THE USE OF SONIC ARTICULATION IN IDENTIFYING CORRELATION IN CAPITAL MARKET TRADING DATA

David Worrall

Sonic Communications Research Group
University of Canberra
worrall@avatar.com.au

ABSTRACT

Despite intensive study, a comprehensive understanding of the structure of capital market trading data remains elusive. The one known application of audification to market price data reported in the 1990 that it was difficult to interpret the results probably because the market does not resonate according to acoustic laws. This paper reports on a technique transforming the data so it does resonate, so audification can be used as a means of identifying autocorrelation in capital market trading data. The results obtained indicate that the technique may have a wider application to other similarly structured time-series data.

1. INTRODUCTION

The statistical analysis of trading data, and stochastic modelling using that analysis, is an area of ongoing quantitative research in finance. Two principal concerns are to find techniques to accurately describe the way prices are distributed statistically and to what extent or auto-correlation exists and can be detected, even pre-empted, as market prices evolve. Understanding the first is important for the risk analysis of various trading instruments in the longer term, and understanding the second is important in attempts to predict future, especially catastrophic, events.

The power of visual representation to enhance and deepen an understanding of phenomena through their data abstractions is undisputed. Yet, as with many time-domain processes, a visual representation does not always reveal the structure of the data. Mandelbrot, arguably the most well known Quant[1], argues that is not possible to *visually* distinguish graphs of real market data and those of Brownian-motion-generated data[2] [30]. This leads to a data sonification question, which the research reported in this paper seeks to answer: Can trading data be presented in a way that permit its autocorrelation to be aurally distinguished from similar, but uncorrelated, data?[3]

This investigation was conducted during the development of software solutions in the *SoniPy* environment for the sonification of large multidimensional datasets [44]. As such, the primary emphasis was tool development and perceptual testing was informal.

2. THE DATA

The importance of understanding the data itself before attempting to sonify it has been long emphasised [5]. As will be discussed later, it is important in this particular circumstance, to distinguish between real trading data, simulated trading data and financial data such as general economic indicators, as they usually, arise out of different processes, have quite different structural characteristics.

2.1. Context

Capital markets are (increasingly virtual) places where companies and traders converge around an exchange to raise new investment capital, and investors and speculators trade exchange-registered securities such as stocks (shares), bonds, currencies, futures and other derivatives contracts. These exchanges have strict government-regulated mechanisms for such activity and the community of freely-participating individuals around them communicate more-or-less informally with each other and formally through exchange-registered brokers who themselves provide information to their clients about specific trading activity as well as about other more general environmental (financial, political, meteorological etc) conditions that may affect an individual's trading decisions. Such decisions, enacted by the brokers, cause excitations of the trading system, known colloquially as a 'trading engine', which in turn produces data records of its activities. Some of that data, and various summaries of it, are fed back for the information of market participants. In turn, these marketplaces operate as systems within national and global economies and international companies may be listed on more than one exchange. Each exchange's trading

[1] The term 'Quant' is used in the field to identify those who and quantitative analyse capital market data, or use such analysis to construct investment portfolios with a specific risk profile. See §2.2.

[2] Brownian motion is an independent (that is uncorrelated) random walk in which the size and direction of the next (price) move is independent of the previous move(s). A statistical analysis of time series data is concerned with the distribution of values without taking into account their sequence in time.

[3] Decorrelation can be achieved by changing *the sequence* of values in a time series. In so doing, any spectral information in the series is destroyed, while its statistical properties remain invariant.

system is designed to be acephalously appropriate for the types of securities that are traded on it.

Trading engines need to be fast, efficient and accurate[4]. They generate large quantities of data, reflecting the moment-to-moment shifting situation of the order book of each of their trading securities as potential buyers and sellers adjust their declared positions, and then eventually undertake trades. Security Trading datasets are sets of time-ordered trading events having a number of empirical dimensions, such as price and volume[5], depending on the particular type of security being traded (share, futures, options etc) and from which other data may be derived. A medium-sized exchange such as the Australian Securities Exchange (ASX) processes approximately two million trades a month: an average of 100,000 trades a day[6].

2.2. Quantitative analysis

As a discipline in finance, quantitative analysis begins with Bachelier's speculation that the market is efficient and thus price movement must be a random walk theory [1]. The mathematics of random walks was well known and an analysis of markets in terms of it enabled the construction of portfolios of stocks with defined risk profiles. Benoir Mandelbrot's study of price action in cotton led him to question this received wisdom, and to develop another mathematics, which he called "fractal" to accommodate his analytic findings. Fractal mathematics has become a widely applicable tool in many fields, both analytic and generative, and continues to be the basis of contemporary quantitative analysis. Quantitative analysis, especially of high-frequency data, remains an area of active research [4] and readable introductions are available in the popular science press for those less mathematically inclined [38][30].

Quantitative analysts generally use statistical analysis and stochastics to model the risk profiles of market indices, segments and individual securities as accurately as possible so as to assist in the construction of investment portfolios with certain characteristics, such as risk exposure, for example. To underestimate the risk of a portfolio is to court calamity, while overestimating it invites lower returns than might have otherwise been possible.

3. PREVIOUS WORK

3.1. Survey of the sonification of financial data

The first users of technology-enabled financial market data sonification[7] were probably the bucket shop traders in the early years of the twentieth century, who were reputed to be able to differentiate the sounds of stock codes, and prices that followed, from the sounds made by their stock-ticker machines as they punched recent trading information, telegraphed from an exchange, into a strip of rolling paper tape [28]. Janata and Childs suggest that Richard Voss may have been the first to experiment with the sonification of historical financial data: stock prices of the IBM corporation [22]. This is possible, as Voss and Mandelbrot were research collaborators in fractal mathematics at IBM's Thomas J. Watson Research Center and Voss played an early seminal role in the visualisation of fractal structures and in the analysis of the fractal dimensions of music and speech [41][42].

Kramer and Ellison used financial data in the early 1990's to demonstrate multivariate sonification mapping techniques [24]. This work was later summarized and published with sound examples [25]. The trading data used included four–and–a–half years of the weekly closing prices of a US stock index, a commodity futures index, a government T-bond index, the US federal funds interest rates, and value of the US dollar. Mappings were to pitch, amplitude and frequency modulation (pulsing and detuning), filter coefficients (brightness) and onset time (attack). Mapping concepts included redundant mapping and datum highlighting (beaconing).

Ben-Tal et al. sonified up to a year's end–of–day data from two stocks simultaneously by mapping them to perceptually distinct vowel-like sounds of about one second duration [3]. A single trading day was represented as a single sound burst. The closing price for the day was mapped to the center frequency, and the volume of trade to the bandwidth. These values were scaled such that the parameters for the last day of trade in each period corresponded to a reference vowel. Closing price was mapped to the number of sound bursts and volume (the number of trades) to duration. They informally observed that they could categorise high volume, high price trading days as loud, dense sounds, while low volume, low price days were heard as pulsed rhythmic sounds.

Brewster and Murray tested the idea that traders could use sounds instead of line-graphs to keep track of stock trends when they are away from the trading floor [6]. Using

[4] It is somewhat ironic that, in an enterprise that relies on 'clean' data, financial data often requires considerable 'washing' before its sonification can be undertaken. This situation is exacerbated by the trait that, with datasets over a certain size, the use of metadata tagging is uncommon, principally because it significantly increases the overall size of the dataset, even though the omission increases the likelihood of error. In any event, any cleaning has to be undertaken algorithmically and so it is expedient to have the tools for doing so integrated with the sonification software being used.

[5] The term 'volume' is used throughout to mean 'trading volume' not as a psychoacoustic parameter.

[6] A breakdown of recent ASX trading volumes is available from their website: www.asx.com.au/asx/statistics/TradingVolumes.jsp

[7] The presence of auditing (hearing of accounts from the Latin *auditus*) has been inferred from records of Mesopotamian civilizations going back as early as 3500 BCE. To ensure that the Pharaoh was not being cheated, auditors compared the 'soundness' of strictly independently scribed accounts of commodities moving in, out and remaining in warehouses [7]. In the alternating intoning of such lists, differences can be easily identified aurally. A faster and more secure method that eliminates any 'copy-cat' syndrome in such alternation, is to have the scribes read the records simultaneously–a type of modulation differencing technique. While we have no evidence that this specific technique was practiced in ancient times, such a suggestion does not seem unreasonable, and would represent possibly the earliest form of data sonification.

Personal Digital Assistants with limited screen space over a wireless network, one month of (presumably intraday) price data for a single share was mapped to pitch via MIDI note numbers. Participants, all students whose previous trading experience was unreported, were required to try to make a profit by buying and selling shares while monitoring price movement using either line or sound graphs. As trade transaction costs appear not to have been factored into the calculations, profits or losses were presumably gross. The experimental results showed no difference in performance between the two modes, but participants reported a significant decrease in workload when they used the sonification as it enabled them to monitor the price aurally while simultaneously using the visual display to execute trades.

Nesbitt and Barrass also undertook a multimodal sonification and visualisation study, this of market depth[8] to test whether subjects could predict the price direction of the next trade [34]. They used real data from a single security's order book. The visualisation used a landscape metaphor in which bid and ask orders (to buy and sell), were 'banked' on either side of a river, the width of which thus represented the size of price gap between the highest bid and the lowest ask, known as the 'bid–ask spread'. A wider river implied slower flow (fewer trades) and so on. The sonification employed the metaphor of an open-outcry market. A sampled male 'buy' and a female 'sell' voice displaying a discretely partitioned dataset (price, volume, price-divergence) was mapped into a discretely partitioned three-dimensional 'importance space' (pitch, loudness, stereo-location). This experimental design illustrates how sonification can be used to assist the apprehension of data segmentation such as where the trajectory of a parameter under focus changes.

Janata and Childs developed *Marketbuzz* as an add-on to conventional trader's terminals, such as those by Bloomberg, for the sonification of real-time financial data [23]. They used it to evaluate tasks involving the monitoring of changes in the direction of real-time price movements, with and without auditory or visual displays. A significant increase in accuracy using auditory displays was reported, especially when traders were visually distracted by a simultaneous diversionary "number-matching" task. Further, Childs details the use of sonification to highlight significant price movements relative to opening price, as well as continually changing features of Stock Options [8].

Mezrich, Frysinger and Slivjanovski developed a dynamic representation, employing both auditory and visual components, for redundantly displaying multiple multivariate time-series [32]. Each variable was represented by a particular timbre. The values of the variable were mapped to pitch. The analyst could focus on a subset of the data by interactively brightening or muting individual variables and could play the data both forwards and backwards. Subsets of the data could be saved and juxtaposed next to each other in order to compare areas where the data might be similar. In almost all cases, the sonified data performed as well as or better than the static displays.

Two other types of sonifications of securities data demonstrate different motivations but are mentioned here for completeness. The first is Ciardi's *sMax*, a toolkit for the auditory display of parallel internet-distributed stock-market data [9]. *sMax* uses a set of Java and Max modules to enable the mapping and monitoring of real time stock market information into recognizable musical timbres and patterns. The second is Mauney and Walker's rendering of dynamic data specifically for peripheral auditory monitoring. The system reads and parses simulated real-time stock market data that it processes through various gates and limiters to produce a changing soundscape of complementary ecological sounds [31].

There are a number of studies, principally those whose purpose was the study of parameter-mapping and auditory graphs, which have been omitted from this survey because it is not clear that there is anything in the findings specific to the structure of financial data; unless it is generated using advanced modelling techniques, fictional data is unlikely to exhibit the same structural characteristics as real financial time series data.

3.2. Audification

In the 1990's Frysinger experimented with playing back market price data directly as a sound waveform. He reported that he found that *the results proved difficult to interpret, probably because the stock market does not follow physical-acoustic resonance laws* resulting in natural or 'ecological' sounds that can be understood from everyday listening experience [16][17]. There appears to be no further reports of security data audification prior to the work reported in this paper. Hayward also suggested that another reason audification fails for arbitrary data such as stock market figures or daily temperatures is that the amount of data required: even at low sampling rates, it is difficult to make a sound with a duration long enough to reveal valuable information to the listener [21].

In summarizing previous work on seismic audification, Hayward reported both Speeth's original experiment on discriminating the seismic sounds of earthquakes from atomic explosions and Frantti's repeat of it with a larger number of participants with less training and data from diverse locations. Frantti found a lower average accuracy and a wider variance in participant performance, which was also critically affected by the audification's time-compression ratio and the number of repeat audits. Neither study used trained seismologists, nor did participants have any interactive control [15][11]. Concentrated on single wavelett and quantitative questions, Hayward indicated a number of solutions to the difficulties encountered as well as some strategic extensions to planetary seismology in general. Dombois reported that he could hear seismological station-specific characteristics in his time-compressed audifications [12]. He informally found that, over time, overall information of a dynamic state was better comprehended with audification, whereas visualization was more effective when a detailed analysis of a single wavelett was required. He developed a unified acceleration method to make records taken under different meteorological and seismic conditions more compatible and in a later report on the state of research in auditory seismology, documented several other investigations in the field in the 1990s and much earlier [13]. He reported an increase in interest among seismologists, and this is also evidenced by the recent reporting in the popular media of the audification of stellar seismology [19].

[8] *Market depth* is a term used to denote the structure of potential buy and sell orders clustered around the most recently traded price.

Using data from a helicopter flight recorder, Pauletto and Hunt showed that audification can be used as an equally effective alternative to spectrograms for the discernment of complex time-series data attributes such as noise, repetitive elements, regular oscillations, discontinuities, and signal power [35]. However, another empirical experiment found that the use of audification to represent data related to the rubbing of knee-joint surfaces was less effective at showing the difference between normal and abnormal signals than other sonification techniques [27].

3.3. Sonification of stochastic functions

Aside from their use in algorithmic music composition, stochastic functions have received little attention in sonification research. Perhaps the first was a study of use of parameter-mapping and physical model sonification is used in a series of experiments in monitoring the performance of Markov chain Monte-Carlo simulations for generating statistical data from higher dimensional probability density functions [22]. The inclusion of some sound-generating tools in the statistical package R has the potential to generate wider interest, as exemplified by its use in tuning a parameter in the Hybrid Monte-Carlo algorithm [20]. Informal auditing of a technique to sonify, using amplitude modulation, cross-correlations in irregularly spiking sequences that resemble a Poisson process led to the postulation that the use of sonification for time series analysis is superior to visualisation in cases where the intrinsic non-stationarity of an experiment cannot be ruled out [2]. Time series data was generated by shaping a uniform distribution (white noise) with a cumulative probability density function, (similar to that used by Xenakis for his *ST* series of compositions [45]), in a differentiation study of the perceptualisation of some statistical properties of time series data generated using a Lévy skew alpha-stable distribution of interest to modellers of financial time series [38]. The study found no evidence that skewness in their data was perceivable, but participants were able to distinguish differences in kurtosis, which correlated with roughness or sharpness of the sound. This research provided empirical support for a part of the earlier initial findings of the experiments outlined below [44].

4. INFORMAL EXPERIMENTS

The experiments described here sought to (a) discover a way to directly audify a Capital Market Trading Dataset that preserved its autocorrelation characteristics and (b) ascertain informally whether such a dataset can be aurally discriminated from an audification of a statistically equivalent uncorrelated dataset. The null hypothesis in each case was that no distinction could reliably be made between the audifications.

4.1. The dataset

The dataset chosen is twenty-two years of the daily closing price of All Ordinaries Index (ticker XAO) of the

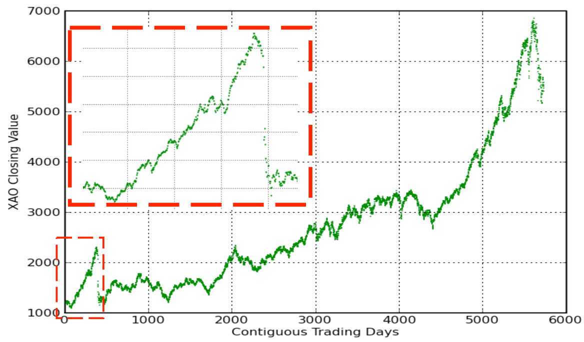

Figure 1. *A plot of the 22 years of daily closing values of the ASX's All-Ordinaries Index (XAO).*

Australian Securities Exchange (ASX)[9] as illustrated by the plot in Figure 1.

The first task was to find a way to overcome the non-resonance problem referred to earlier as discussed by Hayward [21]; one that transformed the dataset to be suitably oscillatory while preserving its correlational integrity. An equivalent problem is to be found in quantitative analysis, as observed by Stony Brook computer scientist Steven Skiena:

> *The price of an asset as a function of time is perhaps the most natural financial time series, but it is not the best way to manipulate the data mathematically. The price of any reasonable asset will increase exponentially with time, but most of our mathematical tools (e.g. correlation, regression) work most naturally with linear functions. The mean value of an exponentially-increasing time series has no obvious meaning. The derivative of an exponential function is exponential, so day-to-day changes in price have the same unfortunate properties.* [38]

The Net Return, or simply, the Return, is a complete and scale-free summary of investment performance that oscillates from positive values (increase) around zero (no change). The XAO dataset was converted to market returns. For an asset whose price changed from p_t at time t to $p_{t+\partial t}$ at time $t+\partial t$, the simple linear return R_{lin} is defined as

$$R_{lin} = p_{t+\delta t} - p_t \qquad (1)$$

Because prices tend to move exponentially over longer timeframes time, that is, in percentage terms, a better measure than R_{lin} is the ratio of successive price differences to the initial prices. These are known as net linear returns [3]:

$$R_{net} = \frac{p_{t+\delta t} - p_t}{p_t} \qquad (2)$$

Figure 2 is a plot of net returns of the XAO dataset. The insert is of the first 500 samples, similarly to the Figure 1 insert. Table 1 summarises the statistical properties of these returns, clearly showing that they are not a normally

[9] The XAO is the broad Australian market indicator, a composite of the 500 largest companies, weighted by capitalisation, which are listed on the exchange. Contextual details are available at http://www.asx.com.au/research/indices/description.htm

distributed. The single largest Net Return, clearly visible in Figures 1 and 2, was on 20 October 1987 ("black" Tuesday) the largest one-day percentage decline in stock market history. The difference between this largest minimum and the second-largest minimum is 62% of the total returns space. This is shown in the histogram of Figure 3, which illustrates the frequency of net returns. The single minimum and second-largest minimum are circled, but barely visible at this scale.

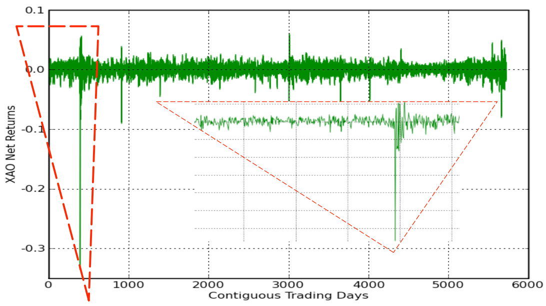

Figure 2. *A plot of XAO net returns*

number of samples	5725
minimum sample	0.333204213
maximum sample	0.05886207483
arithmetic mean	2.1748845e04
variance	9.5685881e-05
skewness	7.6491182
kurtosis	241.72988

Table 1: *Basic statistics for XAO net returns.*

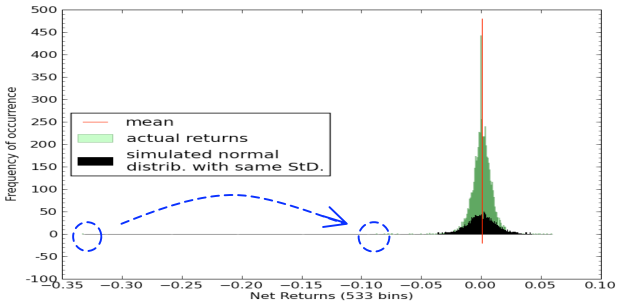

Figure 3 *An histogram of net returns that illustrates the proportion of the dataspace allocated to a single negative outlier.*

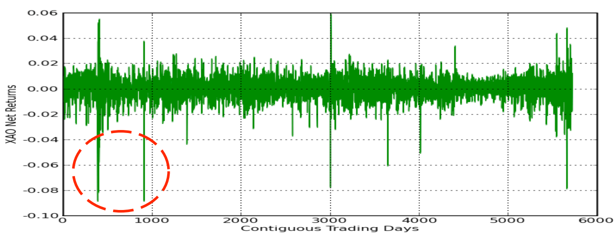

Figure 4. *Plot of XAO net returns, clipped so as to be suitable for audification.*

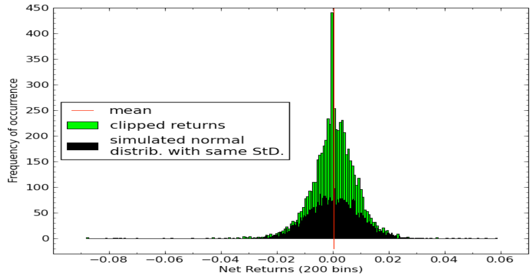

Figure 5. *Histogram of the clipped net returns overlayed with a simulated normal distribution with the same standard deviation and number of datum for comparison.*

So, despite its anecdotal interest, an 'audacious' clipping, or limiting, of the largest minimum sample to that of the second-largest minimum, was performed and the resulting returns plotted in Figure 4. Its histogram is show in Figure 5, which illustrates both the skewness and kurtosis of the dataset, when compared to a normal distribution. The size of the sample-bins of this histogram is kept constant to those of the Figure 4 histogram by decreasing the number of bins. Of interest is the asymmetry of the outliers (the data at the extremities): there are more of the negative variety than positive, and negative ones exist further from the mean; even more so when considering this is the clipped dataset.

For comparison, the net returns were we decorrelated. Figure 6 shows the plots of both the correlated and decorrelated datasets. A number of features are visually apparent. Both have long tails but they appear more evenly distributed throughout the decorrelated dataset, contributing to its more chaotic visual appearance, whilst the correlated dataset appears to have periods of increasing (trapezoid), low (circle), and ramped (diamond) volatility.

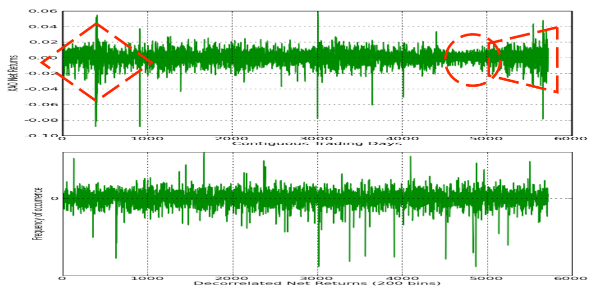

Figure 6. *A plot of the correlated (top) and decorrelated net returns.*

4.2. Experiment 1: Audification

In order to test whether the raw and decorrelated data sets could be distinguished aurally, a number of chunks of audio were prepared with the same number of samples as the Net Returns.

A. Uniformly distributed
B. Normally distributed (Gaussian)
C. Decorrelated Returns
D. Raw Returns

In Audio Example 1 these four chunks can be heard four times at a sample rate of 8 kHz in the order A-B-C-D. Each audio chunk is approx. 0.7 seconds duration. There is a one-second gap between each chunk and a further one-second gap between repeats. The following informal observations can be made:

- The uniformly distributed noise (A) is clearly distinguishable from the Gaussian (B). This distinction is unsurprising: it is that between white and band-limited noise in electronic music parlance. As would be expected, the uniformly random noise sounds "brighter" because of the comparatively greater prevalence of higher frequencies.
- The raw and decorrelated returns (D and C) are clearly distinguishable from A and B: Qualitatively, they sound rougher or grainier, and they have less evenly distributed spectral energy than A and B. This can be interpreted as a result of the increase kurtosis, as reported in an empirical study by Baier et al. [2].
- The 8 kHz sampling rate was settled on after some initial heuristic experimentation with higher and lower values. There appears to be an optimal compromise between durations long enough for possible temporal patterns to be perceptible and sampling rates high enough to make shorter-term correlations perceptible. No formal method for determining the optimization seems to be currently known, yet the choice clearly influences the perceptibility of pattern, as was also observed by Dombois in his seismic audification studies [12][13].

4.3. Experiment 2: Audification

Having ascertained that the Returns were clearly distinguishable from uniform and Gaussian noise, a second experiment was conducted to ascertain whether or not the raw Returns and the decorrelated Returns could be aurally distinguished from each other. An additional Decorrelated Return (E) was generated in the same manner described for C, in Experiment 1, and three files were prepared with the following sequences in which the original raw returns (D) was placed in first second and third place respectively:

Audio Example 2. D-C-E
Audio Example 3 C-D-E
Audio Example 4 C-E D

Audio Examples 2-4. *Three sequences each of three audio chunks, two of which (C & E) are decorrelated versions of the Net Returns (D).*

The listening task, on multiple random presentations of these audio files, was to try to determine, in each case, which one of the three chunks sounded different from the other two. The informal findings of several listeners, all of who had musical training, can be summarised as follows:

- The task was a more cognitively demanding than those in Experiment 1.
- Distinguishability was dependent on a narrower band of sampling rates. Above 8 kHz the characteristics described earlier seem to disappear. Below 3-4 kHz the roughness created by the individuation of large-valued samples meant that the principal means of identifying the raw returns was probably more by its invariance across all chunk presentations than by direct chunk comparison.
- Between 4 kHz and 8 kHz sampling rate, a distinct, though subtle, amplitude modulation was observable in the Net Return chunks that seems not to be present in the decorrelated ones. This amplitude modulation effect required attentive listening, probably, in part, due the relatively short duration of the audio chunks (less that 700 ms).

This last observation pointed to the need for more data to enable longer durations or the application of a technique other than audification that enables a slower sample presentation rate. As no intraday data was available for the dataset in question, the latter approach was chosen in Experiment 3.

4.4. Experiment 3: Homomorphic Modulation Sonification

This experiment was designed to test a simple proposition: That the four datasets A, B, C and D of Experiment 1 could be distinctly identified under homomorphic mapping into a pitch-time auditory space. A homomorphic mapping is one in which the changes in a dimension of the auditory space track changes in a variable in the dataset, with only as few mediating translations as are necessary for comprehension [26]. A narrow interpretation, called Homomorphic Modulation Sonification is used, in which time in the dataset was mapped to time in the auditory display and sample value was mapped to pitch deviation (both positive and negative) from a centre frequency.

There is a subtle but important distinction between Homomorphic Modulation Sonification and the type of parametric mapping in which each datum is played as, or contributes to, a separate tone with its own amplitude envelope. In the separate-tones case, the audio-amplitude profile of the resulting audible stream fluctuates from–and–to zero, resulting in a sequence of auditory objects individuated by more–or–less rapid onset transients. With modulation however, a single continuous pulsed waveform results, affording the opportunity for the amplitude formant to be held relatively constant, resulting in a lower perceptual loading [36][4]

A csound [9] instrument, illustrated in Figure 7, was constructed to produce this homomorphic mapping. Structurally, this is a basic frequency modulator in which an ADSR[10] for controlling modulation index is replaced by a sample buffer of the AIFF samples. These samples, which can be read directly from an audio file, are then used in sequence to control the frequency deviation from a user-defined centre 'reference' carrier frequency.

[10] Computer music parlance: An ADSR is an Attach-Delay-Sustain-Release envelope shaper, a common tool for synthetically controlling the amplitude evolution of computed sounds.

Control was implemented to facilitate heuristic investigation of the perceptual space in *SoniPy* (via MIDI and Csound's *Python* API)[10] and in MacCsound. Figure 8 shows the MacCsound controller interface [29]. The setting shown is for frequency-modulating a 300Hz tone according to the values of successive samples within a range of four

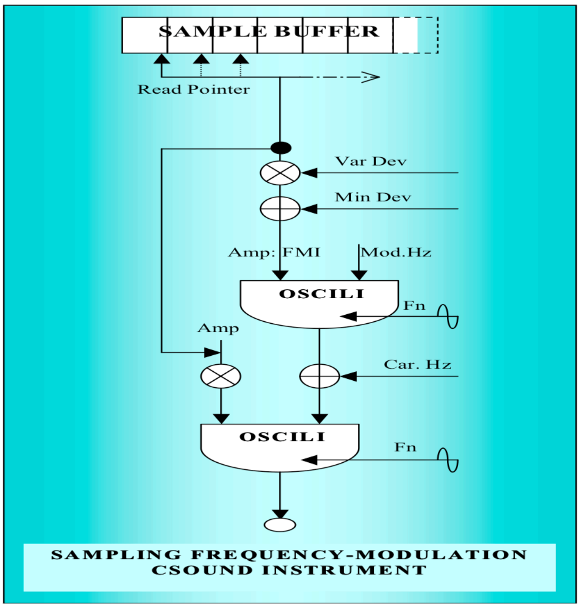

Figure 7. *A graphic illustration of the csound instrument used for the homomorphic mappings experiment.*

octaves at the rate of 480 modulations per minute (8 Hz) from the Net Returns distribution. With this controller, it is possible to dynamically adjust the pitch spread and centre frequency during audition.

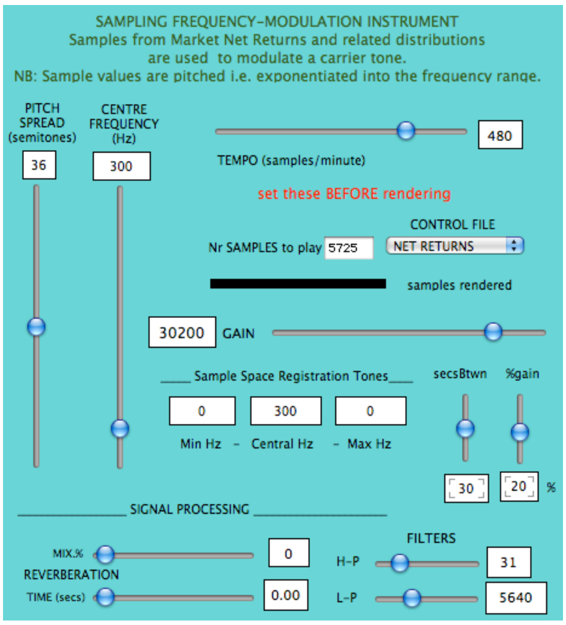

Figure 8. *User interface of the sampling FM instrument.*

In addition to the sample modulator, the sonification also uses has a simple audio 'tick' registration tone generator that acts as an auditory reminder of the upper, lower and central limits of the current render. Both its frequency-of-occurrence (*secsBtwn*) and relative loudness (*%gain*) is adjustable.

Audio Examples 5–9 provide various sonic realisations of the homomorphic mappings of the Net Returns samples generated for Experiments 1 and 2. For consistency, all are rendered with the settings illustrated in Figure 8.

- Audio Example 5 is a series of twenty-second 'snapshots' of each of the four sample sets A, B C and D.
- Audio Example 6 is of A.
- Audio Example 7 is of B
- Audio Example 8 is of C.
- Audio Example 9 is of D.

Audio Examples 5-9 *Examples of homomorphic mapping sonifications of the four distributions: uniform, normal, net returns and decorrelated net returns.*

The informal findings of Experiment 3 can be summarised as follows:

- ☐ The difference between homomorphic mapping sonifications of A, and B is easily noticeable, at least to a musically trained listener, as it was in the audifications of Experiment 2. The homomorphic mapping of A can be observed to be evenly spread across the pitch gamut, while the Normal distribution of B can be observed to be more closely clustered around the centre of the gamut, the mean.
- ☐ Again, C and D are noticeably different to A and B. Whilst both C and D appear to have short trending auto-correlative sequences, those of C (the Net Returns) appear more consistently and when they do they appear to last for longer periods of time. This is particularly noticeable in frequency of sequences of consecutive zero or small Net Returns, a characteristic consistent with the observation by chartists and technical analysts that securities prices frequently shift to a new price 'zone' quite quickly interspersed with longer times consolidating those zones before moving again [33].

5. CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK

It is apparent from these experiments that the simple technique of using net returns is applicable to the sonification of capital market trading data and so it may be possible to apply the same techniques to other similarly structured time-series datasets such as electoencephalography data, trans-synaptic chemical transmitters, and the hierarchical networks arising from social affiliations. A further interesting extension study would be the application of the techniques to the functional simulations of financial-market-like time-series, an active field of econometric investigation in which large datasets can be generated when needed.

Although other techniques can be applied, the directness of audification makes it appealing and controlled empirical experiments to determine which features of the dataset are perceivable under those conditions would be worthwhile. The observation of amplitude modulation in the raw returns in Experiment 2 suggests that an empirical study to isolate the aural characteristics of cross-correlation, such as the spectral modulation suggested by the study, may be useful. This would require the preparation of additional, unrelated, raw returns datasets of the same sample size.

The choice of sampling rate clearly influences pattern perceptibility, as was also observed in seismic audification studies [40] but apart from limiting the resulting frequency band imposed, no reliable formal optimization method is known and this deserves empirical attention, perhaps by using the using the fractal dimension of the dataset as a potential correlation index.

The effect of the size of the dataset on the sampling rate also needs to be isolated, as, whether or not higher sampling rates on larger datasets reveal other distinguishing features is currently not known.

This investigation was conducted during the development of software solutions in the *SoniPy* environment for the sonification of large multidimensional datasets [43]. As such, the primary emphasis was tool development and perceptual testing was informal. A heavily commented version of the script developed for these experiments, along with the audio examples is available for download from http://www.sonification.com.au/securities.

6. ACKNOWLEDGEMENTS

This work was supported by funding from the University of Canberra and the Capital Markets Cooperative Research Centre. Useful feedback was also received from Roger Dean of the Marcs Audio Laboratory, University of Western Sydney.

7. REFERENCES

[1] L. Bachelier. Théorie de la speculation. *Annales de l'Ecole Normale Superieure* 17 pp. 21-86, translated into English by A.J. Boness in *The random character of stock market prices*, edited by P.H. Cootner. MIT Press. pp. 17-78. 1900/1967.

[2] G.T. Baier, T. Hermann, O. M. Lara and M. Müller. "Using sonification to detect weak cross-correlations in coupled excitable systems," in *Proceedings of The Eleventh Meeting of the International Conference on Auditory Display*, Limerick, Ireland, July 6-9. 2005.

[3] O. Ben-Tal, J. Berger, B. Cook, M.Daniels, G. Scavone and P. Cook. "SONART: The sonification application research toolbox," in Proceedings of the 2002 International Conference on Auditory Display, Kyoto, Japan, July 2-5. 2002.

[4] V. Best, E.J. Ozmeral, N. Kopco and B.G. Shinn-Cunningham. Object continuity enhances selective auditory attention. In Proceedings of the National Academy of Sciences of the United States of America, Volume 105, Issue 35. Pages 13174-13178. 2008.

[5] S. Bly "Multivariate Data Mappings", in *Auditory Display: Sonification, Audification, and Auditory Interfaces*, edited by G. Kramer. Santa Fe Institute Studies in the Sciences of Complexity, Proceedings. Vol. XVIII Addison Wesley Publishing Company, Reading, MA, USA. 1994.

[6] E. Boyd. History of auditing in R. Brown (ed.). *A History of Accounting and Accountants*. Edinburgh: T.L & E.C. Jack. 1905.

[7] S. Brewster and R. Murray. "Presenting dynamic information on mobile computers," in *Personal Technologies*, Volume 4, Number 2. pp. 209-212.

[8] E. Childs. "Auditory graphs of real-time data," in *Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display*, Limerick, Ireland, July 6-9. 2005.

[9] F.C. Ciardi. "*sMAX*. A multimodal toolkit for stock market data sonification, in *Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display*, Sydney, Australia, July 6-9. 2004.

[10] http://www.csounds.com

[11] F. Dombois. "Using Audification in Planetary Seismology," in *Proceedings of Eighth International Conference on Auditory Display*, Espoo, Finland, July 29- August 1. 2001.

[12] F. Dombois. "Auditory Seismology on free Oscillations, Focal Mechanisms, Explosions and Synthetic Seismograms," in *Proceedings of the 2002 International Conference on Auditory Display*, Kyoto, Japan, July 2-5. 2002.

[13] F. Dombois. "Underground Sounds. An approach to earthquake prediction by auditory seismology," in *Geophysical Research Abstracts*, Volume 4. EGS02-A-02741. 2002.

[14] J.D. Farmer, P. Patelli, and I.I. Zovko. "The predictive power of zero intelligence in financial markets" in *Proceedings of the National Academy of Sciences of the United States of America*, Volume 102, Number 6, pp. 2254-2259. Accessed on 23 October 2007 from http://www.pnas.org/content/102/6/2254.full.pdf+html. Includes links to extensive range of supplementary material. 2005.

[15] G.E. Frantti and L.A. Leverault. "Auditory discrimination of seismic signals from earthquakes and explosions," in *Bulletin of the seismic society of America,* Volume 55, pp. 1-25. 1965.

[16] C. Frauenberger, A. de Campo and G. Eckel. "Analysing time series data," in *Proceedings of the 13th International Conference on Auditory Display*, Montreal, Canada, June 26 - 29. 2007.

[17] S.P. Frysinger. "Applied research in auditory data representation," in *Extracting Meaning from Complex Data:Processing, Display, Interaction*, Santa Clara, California, SPIE-The International Society for Optical Engineering. 1990.

[18] S.P. Frysinger, S.P. "A brief history of auditory data representation to the 1980s," in *First Symposium on Auditory Graphs*, Limerick, Ireland, July 10, 2005.

[19] P. Gosh. "Team records 'music' from stars". BBC news website. http://news.bbc.co.uk/2/hi/science/nature/7687286.stm. 2008. Accessed on 24 October, 2008.

[20] M. Heymann and M. Hansen. "A new set of sound commands for R; sonification of the HMC algorithm," in *ASA Proceedings, Statistical Computing Section*.

Alexandria, VA: American Statistical Association. pp. 1439-1443. 2002.

[21] C. Hayward. "Listening to the Earth sing", in Kramer G. (ed) *Auditory Display: Sonification, Audification, and Auditory Interfaces*, Addison-Wesley, Santa Fe Institute Studies in the Sciences of Complexity, Proceedings. Vol. XVIII Addison Wesley Publishing Company, Reading, MA, USA. pp. 369-404. 1994.

[22] T.M. Hermann, H. Hansen and H. Ritter. "Sonification of Markov Chain Monte Carlo simulations," in *Proceedings of the 7th International Conference on Auditory Display*, Helsinki University of Technology. pp. 208-216, 2001.

[23] P. Janata and E. Childs. "*Marketbuzz*: Sonification of real-time financial data," in *Proceedings of the 2004 International Conference on Auditory Display.* Sydney Australia July 5-9. 2004.

[24] G. Kramer and S. Ellison. "Audification: The use of sound to display multivariate data," in *Proceedings of the International Computer Music Conference*, San Francisco, C.A. ICMA. pp. 214-221. 1991.

[25] G. Kramer. "Some Organising Principles for Representing Data with Sound", in Kramer G. (ed) *Auditory Display: Sonification, Audification, and Auditory Interfaces*, Addison-Wesley, Santa Fe Institute Studies in the Sciences of Complexity, Proceedings. Vol. XVIII Addison Wesley Publishing Company, Reading, MA, USA. 1994.

[26] G. Kramer. An Introduction to Auditory Display. In G. Kramer (ed.). 1994. Auditory display: Sonification, Audification, and Auditory Interfaces. Santa Fe Institute Studies in the Sciences of Complexity, Proceedings, Volume XVIII. Reading, MA: Addison Wesley Publishing Company. Pages 1-77. 1994.

[27] S. Krishnan, R.M. Rangayyan, B. G. Douglas and C.B. Frank. "Auditory display of knee-joint vibration signals," in *Journal of the Acoustical Society of America*, Volume 110, Number 6, December 2001, pp. 3292-3304. 2001.

[28] E. Lefèvre. *Reminiscences of a Stock Operator*. John Wiley & Sons. NJ, USA. 1923/1994. 7 E. Boyd. History of auditing in R. Brown (ed.). A History of Accounting and Accountants. Edinburgh: T.L & E.C. Jack. 1905.

[29] http://www.csounds.com/matt/MacCsound/

[30] B.B. Mandelbrot and R.L. Hudson. *The (mis)behaviour of markets*. Basic Books, NY, USA. 2004. pp 16-18.

[31] B.S. Mauney and B. N. Walker. "Creating functional and livable soundscapes for peripheral monitoring of dynamic data," in Proceedings of ICAD 04 - The Tenth International Conference on Auditory Display, Sydney, Australia, July 6-9, 2004.

[32] J.J. Mezrich, S. P Frysinger, and R. Slivjanovski. "Dynamic representation of multivariate time-series data", in *Journal of the American Statistical Association*, Volume 79, pp. 34-40. 1984.

[33] J. Murphy. Technical Analysis of the Financial Markets. New York: New York Institute of Finance. 1999.

[34] K.V. Nesbitt and S. Barrass. "Evaluation of a multimodal sonification and visualisation of depth of market stock data," in *Proceedings of the 2002 International Conference on Auditory Display*, Kyoto, Japan, July 2-5. 2002.

[35] S. Pauletto and A. Hunt. "A comparison of audio & visual analysis of complex time-series data sets," in *Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display*, Limerick, Ireland, July 6-9. 2005.

[36] R.D. Patterson. Guidelines for the design of auditory warning sounds. Proceeding of the Institute of Acoustics, Spring Conference, Volume 11, Number 5. Pages 17-24.1989.

[37] E. E.Peters. *Chaos and order in the capital markets*. John Wiley & Sons, Inc. NY. 1991.

[38] S.S. Skiena. *Financial time series data* http://www.cs.sunysb.edu/~skiena/691/lectures/lecture6/lecture6.html 2004. Accessed on 20 January 2008.

[39] D. Sornette, *Why stock markets crash. Critical events in complex financial systems*. Princeton University Press. Princeton, NJ, USA. 2003.

[40] S.D. Speeth. "Seismometer sounds". *Journal of the Acoustic Society of America*, Volume 33, pp. 909-916. 1961.

[41] R.F. Voss and J. Clarke. "*1/f noise* in music and speech," in *Nature 258*, pp. 317-318. November 1975.

[42] R.F. Voss and J. Clarke. "*1/f noise* in music: Music from 1/f noise," in *Journal of the Acoustical Society of America*, Volume 63, Number 1, pp. 258-263. 1978.

[43] D. Worrall. "Audification experiments using XAO returns, in Auditory display as a tools for exploring emergent forms in exchange-trading data". *Report to the Sonic Communication Researcch Group*, University of Canberra, 28 October. www.avatar.com.au/sonify/research/sonsem1/4audExperi.html. 2004.

[44] D. Worrall. M. Bylstra, S. Barrass and R.T. Dean. "SoniPy: The design of an extendable software framework for sonification research and auditory display," in *Proceedings of the 13th International Conference on Auditory Display*, Montréal, Canada, June 26-29, 2007.

[45] I. Xenakis, I. *Formalized Music: Thought and Mathematics in Music*. Indiana: Indiana University Press. pp136-143. 1971.