Complementarity of Hydro and Wind Power: Improving the Risk Profile of Energy Inflows

Michel Denault*† Debbie Dupuis‡

Sébastien Couture-Cardinal§

July 27, 2009

Abstract

The complementarity of two renewable energy sources, namely hydro and wind, is investigated. We consider the diversification effect of wind power to reduce the risk of water inflow shortages, an important energy security concern for hydropower-based economic zones (e.g. Québec and Norway) Our risk measure is based on the probability of a production deficit, in a manner akin to the value-at-risk, simulation analysis of financial portfolios. We examine whether the risk level of a mixed hydro-and-wind portfolio of generating assets improves on the risk of an all-hydro portfolio, by relaxing the dependence on water inflows and attenuating the impact of droughts. Copulas are used to model the dependence between the two sources of energy. The data considered, over the period 1958–2003, are for the province of Québec, which possesses large hydro and wind resources.

Our results indicate that for all scenarios considered, any proportion of wind up to 30% improves the production deficit risk profile of an all-hydro system. We can also estimate the value, in TWh, of any additional one percent of wind in the portfolio.

Keywords: Hydropower, wind power, energy inflows, risk of shortage, energy security.

*HEC Montréal (Management Sciences) and GERAD.

†Corresponding author

‡HEC Montréal (Management Sciences) and GERAD.

§Risk Management, National Bank of Canada

1 Introduction

The variability of energy inflows, i.e. water inflows to the dams and power plants, is the main item on a hydropower company's risk list. A water inflow shortage becomes a loss of profit, or, if severe enough, an energy security concern. Large water reservoirs can smooth out seasonal variability, but cannot replace water that is not there. The diversification of the primary source of energy, e.g. adding thermal plants or, as we suggest, wind farms to the hydro system, can be an interesting solution avenue to lower the risk of hydro energy shortages.

Worldwide wind power production, both planned and installed, is growing rapidly and wind is often seen as the best next step in renewable energy. Many industrialized countries have established national plans to increase wind power penetration on their territories. In the province of Québec (Canada), current installations and calls for tenders by Hydro-Québec, the state-owned and by far largest power producer, should bring the total wind power capacity to 4000 MW by 2015, a level largely constrained by the current grid configuration.

With research interest in wind power growing apace, it is not surprising that the problem of wind intermittency has attracted much attention. Several studies have appeared in the past years, specifically on coupling hydro and wind systems to smooth out the output pattern, essentially using the great flexibility of hydropower to compensate for the stochastic wind speed behaviour. Some papers focus on isolated markets, i.e. small islands, (Bueno and Carta, 2006; Kaldellis and Kavadias, 2001; Kaldellis, 2002) while others explicitely model network constraints and/or market prices (Benitez et al., 2008; Maddaloni et al., 2008; Korpaas et al., 2003; Angarite and Usaola, 2007). See also Castronuovo and Peças Lopes (2004a,b) and Jaramillo et al. (2004). Modeling of the markets and of the generation assets is usually simplified, though Førsund et al. (2008) and Lafrance et al. (2002) rely on the full-scale, industrial optimization models used on their respective grids. Reference (Lafrance et al., 2002) is a commissioned technical study assessing the impact of introducing large-scale wind capacity introduction on the Vermont power grid, with and without integration with the neighboring Québec hydro capacity. The paper by Bélanger and Gagnon (2002) is to our knowledge the only other recently published contribution to discuss the integration of wind and hydro power in Québec. Its focus is different from ours, in that it concentrates on the short-term problem of backing up wind power with flexible (i.e., hydro) generation when wind is down; we focus on the long-term question of energy deficits and surpluses.

In all these papers, the common thread is the control of power fluctuations; short term operations, or even building of new reservoirs, are, one way or another, optimized to achieve this goal. Given the considerable operational flexibility of hydropower, and the inflexibility of wind power, it is clear that from this point of view, hydropower always lends the helping hand to windpower, never the other way around.

The focus of our research is quite different from the previous literature on wind and hydro complementarity, in that we investigate the uncertainty in *annual energy inflow*. We are concerned with water inflow deficits in the long run, and consider wind power as a diversification tool. The time series we consider are comparatively long (46 years), though still fine-grained (hourly data for the wind). We knowingly omit short-term, operational constraints; naturally, the pertinence of our results is greater for lower wind penetration, where this omission is easier to accept.

Our main purpose is to investigate the diversification effect of the energy sources, from all-hydro to a mix of hydro and wind, to lower the risk of energy deficits. Our approach is reminiscent of portfolio optimization, which adjusts the percentage invested in each available asset to minimize risk under a fixed average return constraint. We will indeed call "portfolio" any mix of wind and hydro generation assets that receives some fixed average annual energy inflows.

The risk measure we use is a low quantile ($\alpha = 2\%$) of the probability distribution of combined energy inflows, which is in line with the hydro inflows risk management policy at Hydro-Québec. In other words, our measure of risk is the energy deficit that occurs once every fifty years. Quantile-based risk measures, such as value-at-risk, expected shortfall and variants, are also the cornerstone of modern financial risk management (see e.g. McNeil et al. 2005).

Using statistical models calibrated to observed hydro and wind data, we simulate the annual energy inflows to the hydro system of Hydro-Québec, and to a wind power system dispersed at ten locations in Québec. We then simulate the energy inflows to portfolios that combine together the two "pure" systems, with various weighting. These simulations are the base of our risk analysis.

Since the diversification effect hinges on statistical independence between the two energy sources, some care is taken in its modeling and we use a copula approach. The merits of the copula approach for multivariate modeling in hydrology are described in Favre et al. (2004).

The following Section 2 describes the data and their statistical modeling, including the copula treatment. Section 3 explains our risk analysis approach and presents a discussion of the results. Section 4 concludes the paper. A technical appendix provides the basic definitions and tools of the copula approach.

2 Modeling of water and wind inflows in Québec

Our study is focussed on the province of Québec, in eastern Canada. Québec has abundant hydropower resources, and equally abundant, —but essentially untapped—, wind-power resources.

Our modeling approach, i.e., copulas, separates the treatment of the marginal distributions —of water inflows, and wind inflows— from the modeling of their dependence. We therefore provide data and model details on the two inflows and the dependence in three separate subsections.

2.1 Water inflows

Gathering water inflows data to a specific set of hydropower plants can be a difficult task as they are rarely publicly available. It is even more difficult if one needs to infer past inflows to dams before they were even built, i.e. on the basis of meteorological and terrain considerations. We were fortunate that such a time series of inflows had become publicly available through the hearings of the Régie de l'énergie du Québec in 2004, see Régie de l'énergie du Québec (2004). The series was computed by Hydro-Québec and gives the annual inflows for 1954 to 2003, in terawatt-hours, to all its hydro-electric plants running in 2003.[1] Note that an annual inflow does not necessarily correspond to the energy produced that year: water can be used from the previous year, or kept for the next, for example. Note also that the hearings were open specifically to address Hydro-Québec's energy security concerns and proposed solutions. Higher frequency data were not available; however, the dams are large enough to considerably attenuate the interest for seasonal, intra-year complementarity.

Figure 1 show the series as surpluses and deficits against the series average of almost 189 TWh. We will refer to the set of hydro-electric plants just mentionned as the "hydropower system".

Two periods stand out: the years 1965 to 1983 average more than 203 TWh, while the years 1984 to 2003 average less than 180 TWh. We will pay some attention to these two subperiods, to which we shall refer as the

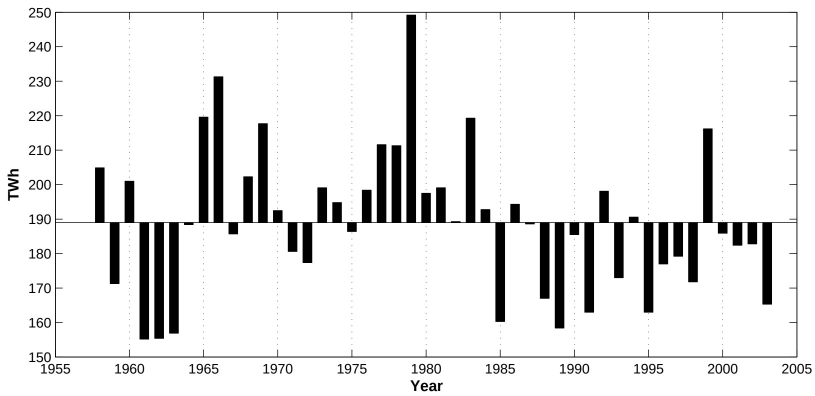

Figure 1: Water inflows, in TWh and for the period 1958-2003, to the hydropower assets of Hydro-Québec running in 2003. Displayed as surpluses and deficits against the series arithmetic average of 189 TWh.

"high inflows" and "low inflows" periods. In fact, we use the data to define three scenarios: one based on "all data" (1958–2003), one on "high inflows" (1965–1983), and one on "low inflows" (1984–2003).

A statistical analysis was performed on the data to find an appropriate distribution and its parameters. We used a standard maximum likelihood approach, and fitted the complete series as well as the two sub-periods mentioned above. The distributions considered were the normal, lognormal and gamma distributions. In all cases, the lognormal distribution provided the best fit; the parameters for each period are provided in Table 1.

Period	**All data** 1958–2003	**High inflows** 1965–1983	**Low inflows** 1984–2003
Average inflow	188.9 TWh	203.3 TWh	179.7 TWh
Std deviation	20.8 TWh	17.8 TWh	14.7 TWh
μ	5.235	5.311	5.188
σ	0.1099	0.0875	0.0818

Table 1: Basic statistics and fitted lognormal distribution parameters for the water inflows

We confirmed the adequacy of the fits via quantile-quantile plots (not shown). The probability density functions for 1958–2003 and the two subperiods are illustrated on the left-hand side of Figure 2. A question that

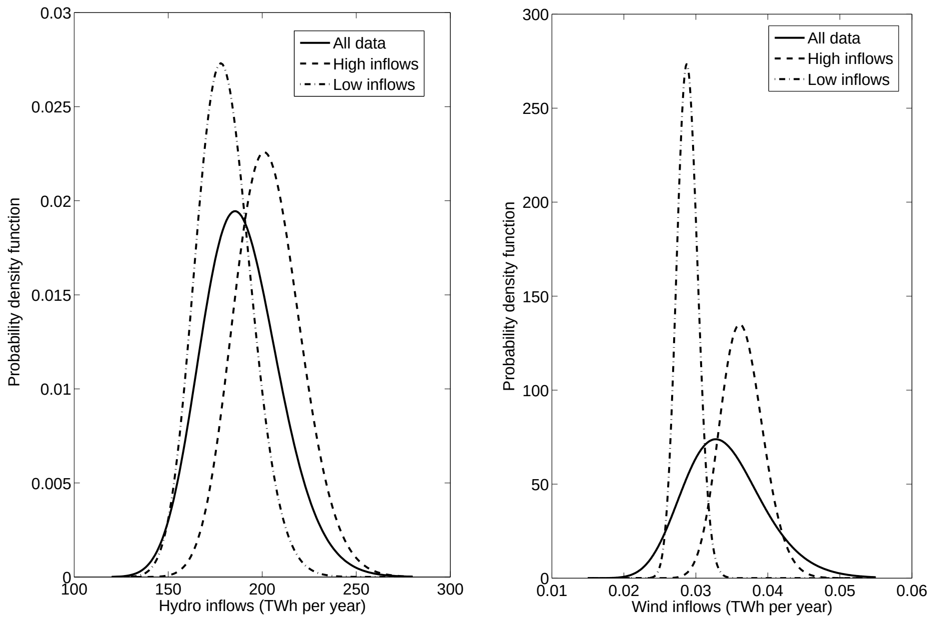

Figure 2: Fitted lognormal probability densities for hydro and wind inflows

arises naturally is why the data were modelled as static and not as a time series. Autocorrelation tests were run on the series; only the lag 1 Pearson test was significant at the 5% level, with a p-value of 3.6%. Kendall's and Spearman's tests for lag 1 autocorrelation both had a p-value of 6%. Note that the 1996 study of Perreault et al. reported similar results with (slight) significant lag 1 autocorrelation for the aggregated inflows to eight watersheds in Québec. In any case, a short series with little autocorrelation does not lend itself to a meaningful time series model. The same argument also steered us towards the static modelling of wind energy inflows, which now follows..

2.2 Wind inflows

Our model is based on hourly wind speed data over the period from 1958 to 2003. For the sake of simplicity, we accounted for no other factor than hourly speed; direction, humidity, temperature, etc. were not included in the model. The meteorological convention followed by Environment Canada

Sites	Years	Climate ID	WPD
Bagotville	1953–2003	7060400	166
Kuujjuarapik	1957–2003	7103536	269
La Grande			
Nitchequon	1953–1985	7095480	176
La Grande IV	1985–2003	7093GJ3	—
Maniwaki			
Maniwaki	1953–1993	7034480	20
Maniwaki	1993–2003	7034482	20
Mont-Joli	1953–2003	7055120	356
Roberval	1958–2003	7066685	144
Rouyn	1954–2003	7086720	76
Sept-Iles	1953–2003	7047910	—
Schefferville	1953–2003	7117825	—
Val-d'Or	1955–2003	7098600	—

Table 2: Wind measurement locations, years for which data are available, Environment Canada Climate Identification number, and wind power density.

is to measure wind speed at a height of 10 m, and we found no evidence of departure from this policy. The data for a specific hour at a specific site are usually the mean of at least a few observations during that hour; our model assumes that the wind was blowing constantly at that average speed, for the duration of the hour. Note that wind speeds were then adjusted to account for turbine height, see details below.

Ten sites were chosen over the province's territory, on the basis of availability of data, quality of data, and relevance of the geographic area for wind power production. The sites are not windfarm sites, most are in fact airports; but long statistical series are simply not available elsewhere. In two cases, we had to bridge the data of two different sites to complete the series. Holes in the series, unavoidable in such large databases, were replaced by their nearest (earlier or later) value[2]. Details appear in Table 2, which includes the years of availability of the data, the Environment Canada climate identification numbers and the wind power density (WPD) in watts per square meter, as computed in Ilinca et al. (2003). Wind power density is not available for all sites.

Energy inflow for each site was derived from the power curve of the Vestas V80 IEC Class I turbine, a 1.8 MW device that has been installed in Québec in recent years. Assessing the best turbine for each site is a complex task, requiring data we do not have; we simply used the same turbine for each site. The V80's power curve is easily available on Vestas's website. Cut-in and cut-out speeds are respectively 4 and 25 meters per seconds (m/s). The density of air was assumed constant at 1.225kg/m^3. Wind speed was extrapolated at a hub height of 67 m, one of the standard heights for the V80. The extrapolation was done with the power law[3] with exponent 1/7, i.e. wind speeds at 67 m are the wind speeds at 10 m multiplied by $(67/10)^{\frac{1}{7}} \approx 1.312$. See, for example, Ilinca et al. (2003) for more details on height extrapolation.

The yearly inflows of the various sites, averaged over the 46 years, are reasonably comparable between themselves, with the exception of Maniwaki perhaps. The Mont-Joli site provided the highest inflow, and Maniwaki the weakest, with a ratio of thirteen to one; other inflows were much closer to that of Mont-Joli, with the highest ratio at 3.6.

The total energy inflow of the 10-turbine system, in TWh, is displayed in Figure 3 as surpluses and deficits against the series average of 0.0341 TWh. We will refer to this system as the "wind-power system".

The inflow to the wind-power system is obviously of a much smaller scale than that of the hydropower system (compare Figure 1 and Figure 3); we deal with this later, as for now it does not affect the copula dependence analysis of Section 2.3.

In Figure 3, notice the sharp apparent "regime change" occuring in 1980–1981: the average yearly inflow for the 1958–1980 period is 0.039 TWh, compared to 0.029 TWh for the 1981-2003 period. This result has been observed before, but not explained (Lauzon, 2003). Just as surprising, is the almost exact parity of the wind "high–low" periods with their hydro counterpart. While we lack the expertise to interpret these statistics from a meteorological point of view, we will definitely investigate this phenomenon in our risk analysis. We decided to draw a common cutting point after 1983 for both hydro and wind inflows series, and use the previously introduced "high inflows" (1965–1983) and "low inflows" (1984–2003) for both series. "All data" always refers to 1958–2003, which adds seven early years to the two subperiods.

The wind energy inflows data were analyzed, again with a maximum likelihood approach, which fitted the normal, lognormal, and gamma distributions. All three distributions yielded very similar likelihood values. Because it more naturally accounts for the necessarily positive values and for the sake of simplicity, the lognormal distribution was used for all three cases. Parameters are as displayed in Table 3.

The probability density functions derived from the fitted distributions are illustrated on the right-hand side of Figure 2.

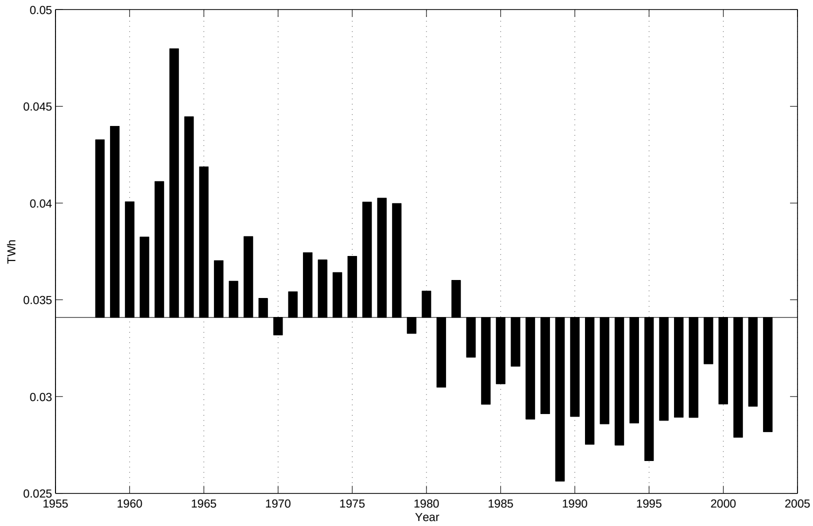

Figure 3: Wind inflows, in TWh, and for the 1958-2003, to a set of ten wind turbines at ten locations in Québec. Displayed as surpluses and deficits against the series average of 0.034 TWh.

2.3 Dependence of hydro and wind energy inflows

A first look at the bivariate data is provided in Figure 4, where the periods of high and low inflows are also displayed.

Copulas allow one to better capture the dependence structure of random variables that do not follow gaussian distributions. They also allow us to characterize the dependence separately from the marginal distributions so that the two steps are carried sequentially.

Basic technical results on multivariate modeling with copulas appear in Appendix A. We discuss in this section only what is specific to our study.

Five classical copulas were considered: the Clayton, Frank, Gumbel, Normal and t copulas. Together, they cover a wide range of dependence structures. Favre et al. (2004) use the Clayton and Frank copulas in their modeling of hydrological flow and volume, and Coles and Tawn (1994) use the Gumbel copula. This latter copula is also a member of the bivariate extreme value class, the only one among our five. The Normal and t copulas are examples of elliptical copulas, and have been found to be quite useful in financial applications where some modeling issues mirror those in hydrolog-

Period	**All data** 1958–2003	**High inflows** 1965–1983	**Low inflows** 1984–2003
Average inflow	0.0341 TWh	0.0365 TWh	0.0288 TWh
Std deviation	0.0056 TWh	0.0030 TWh	0.0015 TWh
μ	-3.392	-3.315	-3.547
σ	0.1626	0.0816	0.0507

Table 3: Statistics and fitted lognormal distribution parameters for the wind inflows

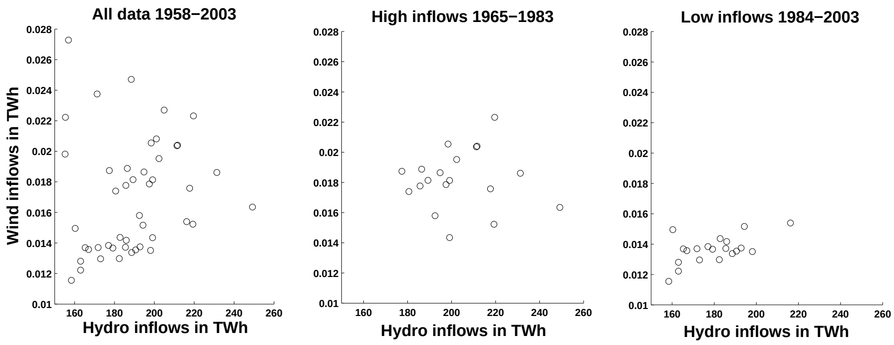

Figure 4: Wind inflows vs hydro inflows

ical applications, see McNeil et al. (2005).

The choice of a specific copula and its parameter(s) was performed by maximum likelihood estimation through the "inference functions from margins" approach (see the appendix for details).

Copulas go beyond the usual correlation (i.e. the Pearson linear correlation) to describe dependence. It is however still useful to have some numerical measure of the degree of dependence that applies equally to all sets of data and is not dependent on the marginal distributions (see Embrechts et al. 2002). Kendall's tau coefficient and Spearman's rank correlation are the most common such measures and both run from 1 (perfect dependence) to 0 (independence) to -1 (perfect anti-dependence).

The results of our analysis were as follows. In terms of likelihood, the t copula was best for both "all data" and the low inflows period. The Normal copula best matched the high inflows period data. Details of the parameters appear in Table 4. The table also provides the values of Kendall's tau. The "all data" period displays moderate positive dependence, and the "low

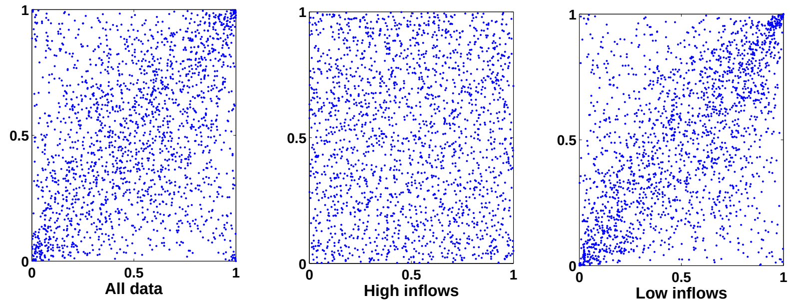

Figure 5: Simulated values for the dependence structures found for each scenario: all data, high inflows, low inflows

inflows" period slightly more. In the high inflows period the series show very small, negative dependence; this near-independence was echoed by the fitted parameters for the Clayton and Gumbel copulae, at 0.0001 and 1.0000 respectively. Some intuition on the dependence structure can be gained from

Period	**All data** 1958–2003	**High inflows** 1965–1983	**Low inflows** 1984–2003
Copula	$t(\rho,\, \nu)$	Normal(ρ)	$t(\rho,\, \nu)$
ρ	0.3722	-0.0350	0.5943
ν	2.37	—	2.35
Kendall's τ	0.243	-0.022	0.405

Table 4: Copulas, fitted distribution parameters

the general look of the dispersion graphs in Figure 5. Each dot represents the simulation of a pair of joint, uniformly distributed random variables, with dependence structure as per Table 4. (Since the effect of the marginal distribution is removed, neither of the two axes corresponds to hydro nor wind). Notice the "X" shape of the t copula (all data and low inflows), and the uniform coverage of the normal copula (high inflows) due to the very low dependence level.

3 Simulation-based risk analysis

The goal of the statistical analysis was to allow a simulation-based risk analysis of portfolios of generating assets, with varying proportions of wind and hydro: all hydro, all wind, or mixed of any weighting. One can also find the proportion of hydro and wind which minimizes any measure of risk. Simulations were performed for the three scenarios derived from the three periods identified in Section 2.1: “all data” (1958–2003), “high inflows” (1965–1983), and “low inflows” (1984–2003).

It is important to stress that all portfolios, irrespective of the wind/hydro weighting, must have the same *average* annual energy inflow. It is the only way to a significant risk comparison: comparing the quantiles of portfolios with different average inflows is nonsensical. Some scaling of the generation assets will be required, to which we return shortly.

Note that we do not account for the cost of changing the portfolio, e.g. the cost of building wind capacity. Our interest is in the (risk management) value of the portfolios; pricing the cost of a change of weights is outside the scope of this paper.

3.1 Simulation of energy inflows for a hydro/wind portfolio

To simulate the energy inflow of a portfolio of hydro and wind generating assets, for one year and one of the scenarios, we proceed as follows. First, two (copula-) dependent, uniformly distributed variables are generated. Second, through the inversion of the respective marginal cumulative distribution functions, the uniforms are given the appropriate marginal distributions associated to hydro and wind (see Sections 2.1 and 2.2). Both steps are easily performed using built-in functions any of a variety of softwares (we used Mathwork’s MATLAB). The procedure provides us with two simulated energy inflow values: one for what was called the hydropower system of all hydro generation assets installed in Québec in 2003, and one for the wind power system consisting of ten wind turbines at ten locations in the province[4]. The two values are “correctly dependent”, that is, in line with historical data. However, a scaling of the wind system is necessary to ensure that any portfolio with wind power provides on average the same energy inflow as the hydro system. Any simulated wind energy inflow is therefore scaled by the appropriate, constant factor which is the ratio of the average inflow of the hydro system by that of the wind system. When all data are considered, the factor is 188.9/0.0341; for the high inflows and low inflows scenarios, it is respectively 203.3/0.0365 and 179.7/0.0288. This effectively ensures that any hydro-wind energy portfolio provides 188.9 TWh (resp. 203.3 TWh and 179.7 TWh) on average. Note the simplifying assumption that more turbines at the windpower sites would produce more energy in exact proportion of the number of turbines.

Adding the hydro inflow and the scaled-up wind inflow, each multiplied by its respective weight in the portfolio, provides one simulation of an energy inflow to the portfolio of generating assets. For all three scenarios, we simulated several one-year energy inflows for various portfolios with hydro weighting 0%, 1%, 2%,... 100%, the balance being wind of course. One thousand one-year simulations per portfolio were sufficient to obtain stable results.

For each set of statistical parameters and each portfolio, we then identified the two-percent quantile, that is, the lowest-in-fifty-years energy inflow; this is our risk measure. Years with high inflows do not constitute a risk management issue, not in our yearly perspetive, since Québec easily sells its surplus to the neighboring states and provinces[5]. The results are presented in Figure 6; the quantiles for the 101 considered weightings appear on the same graph, one graph for each scenario. Included for each curve is a 95% confidence interval[6]

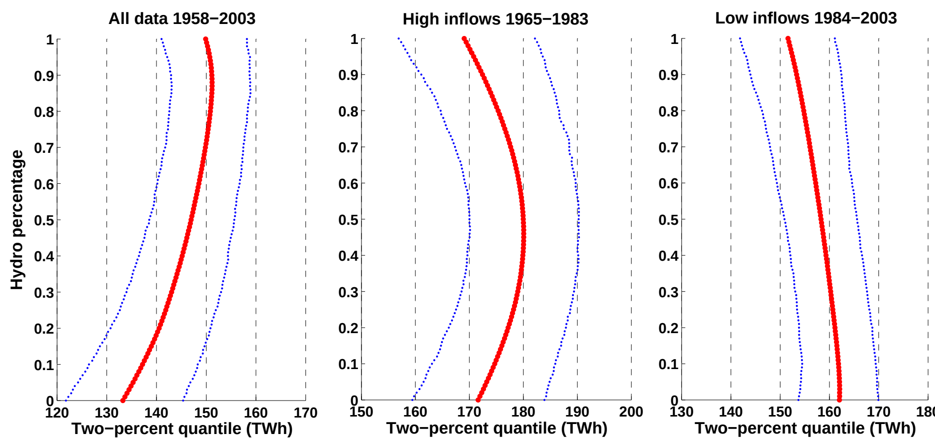

Figure 6: Two-percent quantiles of the energy inflows distribution VS hydro percentage in the generating assets portfolio. Results for all data, high inflows and low inflows

3.2 Discussion of the results

We first give some intuitive insights on the results and follow up with more detailed points.

Recall that the quantile is our measure of risk: the higher, i.e., better, this worst-in-fifty-years inflow, the lower the risk of the generation portfolio. In other words, the quantiles represent energy inflow thresholds: there is a 2% probability that on any given year the actual energy inflow will be lower than the quantile, which would be considered a significant energy deficit with respect to the average annual inflow. The higher this quantile is, the closer it is to the average inflow (which is fixed by construction), the more certain it is that “acceptable” levels of energy inflows will be observed.

In all three scenarios, the minimum risk (i.e. highest value of the quantile), is obtained for a mixed portfolio which includes both wind and hydro generation. More specifically, using the coefficients of variation (i.e. the standard deviation divided by the average, in Tables 1 and 3) and the dependence measure in Table 4, we can say that:

All data The relatively weak dependence between wind and hydro favors the diversification effect, but the longer tails of the wind energy inflow distribution quickly bear on the riskiness of any portfolio with a sizable wind component.

High inflows Both wind and hydro inflows are similarly uncertain, and their near-independence favours the diversification effect, so that the less risky portfolios include equal parts of wind and hydro.

Low inflows The main effect is that wind is relatively less uncertain than hydro, so that portfolios with more wind are less risky; the diversification effect is smaller because of the relatively greater positive dependence (recall Kendall’s τ at 0.40), so less convexity is observed.

Note that the risk associated to the “all data” scenario (1958–2003) is greater than for the other two scenarios, as the low values of the quantile (—mostly below 150 TWh—) indicate. This is perfectly expected, given the nature of the two subperiods, which were chosen precisely on the fact that the inflows are relatively similar.

The nonnegligible widths of the confidence intervals prevent any definitive numerical analysis. However, within these constraints, the following observations can be made.. Note that any portfolio with more than 20% or 25% of wind is hardly realistic for any reasonable time horizon, given the actual situtation in Québec; we will return to this point shortly.

All data Risk is minimized around 12% wind, with the two-percent quantile improved by 1 TWh versus no wind at all. Any percentage of wind between 1% and 29% is better than no wind at all, with respect to the risk measure. More than 30% of wind can make the portfolio riskier or much riskier.

High inflows Risk is minimized for portfolios with 52% to 56% wind, with the two-percent quantile improved by more than 10 TWh. Any percentage of wind is better than no wind at all, with respect to the risk measure.

Low inflows Risk is minimized for portfolios with 95% to 100% wind, with the two-percent quantile improved by more than 10 TWh. Any percentage of wind is better than no wind at all, with respect to the risk measure.

A point can and should be made that some of the above results are useless, in the sense that portfolios with high proportions of wind are unfeasible, for operational and other reasons. It is then more enlightening to compute slopes of the graphs in Figure 6, i.e. the rates of improvement of the risk measure over the lower end of wind energy penetration, say 0%–20%. The slopes (of the secant lines through each subinterval) are as follows.

All data For every percentage point of wind between 0% and 10%, the quantile improves by 0.12 TWh. The quantile is stable between 10% and 15%. Between 15% and 20% wind, every percent of extra wind makes the quantile 0.05 TWh worse.

High inflows For every percentage point of wind between 0% and 20%, the quantile improves by 0.31 TWh.

Low inflows For every percentage point of wind between 0% and 20%, the quantile improves by 0.16 TWh.

Such numbers could be of use when drafting an energy policy on wind penetration in a hydropower area. The main point is the following. Under a quantile-based risk management policy, a certain amount of water must be kept in reserve to face an eventual drought period; the rest can be sold for a profit. This reserve amount is based on the volatility of inflows (and other factors eventually, like demand). By moving to a mixed portfolio through the installation of some wind power, the volatility is lower, the necessary reserve is lower, and the extra energy can be sold on the market. A second point is that less volatile energy inflows make it possible to take more lucrative selling positions.

4 Conclusion

In this study, we performed a long-term statistical analysis of wind power variability for a set of fictive wind farms in Québec. Using hydro energy inflows provided by state-owned Hydro-Québec, we were able to simulate the energy inflows to portfolios of energy assets that include both wind and hydro. Of particular interest was the variability of these inflows with respect to the proportion of wind and hydro.

Somewhat surprisingly, wind and hydro inflows presented a common cyclical pattern with two periods of about twenty years each, during which both wind and hydro were much above average ("high inflows", 1965–1983) then much below ("low inflows", 1984-2003). We analysed three scenarios, two based on the latter periods as well as a scenario based on all available data (1958–2003). One unexpected conclusion of our study is that relying on the "last twenty years" of inflows data may be quite misleading, if long term cycles do exist.

In all three scenarios, wind power provided a substantial diversification effect, and improved the risk profile for any wind penetration up to 30%. The "risk value" of wind, up to 15% wind penetration and at the 2% quantile of annual energy inflows, ranged from 0.12 TWh to 0.31 TWh per percentage point, depending on the scenario. (The values of average energy ouput ranged between 180 and 203 TWh, depending on the scenario)

The intermittency of wind power puts it at an important operational disadvantage in the face of hydro's great flexibility, when it comes to providing capacity at each moment in time. However, it appears that wind can help improve energy inflow volatility, even in cases where it's own volatility is higher than that of hydro. It is suggested that the value of the diversification effect should be taken into account when the net cost of wind implementation is computed.

Further topics along this line of research would include the possibility of unequally distributed wind generating assets with relatively more turbines at some sites than others. The seasonal complementarity (high winds during the cold season which has low water inflows) could be investigated, if hydro data were available.

5 Acknowledgments

The authors are especially thankful to Adrian Ilinca (Université du Québec à Rimouski) for his advice on extrapolating wind speeds, Adrien Julien (Environment Canada) for his input on the wind data, Polynice Oyono (HEC Montréal) for his computing assistance and Pierre-Olivier Pineau (HEC Montréal) for his critical review of the paper. Funding support from NSERC (Natural Sciences and Engineering Research Council of Canada) is gratefully acknowledged by the first two authors, as is scholarship funding by IFM2 (Institut de finance mathématique de Montréal) by the third author.

A Fundamentals of (modeling multivariate joint distributions with) copulas

Our main tool in establishing the dependence structure between wind and water inflows is the copula. A complete presentation of the topic is clearly out of scope, but basic results on the copula formulation to multivariate distributions are presented below. Further details are readily available in recent publications, see Joe (1997); McNeil et al. (2005) or Genest and Favre (2007).

A.1 The copula function

A copula is a joint distribution function of standard uniform random variables. That is,

$$C(u_1, \dots, u_d) = \Pr\{U_1 \leq u_1, \dots, U_d \leq u_d\},$$

where $U_i \sim U(0,1)$ for $i = 1, \dots, d$. Let $X_1, \dots, X_d$ be random variables with joint distribution function F and continuous marginal distribution functions F_i, $i = 1, \dots, d$. Sklar (1959) showed that, for any multivariate distribution F, there exists a unique copula C which can be written as

$$C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$$

where the quantile function F_i^{-1} is defined by $F_i^{-1}(u) = \inf\{x : F_i(x) \geq u\}$. It is also easy to see that, if C is a copula function, and $F_1, \dots, F_d$ are arbitrary distribution functions, then F defined by

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \tag{1}$$

is a multivariate distribution function with marginal distribution functions $F_1, \ldots, F_d$.

For the sake of simplicity, we now restrict ourselves to the copulas that were chosen to model our data, i.e. the bivariate ($d = 2$) versions of the Normal, and t copulas. Definitions for the other copulas used in this paper (the Clayton, Frank, and Gumbel) are easily found in the literature. The copulas are defined as:

the Normal copula,

$$C(u_1, u_2) = \Phi_\rho(\Phi^{-1}(u_1), \Phi^{-1}(u_2)), \tag{2}$$

where Φ_ρ is the distribution function of a bivariate standard normal distribution with correlation ρ, and Φ is the $N(0,1)$ distribution function, and

the t copula,

$$C(u_1, u_2) = \int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\pi\nu\sqrt{1-\rho^2}} \left(1 + \frac{\mathbf{x}' P^{-1} \mathbf{x}}{\nu}\right)^{-\frac{\nu+2}{2}} dx_1 dx_2 \tag{3}$$

where t_ν^{-1} is the quantile function of a univariate (Student-) t_ν distribution, $\mathbf{x} = (x_1, x_2)'$, and P is the correlation matrix

$$P = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

A.2 Measures of dependence

It is useful to have some measure of the degree of dependence provided by a given copula. The usual Pearson linear product-moment correlation depends on the marginal distributions and is not a desirable measure of association for non-normal multivariate distributions. Kendall's tau coefficient and Spearman's rank correlation are the most widely used measures of dependence for non-normal multivariate distributions. Spearman's rank correlation is defined as the (usual, linear) correlation applied to the ranks of the data set

$$\rho_S = \rho\left(F_1(x_1), F_2(x_2)\right)$$

and Kendall's tau coefficient is defined as a difference of probabilities that the random variables "move together":

$$\tau = \Pr\left((X_1 - X_1^*)(X_2 - X_2^*) > 0\right) - \Pr\left((X_1 - X_1^*)(X_2 - X_2^*) < 0\right)$$

where (X_1^*, X_2^*) is an independent copy of (X_1, X_2). We chose Kendall's tau as our measure of dependence and have $\tau(\rho) = 2/\pi \arcsin \rho$ for the Normal and t copulas.

A.3 Likelihood and estimation: choosing a model and its parameter(s)

Consider a copula-based parametric model (1) for the random vector Y where F_i are marginal models, e.g. normal or lognormal, and C as in one of (2)-(3). Let θ be the $p \times 1$ dimensional vector of unknown parameters, e.g. $\theta = (\mu_1, \sigma_1, \mu_2, \sigma_2, \rho, \nu)$ in the case of two lognormal margins linked by a t copula, and Θ be the parameter space. A fundamental tool in estimation is the likelihood for θ based on the observed data y. We write

$$L(\theta) = f(y; \theta), \quad \theta \in \Theta \tag{4}$$

where $f(y; \theta)$ is the density function associated with (1), the density being a function of θ and y, but is regarded as a function of θ for fixed y in (4). The maximum likelihood estimate (MLE) of θ is a value of θ that maximizes the likelihood $L(\theta)$, or equivalently the log likelihood $\ell(\theta) = \ln L(\theta)$. We use $\hat{\theta}$ to denote the MLE, which means that

$$L(\hat{\theta}) \geq L(\theta) \quad \theta \in \Theta.$$

(the likelihood is better with $\hat{\theta}$ then with any other $\theta \in \Theta$)

Given n independent observations and applied to distribution (1), the expression for the log likelihood becomes

$$\ell(\theta) = \sum_{j=1}^{n} \ln c\left(F_1(x_1^j), \ldots, F_d(x_d^j)\right) + \sum_{j=1}^{n} \sum_{i=1}^{d} \ln f_i(x_i^j) \tag{5}$$

where c is the density of the copula C and f_i, $i = 1, \ldots, d$, are the densities of the marginal distributions F_i, $i = 1, \ldots, d$. Maximizing (5) is known as the *exact maximum likelihood method*.

The exact maximum likelihood approach jointly estimates the parameters of the margins and the parameters of the dependence structure. However, the copula representation splits the parameter vector θ into *marginal* parameters and *dependence* parameters, say $\theta = (\theta_{m1}, \ldots, \theta_{md}, \alpha)$. The log likelihood (5) could then be written as

$$\ell(\theta) = \sum_{j=1}^{n} \ln c\left(F_1(x_1^j; \theta_{m1}), \ldots, F_d(x_d^j; \theta_{md}); \alpha\right) + \sum_{j=1}^{n} \sum_{i=1}^{d} \ln f_i(x_i^j; \theta_{mi}). \tag{6}$$

In a first instance, we could perform d separate estimations, one for each univariate marginal distribution, i.e. obtain

$$\hat{\theta}_{mi} = \text{argmax} \sum_{j=1}^{n} \ln f_i(x_i^j; \theta_{mi})$$

for $i = 1, \ldots, d$ and then estimate α given the previous estimates

$$\hat{\alpha} = \text{argmax} \sum_{j=1}^{n} \ln c \left(F_1(x_1^j; \hat{\theta}_{m1}), \ldots F_d(x_d^j; \hat{\theta}_{md}); \alpha \right).$$

This approach is known as the method of *inference functions for margins* or IFM, and it is the approach taken in this paper.

Notes

[1]The series in fact runs back to 1943, but since our wind data are sufficient from 1958 only, we did not use the earlier hydro data.

[2]Overall, less than 3.5% of the data are missing. This missing data are however largely concentrated on site "Rouyn" and to a lesser extent, site "Maniwaki". In both cases, the "holes" are sufficiently well dispersed; moreover, these sites are two of the lower producers. To confirm the minimal impact of these two sites which concentrate most of the missing data, we re-ran all tests without "Rouyn", and without both "Rouyn" and "Maniwaki", and while the estimated parameters changed slightly, the quantiles (figure 6) were virtually identical

[3]Exponent 1/7 comes from fluid mechanics theory, and applies to the theoretical case of a plane (i.e. an abstract, two-dimensional surface). In an actual analysis of wind capacity, the terrain characteristics must be accounted for, but without such details, exponent 1/7 is considered the only reasonable value. In wind power applications, it is usually a conservative choice (i.e. slightly smaller than the site-specific value).

[4]The choice of identical capacity at each site was made for the sake of simplicity. A recent paper (Drake and Hubacek, 2007) on the topic of dispersion of wind facilities and the allocation of capacity (MW) to them addresses these issues directly.

[5]In a day-to-day perspective, the Spring meltdown of snow and ice must be correctly managed however.

[6]Maximum likelihood estimators (MLE) converge in distribution to a normal distribution with mean equal to the true parameter value and covariance matrix equal to the inverse of the Fisher information matrix. The latter is easy to estimate as the Hessian of the (log) likelihood function evaluated at the MLE values. Confidence intervals are based on a parametric bootstrap (resimulations) using this asymptotic normal distribution to account for the variability in our marginal and copula parameter estimates.

References

Angarite, J.M., Usaola, J.G., 2007. Combining hydro-generetion and wind energy: Biddings and operation on electricity spot markets. Electric Power Systems Research, 7, 393–400.

Bélanger, C., Gagnon, L., 2002. Adding wind energy to hydropower. Energy Policy, 5, 1279–1284.

Benitez, L.E., Benitez, P.C., van Kooten, G.C., 2008. The economics of wind power energy storage. Energy Economics, 16, 1973–1989.

Bueno, C., Carta, J.A., 2006. Wind powered pumped hydro storage systems, a means of increasing the penetration of renewable energy in the Canary Islands. Renewable and Sustainable Energy Reviews, 28, 312–340.

Castronuovo, E.D., Peças Lopes, J.A., 2004a. On the Optimization of the Daily Operation of a Wind-Hydro Power Plant. IEEE Transactions on Power Systems, 8, 1599–1606.

Castroduovo, E.D., Peças Lopes, J.A., 2004b. Optimal operation and hydro storage sizing of a wind-hydro power plant. Electric Power and Energy Systems, 7, 771–778.

Coles, S.G., Tawn, J.A., 1994. Statistical methods for multivariate extremes: an application to structural design (with Discussion). Applied Statistics, 43, 1–48.

Drake, B., Hubacek, K., 2007. What to expect from a greater geographic dispersion of wind farms?— A risk portfolio approach. Energy Policy, 35, 3999–4008.

Embrechts, P., McNeil, A.J., Straumann, D., 2002. Correlation and dependence in risk management: properties and pitfalls, in Dempster, M. (Ed.), Risk management: value at risk and beyond. Cambridge University Press, Cambridge.

Favre, A.-C., El Adlouni, S., Perreault, L., Thiémonge, N., Bobée, B., 2004. Multivariate hydrological frequency analysis using copulas. Water Resources Research, 40, W01101, doi:10.1029/2003WR002456.

Førsund, F.R., Singh, B., Jensen, T., Larsen, C., 2008. Phasing in wind-power in Norway: Network congestion and crowding-out of hydropower. Energy Policy, 6, 3514–3520.

Genest, Ch., Favre, A.-C., 2007. Everything you always wanted to know about Copula Modeling but Where Afraid to Ask. Journal of Hydrological Engineering, 12-4, 347–368.

Lauzon, L., 2003. Climatological adjustment of wind speed: the importance of selecting an appropriate reference period. Presentation by Hélimax Energy at the Canadian Wind Energy Association conference.

Ilinca, A., McCarthy, E., Chaumel, J.-L., Rétiveau, J.-L., 2003. Wind potential assessment of Quebec Province. Renewable Energy, 28, 1881–1897.

Jaramillo, O.A., Borja, M.A., Huacuz, J.M., 2004. Using hydropower to complement wind energy: a hybrid system to provide firm power. Renewable Energy, 22, 1887–1909.

Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman and Hall, London.

Kaldellis, J.K., 2002. Parametrical investigation of the wing-hydro electricity production solution for Aegean Archipelago. Energy Conversion and Management, 16, 2097–2113.

Kaldellis, J.K., Kavadias, K.A., 2001. Optimal wind-hydro solution for Aegean Sea islands' electricity-demand fulfilment. Applied Energy, 21, 333–354.

Korpaas, M., Holen, A.T., Hildrum, R., 2003. Operation and sizing of energy storage for wind power plants in a market system. Electric Power and Energy Systems, 7, 599–606.

Lafrance, G., Krau, S., Saulnier, B., 2002. Assessment of the impact of wind power penetration on the Vermont electricity grid. Technical Reference 1.

Maddaloni, J.D., Rowe, A.M., and van Kooten, G.C., 2008. Network constrained wind integration on Vancouver Island. Energy Policy, 11, 591–602.

McNeil, A., Frey, R., Embrechts, P., 2005. Quantitative Risk Management. Princeton University Press, Princeton.

Perreault, L., Haché, M., Bobée, B., 1996. Rapport final du Projet C2 – Analyse statistique des séries d'apports énergétiques. INRS-Eau, Research Report No. R-474.

Régie de l'énergie du Québec, 2004. Réponses d'Hydro-Québec Production à la demande de renseignements no. 1 de la Régie. Available on `http://www.regie-energie.qc.ca/audiences/3526-04/mainRepDemRensRegie3526.htm`.

Sklar, A., 1959. Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Université de Paris, 8, 229–231.