The basics of data acquisition systems and the need to over-sample
or
Is Michael Klemen misguided?
An Opinion Paper
by

Mel Tyree
Professor, Department of Renewable Resources,
University of Alberta, Edmonton, Canada T6G 2E3

Introduction

Let us start with a quote from Michael Klemen (Small-Wind-Home # 25857)
"I am concerned with the accuracy and meaning of whatever data you are collecting. I record 5 Hz data to attain good resolution and for data correlation, and I also sample at 2100 Hz for accuracy. By data correlation, I know that my wind data is meaningful relative to the power data, because most of the time, over 90% of the wind that passes by the turbine is the same wind passing by the anemometer, and I measure ALL of it. I don't wonder about high speed sampling, or storing short time frame data. I do it as a matter of course. But, it can be done wrong, and/or poorly."

Michael still does not believe some of the messages I am posting about my 10 kW turbine made by Bergey WindPower Corp (BWC model Excel-s). He is still trying to discredit me through saying that the accuracy of my measurements are not up to his standards. I have posted several items to address this issue (see SWIEP postings R#40 thru 50). SWIEP reports can be downloaded from
http://www.ualberta.ca/~mtyree/SWIEP/Publications.html
In some cases he is impressed with my system (with reservations). For example in Small-Wind-Home # 25839

"I appreciate having some information regarding what you are doing. I must say that your data acquisition board is quite a nice one. :) Maybe someday I can have something like that!

If you are using the default configuration for the counter/timer, is not really a 1 second average wind speed, but just a 1 Hz sample of the wind speed. So, if the frequency at the end of the second is 20Hz, you only know what happened during 1/20th of the second. You can expect more scatter than if you actually measured the average wind speed. It's not a huge deal, just so you know."

Analysis and discussion

The general theme here is the statement that sampling more frequently (up to 2100 readings a second) might increase the accuracy of ones measurements of, say, a power curve. Michael, might I ask that you do an experiment to prove this in your own system? What you are after is an accurate power curve. On your website you often give a power curve with statistics. That is good; not everyone does that! So you quote bin values of wind speed and power and quote a mean power ± a standard deviation (sd) from which you can get a measure of relative 'noise' from the coefficient of variation (CV) = sd/mean. This coefficient of variation is running around 5 to 10% of the mean. That is a fairly large statistical scatter relative to the likely accuracy AND resolution of your equipment. I would guess the accuracy of your system will be about 0.5% for power and wind speed. You can tell me the exact figures if you like; I could not find this from your website. So the CV is 10 x the basic accuracy. That suggests to me that the high CV is caused by stochastic processes in wind rather than ‘errors’ in data acquisition. So here is an experiment you can do, instead of sampling at 2100 Hz you can sample at 210 Hz or even 21 Hz and see if this affects the CV of your power curves. I don’t think it will, but even if noise (= high CV) results from lower sampling frequency, it proves the necessity of over-sampling only for your system; so it does not necessarily apply to mine!

Why to I think high sampling rate is unimportant even though I do sample power at about 500 Hz? I measure lots of signals in my home: temperature at 4 places, power consumption of my geothermal heat pump, my water pump, my hot water tank, and power output of my GridTek10 inverter, and AC line voltage at 4 places…and more! In some cases high sampling will help and in others probably not.

I measure temperature using an LM335 chip (= the transducer) that puts out 10 mV per deg K, so room temperature which is about 20 C = 293 K equals 2.930 volts transducer output. To know when over-sampling (high Hz) is helpful or not you need to understand how your sensors and A/D system works together. Read appendix 1 to get an overview of the issues, otherwise just take my word for the following: In my A/D system, the basic voltage measurement resolution is around ±10 mV (= ±1.0 ^{o}C for temperature) when I am measuring temperature. This is because of the typical electrical noise level of my system-bus and I get that even though the theoretical resolution of a 16 bit A/D ought to be ±0.3 mV. In this case the noise is random due to electrical noise in the computer bus that is an integral part of the Daq3005 system I use to measure things as well as the 60 Hz noise in my house wiring. So I found by trial & error that I could get a mean temperature to about ±3 mV (= ±0.3 ^{o}C) if I sample at about 500 Hz and spit out the mean once per s. I use this sampling rate for all channels even though it doesn’t make sense for most other power or AC voltage measurements.

But this high sampling rate does not give me the true air temperature because of the slow response time of the transducer. I have worked lots with the LM335 chip and know that if you rapidly transfer it from cold to warm air it has a response time-constant of about 2 s, i.e., it takes time to warm up or cool down to be at the same temperature as the air. This response time-effect is illustrated in Fig. 1 below. The blue line (= the signal, air temperature) increases by 5 ^{o}C at time = 5 sec. The transducer has a response-time constant of 2 s, hence it takes a while to register the new air temperature (yellow line) Increasing the sampling rate from 1 Hz to 500 Hz (with means saved every second) does reduce the ‘error’ in measurement of the transducer output, but the transducer does not follow the air temperature change because of the rather long equilibrium response-time constant. Keep this example in mind when we talk about wind speed and power measurements below!

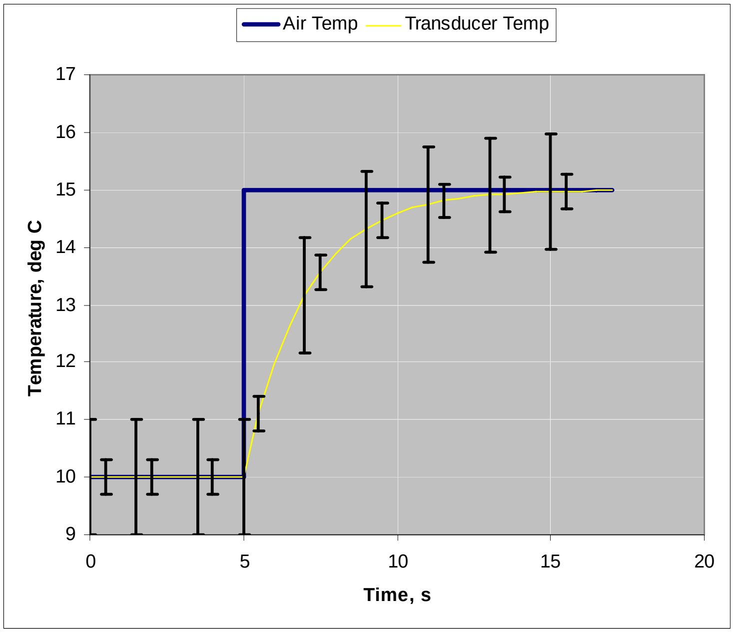

Figure 1: This illustrates the response time of a temperature transducer (LM335). The blue line represents a step change in air temperature. The yellow line shows how long it takes the transducer to reach the new temperature which has changed from 10 to 15 ^{o}C at time = 5 s. This plot assumes a time constant of τ = 2 s and an inverse exponential approach to equilibrium. The large error bars show the reproducibility of transducer temperature measurements I would get if I sampled at just 1 Hz. The smaller error bars indicate the mean error if I sample at 500 Hz and save the mean once per second.

Why is a high sampling rate "meaningless" for power measurements or wind speed measurements? That is because you have to take into account the basic time resolution or time-constant of the transducer of the signal. The time-constant is a measure of how fast the transducer can respond to a change of signal. The time constant is defined as the time to get to 1/e of the true value after as step change in the signal, where e = Napier's number, so after one time constant the reading is 1/e = about 36.8% away from the true value, after two time constants it is $(1/e)^2$ away = 13.5% and after 4 time constants it is $(1/e)^4$ away = 1.8% away. In the case of a cup anemometer, the time constant for speed-up is about 0.1 s and for slow down is about 0.25 s. So if it takes about 1 s for a cup anemometer to register most of a change in wind speed, very little is gained by sampling at faster than 2 Hz. The same is true for the Ohio Semitronics transducer that is commonly used to measure power. The manufacturer claims a response time of 0.25 s. So you need about 1 s for it to register the AVERAGE power output within 1.8% after a step change in power. Why the long time constant? This is because the sensor is trying to send a DC voltage to the A/D system that represents the TRUE RMS value of power. Real power changes from instant to instant in a 60 Hz power grid. You have sine waves of voltage (V) and current (I). At any instant, the instantaneous power in a single phase system = VxI. And this value is zero when V or I = 0 and maximum when V or I are at the extremes of the sine-wave cycle. So what is needed is the 'average' power output over several cycles. This averaging is usually done with resistors and capacitors that electrically average out the dynamically changing power readings.

I don't use an Ohio Semitronics power transducer, although I have ordered one to play with. Instead I built a custom circuit (transducer) to measure RMS line voltage on my service panel (breaker box). I calibrated this system against a DVM accurate to 0.1%. I also measure RMS current using a CT coil transducer (see R#24), which probably has an accuracy of 0.4% (although the manufacturer claims only 1%). So I think I can compute watts from V x I with an overall accuracy of about 0.5%. So how well am I doing? Figure 2 below nicely illustrates this and allows me to discuss the difference between accuracy and resolution in an A/D system. My system is capable of measuring about 22 kW at an accuracy of, say, 0.5%, which means an uncertainty of 110 watts. This uncertainty could come from non-linearity in the calibration of the two transducers and/or from zero offsets. But the resolution of measurements is much better, i.e., a general 'noise' of ± 0.8 watts! Note that the noise has 'quantum jumps'. Why is that? To answer that we have to look at the general resolution of the current measurement, which is worse than the resolution of voltage measurement. The current jumps by about 0.003 amps which corresponds to an A/D resolution of ± 1 in the integer count out of a range of 32,767 for the positive half of the digital range of my 16-bit A/D system. (SEE APENDIX 2) The resolution of the voltage measurements, in terms of integer values is better, so gradual changes in V cause small changes in power = V x I, but the incremental jumps in power are due to resolution limitation in the measurement of current (I).

So having said all this, Michael, allows me to critique a graph of yours I have seen twice. Perhaps you can help me find it again. I saw it once on your web site and once again in your talk at the recent NREL conference we both attended in the fall of 2008. In this graph, you claimed you could detect high frequency response in turbine power output due to high frequency changes in wind speed. I think the frequency response was 2 kHz or 20 kHz? I do not believe your interpretation because of the long response-time of the power and wind speed transducers you probably use. Perhaps you can tell me what you use? I think what you are measuring is the basic integer noise level of you're A/D system!

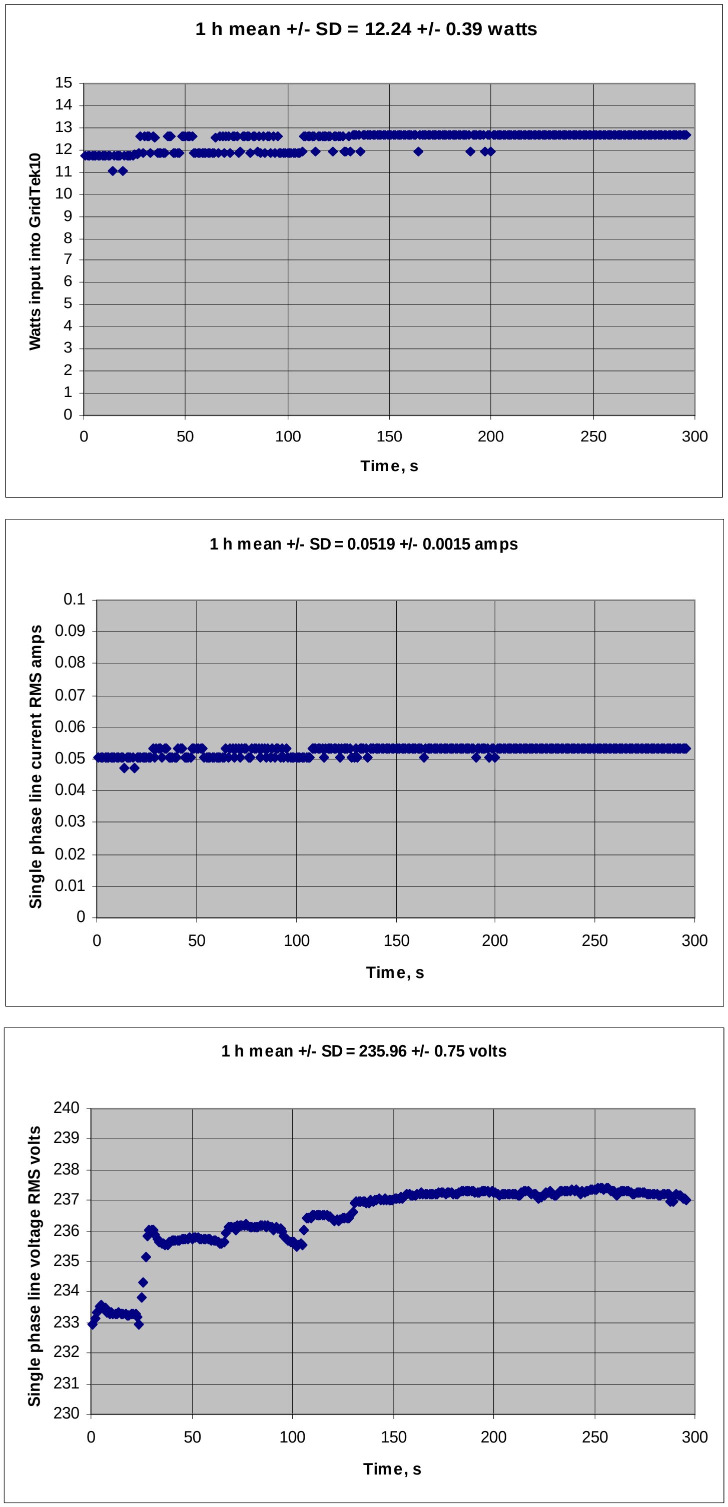

Figure 2: Power measurements when the GridTek10 is in standby mode. Plotted is the first 5 min of 1 h of data. Upper = power, middle = RMS current and lower = RMS line voltage. Line voltage goes up and down as loads in the house come on and off. See text for more details.

So the bottom line is that the cup anemometers ‘mechanically averages’ the wind speed and the power transducer electrically averages the power. So virtually nothing is gained by a high sampling rate unless your particular data-logger has a lot of internal noise.

Maybe you do not know that I have some experience in designing A/D systems, I have built such systems from the chips up. That means I have to understand the electrical characteristics at the analog end and the digital characteristics at the computer bus, i.e., I have to know: (1) how to interface the computer I/O address lines so a computer program can communicate with the chip, (2) how the data lines are used to get the integer values off the chips etc, (3) how to build amplifiers to convert transducer signals to usable A/D voltage ranges, and (4) how to write programs to make the A/D system come to life and deliver data. I have commercialized one of my designs, that uses a computer printer port to provide the I/O and data lines. The A/D board is sold as part of a system designed to measure hydraulic conductivity of plants, so you probably would not want one and it would cost $12,000 if you want to buy it from Dynamax Corp, who manufactures the complete system based on my design. So I think I have an above-average appreciation for error analysis in data-loggers.

What internal checks have I taken to insure that my power measuring systems have reasonable small noise-levels? One way is to record the noise level when the power is known to be zero or constant (e.g., turbine is not turning or my water heater is off). The other way is to measure the noise on the power consumption of a constant load, e.g., of the electric heating elements of my water heater or the compressor motor on my geothermal heat pump. From these measurements, I am satisfied that the resolution (repeatability) of my measurements is quite high (see Fig 2) and the accuracy in the range of 0.1 to 0.5 % of full scale for all power-transducer systems in my house.

In my ‘research project’ my main aim is to estimate how much energy is lost when the GridTek10 is off-line. This is now well documented in several reports. I would like to point out, however, that I do not need an accurate anemometer measurement of wind speed to do these calculations, rather I need only a reproducible anemometer. As long as my anemometer spins at a reliable RPM at any given wind speed, I can come to the same conclusions, i.e., that the GridTek10 squanders a large percentage of the energy it could gain if it worked properly (25.2% as of month 5 of my study). The sad thing is that BWC (and perhaps you) do not yet believe this fundamental fact. This is too bad, because if BWC does not believe there is room to improve their product then they might not try to improve it. BWC is beginning to ship a new inverter (the PowerSync II), and I wonder now if they really fixed the problem if they don’t admit there is one needing fixing.

I would like BWC to have an even better product. It is already quite good, but it could be better. I am trying to get them to make a better product. A better product would be good for BWC and for their customers, so that is why I keep at this. Some people wonder why I keep bashing Mike Bergey and BWC. That is only because they have refused to talk with me seriously and hence they neither believe my message nor the data I have collected so far.

Please feel free to challenge me in the future, Michael. I encourage you to call me to seek details or make suggestions at any time. If you convince me of the need for improvements, I will make them. It is much more efficient if you just talk to me privately first. When you contact me publicly you do so with a large measure of anger, which is counter-productive. When I am angry I do not think clearly, I am sure it is the same with you. That is why I frequently ignore your public displays of anger and wait until things ‘cool down’ before I respond. But I promise that eventually I will reply.

I welcome comments and suggestions for improvement to this and other SWIEP reports. I promise to correct mistakes promptly thru revised reports with acknowledgement to the person finding the mistake.

Mel Tyree

Appendix 1

Some basics on data acquisition systems.

Carefully read Figure 3 & caption on next page then start reading below.

A/D chips (stage 3A in Figure 1) usually convert DC voltages between +10 to -10 V DC to an integer. The size of the integer (a binary number) is determined by the number of bits. A 16 bit integer can represent values from 0 to 2^N-1 where N = the number of bits, which equals the number of binary digits, hence, integer values of 0 to 65,535 equals binary (0000 0000 0000 0000) to (1111 1111 1111 1111). If you want both positive and negative values then the integer range is -32767 to + 32767 where the most significant bit of the binary number indicates sign, so binary (0000 0000 0000 0001) = 1 and binary (1000 0000 0000 0001) = -1, and binary (1111 1111 1111 1111) = -32767 and binary (0111 1111 1111 1111) = +32767 unless you are using two's complement binary math but let's skip that! If you are not comfortable with binary number the main thing to realize is that voltage is converted to an integer and the count of "1" is the smallest voltage that can be measured. So for an A/D that can measure ± 10 volts (a range of 20 volts) the smallest absolute voltage than can be detected (counted) = 20/65,535 = 0.000305 V (0.305 mV) for a 16 bit A/D. Fast 16 bit A/D systems used to be expensive so many A/D systems purchased 10 to 20 years ago are 12 or 10 bits which corresponds to maximum counts of 4,096 or 2,048 respectively. Hence their measuring resolution is just 4.88 or 9.76 mV respectively. So what is the maximum theoretical resolution of a power transducer, such as the Ohio Semitronics system?

Lets take, for example, a 20 kW module that can measure +20 kW (power out) to -20 kW (power in) and converts it to +10 to -10 V, respectively. So 1 V = 2 kW. The Ohio Semitronics unit combines stages 1 & 2 into a single unit. If you measure the output of stage 2 with a 16 bit A/D the smallest level of power you can resolve is ± 0.61 W, or for 12 or 10 bit A/D systems the maximum theoretical resolution is ± 9.76 or 19.5 W. This looks fairly impressive until you consider two things: (1) A/D chips rarely achieve their maximum potential resolution in a real system (2) The maximum resolution might exceeds the factory guaranteed accuracy for most transducers.

A/D chips rarely achieve maximum accurately because sometimes they cannot count correctly, i.e., instead of counting 1, 2, 3, 4 (binary 0001, 0010, 0011, 0100) they count 1, 4, 2, 3 as signals increase. I have put some of the cheaper A/D chips on known voltage sources to confirm this. Hence a 16 bit chip might really be a 13 bit chip. More commonly, however, the A/D chip (stage 3A) and other electronics (stage 2) are in a noisy electrical environment. This makes the least significant bits useless. Computer TTL buses are always noisy environments and all A/D chips must be connected to a computer thru some kind of TTL bus. Additional noise comes from the stage 2 amplifiers. The bottom line is that a 16 bit A/D might have a real resolution of an 'ideal' 10 or 11 bit A/D. In terms of voltage this means a likely random error of ± 10 mV out of ± 10 V, so now we are down to a measuring resolution of ± 1 out of 1000 or 0.1%. Still, this is not too bad. That is because measurement resolution usually exceeds measurement accuracy.

What is the typical accuracy of a voltage/power measuring system? An overall accuracy of 0.2% to 0.5 % is fairly typical for voltage or power measurement. The 20 kW Ohio Semitronics system is factory guaranteed to be about ±0.5% of full scale = ± 100 W on a 20 kW system. So if you have a 5 kW system you are probably good to ± 25 W.

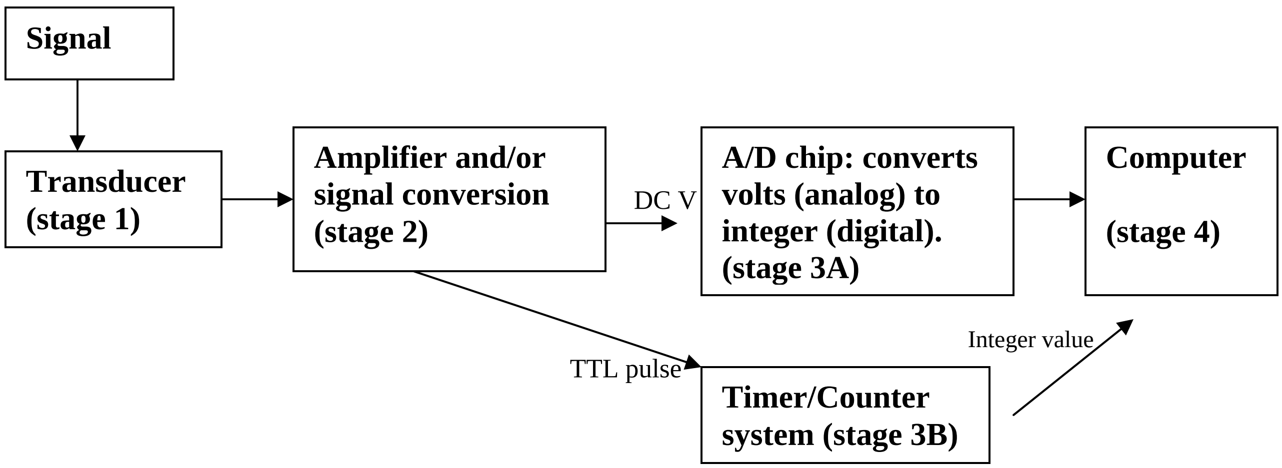

Figure 3: A block diagram showing the components of a data acquisition system to get from ‘signal’ to ‘numbers’. The signal can be temperature, pressure, wind speed, AC voltage, AC current, AC power etc. Stage 1 (the transducer) converts the signal to something else that can be measured THAT IS PROPORTIONAL TO THE SIGNAL, e.g., AC volts, AC pulses (frequency), ohms resistance, DC volts etc. The A/D chip usually requires DC volts in the range of +10 to -10 V. So if the transducer puts out DC volts of the correct magnitude then the output can go directly to the A/D chip (stage 3A). Stage 2: However, if the transducer puts out volts that are too small, AC volts, or pulses, or resistance etc, then the transducer signal has to go through an amplifier to boost the DC volts or converts AC volts, resistance or pulse to a DC volts (e.g., frequency to voltage circuit). Or the signal conversion in stage 2 can convert small pulses to TTL level (computer logic pulses of 0 to 5 V DC). Stage 3A: If stage 2 puts out a ‘good’ DC voltage it goes to Stage 3A (the A/D). This converts the analog DC volts to an integer value. Stage 3B: If stage 2 makes clean TTL pulses whose frequency is proportional to the signal (wind speed) then it can go to a counter/timer (Stage 3B), which can either measure the number of pulses in a desired time-interval or can measure the time between pulses (= period) from which frequency can be computed in Hz = 1/(period in seconds) depending on your needs. Stage 4: The computer reads the digital value, which is usually an integer and converts to it the nearest decimal value. Note the computer has to be programmed with the conversion factor between an integer-count from stage 3A or 3B to the units of the signal. The computer manipulates the values as floating point (decimal) values that appear to have more resolution than is really there. See text for more details.

Appendix 2

The astute reading might notice an apparent contradiction here. When I connect my LM335 temperature transducer to my 16 bit A/D I get only about 10 bits of resolution even with over-sampling at 500 Hz. That means there is electrical noise somewhere. I been using LM335’s for nearly 30 years with various data-loggers. I have always seen this noise level and I have assumed the noise is in the A/D computer bus or 60 Hz wiring noise. However, when I connect the CT-Coil current transducer I get full resolution (all 16 bits!) at least when readings are near zero. Why? I am thinking about this one a lot recently. Both transducers put out a DC signal, and both ought to have some inherent noise, but why is the noise from the LM335 so much more than from the CT-Coil? Perhaps when the CT-coil output rises the noise rises with the signal. Anyway, I only noticed this difference while writing and correcting this report. Appendix 2 was written as my first revision of the Feb 2nd report. I intend to investigate this some more, but if anyone has any ideas please Email me! mtyree@ales.ualberta.ca .