections, we will use this approach for comparison and validation. Our main objective is to develop methods not based on sampling.

4 SVRPTW – Single Chance Constraints

In this section, we assume that all demands are deliveries, deterministic and known before the optimization process, and not split. A typical application for this scenario are deliveries from a post office. An important characteristic of this kind of problem is that if a customer in a route happens not to be served for some reason, the next customer in this route can still be served by the same vehicle. This is because the service failure at one customer does not undermine the service at the next customer. Therefore, in the chance constraint approach to be developed, a risk level of service failure is fixed for each customer independently. In other words, the feasibility of a route is determined considering the union of the single chance constraints in the route. In the following, we describe our approach in mathematical and algorithmic terms and show experimental results concerning solution quality and running times. We first address the issue of correlations in Section 4.1, then we consider time-dependent travel times in Section 4.2.

4.1 Including correlations

Our first aim is to extend the approach of Ehmke et al. [7] in order to deal with correlated travel times. As we will show in our experiments in Section 4.1.2, taking correlations into account leads to a much more precise assessment of the feasibility of a route.

4.1.1 Algorithm

Assume we are given a route r with arc set $E_r = \{e_1, \ldots, e_k, e_{k+1}\}$ (in this order). As already mentioned, the chance constraint consists in placing a restriction on the probability that a given customer time window is missed. For the first customer, the constraint is easily modeled as $P(X_{e_1} > b_{e_1}) \leq \epsilon$, which is equivalent to $b_{e_1} \geq \mu_{e_1} + \Phi^{-1}(1-\epsilon)\sigma_{e_1}$ where Φ denotes the cumulative distribution function of the standard normal distribution. However, when arriving before a_{e_1}, the driver has to wait. This may lead to other costs than driving. More importantly, the potential waiting time will influence the arrival times at the following costumers in the route. In particular, the distributions of the arrival times at the subsequent customers in the route have to be re-modeled.

As already discussed by Ehmke et al. [7], the adapted arrival times at a customer c_i do not follow a normal distribution any more. Anyway, they propose to approximate the resulting distributions by normal distributions again, iteratively at every customer, and show experimentally that the resulting error is negligable. More precisely, the idea is to compute means and variances of the distributions of every arrival time and to replace the random arrival times by the corresponding normal distributions.

In this section, we follow the same idea, relaxing however the assumption of independent travel times and instead taking into account the information of the correlations between routes

and arcs. This is more complicated to do because we now also have to calculate the covariances between routes and arcs. Formally, we replace the vector containing the arrival time at the current node and the travel times of the subsequent arcs in the route by a jointly normally distributed vector having the same means and covariances. Clearly, this generalization leads to a higher running time due to the quadratic input in terms of the covariances, but we think it is worth to follow this approach for having a more realistic formulation and better solutions, as we will show in detail in the following Section 4.1.2.

Algorithm 2 Feasibility Check with Single Chance Constraint, taking correlations into account

---

Input: route $r = (0, r_1, \ldots, r_t, 0)$
Output: decision if r is feasible with high probability; expected driving time D; expected waiting time W

1: $k = 0$
2: $exp_k = 0,\ var_k = 0$
3: **for** $e \in E_r$ **do**
4:      $cov_{k,e} = 0$
5: **end for**
6: $W = 0$

7: **for** $e \in E_r \setminus (r_t, 0)$ **do**
8:      $exp_k \mathrel{+}= \mu_e$
9:      $var_k \mathrel{+}= \sigma_e^2$
10:     $var_k \mathrel{+}= 2cov_{k,e}$
11:     **for** $f \in E_r$ after e **do**
12:         $cov_{k,f} \mathrel{+}= \sigma_{e,f}$
13:     **end for**
14:     assume $\tilde{X}_k \sim \mathcal{N}(exp_k, var_k)$

15:     **if** $P(\tilde{X}_k > b_e) > \epsilon$ **then**
16:         discard route r
17:     **end if**

18:     $exp_{k+1} = E[max\{\tilde{X}_k, a_e\}]$ ▷ Formula (A1)
19:     $var_{k+1} = Var[max\{\tilde{X}_k, a_e\}]$ ▷ Formula (A2)
20:     **for** $f \in E_r$ after e **do**
21:         $cov_{k+1,f} = Cov[max\{\tilde{X}_k, a_e\}, X_f]$ ▷ Formula (A3)
22:     **end for**

23:     $W \mathrel{+}= E[max\{a_e - \tilde{X}_k, 0\}]$ ▷ Formula (A1)
24:     $k \mathrel{+}= 1$
25: **end for**
26: $D = exp_k + \mu_{(r_t,0)}$
27: accept route r and return D and W

Our approach is described in Algorithm 2. After initialization in Lines 1–6, it loops through the costumers of the given route r. It first updates the distribution information of the arrival time at the next customer (exp_k and var_k) as well as its covariance with all arcs f coming later in the route ($cov_{k,f}$); see Lines 8–14. Next, the route is discarded in case the chance constraint is violated (Line 15–17). Finally, the distribution information is updated once again due to a possible waiting at the current costumer (Lines 18–22). All formulas needed for updating the distributions are derived in Appendix A.

Compared to the method proposed in [7], Lines 3–5, 10–13, 20–22, and 23 are new. All but the latter concern the calculation of the correlations between arcs and routes that [7] do not consider because of the assumption of independent travel times. Line 23 calculates the expected waiting time W of the vehicle at every customer. Together with the expected driving time D, it can be used to calculate the expected cost of route r, using any function in the two values D and W.

A first comparison between the method involving dependent travel times proposed in this paper (Algorithm 2) and the method with the assumption of independent travel times of [7] can be already made. By the introduction of covariances it is possible to estimate better the variance of the arrival time at every customer and, consequently, of the travel time $\tilde{X}_k$ of the route at every step k. Without considering the covariances between arcs and routes, the variance is underestimated by [7] in case of positive correlation (see Line 10). For $\epsilon < 0.5$ (and particularly for small ϵ), a smaller value of the variance leads to accepting as feasible a higher number of routes. A deeper analysis on the comparison between the solution methods is given in the following Section 4.1.2.

4.1.2 Experimental Results

In order to investigate the practical relevance of the approach devised above and to illustrate the importance of taking correlations into account, we performed a case study using a set of realistic instances. The instances used for our experiments are based on real traffic data for the surroundings of the port of Duisburg. The port itself is chosen as depot and 19 nearby locations are picked as customer positions. The expected values and covariances of the travel times were calculated by a sample of data taken on 25 consecutive days at 3 pm from Google Maps [11]. If necessary, we added a value of 10^{-4} to all diagonal entries of the resulting covariance matrix in order to guarantee positive definiteness and to avoid numerical problems. Note that, if any, this has the effect of making the covariances slightly less relevant with respect to the variances.

Each arc of the graph corresponds to the shortest path between two customers or the port and a customer at the time point the data is taken. Therefore, an arc does not necessarily refer to the same path in all of the samples. This seems more realistic because a driver would always choose the shortest path at a given time. The considered network of customers and arcs is directed with an asymmetric matrix of the costs that satisfies the triangle inequality.

The arcs lengths observed in our samples ranged between 5 and 167 minutes, with an average of 60 minutes. In the resulting instance, the average correlation coefficient is 0.64 between two adjacent arcs and 0.60 between two non-adjacent arcs. The maximum correlation coefficient is 1.00 in both cases. This confirms our assumption that it is important to take the covariances into account in real life instances.

We created 10 different instances that only differ from each other in the time windows. The time windows of each instance were computed randomly in the following way: for each customer, the lower bound a_e of the time window is chosen uniformly at random as an integer between 0 and 7. The length of all time windows is 1 hour. If the time windows do not allow a feasible solution for the whole VRP for any of the algorithms considered, the time windows are recomputed randomly until they do.

The purpose of these experiments is to show, based on a real world example, that considering covariances is crucial for an accurate feasibility check. Additionally, we show that Algorithm 2 is reliable as a feasibility check. As long as we do not consider time-dependent travel times (as in Section 4.2), the feasibility check could as well be performed by sampling, which would even be faster than our approach. However, we ultimately aim at combining correlations and time dependency, in which case sampling will not be a practicable approach any more; see our discussion in Section 4.2.2. In the remainder of this section, we are thus interested in a comparison in terms of quality rather than running time.

In the following, we compare our new Algorithm 2 to the same algorithm using zero covariances (Algorithm E), which essentially agrees with the algorithm of Ehmke et al. [7]; to our adaptive sampling algorithm of Section 3 (Sampling 1) with $s_{\max} = 10,000$ and $\delta = 0.01$; and to standard sampling (Sampling 2) with 10,000 samples, by solving the VRP problem on all 10 instances with each $\epsilon \in \{0.01, 0.05, 0.1\}$; see Table 2. We then evaluate the solutions of the algorithms by sampling with 100,000 samples using the covariances, which enables us to calculate the "real" objective value and count how many of the chosen routes are actually infeasible. The

ϵ	Algorithm 2			Algorithm E			Sampling 1			Sampling 2		
	time	# inf.	obj.	time	# inf.	obj.	time	# inf.	obj.	time	# inf.	obj.
0.01	61.8	1	1.00	49.2	4	1.00	28.5	0	1.00	234.2	0	1.00
0.05	74.8	0	1.00	58.4	0	1.00	34.6	0	1.00	283.7	0	1.00
0.1	81.0	0	1.00	63.0	1	1.00	39.9	1	1.00	325.0	1	1.00

Table 2: *Comparison for Single Chance Constraint with correlations*

objective function is $D + \frac{1}{2}W$, i.e., waiting is half as expensive as driving. For every sample with covariances, we calculated a vector of standard gaussians z with dimension $|E|$, multiplied it with a matrix L calculated by Cholesky decomposition of Σ, and added the vector of expected values. For solving the exact VRP we used the Set Partitioning formulation and solved it with CPLEX 12.9. All algorithms are implemented in Java version 1.8.0_212 on an Intel(R) Xeon(R) CPU E5-2680 0 with 2.8 GHz.

Table 2 consists of five main columns. In the first column, the value of ϵ is specified. The second column describes the results for the algorithm with covariances and the third for the algorithm without covariances; the fourth and fifth column contain the results for sampling. The second to fifth column are divided into three subcolumns each: the first presents the average cpu time in seconds, the second the total number of "optimal" solutions containing at least one infeasible route, and the third the objective of the method divided by the objective of the algorithm with covariances in the cases in which both methods have produced feasible solutions. In other cases a comparison would be unfair because the algorithm with more infeasible routes clearly has an advantage in terms of the objective value.

We can see that our Algorithm 2 outperforms the algorithm of Ehmke et al. [7] in terms of feasibility in 13% of the settings. In 4 instances it had less infeasible routes in the solution than the algorithm of [7] and there was only one instance in which both algorithms were infeasible. For all other instances, both algorithms returned the same solution and therefore the objective is the same in all feasible settings. The algorithm of [7] needs slightly less running time (between 77% and 79%), which is not surprising because it has to perform less calculations for every route. On the other hand, it computes more feasible routes and therefore the gap is not significant. We can conclude that the advantages of the algorithm considering covariances in terms of feasibility clearly outweigh the slightly higher running time. The adaptive sampling saves in average around 90% of the running time compared to the normal sampling while keeping the same accuracy. It has the same amount of infeasible solutions as our Algorithm 2 but is 2 times faster.

4.2 Including time dependency

We next address time dependency, i.e., we now allow that the traveling time of every arc e in the network varies depending on the time of the day. More precisely, we assume that the expected value and the variance of the travel time needed for an arc are functions of the point in time when the arc is entered. In practice, the traffic situation and hence the travel times strongly depend on the time of the day.

The difficulty here is that the time in which an arc is entered is itself a random variable, so that we have to deal with normally distributed random variables having expectations and variances that are implicitly defined by random variables again. An important modelling issue is how to define the dependency, i.e., which type of functions to allow. We decided to use a piecewise constant model, as it keeps the definition of instances easy and at the same time allows to efficiently update expected values and variances in our algorithm. Alternative approaches could use piecewise linear models or polynomials, splines, or even trogonometric approximations.

4.2.1 Algorithm

We assume to have information on travel times for a fixed set $t_0, ..., t_\Omega$ of time points during the day. We produce two piecewise constant functions describing the distribution at time t for arc e by the expected value $\mu_e(t)$ and the variance $\sigma_e^2(t)$, given the values $\mu_e(t_i)$ and $\sigma_e^2(t_i)$ for $i = 0, ..., \Omega$, as $\mu_e(t) := \mu_e(t_i)$ if $t \in [t_i, t_{i+1}]$ and analogously for $\sigma_e(t)$ (setting $t_{\Omega+1} = \infty$).

In our algorithm, we do not know the exact time when arc e is entered, it is given by a normal distribution. Hence, we have to consider t a random variable. Given its distribution function F_t, we can obtain the expected parameters for the distribution of travel time for e as

$$E[\mu_e(t)] = \sum_{i=0}^{\Omega} \mu_e(t_i) P\left(t \in (t_i, t_{i+1}]\right) = \sum_{i=0}^{\Omega} \mu_e(t_i)\left(F_t(t_{i+1}) - F_t(t_i)\right)$$

and analogously for $\sigma_e^2(t)$. This formula is used in Algorithm 3, Lines 6–7. The remaining parts of Algorithm 3 are analogous to Algorithm 2 except that no correlations are taken into account.

Algorithm 3 Feasibility Check with Single Chance Constraint, taking time dependency into account (but no correlations)

Input: route $r = (0, r_1, \ldots, r_t, 0)$
Output: decision if r is feasible with high probability; expected driving time D; expected waiting time W

1: $k = 0$
2: $exp_k = 0, var_k = 0$
3: $W = 0$

4: **for** $e \in E_r \setminus (r_t, 0)$ **do**
5:     assume $S \sim \mathcal{N}(exp_k, var_k)$
6:     $exp_k \mathrel{+}= E[\mu_e(S)]$
7:     $var_k \mathrel{+}= E[\sigma_e^2(S)]$
8:     assume $\tilde{X}_k \sim \mathcal{N}(exp_k, var_k)$
9:     **if** $P(\tilde{X}_k > b_e) > \epsilon$ **then**
10:         discard route r
11:     **end if**
12:     $exp_{k+1} = E[max\{\tilde{X}_k, a_e\}]$ ▷ Formula (A1)
13:     $var_{k+1} = Var[max\{\tilde{X}_k, a_e\}]$ ▷ Formula (A2)
14:     $W \mathrel{+}= E[max\{a_e - \tilde{X}_k, 0\}]$ ▷ Formula (A1)
15:     $k \mathrel{+}= 1$
16: **end for**

17: $D = exp_k + \mu_{(r_t, 0)}$
18: accept route r and return D and W

4.2.2 Experimental Results

Because we do not have hourly data for more than one day, we artificially extended the instances described above by using data of one day from the same streets hourly taken from 8 am to 5 pm. For every edge and for the value of every hour we computed the quotient to the value of 3 pm. These quotients were multiplied with the expected value of the edge to generate the expected value for that edge in that time. The variances remained unchanged, which is a disadvantage to our algorithm with time dependencies because in reality the variances also tend to be higher in times of the day with more traffic, and our algorithm could exploit this information while the other algorithm cannot.

Results are shown in Table 3. We compare our new Algorithm 3 with Algorithm 2 described above, however without using correlation information, i.e., with the approach proposed by Ehmke

et al. [7]; with our adaptive sampling algorithm (Sampling 1) with $s_{\max} = 10,000$ and $\delta = 0.01$; and with standard sampling (Sampling 2) with 10,000 samples. The results show that in 6 out of 30 settings the algorithm without time dependencies produced infeasible routes, whereas the algorithm taking them into account always returned feasible solutions. The feasibility was again checked using sampling with 100,000 samples. Also in terms of solution value the algorithm without time dependency is less efficient and returns solutions with a value between 1 and 3 percent higher. On the other hand, it uses only 12 to 14 percent of the running time compared to the new algorithm. The adaptive sampling algorithm outperforms the normal sampling by far in running time and is nearly as fast as the algorithm that ignores the time dependencies. Both sampling approaches did not produce infeasible solutions.

	Algorithm 3			Algorithm 2/E			Sampling 1			Sampling 2		
ϵ	time	# inf.	obj.	time	# inf.	obj.	time	# inf.	obj.	time	# inf.	obj.
0.01	256.6	0	1.00	34.0	1	1.01	56.1	0	1.00	541.7	0	1.00
0.05	308.1	0	1.00	41.9	2	1.03	68.5	0	1.00	654.6	0	1.00
0.1	338.9	0	1.00	45.6	3	1.03	75.5	0	1.00	695.2	0	1.00

Table 3: *Comparison for Single Chance Constraint with time dependencies (without covariances)*

4.3 Combining correlations and time dependency

In the preceding sections, we have explained how to take correlations and time dependency into account separately. It is also possible to combine both in one algorithm. For this, every time we need to know a covariance $\sigma_{e,f}(t_e, t_f)$ between two edges e and f, we need to compute an expected covariance of a two-dimensional piecewise constant function over two jointly normally distributed random variables t_e and t_f (the starting times of the two edges). As before, the starting time of an edge is just the random variable describing the length of the route up to that time. Therefore, we also need the covariances between subroutes. However, we only need to consider subroutes that start from the depot, so that the number of covariances to be computed remains quadratic in the route length. These covariances can be computed analogously to the covariance between the entire route and a single edge. All necessary data can be computed in the moment when it is required. When using piecewise constant distributions, as above, we need the joint distribution function of two normally distributed random variables to calculate the expected covariance. Computing this is numerically challenging but possible. It is also possible to sample the expected covariance with the given expected values and covariance matrix for the two starting points. This would lead to a hybrid approach using a combination of sampling and a deterministic algorithm.

We would like to emphasize that any sampling approach is impracticable when considering correlations and time dependencies. In both approaches discussed so far, the same samples could be used for all routes, and the covariance matrix Σ was fixed. In particular, the Cholesky decomposition, needed to sample travel times in the dependent case, had to be computed only once. When considering the time-dependent case, we need the knowledge about the order of customers in the route already during the process of sampling. Therefore, we need to consider every route separately and cannot create one sample used for all routes. Even worse, since the expected covariance $\sigma_{e,f}(t_e, t_f)$ depends on the starting times t_e and t_f now, the Cholesky decomposition cannot be computed in the preprocessing phase any more, since the matrix Σ now depends on t_e and t_f. This computation is further complicated when taking waiting into account.

As a consequence, sampling is not a practical approach here, so that we cannot present a meaningful comparison with our approach. Moreover, without a sampling-based method, we cannot evaluate feasibility and quality of the solutions computed by our approach as in the last sections. For this reason, we do not present an experimental evaluation for the case of time dependent and correlated travel times.

5 SVRPTW with Backhauls and Linehauls – Joint Chance Constraints

So far, we have assumed that feasibility of a route is determined by a high enough probability to reach each customer within its time window. An implicit assumption in our model was that the failure of serving a customer does not influence the feasibility of the rest of the route directly. We now consider the case where failing to serve a customer may directly imply the failure of serving one of the remaining costumers on the route, motivated by problems where customer demands can be either deliveries or a pickups. We will refer to this problem as a Stochastic VRPTW with Backhauls (pickup) and Linehauls (deliveries), SVRPTW-BL. Note however that our problem differs from the classical VRP with Backhauls [28], because no precedence constraints on the deliveries are assumed, and it differs from the VRP with Pickup and Delivery [28], in that the latter assumes a delivery and a pickup for each customer.

Our investigation of this VRP variant is motivated by problems in intermodal freight transport, for example freight transportation in intermodal containers using trucks, which involves the distribution of loaded and empty containers between an intermodal facility or depot and fixed export and import customers. In these applications, a large number of different customer request types and container constraints may arise [3]. In particular, if a customer in a route is not served for some reason, the feasibility of the rest of the route is not guaranteed anymore, e.g., a failed pickup of an empty container leads to the failure in loading the freight at the next customer or a failed delivery leads to a lack of space in the truck for a subsequent pickup.

We thus have to guarantee a service level for the whole route and not for each single customer, that is, we need to consider a joint chance constraints in the place of single chance constraints. In other words, we have to deal with the risk of infeasibility of the whole route rather that considering the risk of failing to serve any of the customers in the route independently. For simplicity, we assume here that failing to serve any customer within its time window makes the entire route infeasible. In the remainder of this section, an algorithm for this case, using joint chance constraints in an approximative way, is provided and computational results are presented. In principle, by a combination with the approach presented in the preceding section, our approach could be extended to the case where not all failures render the rest of the route infeasible. Besides having to deal with joint chance constraints now, another complication is that, after passing a customer, distributions have to be truncated, as we are only interested in travel times under the assumption that the customers visited so far have been served in time.

5.1 Algorithm

We now present an algorithm to address the SVRPTW-BL problem described above; see Algorithm 4. It takes correlations into account, but does not deal with time dependency. With respect to Algorithm 2, two major changes arise. Firstly, instead of checking a chance constraint for every customer, we have to check one chance constraint for the whole route. Unfortunately, dealing with joint chance constraints is hard, and there is no compact formula modeling the joint risk, therefore it is often approximated in the literature by calculating the risk of every single event (here a failure of one customer), summing it up (using ϵ_{total} in Algorithm 4), and comparing it to the maximum allowed risk for the joint chance constraint ϵ. We follow this approach; the corresponding changes in Algorithm 4 concern Lines 19-22.

When Algorithm 4 ends with a feasible route, i.e. when $\epsilon_{total} \leq \epsilon$, the value ϵ_{total} represents the probability of the route failure. This information could also be used in a bicriteria-style approach: instead of discarding all routes with $\epsilon_{total} > \epsilon$, one could sort the routes by their risk and solve the optimization problem for different risk levels ϵ without having to recompute the feasible routes. E.g., in a column generation approach, increasing ϵ would then just lead to the addition of more columns.

The second change concerns the truncation discussed above: for calculating the expected value and variance for later arrival times in the route, only the scenarios being feasible so far should be considered. E.g., if a driver misses the time window of the first customer and thus cannot serve it, he cannot continue to the second customer, and therefore the arrival time in this scenario does not influence the arrival time at the second customer, and so on. To model this, we use a one sided truncated normal distribution of the upper tail (using $\hat{\mu}$ and $\hat{\sigma}$ to denote expectations and covariances after truncation). After checking the chance constraint, the truncation for the expected value and for the variance are performed in Line 23 and 24. As we consider joint normal distributions, the truncation of one variable also affects the covariances between two other variables, so the whole submatrix of Σ corresponding to the current route has to be updated; see Lines 25–28. Note that the truncated distributions are again replaced by normal distributions. The formulas used for computing the truncated distributions can be found in Appendix A.

Algorithm 4 Feasibility Check for Joint Chance Constraint for a Route r

---

Input: route $r = (0, r_1, \ldots, r_t, 0)$
Output: decision if r is feasible with high probability; expected driving time D; expected waiting time W
1: $k = 0$
2: $exp_k = 0$, $var_k = 0$
3: **for** $e \in E_r \setminus (r_t, 0)$ **do**
4:      $cov_{k,e} = 0$
5: **end for**
6: **for** $e, f \in E_r$ with $e \neq f$ **do**
7:      $\hat{\mu}_e = \mu_e$
8:      $\hat{\sigma}_{e,f} = \sigma_{e,f}$
9: **end for**
10: $W = 0$, $\epsilon_{total} = 0$
11: **for** $e \in E_r$ **do**
12:      $exp_k \mathrel{+}= \hat{\mu}_e$
13:      $var_k \mathrel{+}= \hat{\sigma}_e^2$
14:      $var_k \mathrel{+}= 2cov_{k,e}$
15:      **for** $f \in E_r$ after e **do**
16:          $cov_{k,f} \mathrel{+}= \hat{\sigma}_{e,f}$
17:      **end for**
18:      assume $\tilde{X}_k \sim \mathcal{N}(exp_k, var_k)$
19:      $\epsilon_{total} \mathrel{+}= P(\tilde{X}_k > b_e)$
20:      **if** $\epsilon_{total} > \epsilon$ **then**
21:          discard route r
22:      **end if**
23:      $exp_k = E[\tilde{X}_k \mid \tilde{X}_k < b_e]$ $\triangleright$ Formula (A5)
24:      $var_k = Var[\tilde{X}_k \mid \tilde{X}_k < b_e]$ $\triangleright$ Formula (A7)
25:      **for** $f, g \in E_r$ after e with $f \neq g$ **do**
26:          $\hat{\mu}_f = E[X_f \mid \tilde{X}_k < b_e]$ using $cov_{k,f}$ $\triangleright$ Formula (A5)
27:          $\hat{\sigma}_{f,g} = Cov[X_f, X_g \mid \tilde{X}_k < b_e]$ using $cov_{k,f}$ and $cov_{k,g}$ $\triangleright$ Formula (A6)
28:      **end for**
29:      assume $\tilde{X}_k \sim \mathcal{N}(exp_k, var_k)$
30:      $exp_{k+1} = E[max\{\tilde{X}_k, a_e\}]$ $\triangleright$ Formula (A1)
31:      $var_{k+1} = Var[max\{\tilde{X}_k, a_e\}]$ $\triangleright$ Formula (A2)
32:      **for** $f \in E_r$ after e **do**
33:          $cov_{k+1,f} = Cov[max\{\tilde{X}_k, a_e\}, \hat{X}_f]$ $\triangleright$ Formula (A3)
34:      **end for**
35:      $W \mathrel{+}= E[max\{a_e - \tilde{X}_k, 0\}]$ $\triangleright$ Formula (A1)
36:      $k \mathrel{+}= 1$
37: **end for**
38: $D = exp_k + \mu_{(r_t,0)}$
39: accept route r and return D and W

ϵ	Algorithm 4			Algorithm 5			Algorithm 6			Sampling 1			Sampling 2		
	time	# inf.	obj.	time	# inf.	obj.	time	# inf.	obj.	time	# inf.	obj.	time	# inf.	obj.
0.01	303.5	0	1.00	55.1	2	1.00	45.5	0	1.00	23.1	0	1.00	179.7	0	1.00
0.05	374.0	1	1.00	65.4	2	1.00	53.3	1	1.00	31.2	1	1.00	248.0	1	1.00
0.1	419.8	0	1.00	71.0	1	1.00	60.9	0	1.00	36.2	1	1.00	325.1	1	1.00

Table 4: *Comparison for Joint Chance Constraint with correlations*

5.2 Experimental Results

In our experimental evaluation, we create instances in exactly the same way as described in Section 4.1.2, except that different sets of time windows may now be discarded due to infeasibility. Table 4 shows the results of a comparison between Algorithm 4 considering correlations and truncations; the variant of it not considering correlations (Algorithm 5); another variant not performing truncations but considering correlations (Algorithm 6); our adaptive sampling algorithm (Sampling 1) with $s_{\max} = 10,000$ and $\delta = 0.01$; and normal sampling (Sampling 2) with $10,000$ samples. The algorithm considering correlations and truncations returned in only one of thirty settings an infeasible route. Ignoring the truncations did not lead to more infeasible settings, but decreased the running times by around 85 percent. Ignoring the correlations, on the other hand, led to five times more infeasible settings. In fact, every sixth setting resulted in an infeasible route in this approach. At the same time, it did not even run faster than the algorithm ignoring truncations, because it allows more feasible routes, which outweighs the additional running time per route. Again adaptive sampling is much faster than the normal sampling by keeping the same accuracy. Both sampling methods made two mistakes, which is double as much as Algorithm 4. As a conclusion, the recommended algorithm for solving instances of SVRPTW-BL considers correlations but ignores truncations.

6 Conclusion

We devised several algorithms for checking the feasibility of a given route in the Stochastic Vehicle Routing Problem with time windows and with correlated and time-dependent travel times. They are either based on closed formulas or sampling. We consider either single or joint chance constraints depending on whether missing a customer's time window makes the entire route infeasible or not. These algorithms can be embedded into most algorithms for solving the Vehicle Routing Problem with hard time windows and waiting times as a feasibility check. Due to the fact that the algorithms also compute expected waiting times, they can be easily adapted to variants of the problem with soft time windows and penalties. The experimental results obtained in our case study show that the algorithms assess the feasibility of a given route with a reasonable precision, in particular when correlations are taken into account. However, for a large number of samples, the latter approach is time consuming.

Considering the complementary strengths of the deterministic and the sampling approach, one could also investigate combinations of both approaches. E.g., one could check for a given route how likely it is to become infeasible, and if the result is very close to the value of acceptance, do a second check performing sampling with a high number of samples – possibly only for routes that were accepted, but very close to be not accepted, because adding an infeasible route is in general more problematic than missing a feasible one. Another hybrid approach would be to use sampling with a high number of samples just as a callback inside the optimization process: first compute an optimal solution using our approach, then check the feasibility of all routes used in this solution by sampling, and if some route turns out to be infeasible, remove it and re-optimize. The advantage of this approach is that the sampling has to be performed for a very small set of routes and can thus be performed with a large number of samples in reasonable time.

A Computation of Updated Moments

For the convenience of the reader, we explain in the following how to compute the expected values, variances and covariances of the random variables appearing in Algorithms 2 and 4. For this, we use the assumptions and notation of Section 2, in particular, we assume $X \sim \mathcal{N}(\mu, \Sigma)$ and thus $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$. For indices $i_1, \dots, i_k$, we consider the joint density function of $X_{i_1}, \dots, X_{i_k}$, given by

$$f_{X_{i_1}, \dots, X_{i_k}}(s_1, \dots, s_k) := \frac{1}{\sqrt{(2\pi)^k \det(\bar{\Sigma})}} e^{-\frac{1}{2}(s-\bar{\mu})^\top \bar{\Sigma}^{-1}(s-\bar{\mu})},$$

where $\bar{\mu}$ denotes the vector with entries $\mu_{i_1}, \dots, \mu_{i_k}$ and $\bar{\Sigma}$ the corresponding submatrix of Σ. We again assume here that Σ and hence $\bar{\Sigma}$ is positive definite.

A.1 Rectified Gaussian Distributions

We first consider the random variable $\max\{c, X_i\}$ for a constant c. For the expected value, we obtain

$$
\begin{aligned}
E(\max\{c, X_i\}) &= \int_{-\infty}^{\infty} \max\{c, s\} f_{X_i}(s) \, ds \\
&= \int_{-\infty}^{c} c f_{X_i}(s) \, ds + \int_{c}^{\infty} s f_{X_i}(s) \, ds \\
&= c P(X_i \leq c) + \mu_i P(X_i \geq c) + \sigma_{ii} f_{X_i}(c)
\end{aligned}
\tag{A1}
$$

To compute the variance

$$Var(\max\{c, X_i\}) = E(\max\{c, X_i\}^2) - E(\max\{c, X_i\})^2 \tag{A2}$$

we can use (A1) and

$$
\begin{aligned}
E(\max\{c, X_i\}^2) &= \int_{-\infty}^{\infty} \max\{c, s\}^2 f_{X_i}(s) \, ds \\
&= \int_{-\infty}^{c} c^2 f_{X_i}(s) \, ds + \int_{c}^{\infty} s^2 f_{X_i}(s) \, ds \\
&= c^2 P(X_i \leq c) + (\mu_i^2 + \sigma_{ii}) P(X_i \geq c) + \sigma_{ii}(c + \mu_i) f_{X_i}(c)
\end{aligned}
$$

Finally, we have

$$
\begin{aligned}
E(X_j \cdot \max\{c, X_i\}) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t \max\{c, s\} f_{X_i, X_j}(s, t) \, ds \, dt \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{c} t c f_{X_i, X_j}(s, t) \, ds \, dt + \int_{-\infty}^{\infty} \int_{c}^{\infty} t s f_{X_i, X_j}(s, t) \, ds \, dt \\
&= c \mu_j P(X_i \leq c) + (\mu_i \mu_j + \sigma_{ij}) P(X_i \geq c) + \mu_j \sigma_{ii} f_{X_i}(c)
\end{aligned}
$$

and thus obtain

$$
\begin{aligned}
Cov(X_i, \max\{c, X_j\}) &= E(X_i \cdot \max\{c, X_j\}) - E(X_i) E(\max\{c, X_j\}) \\
&= \sigma_{ij} P(X_j \geq c)
\end{aligned}
\tag{A3}
$$

A.2 Truncated Gaussian Distributions

We next develop formulas for the moments of the random vector X truncated by a condition $X_i < b$, where b is again a constant. For the expected values, we have

$$
\begin{aligned}
E(X_j \mid X_i \leq b) &= \frac{1}{P(X_i \leq b)} \int_{-\infty}^{+\infty} \int_{-\infty}^{b} t f_{X_i, X_j}(s, t) \, ds \, dt \\
&= \mu_j - \sigma_{ij} \frac{f_{X_i}(b)}{P(X_i \leq b)}
\end{aligned}
\tag{A4}
$$

and, in particular,

$$E(X_i \mid X_i \leq b) = \mu_i - \sigma_{ii} \frac{f_{X_i}(b)}{P(X_i \leq b)} \tag{A5}$$

For the covariances, we have

$$Cov(X_j, X_k \mid X_i \leq b) = E(X_j X_k \mid X_i \leq b) - E(X_j \mid X_i \leq b)E(X_k \mid X_i \leq b) \tag{A6}$$

which can be computed using (A4) and

$$E(X_j X_k \mid X_i \leq b) = \frac{1}{P(X_i \leq b)} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{b} tu f_{X_i, X_j, X_k, X_i}(s, t, u) \, du \, ds \, dt$$

$$= \mu_j \mu_k + \sigma_{jk} - \left(\frac{\sigma_{ij}\sigma_{ik}}{\sigma_{ii}}(b - \mu_i) + \sigma_{ik}\mu_j + \sigma_{ij}\mu_k \right) \frac{f_{X_i}(b)}{P(X_i \leq b)}$$

As a special case, we obtain a formula for

$$Var(X_j \mid X_i \leq b) = Cov(X_j, X_j \mid X_i \leq b) \tag{A7}$$

which also applies to the case $i = j$.

References

[1] Bastian C., Rinnooy Kan A.H.G., *The stochastic vehicle routing problem revisited*, European Journal of Operational Research, vol. 56, issue 3, 407-412, 1992.

[2] Bebu I., Mathew T., *Confidence intervals for limited moments and truncated moments in normal and lognormal models*, Statistics & Probability Letters, vol. 79, issue 3, 375 - 380, 2009.

[3] Bomboi F., Pruente J., *A new drayage problem with different customer services and container requirements*, Technical Report, 2018.

[4] Clark C. E., *The Greatest of a Finite Set of Random Variables*, Operations Research, vol. 9, issue 2, 145 - 162, 1961.

[5] Clarke G., Wright, J. W., *Scheduling of Vehicles from a Central Depot to a Number of Delivery Points*, Operations Research, vol. 12, issue 4, 568-581, 1964.

[6] Dinh T., Fukasawa R., Luedtke J., *Exact algorithms for the chance-constrained vehicle routing problem*, Mathematical Programming, vol. 172, issue 1, 105 - 138, 2018.

[7] Ehmke J. F., Campbell A. M., Urban T. L., *Ensuring service levels in routing problems with time windows and stochastic travel times*, European Journal of Operational Research, vol. 240, issue 2, 539 - 550, 2015.

[8] Errico F., Desaulniers G., Gendreau M., Rei W., Rousseau L. M., *A priori optimization with recourse for the vehicle routing problem with hard time windows and stochastic service times*, European Journal of Operational Research, vol. 249, issue 1, 55 - 66, 2016.

[9] Gendreau M., Laporte G., Séguin R., *Stochastic vehicle routing*, vol. 88, issue 1, 3-12, 1996.

[10] Golden B. L., Yee J. R., *A Framework For Probabilistic Vehicle Routing*, A I I E Transactions, vol. 11, issue 2, 109-112, 1979.

[11] Google Maps, `maps.google.com`.

[12] Guo Z. G., Mak K. L., *A heuristic algorithm for the stochastic vehicle routing problems with soft time windows*, Proceedings of the 2004 Congress on Evolutionary Computation (IEEE Cat. No.04TH8753), vol. 2, 1449-1456, 2004.

[13] Hoeffding, W., *Probability Inequalities for Sums of Bounded Random Variables*, Journal of the American Statistical Association, vol. 58, issue 301, 13 - 30, 1963.

[14] Hou L., Zhou H., *Stochastic Vehicle Routing Problem with Uncertain Demand and Travel Time and Simultaneous Pickups and Deliveries*, 2010 Third International Joint Conference on Computational Science and Optimization(CSO), vol. 01, 32-35, 2010.

[15] Lambert V., Laporte G., Louveaux F., *Designing collection routes through bank branches*, Computers & Operations Research, vol. 20, issue 7, 783-791, 1993.

[16] Li X., Tian P., Leung S.C.H., *Vehicle routing problems with time windows and stochastic travel and service times: Models and algorithm*, International Journal of Production Economics, vol. 125, issue 1, 137-145, 2010.

[17] Miranda D. M., Conceicao S. V., *The vehicle routing problem with hard time windows and stochastic travel and service time*, Expert Systems with Applications, vol. 64, 104-116, 2016.

[18] Nadarajah S., Kotz S., *Exact Distribution of the Max/Min of Two Gaussian Random Variables*, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 16, issue 2, 210 - 212, 2008.

[19] Nahum O. E., Hadas Y., *Developing a model for the stochastic time-dependent vehicle-routing problem*, 2009 International Conference on Computers Industrial Engineering, 118-123, 2009.

[20] Novoa C., Berger R., Linderoth J., Storer R., *A set-partitioning-based model for the stochastic vehicle routing problem*, technical report 06T-008, 2006.

[21] Oyola J., Arntzen H., Woodruff D. L., *The stochastic vehicle routing problem, a literature review, Part II: solution methods*, EURO Journal on Transportation and Logistics, vol. 6, issue 4, 349 - 388, 2017.

[22] Oyola J., Arntzen H., Woodruff D. L., *The stochastic vehicle routing problem, a literature review, part I: models*, EURO Journal on Transportation and Logistics, vol. 7, issue 3, 193 - 221, 2018.

[23] Park D.,Laurence R. R., *Forecasting Freeway Link Travel Times with a Multilayer Feedforward Neural Network*, Computer-Aided Civil and Infrastructure Engineering, vol. 14, issue 5, 357-367, 1999.

[24] Sinha D., Zhou H., Shenoy N. V., *Advances in Computation of the Maximum of a Set of Gaussian Random Variables*, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 26, issue 8, 1522-1533, 2007.

[25] Tas D., Dellaert N., van Woensel T., de Kok T., *Vehicle routing problem with stochastic travel times including soft time windows and service costs*, Computers & Operations Research, vol. 40, issue 1, 214-224, 2013.

[26] Tas D., Dellaert N., van Woensel T., de Koktitle T., *The time-dependent vehicle routing problem with soft time windows and stochastic travel times*, Transportation Research Part C: Emerging Technologies, vol. 48, 66-83, 2014.

[27] Tas D., Gendreau M., Dellaert N., Van Woensel T., De Kok A.G., *Vehicle routing with soft time windows and stochastic travel times: A column generation and branch-and-price solution approach*, European Journal of Operational Research, vol. 236, issue 3, 789 - 799, 2014.

[28] Toth P., Vigo D., *The Vehicle Routing Problem*, Siam, 2002.

[29] Van Woensel T., Kerbache L., Peremans H., Vandaele N., *A Vehicle Routing Problem with Stochastic travel times*, In Fourth Aegean International Conference on Analysis of Manufacturing Systems location, Samos, 2003.

[30] Zhang J., Lam W. H. K., Chen B. Y., *A Stochastic Vehicle Routing Problem with Travel Time Uncertainty: Trade-Off Between Cost and Customer Service*, Networks and Spatial Economics, vol. 13, issue 4 471-496, 2013.