

Tracking a rainfall index constrained by Conditional Value-at-Risk

Brian Webby¹

John Boland²

Phil Howlett³

Andrew Metcalfe⁴

(Received 10 February 2010; revised 27 April 2010)

Abstract

Portfolio replication is a well studied problem from finance, whereby a subset of instruments is selected to track the average movement of all such instruments. Using expected monetary value as the decision criterion for selection is a trade-off of over and under estimators. An alternate criterion developed in finance is Conditional Value-at-Risk, employed to measure and restrict excessive deviation either above or below the centre. That same objective may be useful in monitoring environmental variables where the exclusion or inclusion of extreme values may be critical to representing the features of interest of the variable. We demonstrate the application of the risk measure to meteorological data in Australia, selecting a weighted combination of stations to best track an index of average annual rainfall.

<http://anziamj.austms.org.au/ojs/index.php/ANZIAMJ/article/view/2607>
gives this article, © Austral. Mathematical Soc. 2010. Published May 11, 2010. ISSN
1446-8735. (Print two pages per sheet of paper.)

Contents

1 Introduction	C202
1.1 Application in a rainfall problem	C203
2 Model specification	C204
2.1 Conditional Value-at-Risk	C204
2.2 Mathematical formulation of the problem	C206
3 Results	C208
4 Conclusion	C212
References	C213

1 Introduction

A common problem in finance is to construct a portfolio to track a financial index. For example, the ASX 200 is the best known index in Australia. Fund managers investing in ASX securities often attempt to have their fund closely track any movements in the index, preferably out-performing the ASX 200, or at least performing no worse than the index. Such a fund acts as a replicating portfolio.

A replicating portfolio consists of shares in a subset of the companies listed in the index. The portfolio replication problem is to identify which companies to include in the subset along with the proportions of holdings to take in each of these companies. Typically, the number of companies in the replicating portfolio and the proportion of funds held in each company would be bounded. The agreement of the portfolio with the index is typically measured as the average absolute relative difference in their respective values. The agreement is tested at periodic intervals over a specified time period.

This is the problem that Rockafellar and Uryasev address in their 2002 paper “Conditional value-at-risk for general loss distributions” [2]. Their objective is to minimise the average absolute relative difference between the portfolio and the index; however, they also include a CVaR constraint in order to limit the degree of undershooting by the portfolio. In the financial context, the replicating portfolio is intended to perform at least as well as the index and so a one-sided constraint is apt.

An analogous problem in environmental monitoring is that of tracking a weather variable. For example, rainfall is distributed stochastically in time and space across an area and monitoring stations are restricted in number. A subset of stations can be used to track average rainfall across an area with undershooting of the index restricted so that the portfolio accurately represents the minimal average rainfall. A similar use for the model could be in tracking average water holdings (with minimal undershooting) in all small and large dams across an area by monitoring the holdings of a few. Another potential application is in the selection of a number of meteorological stations across Australia to best represent the average temperature of the nation, minimising over-reporting. Given restricted budgets and personnel, the number of monitoring stations and the amount of data collected might be reduced, with the most representative stations identified and retained. The technique for solving the portfolio replication problem provides a method for identifying the stations or evaluating such monitoring programs.

1.1 Application in a rainfall problem

The example we use to illustrate the technique is that of tracking average annual rainfall for the six Australian state capitals, from the annual rainfall records of any two of the capitals. The focus here is on the definition and implementation of the technique. The problem is of interest since websites and newspapers which publish world climatic data need to be selective on sites reported. CVaR was developed in finance but has been applied to water

resources problems [3, 4, 5].

We used rainfall data for the regional Bureau of Meteorology offices in the six state capitals, in a contemporaneous data set from 1882 to 1991. A time series of the data is given in Figure 1. The data record has 110 years of six-variate data. We estimate parameters for the model using the first 80 years and assess the model with the remaining 30 years.

2 Model specification

2.1 Conditional Value-at-Risk

Value-at-Risk is a downside risk measure, defined as the maximum loss likely to be incurred in a given time period at a specified probability. For a continuous loss distribution and a probability of $\alpha = 0.95$, VaR is the 95th percentile of the distribution. Conditional Value-at-Risk is the expected loss of those losses greater than or equal to VaR . In other words, CVaR is the mean of the upper tail of the loss distribution. In the situation considered here, the loss distribution is a discrete one, comprised of the losses associated with the finite number of various combinations of capitals and proportions of those capitals. For this scenario-based optimisation, the special function of Rockafellar and Uryasev provides a convex function which, when minimised, gives the minimum value of CVaR [1]. For a loss $\zeta = f(\mathbf{x}, \mathbf{y})$ where \mathbf{x} is a vector of possible decisions and \mathbf{y} is a variable influencing the loss associated with a particular decision, the special function is

$$F_{\alpha}(\mathbf{x}, \zeta) = \zeta + \frac{1}{1 - \alpha} \mathbb{E} \{ [f(\mathbf{x}, \mathbf{y}) - \zeta]^+ \}, \quad \text{where } [t]^+ = \max(0, t) .$$

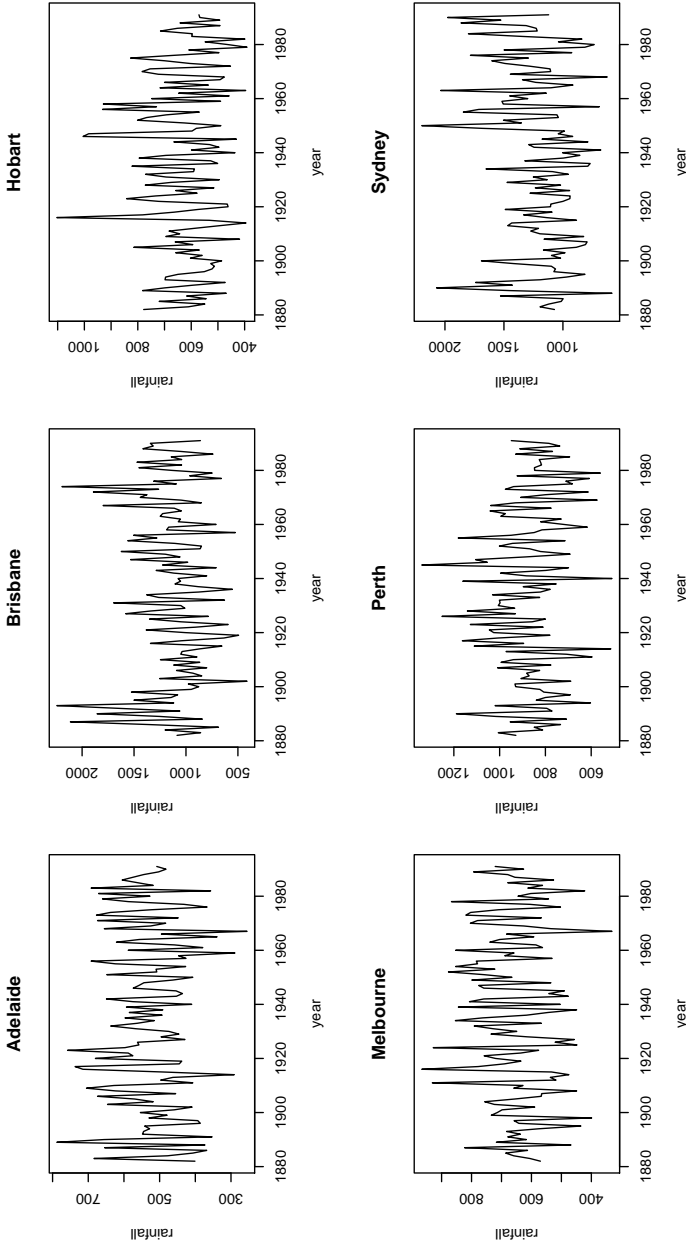


FIGURE 1: Time series of annual rainfall (mm) for the six Australian state capitals, 1882–1991.

2.2 Mathematical formulation of the problem

Below we list the terms appearing in the formulation, following the notation used by Rockafellar and Uryasev [2].

- I_t is the index that is tracked, that is, I_t is the average annual rainfall of the six Australian state capitals at time t , $t = 1, \dots, T$, with $t = 1$ representing the year 1882.
- S_j , $j = 1, \dots, n$, are the capitals. We restrict the number of capitals in the portfolio to two, giving $\binom{6}{2} = 15$ possible combinations of capitals.
- k , $k = 1, \dots, 15$, is a list of the permitted combinations of capitals.
- p_{tj} is the annual rainfall of capital j at time t .
- x_j is the proportion of capital j 's rainfall in the replicating portfolio. The value of the portfolio at time t is $\sum_{j=1}^m p_{tj}x_j$, where $m = 2$.

The deviation of the value of the portfolio from the index or target value is a loss, $z = (x, p)$, where the decision, x , is the proportion of a capital's rainfall to include in the portfolio, and p , the annual rainfall of that capital, is a random variable influencing the loss. The loss associated with a decision x is the relative shortfall

$$f(x, p) = \left(I_t - \sum_{j=1}^m p_{tj}x_j \right) / I_t.$$

The objective of the problem is to identify the combination of capitals which minimises the mean of the absolute value of the relative shortfalls across the specified time period. The positive relative shortfalls are measured and constrained through a bound on CVAR. Proportions must be non-negative and are restricted above. We require the value of the portfolio at time T to equal the value of the index, v , at time T . The minimisation problem is

$$\min_k g(x) = \frac{1}{T} \sum_{t=1}^T \left| \left(I_t - \sum_{j=1}^m p_{tj}x_j \right) / I_t \right| \quad (1)$$

$$\begin{aligned} \text{subject to } & \sum_{j=1}^m p_{Tj} x_j = v, \\ & \zeta + \frac{1}{(1-\alpha)T} \sum_{t=1}^T \left[\left\{ \left(I_t - \sum_{j=1}^m p_{tj} x_j \right) / I_t \right\} - \zeta \right]^+ \leq \omega, \quad (2) \\ & 0 \leq x_j \leq \gamma_j, \quad j = 1, \dots, n. \quad (3) \end{aligned}$$

The minimisation takes place with respect to x and ζ . Constraint (2) is equivalent to requiring CVaR to be less than or equal to a loss tolerance ω for deviations below the target value. Convert to linear conditions the non-negativity condition in (2) and the absolute value condition in (1) by introducing variables η_t such that

$$\eta_t \geq 0 \quad \text{and} \quad \left[\left(I_t - \sum_{j=1}^m p_{tj} x_j \right) / I_t \right] - \zeta - \eta_t \leq 0.$$

The problem then is

$$\begin{aligned} \min_k g(x) &= \frac{1}{T} \sum_{t=1}^T \eta_t \\ \text{subject to } & \sum_{j=1}^m p_{Tj} x_j = v, \\ & \eta_t \geq \left(I_t - \sum_{j=1}^m p_{tj} x_j \right) / I_t \quad \text{and} \quad \eta_t \geq - \left(I_t - \sum_{j=1}^m p_{tj} x_j \right) / I_t, \\ & \zeta + \frac{1}{(1-\alpha)T} \sum_{t=1}^T \eta_t \leq \omega, \\ & 0 \leq x_j \leq \gamma_j \quad \text{for } j = 1, \dots, n, \quad \text{and} \quad \eta_t \geq 0. \end{aligned}$$

3 Results

We restrict undershooting of the index. In practice this means that the portfolio guarantees, to a certain probability level, a minimum amount of average annual rainfall across the six capitals, accepting rainfall above the index to be a benefit. With $\omega \leq 0.3$, seven out of the fifteen combinations meet the CVaR constraint. The optimal proportional combination to track average state capital city rainfall is $1.4297 \times \text{Adelaide} + 0.23198 \times \text{Brisbane}$, with a mean absolute deviation over the training period of 0.2213 (Figure 2). Setting $\omega \leq 1.5$, there are now only two combinations from fifteen which are infeasible and mean absolute deviation has reached its minimum for the conditions (CVaR_{0.95} and $T = 80$). The deviation is 0.0830 from a combination of $0.2444 \times \text{Brisbane} + 0.6438 \times \text{Perth}$ (Figure 3). Relaxing the CVaR limit further, (in effect, solving the linear program using an expected value criterion) does not find a combination with a lower deviation, or, in other words, does not improve the tracking of the index. The last combination to become feasible as the CVaR constraint is relaxed, $\omega \leq 2.5$, is that of Brisbane and Sydney, the two cities with the highest rainfall.

Figure 2 shows the index and the optimal tracking portfolio ($\omega \leq 0.3$) over the 80 years of the training data. The combination of Adelaide and Brisbane rainfall is seldom below the index but often considerably overshoots the index. The optimal portfolio combination ($\omega \leq 1.1$) of Brisbane and Perth is seen to track the index better, Figure 3. A good compromise in minimal undershooting of the index and minimal deviation from the index is the combination of $0.2776 \times \text{Brisbane} + 0.8774 \times \text{Melbourne}$, which meets a CVaR_{0.95} criterion of $\omega \leq 0.6$ and has a mean absolute deviation of 0.0924 (Figure 4).

Altering the probability level to CVaR_{0.90}, with other parameters as before, that is, $\omega \leq 1.1$ and $\gamma = 3$, does not find an improved tracking combination but does produce the result that several of the combinations now have a coefficient of zero against one city. That is, average annual rainfall for the six capitals is represented, somewhat imprecisely, by the rainfall of just one

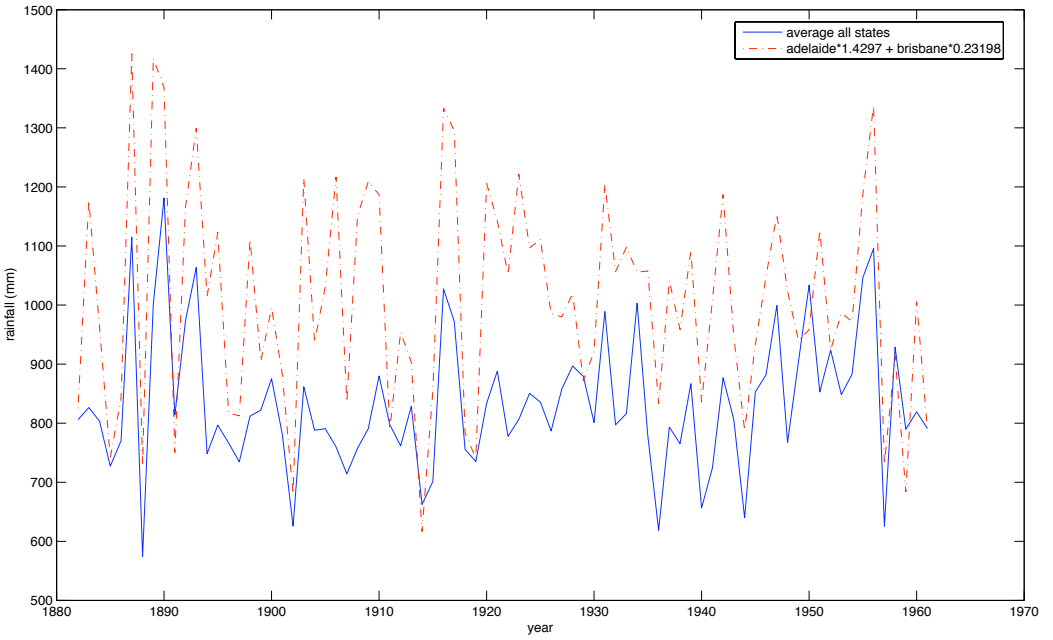


FIGURE 2: Time series of index and tracking portfolio, 1882–1961, with Adelaide and Brisbane rainfall making up the portfolio.

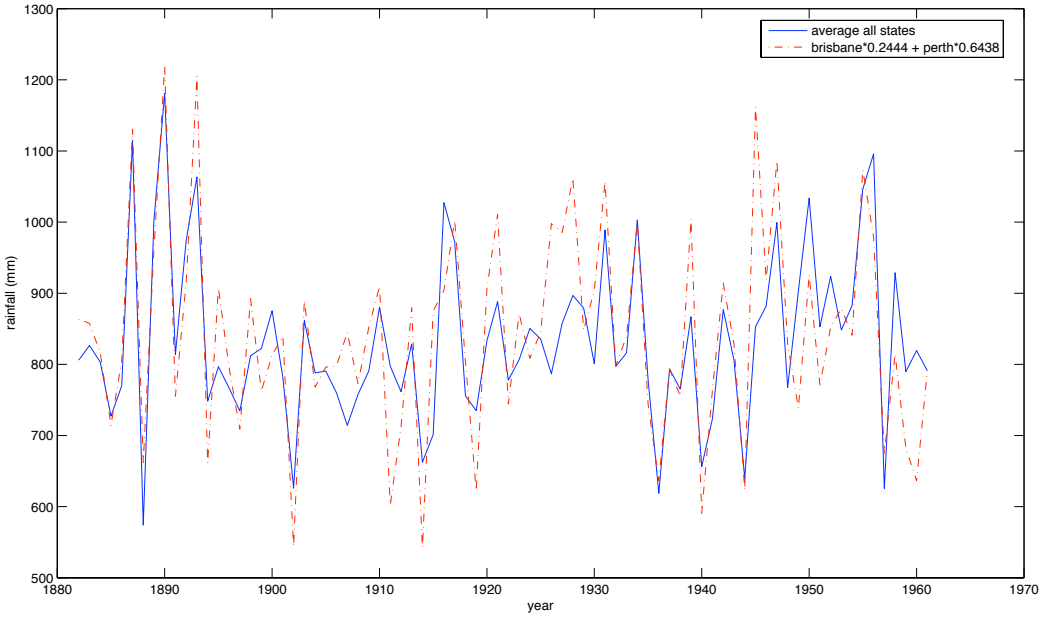


FIGURE 3: Time series of index and tracking portfolio, 1882–1961, Brisbane and Perth rainfall making up the portfolio.

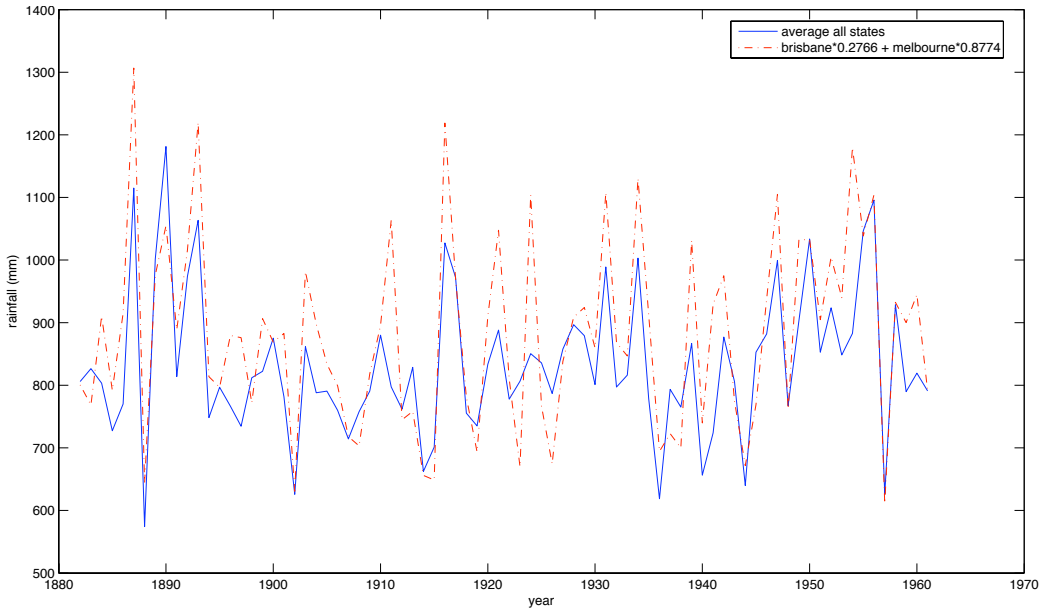


FIGURE 4: Time series of index and tracking portfolio, 1882–1961, Brisbane and Melbourne rainfall making up the portfolio.

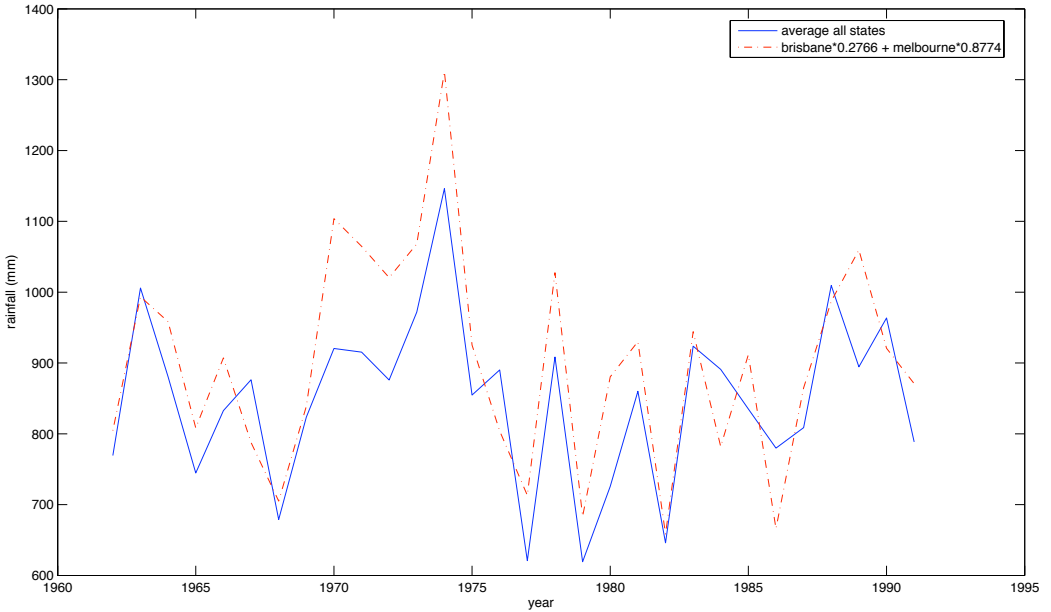


FIGURE 5: Time series of index and tracking portfolio, 1962–1991, the selected portfolio of $0.2766 \times$ Brisbane + $0.8774 \times$ Melbourne rainfall plotted against the last 30 years of the data set.

of them. That with the lowest deviation, of 0.1358 , is $0.9635 \times$ Perth.

Figure 5 shows the performance of the selected replication model over the last 30 years of the data. Mean absolute deviation was 0.1007 , somewhat higher than the combination achieved on the training set; again the pattern of undershooting of the index by this combination was restricted.

4 Conclusion

The objective of a decision criterion, and a linear programming algorithm, is often that of minimising expected loss (conversely, maximising expected

gain). We have extended this paradigm to include restriction of values to one side of the central measure, and demonstrated how to incorporate this restriction as a constraint in a linear program. The optimal model under these conditions had a higher overall deviation (0.0924) than the model found by minimising expected loss (0.0830), that is, by relaxing the CVaR constraint. However, the loss of accuracy was not great, and is readily acceptable where there is reason to prefer a model that restricts undershooting of a target.

References

- [1] Howlett, P. G. and Piantadosi, J. (2007). A note on Conditional Value at Risk. *Optimization*, **56**, 629–632. doi:10.1080/02331930701617080 C204
- [2] Rockafellar, R. T. and Uryasev, S. (2002). Conditional Value-at-Risk for General Loss Distributions. *Journal of Banking and Finance*, **26**, 1443–1471. doi:10.1016/S0378-4266(02)00271-6 C203, C206
- [3] Webby, R. B., Adamson, P. T., Boland, J., Howlett, P. G., Metcalfe, A. V. and Piantadosi, J. (2006). The Mekong—applications of Value-at-Risk and Conditional-Value-at-Risk simulation to the benefits, costs and consequences of water resources development in a large river basin. *Ecological Modelling*, **201**, 89–96. doi:10.1016/j.ecolmodel.2006.07.033 C204
- [4] Yamout, G. M., Hatfield, K. and Romeijn, H. E. (2007) Comparison of new conditional value-at-risk management for optimal allocation of uncertain water supplies. *Water Resources Research*, **43**, W07430. doi:10.1029/2006WR005210 C204
- [5] Yum, K.-K., Blackmore, J. and Anticev, J. (2009). Modelling and evaluating water allocation risks using Value-at-Risk. *Proceedings of the 18th World IMACS/MODSIM Congress*.

<http://www.mssanz.org.au/modsim09/G3/yum.pdf> [20/04/2010].

C204

Author addresses

1. **Brian Webby**, Regional Sustainability Centre, University of South Australia, Whyalla, South Australia, AUSTRALIA.
<mailto:brian.webby@unisa.edu.au>
2. **John Boland**, University of South Australia, Adelaide, South Australia, AUSTRALIA.
3. **Phil Howlett**, University of South Australia, Adelaide, South Australia, AUSTRALIA.
4. **Andrew Metcalfe**, The University of Adelaide, Adelaide, South Australia, AUSTRALIA.