

**A Search for Periodic Time Variations in the Solar Neutrino
Data from the Sudbury Neutrino Observatory**

by

Louise Heelan

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of the requirements
for the degree of
Master of Sciences

Ottawa-Carleton Institute for Physics
Department of Physics
Carleton University
Ottawa, Ontario, Canada
September, 2005 ©copyright

Louise Heelan, 2005



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*

ISBN: 0-494-10139-3

Our file *Notre référence*

ISBN: 0-494-10139-3

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Abstract

The Sudbury Neutrino Observatory (SNO) is a real-time Čerenkov radiation detector that detects ${}^8\text{B}$ solar neutrinos via three interactions on a heavy water medium. In the first two phases, spanning four calendar years, the experiment recorded 2924 and 4722 candidate events, respectively.

These data sets are examined with the Lomb-Scargle periodogram to search for sinusoidal periodic time variations in the observed solar neutrino rate. In total, 7300 periods are sampled, from 2 to 3650 days. A Monte Carlo analysis is performed to statistically quantify the results and to test the sensitivity of the method. This study finds no significant periodic behavior in the solar neutrino data taken by SNO.

Acknowledgements

This journey has led me up an incredible path, and in reaching this summit I have come across a group of most fabulous people to whom I would like to thank for their guidance, support, and kindness.

I wish to express my gratitude and admiration for my supervisor Richard Hemingway. He has provided unwavering support and guidance in the development of this thesis. Along the way he has taught me additional lessons on the value of asking ‘why’, necessity for organization, and importance of humour; these are lessons I will carry forward in my academic and personal life. It has been a pleasure and an honour to have worked with you, Richard.

I would like to thank the SNO collaboration for the privilege of analyzing the solar neutrino data; specifically the periodicity working group, topic committee, and editorial board for their feedback on this analysis. I would like to extend a large thank-you to Scott Oser, the leader of the periodicity working group. He has taught me a lot about the ways of statistical analysis, critical thought, and hard work. I hope in the future our paths cross and we once again work together.

I would like to thank the entire Carleton SNO group for their moral and financial support. Thank you to those involved in the Carleton computing facility. And thanks to my fellow graduate students for lightening those stressful times.

Behind the scenes my family and friends have provided support, encouragement, and balance. Thank you to all my family, especially Niall and Annette, and David, Pauline, and Tim. Thank you to my close friends Laura (B and H), and Olivier (who has added another layer to this experience). To Jared and Kristin, my physics buddies from UWaterloo, thank you for those endless nights in the study room where we scratched our heads in confusion.

Working on this experiment over these last two years has been an incredible experience that I will never forget. Thank you.

Statement of Originality

The ideas and results presented in this thesis are the product of the original author, L. Heelan. Any work from people other than this author is fully acknowledged in accordance with standard referencing policies.

The first two chapters of this thesis borrow heavily from theoretical work on solar physics and spectral analysis. All of the original authors have been cited within the text. This author has attempted to weave together a brief review of solar and neutrino physics, along with theoretical motivations for a periodicity analysis, and past work on this topic.

The extensive task of extracting the event and run information is a collaborative effort on the part of all SNO members and the list of authors is provided in the collaboration papers that are referenced within the text. K. Graham, S. Oser, and D. Waller provided this author with the official SNO event and run lists from the D₂O and salt phases.

All analysis code has been written by the author in C++, with use of the CERN ROOT libraries. The core Lomb-Scargle (LS) routine was taken from *Numerical Recipes in C* [1] to make use of the CPU time-saving recursion relations. This code was modified by the author to be written in a C++ version and incorporate weighted data points.

E. Inrig completed her LS study of SNO's D₂O data set in May 2004 at Carleton University [2]. Inrig used the regular LS analysis technique and applied it to an unofficial version of the D₂O run and event lists, making use of a series of Fortran programs she had written. The author of this thesis acknowledges the founding work of Inrig on this topic. In the early days of the analysis, this author reproduced Inrig's results using the new C++ code, and moved forward to analyze the official D₂O, salt, and combined phase data sets using the modified (weighted) LS method.

The concept for the Monte Carlo (MC) data set generation was originally proposed by Oser. The author of this thesis created the routine to generate the null hypothesis and periodic MC files. The generation and analysis of the MC data sets was performed by this author on the computing facility at Carleton University.

The author has posted seven documents on the private SNO manhattan server to document the development of this analysis. These range in subject from preliminary results to systematic checks, some of which are included in this thesis.

This author is a primary contributor to [3], a SNO collaboration paper reviewing the results of several periodicity analyses, which includes the work presented in this thesis.

Contents

Abstract	iii
Acknowledgments	iv
Statement of Originality	vi
List of Tables	xi
List of Figures	xiii
List of Acronyms	xvi
1 Solar and Neutrino Physics: A Brief Historical Review	1
1.1 The Standard Solar Model	2
1.2 Neutrino Properties	7
1.2.1 The Solar Neutrino Problem and Solutions	10
1.3 Decades of Solar Neutrino Experiments and Periodicity Searches	14
1.4 Spin Flavour Precession Revisited	17
1.5 Motivations to Search for a Periodic Solar Neutrino Rate	19
1.6 Synopsis of the Thesis	22
2 Searching for a Periodic Signal	23
2.1 General Introduction to Fourier Analysis	23
2.2 Sampling Data at Even Time Intervals	26
2.3 Sampling Data at Uneven Time Intervals	28

2.3.1	Lomb-Scargle Method	29
2.3.2	Modified Lomb-Scargle Method	34
2.3.3	A Monte Carlo Example of the Regular and Modified Lomb-Scargle Methods	36
3	Detecting Solar Neutrinos at SNO	39
3.1	The Sudbury Neutrino Observatory	39
3.2	Timing System	43
3.3	Run List Selection	44
3.4	Event Selection	47
3.4.1	Trigger System	47
3.4.2	Backgrounds	48
3.4.3	Candidate Solar Neutrino Events	50
4	Lomb-Scargle Method Applied to the SNO Data	52
4.1	Input Variables	52
4.1.1	Bin size	52
4.1.2	Time	53
4.1.3	Rate	54
4.1.4	Weight	55
4.1.5	Frequency Range	56
4.1.6	Summary	57
4.2	Lomb-Scargle Analysis	57
4.2.1	D ₂ O and Salt Data Separately	57
4.2.2	D ₂ O and Salt Data Combined	60
4.2.3	Results	62
5	Monte Carlo Analysis of the Periodogram Results	66
5.1	Monte Carlo Generation	67
5.2	Null Hypothesis Monte Carlo	69
5.2.1	Rejecting the Null Hypothesis	69
5.2.2	Fit for the Number of Independent Frequencies	73
5.2.3	Summary	76
5.3	Monte Carlo Sensitivity Study	77
5.3.1	Generating Periodic Monte Carlo Data Sets	77

5.3.2	Sensitivity Analysis	79
5.3.3	Sensitivity Region	84
5.3.4	Summary	85
6	Systematic Checks	87
6.1	Alternative Bin Size	88
6.1.1	Data Results	88
6.1.2	Monte Carlo Study	90
6.2	Vary Bin Start Boundary	91
6.3	Sampled Frequencies	93
6.3.1	Sampling the Data Against Alternative Frequencies	94
6.3.2	Monte Carlo Event Files	95
6.3.3	Summary	97
6.4	Combining Condition	100
6.4.1	Analysis of Combined Data Set	101
6.4.2	Monte Carlo Sensitivity Study	103
6.4.3	Summary	104
6.5	Experimental Systematics	105
6.5.1	Detector Acceptance	106
6.5.2	Background Systematics	108
6.5.3	Summary	110
7	Summary and Conclusion	111
7.1	Summary	111
7.2	Comments on Two Specific Frequencies	113
7.3	Future Prospects	115
7.4	The Last Word	116
A	Comparison of Adjusting Rates for the Combined Analysis	118
B	Detailed Sensitivity Analysis	121
B.1	D ₂ O	122
B.2	Salt	123
B.3	Combined	124
	Bibliography	125

List of Tables

1.1	Solar neutrino production reactions.	8
1.2	Best fit survival probability parameters for solutions to the solar neutrino problem (2002).	18
1.3	Temperature and radial dependence of solar neutrino production reactions.	20
3.1	Raw clock livetimes.	45
3.2	Background dead time statistics.	46
3.3	Background contributions to neutrino events.	51
4.1	Summary of variables used in periodogram calculation.	58
4.2	Statistics of the SNO data used in this analysis.	60
4.3	Data results of the regular and modified periodogram methods.	63
5.1	Probability of the null hypothesis using Monte Carlo data sets.	74
5.2	Monte Carlo fits for the number of independent frequencies.	76
5.3	Periods and amplitudes generated for the periodic Monte Carlo.	77
5.4	Probability of rejecting the null hypothesis for various periodic signals in the D ₂ O, salt, and combined phases.	86
6.1	Background contributions to the neutrino events.	110
A.1	Statistics of the combined phase analyses.	119
B.1	Probability of rejecting the null hypothesis at the 99% CL for various periodic signals in the D ₂ O phase.	122
B.2	Probability of rejecting the null hypothesis at the 99% CL for various periodic signals in the salt phase.	123

B.3	Probability of rejecting the null hypothesis at the 99% CL for various periodic signals in the combined D ₂ O and salt phases.	124
-----	---	-----

List of Figures

1.1	Layers of the Sun.	3
1.2	Sunspot number as a function of time (1750 - present).	6
1.3	Solar neutrino energy spectrum.	9
2.1	A sample distribution of event rate versus time.	36
2.2	An example of the regular and modified Lomb-Scargle periodograms.	37
2.3	Scatter plot the modified versus regular powers for the sample distribution of data.	38
3.1	Layout of the SNO detector.	41
3.2	Individual run lengths for the D ₂ O and salt phases.	47
4.1	Graphical representation of binning the SNO data.	53
4.2	One day rate versus time for the D ₂ O and salt data sets.	59
4.3	One day event rates fit to a Gaussian function.	61
4.4	Time difference between consecutive bins.	61
4.5	Lomb-Scargle periodograms for the SNO data.	64
4.6	Scatter plot of modified versus regular powers.	65
5.1	Average time between neutrino events for the salt phase data and Monte Carlo.	68
5.2	Mean event rates for 10 000 Monte Carlo data sets with no periodic behavior.	69
5.3	D ₂ O phase Monte Carlo maximum power and frequency distributions.	70
5.4	Salt phase Monte Carlo maximum power and frequency distributions.	71
5.5	Combined phase Monte Carlo maximum power and frequency distributions.	72

5.6	Combined phase Monte Carlo maximum powers on a logarithmic scale.	74
5.7	Combined phase Monte Carlo maximum powers fit to the analytic probability density function.	75
5.8	Demonstration of sinusoidal functions used in the accept/reject method of generating the periodic Monte Carlo data sets.	78
5.9	Maximum powers for a 20 day period at amplitudes of 10%, 15%, 20%, and 25%.	80
5.10	Probability of rejecting the null hypothesis for a 20 day period with amplitudes of 8%, 9%, 10%, and 11%.	81
5.11	Frequencies at the maximum powers for a 20 day period at amplitudes of 8%, 9%, 10%, and 11%.	83
5.12	Maximum power versus frequency at maximum power for 1000 Monte Carlo data sets with a period of 20 days and amplitude of 10%. . . .	84
5.13	Sensitivity contour for the combined phases.	85
6.1	Maximum power for 101 different bin sizes in the combined data sets.	89
6.2	Frequency at maximum power for 101 different bin sizes in the combined data sets.	89
6.3	Maximum power for 100 Monte Carlo data sets sampled with 101 bin sizes.	90
6.4	Data periodograms when the start bin boundary is varied by 24 hours.	92
6.5	Maximum power and frequency at maximum power for 20 bin starting boundaries in a 24 hour period.	93
6.6	Periodograms for the combined data sets when 3650, 7300, and 14 600 frequencies are sampled.	94
6.7	Data periodograms zoomed into a frequency of 0.0505 day^{-1}	96
6.8	Gaussian fits to periodogram peaks at the signal frequency for six periodic Monte Carlo data sets.	98
6.9	Distribution of the frequency widths for six types of periodic Monte Carlo.	99
6.10	Frequency peak widths versus mean frequency for six types of periodic Monte Carlo data sets.	100

6.11	One day rates for the D ₂ O and salt phases when low statistic bins are, and are not, combined.	102
6.12	Scatter of powers when low statistic bins are, and are not, combined.	103
6.13	50% contour sensitivities, for the D ₂ O phase, when low statistic bins are, and are not, combined.	105
6.14	Difference in the combined data sets powers when detector systematics are included in the weighting procedure.	108
7.1	Periodogram from the combined phases, superimposed with probabilities of rejecting the null hypothesis at the 50%, 90%, 95%, and 99% confidence level.	112
7.2	Sensitivity contours for a 50% probability of detecting a periodic signal for all three phases.	113
7.3	Periodogram from the combined data sets, zoomed into the 365.25 day period.	114
7.4	Periodogram from the combined data sets, zoomed into the 9.43 yr ⁻¹ frequency.	115
A.1	The periodograms applied to the combined data, when the rates in the D ₂ O phase are adjusted up and salt phase adjusted down.	120
B.1	Sensitivity contour for the D ₂ O phase.	122
B.2	Sensitivity contour for the salt phase.	123
B.3	Sensitivity contour for the combined D ₂ O and salt phases.	124

List of Acronyms

AV	Acrylic Vessel
CC	Charged Current
CL	Confidence Level
ES	Elastic Scattering
FAP	False Alarm Probability
GPS	Global Positioning System
LS	Lomb-Scargle, creators of the Lomb-Scargle analysis
MC	Monte Carlo
MSW	Mikeheyev - Smirnov - Wolfenstein, creators of matter-enhanced oscillation theory
NC	Neutral Current
NCD	Neutral Current Detector
PDF	Probability Density Function
PGT	Pulsed Global Trigger
PMT	PhotoMultiplier Tube
PSUP	Photomultiplier SUPport structure
(R)SFP	(Resonant) Spin Flavour Precession
SM	Standard Model of Particle Physics
SNO	Sudbury Neutrino Observatory
SNP	Solar Neutrino Problem
SSM	Standard Solar Model
UTC	Coordinated Universal Time

Chapter 1

Solar and Neutrino Physics: A Brief Historical Review

The Sun is a giant ball of power and complexity unmatched in the Solar System. As a whole, it is in a constant state of evolution. A quick glance at the Sun would lead one to believe the mechanisms are random and chaotic, but on the contrary, evidence points to precise systems that power the furious ball.

The Sun has been the subject of interest for centuries. Early civilizations praised the Sun as a god; Ancient Egyptian's named their sun-god Ra, Sumerian's gave the name Utu, and Greek's, Apolla. Through the ages fascination with the Sun can be seen in pictures, ritualistic dances, and structural pieces. For many centuries it was believed that the Sun rotated around the Earth. Copernicus (1543) and Galileo (1610) knew otherwise, and proved the heliocentric nature of the Solar System. From Aristotle to Galileo, and Kepler to Newton, great minds continued to ponder the ways of the Sun.

The last century has seen an exponential rate of scientific progression, and with

it experimental evidence to provide a backbone to solar models. Current predictions of solar models agree well with observations. However, there is a lack of experimental methods and data to confirm predictions of internal structures.

In the last four decades, scientists have been using particles called neutrinos to probe the solar interior. These neutrinos are produced through nuclear reactions, and hence the Sun is a factory, producing these particles. Due to their properties, neutrinos rarely interact with matter, and so they pass through the Sun in a matter of seconds. Detecting these solar neutrinos can provide a real-time picture of the Sun and information about the nuclear processes occurring deep in the solar core.

1.1 The Standard Solar Model

The Sun is composed of several layers, each with different properties and functions. Figure 1.1 shows the approximate layout of these layers. The core goes out to approximately 25% of the Sun's radius, and is the center for nuclear fusion reactions. The net fusion reaction combines four hydrogen atoms to form helium. The mass difference between the initial and final state atoms is carried away as energy.

The radiative zone follows, here the energy produced in the core is constantly absorbed and re-emitted, radiated by the material in this layer. This layer extends out to 85% of the solar radius. Within this layer the temperature and density decrease with increasing distance from the solar center.

The radiative zone is surrounded by the convection zone. Here, the energy is transported through convection currents. The plasma close to the radiative zone will receive the outgoing energy and consequently rise in temperature. This causes the hot plasma to move toward the surface due to a decrease in density. As the hot plasma

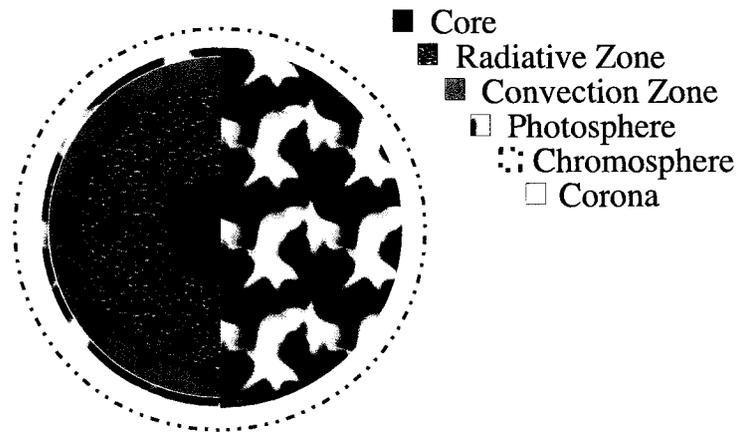


Figure 1.1: The different layers that compose the Sun.

moves towards the surface it begins to cool and become denser, which will send it back toward the center. And the process will continue again; this is convection. The convection currents carry energy away from the core and out into the photosphere.

As photons travel to the surface they are losing energy through these radiative and convection processes. This creates a shift in wavelength into visible light. Most of the visible light seen on Earth comes from the photosphere. The chromosphere and corona are the Sun's atmospheres, and are visually seen as a halo during a total solar eclipse.

The Standard Solar Model (SSM) [4] combines theory with observation to provide a quantitative description of the Sun. It relies on fundamental laws of physics such as conservation of energy and particle interactions, which govern the competing gravitational and thermal forces that maintain the Sun's equilibrium. Conservation of energy implies that the amount of energy brought into the system or produced within the system, must equal all work done by the system plus the amount of energy leaving the system. For the Sun, this translates into energy that is produced through the nuclear reactions that propagates between the different layers. The Sun maintains its

equilibrium state through a balance of gravity threatening to collapse the star, and thermal pressure pushing the molecules apart.

Several approximations are made to simplify these calculations. The Sun is assumed to be spherically symmetric. Magnetic fields and rotations complicate equations describing the motion of charged particles, and are not expected to alter the model greatly, therefore these are usually neglected in solar models. Measured quantities of temperature, luminosity, mass, size, composition, and age constrain these physical predictions and validate their predictions.

Although rotation and magnetic fields are typically neglected in the SSM these phenomena play an important role in stellar activity. One of the earliest visible indications of magnetic activity came in the form of sunspots appearing on the photosphere. These are areas of increased magnetic strength and decreased temperature, that appear as dark patches on the solar surface. These sunspots were first notably observed around 1610, simultaneously by Fabricius, Galileo, Scheiner, and Harriot. In 1843 Schwabe, a German apothecary, noted that the number of sunspots increased and decreased with an approximate 10 year periodicity [5]. This was confirmed in 1852 by Wolf, who demonstrated an 11 year cycle. Detailed records of sunspot numbers have been kept since the mid-1700's, and Figure 1.2 shows the sunspot number as a function of time. The sunspot number, $S.N.$, is calculated according to the following formula:

$$S.N. = k(10g + s), \quad (1.1)$$

where g is the number of sunspot groups, s is the total number of individual sunspots (not in a group), and k is a scaling factor that varies with observation conditions (≤ 1). This calculation is made from observing stations around the globe, from which

average values are calculated.

Sunspots have lifetimes ranging from several hours to weeks. Tracking the rotation rate of these sunspots has shown that surface layers of the Sun rotate with differential motion. The equator rotates with an approximate 25 day period, and the rotation period increases with latitude to approximately 35 days, near the poles [6]. Solar flares, abrupt releases of electromagnetic material, are correlated with solar activity, and typically occur near sunspots.

The 11 year sunspot cycle is a manifestation of a larger 22 year periodicity of the solar poles. With the same 11 year sunspot period, the magnetic poles of the Sun change from north to south, requiring a total of 22 years for the complete global magnetic pole cycle.

Helioseismology is a field of experimental astrophysics that probes the internal structure of the Sun. Experiments observe periodic waves that appear on the photosphere, shifting the observed wavelengths. The waves are believed to be caused by turbulent activity in the convection zone. They act like acoustic waves, with areas of high and low pressure, oscillating inside the convection zone due to the large density gradient at the boundaries. These acoustic waves depend upon internal temperature and density. Since pressure is the restoring force that acts to maintain the oscillations, these waves are called p-waves. Typical p-modes have periods of several minutes and a lifetime of several days.

Scientists now use helioseismology to examine internal structures, like rotation, through their effects on the oscillations. Helioseismology has confirmed the surface differential rotation, and data suggests that the differential rotation continues through the convection zone, varying with radius, until the tachocline, which separates the convection and radiative zones, where this differential rotation stops, leaving the

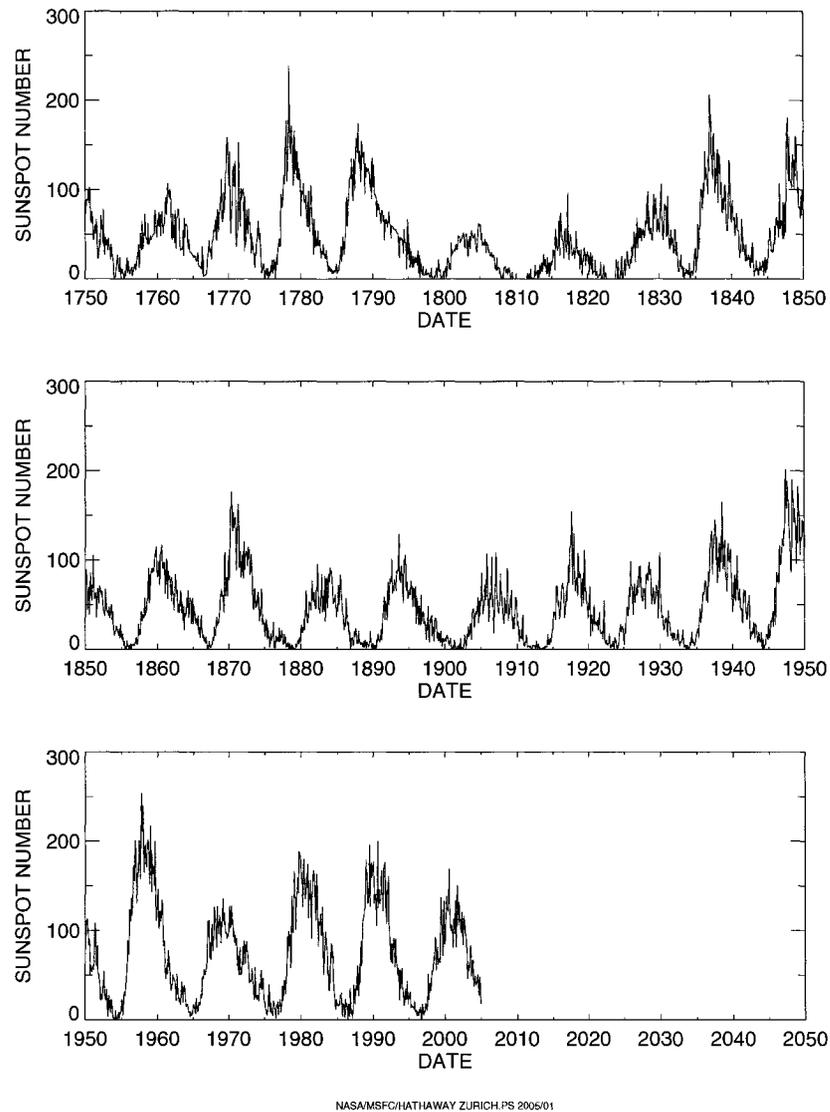


Figure 1.2: Monthly averaged sunspot number versus time from 1750 to present. The 11 year periodic nature of sunspots can be seen from this figure, taken from [7].

radiative zone to move as one piece [8].

Another form of oscillatory motion is postulated to exist deep inside the Sun. This form of vibration is called g-mode oscillation, and is the result of gravitational restoring forces acting against density changes. These are thought to be similar to p-modes, but with longer period oscillations. To date, these g-modes have not been detected.

Various solar processes appear to have a periodic tendency. But to what extent do these rotations and periodic magnetic fields influence the rate of nuclear reactions, and do these periodic tendencies affect the produced neutrinos? Understanding neutrino properties, and observing their solar rates may help answer these questions.

1.2 Neutrino Properties

Neutrinos are neutrally charged, weakly interacting leptons. There are three known active neutrino flavours, ν_e, ν_μ, ν_τ , partners to each of the three known charged leptons. According to the Standard Model (SM) of Particle Physics neutrinos are massless, hence under this model all neutrinos have a left-handed chirality, and all antineutrinos are in the right-handed state.

As previously mentioned, the net solar fusion reaction fuses four hydrogen atoms to create helium, along with positrons, neutrinos, and energy. This net reaction is achieved through a chain of intermediate nuclear reactions. Within this chain there are several neutrino-producing reactions, all of which produce only one neutrino type; this is the electron neutrino, ν_e . Only the ν_e can be produced due to energy and lepton number conservation laws.

Table 1.1 lists the nuclear reactions in the so-called ‘p-p chain’, and associated

Reaction	Termination (%)	ν_e Flux ($10^{10} \text{ cm}^{-2} \text{ s}^{-1}$)	
$p + p \longrightarrow {}^2\text{H} + e^+ + \nu_e$	100	5.95	(<i>pp</i>)
or			
$p + e^- + p \longrightarrow {}^2\text{H} + \nu_e$	0.4	1.40×10^{-2}	(<i>pep</i>)
${}^2\text{H} + p \longrightarrow {}^3\text{He} + \gamma$	100		
${}^3\text{He} + {}^3\text{He} \longrightarrow \alpha + 2p$	85		
or			
${}^3\text{He} + {}^4\text{He} \longrightarrow {}^7\text{Be} + \gamma$	15		
${}^7\text{Be} + e^- \longrightarrow {}^7\text{Li} + \nu_e$	15	4.77×10^{-1}	(<i>⁷Be</i>)
${}^7\text{Li} + p \longrightarrow 2\alpha$	15		
or			
${}^7\text{Be} + p \longrightarrow {}^8\text{B} + \gamma$	0.02		
${}^8\text{B} \longrightarrow {}^8\text{Be}^* + e^+ + \nu_e$	0.02	5.05×10^{-4}	(<i>⁸B</i>)
${}^8\text{Be}^* \longrightarrow 2\alpha$	0.02		
or			
${}^3\text{He} + p \longrightarrow {}^4\text{He} + e^+ + \nu_e$	0.00002	9.3×10^{-7}	(<i>hep</i>)

Table 1.1: The solar fusion reactions that produce ν_e . Termination is the percentage of time each reaction will occur for every iteration of the p-p chain, taken from [4]. The flux values are as seen on Earth, and have been taken from [9]. The abbreviations in soft brackets are the common terms used to refer to each ν_e reaction.

theoretical ν_e production rates. These rates are calculated using experimental values for solar parameters, such as temperature, radius, age, chemical composition, and theoretical models of weak interactions. The so-called pp reaction occurs every time the p-p chain is initiated, with each reaction producing, what is referred to as a pp neutrino.

The different reactions listed in Table 1.1 produce ν_e with a range of energy values. Figure 1.3 shows the energy distribution for the various reactions. These energy values are based on the Coulomb barrier necessary to overcome the repulsion forces of the initiating elements. For example, ${}^8\text{B}$ neutrinos are produced with a range of possible

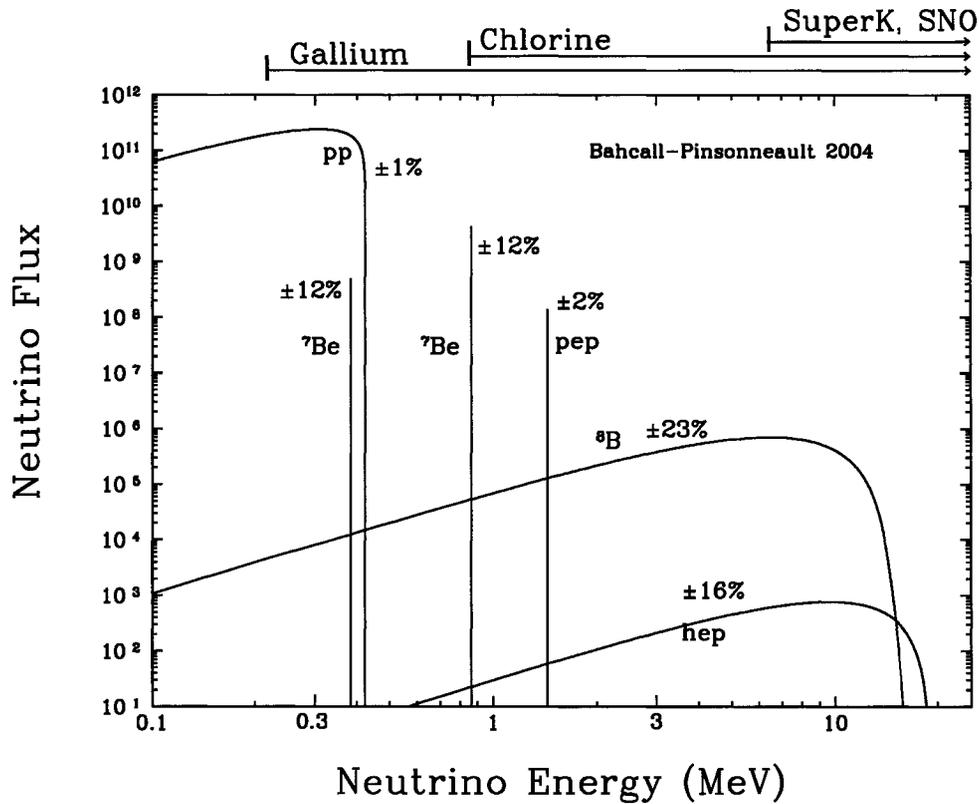


Figure 1.3: Energy spectrum for the various solar neutrino production reactions shown in Table 1.1, [10]. The neutrino flux at the Earth's surface is in units of $\text{cm}^{-2} \text{s}^{-1}$. The four labels on the top axis (Gallium, Chlorine, SuperK, SNO) are references to solar neutrino experiments, which are briefly described in Section 1.3.

energies up to approximately 16 MeV. Since neutrinos are weakly interacting and near massless, they do not lose energy as they propagate through the solar medium.

Due to weak interactions neutrinos produced near the core of the Sun can travel to the surface in a few seconds, compared to the several thousands of years it can take the generated energy to travel the same distance. Observations of ν_e rates and associated energies can provide a snapshot of the Sun's interior, and help confirm solar fusion rates and elemental abundances.

1.2.1 The Solar Neutrino Problem and Solutions

In the 1960's, Davis and Bahcall designed an experiment to measure solar ν_e rates to infer solar properties. In 1968, the first results revealed that the ν_e rate was a factor of two to three times less than that predicted by solar models. This implied that there was either an error with the experiment, or the solar model, or that the SM incorrectly described neutrino properties. In subsequent experiments these results were verified. The SSM had support from other solar observables, such as solar activity and luminosity. The scientific community began to question the validity of the SM description of neutrinos. And so began the solar neutrino problem (SNP).

Theorists jumped to the occasion, and began postulating modifications to the SM to account for the neutrino deficit. Two of the many theories included a matter-enhanced neutrino flavour transformation, and a neutrino magnetic moment spin flip. Both theories require the neutrino to have a non-zero mass. These two theories are discussed below.

The basis for these models is that the produced electron neutrino can change into another state due to some physical mechanism. In both cases, the evolution of the neutrino state is described by a Hamiltonian, that accounts for these state-altering interactions. By solving the Schrodinger equation for this Hamiltonian, the probability for an electron neutrino to remain as an electron neutrino is described by a survival probability.

Matter induced neutrino flavour transformations The first theory is now called the MSW effect, named after Mikeheyev, Smirnov, and Wolfenstein [11, 12], who developed the theory. Although not in the acronym, Pontecorvo in 1957, independently postulated this effect [13]. The MSW effect describes neutrino flavour

oscillations via composite mass eigenstates that oscillate due to interactions with matter and propagation through the vacuum. In the most basic model, there are two mass eigenstates, ν_1 and ν_2 , that compose two flavours states, ν_e and ν_x , where ν_x is some linear combination of the ν_μ and ν_τ .

The survival probability, P_e , of a solar ν_e produced at time t to remain a ν_e at a later time t' is given by:

$$P(\nu_e \rightarrow \nu_e) = \frac{1}{2} + \left(\frac{1}{2} - P_j\right) \cos 2\theta_M \cos 2\theta_V, \quad (1.2)$$

$$\text{where } P_j = \exp \left[-\frac{\pi \Delta m^2}{4E} \left(\frac{\sin^2 2\theta_V}{\cos 2\theta_V} \right) \left(\frac{n_e}{dn_e/dr} \right) \right] \quad (1.3)$$

is the jump probability that assumes a linearly changing electron density n_e with solar radius r [4]. The other parameters in these equations are θ_M , the matter mixing angle that depends upon the vacuum mixing angle θ_V , the ν_e energy E (eV), and the mass squared difference between the mass eigenstates, $\Delta m^2 = m_2^2 - m_1^2$ (eV²). To date, this is the favoured solution, and experimental evidence is proving the validity of this theory.

Neutrino spin flip precession The second theory was first proposed by Cisneros in 1970 [14]. Particles couple to the weak interaction via their left-handed component. This means that all solar neutrinos, which are near massless and moving near the speed of light, produced via the weak interaction, are (primarily) left-handed. However, if neutrinos have mass and a large enough magnetic moment, a large magnetic field could flip the magnetic moment, causing the spin and momentum to be aligned, a right-handed state. This could explain the observed neutrino deficit, since a right-handed neutrino essentially, does not interact with matter. In the late 1980's the theory was

extended to combine the spin flip and neutrino flavour conversion, allowing for flavor-off-diagonal transition magnetic moments. Under this spin flavour precession (SFP) model it has been shown that the neutrino rate would be anticorrelated with solar activity [15–18].

In a two flavor model the probability for ν_e^L to transform into $\overline{\nu_x^R}$ is given by

$$P(\nu_e^L \rightarrow \overline{\nu_x^R}) = \frac{(2\mu B)^2}{D} \sin^2 \left(\frac{1}{2} r \sqrt{D} \right) \quad (1.4)$$

where r is the radial distance and

$$D = (\Delta m^2/2E)^2 + (2\mu B)^2, \quad (1.5)$$

B [T] is the component of the Sun's magnetic field perpendicular to the motion of the neutrino, and μ [eV/T] is the transition magnetic moment [19]. The Δm^2 and E are the same as the MSW case. The theory was problematic, for if $\Delta m^2/2E$ is large compared to $2\mu B$ then the probability amplitude for transition is suppressed, and can not describe the observed neutrino deficit.

An extension to SFP is resonant spin flavour precession (RSFP), a theory that includes matter interactions with the left and right-handed states. There is an additional term that takes into account the neutrino interactions with the electron and neutron components of the solar material. This term gives rise to a resonance condition that enhances the transition probability, which could explain the SNP [20,21]. In a two flavour model, with ν_e and ν_x , Dirac neutrinos could undergo $\nu_e^L \rightarrow \nu_x^R$, and for Majorana neutrinos, assuming lepton number is not conserved, $\nu_e^L \rightarrow \overline{\nu_x^R}$. For RSFP,

the probability for transition, Equation (1.4), is altered such that D is now given by

$$D = (\Delta m^2/2E - \sqrt{2}G_F N_{\text{eff}})^2 + (2\mu B)^2, \quad (1.6)$$

where N_{eff} depends upon the neutrino type, Majorana or Dirac, and is the effective solar electron and neutron number density,

$$\begin{aligned} N_{\text{eff}} &= N_e - N_n/2 \quad \text{Dirac} \\ &= N_e - N_n \quad \text{Majorana,} \end{aligned} \quad (1.7)$$

and G_F is the Fermi constant [19]. The probability amplitude now includes a term related to the density of the matter, and can be constructed such that the amplitude is maximized. For the survival probability to be decreased the argument of the sine function must be large. This sets lower limits on the product μB .

The theory implies that if the neutrino had a large enough magnetic moment, strong magnetic fields could reduce the ν_e^L rate emerging from the Sun.

Both the MSW and SFP-type solutions are similar in mathematical formulation, and in many of their predicted observables, such as energy dependence of the ν_e^L survival probability. One key difference is that SFP-type solutions rely on the presence of a strong solar magnetic field, and a large neutrino magnetic moment. Although knowledge of the interior solar magnetic activity is limited, there is evidence supporting an 11 year periodic solar cycle, in which the magnetic field strength increases and decreases. If an SFP-type solution is correct, and the product μB large enough, then the rate of detected ν_e^L could vary with this solar cycle. The ν_e^L rate could also vary with other solar magnetic rotational periods, such as the inclination of the Earth with

respect to the Sun, since magnetic field strength varies from the equator. Or it could vary with the solar equatorial rotation period of 25 days.

A periodic neutrino rate could suggest a spin flavour flip solution and reveal properties of the internal solar magnetic behavior. With this in mind, solar neutrino experiments began analyzing their data to search for periodic variations in the solar neutrino rates.

1.3 Decades of Solar Neutrino Experiments and Periodicity Searches

The first solar neutrino experiment designed by Davis and Bahcall ran from the late 1960's to 1995 in the Homestake mine, South Dakota. The detector was filled with 615 tons of perchloroethylene, C_2Cl_4 , a cleaning fluid, that was sensitive to neutrinos with energies of 0.814 MeV or greater, through the reaction:



Over 25 years 108 runs were completed. Many periodicity analyses were performed on the long data set to look for correlations between ν_e rates and solar activity. Some authors have claimed correlative behavior for various parts of the data set [22–25]. Others have argued that the statistical methods used were inappropriate, or that only a few points pull the correlation, which may unintentionally induce an apparent correlation [26–31]. There has been no definitive conclusion on whether or not the Homestake solar neutrino data displays any periodic behavior.

Gallex and GNO were other radiochemical experiments that ran from 1992 to

1997, and 1998 to 2003, respectively, in the Laboratori Nazionali del Gran Sasso, Italy. These experiments detected solar neutrinos in a 30 ton gallium tank through the reaction:



With a threshold of 0.233 MeV, the detector was sensitive to pp, pep, ${}^7\text{Be}$, ${}^8\text{B}$, and hep neutrinos. The combined experiments had a total of 123 runs. In June 2004, Pandola, a member of the GNO Collaboration, applied two different methods to the full data set to search for a general periodic signal, with periods in the range of 14 days to 25 years. He used the Lomb-Scargle method, which is described in detail in Chapter 2, and a maximum likelihood technique. He concluded that both methods were consistent, indicating a constant neutrino rate, but based on MC simulations his methods were not sensitive to signals with periods less than 30 days [32]. A similar analysis was performed by Sturrock and Weber; they used the Lomb-Scargle method and applied it to the full Gallex data and limited GNO data. They found the neutrino data displayed a 28 day periodicity, possibly associated with rotations in the Sun's convection zone [33].

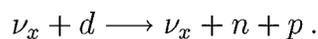
Super-Kamiokande is a relatively new solar neutrino experiment located in the Mozumi mine, Kamioka-cho, Yoshiki-gun, Gifu-ken, Japan. This group has been actively detecting neutrinos since 1996 using 50 kilotons of water. The solar neutrinos are detected via Čerenkov light that is generated by fast moving electrons after being elastically scattered off incoming neutrinos via the reaction:



This reaction is sensitive to all neutrino flavours, $x = e, \mu, \tau$, but is on the order of

six times more sensitive to ν_e due to couplings with both the Z and W bosons. The energy threshold for this experiment is 5 MeV, which makes it sensitive only to neutrinos originating from the ${}^8\text{B}$ and hep solar reactions. Neutrino reactions cannot be determined on an event by event basis because of the large number of backgrounds, instead the signal is distinguished from the background through the distinctive directional quality this elastic interaction bears. In 2003, the Super-Kamiokande collaboration released the results of a search for periodic time variations in their solar neutrino data. Data were binned into 5-day and 10-day intervals, and using the statistical separation of signal from background found an associated rate in each bin. They used the Lomb-Scargle method to search for periodic behavior, and found the data set to be consistent with statistical fluctuations [34]. The data were made public and subsequently similar analyses have been performed, some claiming the neutrino data displays a periodicity of 9.43 yr^{-1} [35–38]. Questions of statistical methods and significance tests used in all of these analyses, has left no definitive conclusion on the time variability of the Super-Kamiokande data set.

The Sudbury Neutrino Observatory (SNO) began detecting solar ${}^8\text{B}$ neutrinos in 1999 from the Creighton mine in Sudbury, Ontario, Canada. This detector uses heavy water as its detecting medium, which provides interactions that are equally sensitive to neutrinos of all flavours, through a neutral current interaction,



Details of this detector are described in Chapter 3. SNO has the unique ability to be able to determine individual neutrino interactions with a high degree of certainty because of the ultra-clean and low background environment. In this experiment the

exact time of every neutrino interactions is known, so the rate can be examined for periodic behavior. This thesis presents the results of the periodicity analysis on SNO's data using the Lomb-Scargle method. At the time of this thesis, the SNO collaboration has submitted a paper for publication [3] reviewing the periodicity analyses that have been applied to its data set.

1.4 Spin Flavour Precession Revisited

In 2001, SNO released results which revealed that the total number of all active solar neutrino flavors observed via the neutral current reaction was approximately equal to the expected rate of ν_e produced in the Sun and that the flux of ν_e from the CC reaction was roughly 1/3 that of the total flux [39]. The results imply that solar ν_e^L are changing into detectable neutrinos, $\nu_x = \nu_\alpha^L$ or $\overline{\nu}_x^R$, where $\alpha = e, \mu, \tau$. If SFP was correct, this would rule out a Dirac-type neutrino which would require the conversion of $\nu_e^L \rightarrow \nu_e^R$, in which SNO's neutral current reaction would still show a neutrino deficit. For SFP to explain the SNP, neutrinos would have to be of Majorana type, undergoing $\nu_e^L \rightarrow \overline{\nu}_e^R$ where the resulting antineutrinos would be actively detected in the Čerenkov detector. However, in 2004 the SNO collaboration released a paper which set an upper limit on the flux of $\overline{\nu}_e^R$, stemming from only two candidate antineutrino events observed in 312.9 live days [40].

With so much solar neutrino data available in 2002, Barranco *et al.*, used the MSW and RSFP survival probabilities, Equation (1.2) and Equation (1.4) to fit for the unknown parameters in both cases [41]. The RSFP survival probability requires values for the product of the neutrino magnetic moment and the solar magnetic field, μB . Based on experimental evidence they fixed these values to optimal choices,

Solution	$\Delta m^2(\text{eV}^2)$	$\tan^2 \theta$	χ^2 Goodness of Fit
MSW	5.6×10^{-5}	0.47	78%
RSFP	7.9×10^{-9}	4.2×10^{-4}	83%

Table 1.2: Using the survival probability formulas, Equation (1.2) and Equation 1.4, J. Barranco *et al.* fit for the unknown parameters using all available data in 2002 [41].

$\mu = 10^{-11} \mu_B$ and $B = 80$ kGauss. The data used in these fits were from the Homestake, and gallium (SAGE, Gallex/GNO) experiments, along with the 1496 Super-K days, and SNO's results [39, 42]. The results of this fit are shown in Table 1.2. The MSW solution favours the large mixing angle with a large Δm^2 , the so-called LMA solution, but for RSFP the Δm^2 fit is on the order of four times smaller. In 2002, the available data allowed both solutions with good fit probabilities.

At the beginning of 2002 a new neutrino experiment began; KamLAND a liquid scintillator detector which began measuring electron antineutrino, $\bar{\nu}_e$, rates from reactor neutrinos via the reaction:



with sensitivity to $\bar{\nu}_e$ with energies greater than 1.8 MeV. This group released its first results in December 2002 [43]. These results reduced the allowed regions of the $\Delta m^2 - \theta$ plane to include only the LMA region, areas that render any dominant SFP solution unlikely, giving further merit to the MSW solution.

Further evidence against an SFP-type solution comes from the general SM neutrino magnetic moment and mass relationship,

$$\mu \simeq 3.2 \times 10^{-19} \left(\frac{m_\nu}{1 \text{ eV}} \right) \mu_B, \quad (1.12)$$

where μ_B is the electron Bohr magneton with units of eV/T, and m_ν is the mass of the neutrino in units of eV. Current limits on the neutrino mass (< 2.2 eV) leave $\mu_\nu = 10^{-18} \mu_B$ [44], but constraints on the solar magnetic field cause SFP theories to require μ_ν to be on the order of $10^{-11} \mu_B$. However, there are alternative mass and magnetic moment dependencies for some non-SM theories.

Experimental evidence indicates that the MSW is likely the dominant mechanism accounting for the SNP. However, some authors suggest SFP could still be a subdominant process, that could produce a time varying neutrino flux [45].

Despite the unlikely dominant SFP mechanism, there are other processes which could give rise to a periodic solar neutrino rate, these are discussed in the next section.

1.5 Motivations to Search for a Periodic Solar Neutrino Rate

The solar neutrino rate could be periodic if either (a) the nuclear processes which produce the neutrinos vary in time, or (b) neutrinos have non-standard interactions with solar material that cause them to change into undetectable states with some periodicity. There is also the possibility that (b) proceeds such that a ν_e converts to another active detectable state. In this case, the total neutrino rate would be constant, only the ν_e rate would vary in time.

With respect to the former case (a), nuclear fusion rates depend upon the temperature, pressure and elemental abundances in the surrounding medium. The reactions listed in Table 1.1 have varying temperature dependences. Bahcall and Ulmer used stellar evolution code, based on SSM, to simulate solar reactions in the Sun, and then calculate the expected neutrino flux on Earth. These simulations were used to

Reaction	Exponential Temperature Dependence, m	Radial Production Region, R_{\odot}
pp	-1.1 ± 0.1	0.04 - 0.17
pep	-2.4 ± 0.9	unknown
${}^7\text{Be}$	10 ± 2	0.03 - 0.10
${}^8\text{B}$	24 ± 5	0.02 - 0.07
hep	unknown	0.06 - 0.21

Table 1.3: Characteristics of the solar neutrino production reactions. The exponential temperature dependence takes the form T^m , where the fit for m is shown, this has been taken from [46]. The radial production ranges are half peak values, they have been taken from [4].

fit for the temperature dependence of each neutrino reaction, Table 1.3 shows these results [46]. Neutrinos produced by the ${}^8\text{B}$ and ${}^7\text{Be}$ reactions depend greatly on the temperature, so if the temperature varies in time due to internal rotations, magnetic activity, or some other mechanism, then the neutrino rate from these reactions would also vary with time. For the SNO experiment, which predominantly detects ${}^8\text{B}$ neutrinos, this variation would be evident in all reactions, since the total rate of all active neutrino flavours would have the same period.

Periodic behavior in the density and pressure could result from solar rotations and magnetic fields which could also cause the nuclear reactions to vary in time. Table 1.3 also lists the production regions for each reaction. The ${}^8\text{B}$ neutrinos are produced very close to the core (within 7% of the solar center), and so density and pressure variations are unlikely in this central region. However, hep and pp neutrinos, which are produced at a wide range of radii within the Sun, could experience pressure and density variations that could ultimately affect their nuclear reaction rate. Up to about $0.75R_{\odot}$ the solar interior is thought to rotate as one piece. If the distribution of material is not uniform, possibly due to magnetic fields, this could cause neutrino

rates to vary with a rotational period of the producing layers. This effect could cause the different neutrino channels to have slightly different production rates, constant versus periodic in time, depending on the production region. Experimentally, data would need to be separated into the various production reactions and periodicity analyses applied to each channel.

In [47], Burgess examined the effects of a fluctuating solar medium on the survival probability of ν_e , assuming the MSW solution. He found that if the correlation length of the density fluctuations is approximately equal to the length necessary for neutrino oscillations to occur, and if the amplitude of the fluctuation is at least 1%, then the survival probability can be affected by the density changing fluctuations. When time varying fluctuations are taken into account the MSW ν_e survival probability, previously shown in Equation (1.2), becomes:

$$P(\nu_e \rightarrow \nu_e) = \frac{1}{2} + \left(\frac{1}{2} - P_j\right) \lambda \cos 2\theta_M \cos 2\theta_V, \quad (1.13)$$

The fluctuation effects are contained in the parameter λ ,

$$\lambda = \exp \left[-2G_F^2 \int_t^t d\tau \alpha(\tau) \sin^2 2\theta_M(\tau) \right], \quad (1.14)$$

here $\alpha(\tau)$ describes the average density changes as a function of time, τ , and G_F is again the Fermi constant. The rate of electron neutrinos could be periodic as a result of this mechanism, if the density in the region of oscillation is periodic. However, with this model the total number of solar neutrinos from all active flavors would be constant in time, since the MSW mechanism is changing ν_e into other active flavors.

The Earth travels in an elliptic plane around the Sun, where the Sun is at one focal point of the ellipse. This means that the Earth-Sun distance varies with a yearly

period. Since the flux of any material originating from the Sun as measured on Earth varies with the squared distance between the two celestial objects, there is an expected natural yearly variation of the neutrino flux. The eccentricity of the Earth is 1.7%, and so the difference in neutrino flux between the perihelion and aphelion is expected to be 7%.

1.6 Synopsis of the Thesis

This chapter has briefly introduced solar and neutrino physics, and presented various mechanisms that could induce a periodic solar neutrino rate. The remainder of this thesis will examine the ^8B solar neutrino rate, as measured by the Sudbury Neutrino Observatory (SNO), to determine if the rate is periodic in time. Chapter 2 sets up the mathematical framework to search for sinusoidal periodic behavior in the given data set. The SNO detector and data acquisition system are described in Chapter 3. A description of how the mathematical method is applied to the SNO data is given in Chapter 4. A Monte Carlo (MC) analysis is employed in Chapter 5 to quantify the results. Chapter 6 describes some systematic tests that were performed to ensure an unbiased and robust method. The thesis ends with a summary and some final thoughts in Chapter 7.

The presence of a periodic solar neutrino rate could imply new properties of the neutrino and/or point to new information about the solar interior. If solar neutrino experiments show that data are consistent with a constant rate, this puts constraints on non-standard theories expecting such an effect, some of which have been described above.

Chapter 2

Searching for a Periodic Signal

In many areas of research data are analyzed to determine if they have some periodic tendency. This is an important study in fields such as astrophysics when analyzing the motion of distant stars, and medicine, when measuring patient heart rates. Many mathematical methods have been developed to find the period, if any, of a signal in a data set. If a period is found, there are recipes to determine the likelihood that this is indeed the result of a periodic signal, and not just the result of random fluctuations in the data.

2.1 General Introduction to Fourier Analysis

Fourier analysis is the method most commonly used to search for a frequency in a time distributed signal. It takes data, measured in the time domain, and transforms the data into the frequency domain. This method is based on the principle that any periodic function, $f(t)$, with period T , which is absolutely integrable over one period ($\int_0^T f(t)dt < \infty$), can be expressed as a linear combination of complex exponential

functions:

$$f(t) = \sum_{n=-\infty}^{\infty} F(n\omega)e^{in\omega t} \quad (2.1)$$

where i is $\sqrt{-1}$, n is an integer, t is time, $\omega = 2\pi/T$ is the angular frequency, and $F(n\omega)$ is the complex Fourier coefficient. Euler's identity states that a complex exponential can be expanded as follows:

$$e^{\pm in\omega t} = \cos(n\omega t) \pm i \sin(n\omega t). \quad (2.2)$$

Hence, the periodic function $f(t)$ can be expressed as an infinite sum of cosine and sine functions. These complex exponentials form a complete and orthogonal basis for all periodic functions, since:

$$\frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} e^{in\omega t} e^{-im\omega t} dt = \begin{cases} 0 & \text{if } n \neq m, \\ 1 & \text{if } n = m. \end{cases} \quad (2.3)$$

As shown by Morrison in [48], multiplying both sides of equation (2.1) by $e^{-im\omega t}/T$, integrating over one period, T , and using the orthogonality relation in Equation (2.3), yields the solution for each of the n Fourier coefficients:

$$F(n\omega) = \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} f(t)e^{-in\omega t} dt. \quad (2.4)$$

This is called the Fourier transform of $f(t)$. In general, these coefficients are complex numbers, defined by an amplitude and a phase. The coefficients determine the contribution of each integer multiple of frequency, $n\omega$, to the signal, and hence define the Fourier series expansion of the observed signal.

The Fourier coefficients also determine the *power* at a given frequency in the signal. The power of a frequency is defined as the square of the absolute value of the Fourier coefficient at that frequency,

$$P(n\omega) = |F(n\omega)|^2. \quad (2.5)$$

It is often necessary to examine a continuous range of frequencies, rather than just integer multiples of one frequency, $n\omega$. In the limit that ω , the value between consecutive sampled frequencies, goes to zero, the period goes to infinity, *i.e.*, in the limit,

$$\omega \rightarrow \delta\omega \quad (2.6)$$

$$\text{and } T \rightarrow \infty, \quad (2.7)$$

$$\text{then let } n\omega \rightarrow \omega_n. \quad (2.8)$$

Now ω_n is the exact frequency being sampled. In this continuous spectrum of sampled frequencies, the discrete sum in Equation (2.1) becomes an integral,

$$f(t) = \int_{-\infty}^{\infty} F(\omega_n) e^{i\omega_n t} d\omega_n. \quad (2.9)$$

To find the Fourier coefficients, $F(\omega_n)$, Equation (2.9) is multiplied by $\int_{-\infty}^{\infty} e^{-i\omega_n t} dt$.

Using properties of the Dirac delta function, it can be shown that

$$F(\omega_n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-i\omega_n t} dt. \quad (2.10)$$

This is called the Fourier integral representation of $f(t)$.

2.2 Sampling Data at Even Time Intervals

Now consider the case when an unknown periodic signal, $f(t)$, is sampled at regular time intervals, $f(t_1), f(t_2), \dots, f(t_N), \dots$, where the constant time interval is $\Delta t = t_{i+1} - t_i$. To extract the frequency of this signal from the sampled data, a range of discrete frequencies must be sampled, and the associated Fourier coefficients analyzed. A plot of $|F(\omega_n)|$ vs ω_n would show the contribution of each sampled frequency to the periodic function.

There is a maximum frequency that can be sampled, or in other words, a minimum period that can be analyzed, related to the time interval between the sampled data, Δt . For example, if an experiment takes a measurement of the Earth's temperature at a time interval of 20 days, it is impossible to determine the contribution of an hourly temperature fluctuation to the signal. In terms of frequency, the maximum frequency that can be accurately sampled is called the Nyquist frequency, ν_C , and is defined as:

$$\nu_C = \frac{1}{2\Delta t}, \quad (2.11)$$

where $\nu = \omega/2\pi$. The sampling theorem states that if the true frequency of the signal is less than the Nyquist frequency, then the periodic function $f(t)$ can be defined. However, if the signal contains frequencies greater than the Nyquist frequency, then aliasing is likely to occur [48]. Aliasing occurs when powers from 'high' frequencies ($> \nu_C$) leak into the powers at the 'low' frequencies ($< \nu_C$), causing a distortion of the Fourier coefficients at these lower frequencies. Aliasing is a result of a lack of information about the data due to the gaps of time between the limited data points.

There is a direct relationship between the signal frequency, ν_{signal} , and the resulting

aliased frequency,

$$\nu_{\text{alias}} = n\nu_{\text{d.s.}} \pm \nu_{\text{signal}}, \quad (2.12)$$

where $\nu_{\text{d.s.}}$ is the frequency of the data spacing $= 1/\Delta t$, and n is an integer. For example, if data are sampled every day, then the Nyquist frequency corresponds to 0.5 day^{-1} . If the data has a signal frequency of 0.8 day^{-1} , then an aliased frequency will occur at 0.2 day^{-1} . To discriminate the aliased frequency from the signal frequency, the data can be re-analyzed with an alternate data spacing. The true signal frequency will have a large power for both spacing choices.

There are N discrete frequencies that can be sampled for evenly spaced data corresponding to

$$\nu_k = \frac{k}{2N\Delta t} \quad \text{where } k = 1, 2, \dots, N, \quad (2.13)$$

where N is the number of data points describing the signal $f(t)$. The first frequency sampled corresponds to a period twice the length of the data span, $T = n\Delta t$. And the last frequency sampled $k = N$ corresponds to the Nyquist frequency, $\nu_{k=N} = \nu_C$.

The discrete Fourier transform is used to estimate the Fourier coefficients from a limited number of measurements taken at evenly spaced discrete time intervals. As before, the signal $f(t)$ can be written as a Fourier expansion of each of the N sampled frequencies,

$$f(t_p) = \sum_{k=1}^N F(\omega_k) e^{i\omega_k t_p} \quad \text{for } (1 \leq p \leq N), \quad (2.14)$$

$F(\omega_k)$ are the corresponding Fourier coefficients, and are found by multiplying the above equation by $\sum_{m=1}^N e^{-i\omega_k t_m}$, and orthogonality relations to give

$$F(\omega_k) = \frac{1}{N} \sum_{p=1}^N f(t_p) e^{-i\omega_k t_p} \quad \text{for } (1 \leq k \leq N). \quad (2.15)$$

As defined in Section 2.1, the power of a periodic function at a frequency ω is defined as the square of the Fourier coefficient,

$$P(\omega_k) = |F(\omega_k)|^2. \quad (2.16)$$

When data are sampled at discrete times, the power spectrum is no longer a smooth, continuous function, instead it has values only at the sampled frequency values, ω_k . For the discrete Fourier transform a plot of the powers is called a periodogram of the sampled data. For a given $\omega_k = \omega_0$, a large value of $P(\omega_0)$ in the absence of other similarly large, or larger, powers at other frequencies, could indicate that the data are periodic with a frequency of $\nu_0 = \omega_0/2\pi$.

2.3 Sampling Data at Uneven Time Intervals

When the sample time between data points is not constant, the data are said to be sampled at uneven time intervals. The solar neutrino data analyzed in this thesis are unevenly spaced in time, hence to determine the presence of a periodic signal in this data set, the methods described above must be re-examined. In 1975, Lomb [49] developed a method to determine the power of a frequency in a periodic signal for unevenly spaced data. This method was updated by Scargle [50, 51] in 1982, where he examined the significance of the periodogram results. The following sub-sections outline this theory of finding periodicities in unevenly spaced data.

2.3.1 Lomb-Scargle Method

The fundamental idea of finding a periodicity in a sampled data set, $x(t_i)$, is to fit the data to cosine and sine functions, and find the frequency which best fits the data. Lomb used a least squares approach to fit a function, $f(t_i)$, with angular frequency ω to the data,

$$f(t_i) = a \cos(\omega t_i) + b \sin(\omega t_i), \quad (2.17)$$

where a and b are coefficients to be determined. The least squares method finds a and b such that the term

$$L(\omega) = \sum_{i=1}^N [x(t_i) - f(t_i)]^2 \quad (2.18)$$

is minimized. A minimum occurs when $\partial L(\omega)/\partial a = 0$ and $\partial L(\omega)/\partial b = 0$. Performing these partial derivatives on $L(\omega)$ leads to the following set of equations:

$$a \sum_{i=1}^N \cos^2(\omega t_i) + b \sum_{i=1}^N \cos(\omega t_i) \sin(\omega t_i) = \sum_{i=1}^N x(t_i) \cos(\omega t_i) \quad (2.19)$$

$$a \sum_{i=1}^N \cos(\omega t_i) \sin(\omega t_i) + b \sum_{i=1}^N \sin^2(\omega t_i) = \sum_{i=1}^N x(t_i) \sin(\omega t_i). \quad (2.20)$$

These two equations are used to solve for the two unknown coefficients, a and b .

To simplify the equations, t_i is replaced by $(t_i - \tau)$, and a solution is found for the constant τ such that

$$\sum_{i=1}^N \cos \omega(t_i - \tau) \sin \omega(t_i - \tau) = 0. \quad (2.21)$$

By use of trigonometric identities this simplifies to:

$$\tan(2\omega\tau) = \frac{\sum_{i=1}^N \sin(2\omega t_i)}{\sum_{i=1}^N \cos(2\omega t_i)}. \quad (2.22)$$

The introduction of τ allows for time translational invariance of the periodogram. Because τ is a constant, dependent only upon sine and cosine terms of the angular frequency and time, if all the time values, t_i are shifted by $t_i + T_0$, then τ becomes $\tau + T_0$ and the sinusoidal arguments $\omega((t_i + T_0) - (\tau + T_0))$ reduce to $\omega(t_i - \tau)$, yet again, [50]. So the periodogram is independent of any phase shift.

Lomb used a reduction in the sum of squares, defined as:

$$\Delta L(\omega) = \left[\sum_{i=1}^N x(t_i) \right]^2 - L(\omega) \quad (2.23)$$

to find a solution for ω which *maximizes* $\Delta L(\omega)$, and is equivalent to finding a minimum solution for $L(\omega)$ in Equation (2.18).

Using Equation (2.22) and some simplifications, Equation (2.23) reduces to

$$\Delta L(\omega) = \frac{\left[\sum_{i=1}^N x(t_i) \cos \omega(t_i - \tau) \right]^2}{\sum_{i=1}^N \cos^2 \omega(t_i - \tau)} + \frac{\left[\sum_{i=1}^N x(t_i) \sin \omega(t_i - \tau) \right]^2}{\sum_{i=1}^N \sin^2 \omega(t_i - \tau)} \quad (2.24)$$

The user provides a value, or a range of values, for ω . $\Delta L(\omega)$ is treated as a power of the data at the given ω . Through this approach Lomb developed a method for finding periodic behavior in a data set sampled unevenly in time [49].

In Scargle's method, he starts with the definition of the classical periodogram for a discrete fourier transform, Equation (2.16). In [50], he critiques the classical periodogram, arguing that $P(\omega)$ is very noisy, even for noise-free data, and that

aliasing can be a large problem. He modifies the classical definition, and redefines it to be:

$$P(\omega) = \frac{1}{2} \left(\frac{\left[\sum_{i=1}^N x(t_i) \cos \omega(t_i - \tau) \right]^2}{\sum_{i=1}^N \cos^2 \omega(t_i - \tau)} + \frac{\left[\sum_{i=1}^N x(t_i) \sin \omega(t_i - \tau) \right]^2}{\sum_{i=1}^N \sin^2 \omega(t_i - \tau)} \right). \quad (2.25)$$

With the exception of the 1/2 term in front, Scargle's periodogram matches that of Lomb's least squares equation. A maximum in Scargle's periodogram occurs at the same frequency that maximizes Lomb's least squares technique [50].

Lomb and Scargle start with different principles to develop their frequency periodogram, but end up with the same formula, hence the method to be used in this analysis will be referred to as the Lomb-Scargle (LS) method.

In his paper, Scargle examines the significance of the powers found in the periodogram. If the data are purely noise, then the resulting spectrum will also have the appearance of noise. Assuming that the data are pure noise, with zero mean and unit variance, then the probability of having a power Z between z and $z + dz$ follows an exponential distribution,

$$\text{Prob}(z > Z > z + dz) = p(z)dz = e^{-z}dz. \quad (2.26)$$

Then the cumulative probability distribution, which is the probability that the power, Z , is less than a given value, z , is given by

$$\text{Prob}(Z < z) = \int_0^z p(z')dz' = 1 - e^{-z}. \quad (2.27)$$

If M independent frequencies are sampled, then $\text{Prob}(Z < z)$ becomes

$$\text{Prob}(Z < z) = (1 - e^{-z})^M. \quad (2.28)$$

This is a result of multiplying together the M independent cumulative probability distribution functions, defined in Equation (2.27). The more frequencies that are sampled, the more likely it is to find a large peak in the power. The significance of a power when M independent frequencies are sampled, is the probability that any power, Z , would be greater than a given power z ,

$$\text{Prob}(Z > z) \equiv F(z) = 1 - \text{Prob}(Z < z) = 1 - [1 - e^{-z}]^M. \quad (2.29)$$

In the context of spectral analysis, this significance is sometimes referred to as the false alarm probability (FAP). If the FAP is small, it could indicate the presence of a sinusoidal signal with a frequency for which this maximum power occurred.

When the variance of the data, σ^2 , is not unity, as was assumed above in the development of the significance, then the periodogram of Equation (2.25) must be normalized. The correct normalization of Equation (2.25) is the variance on the data, [52]

$$P_n(\omega) = \frac{P(\omega)}{\sigma^2}, \quad (2.30)$$

where the subscript n indicates the periodogram has been normalized by its variance.

To apply the LS method to a signal $y(t)$ that is sampled at N uneven time intervals: $y(t_1), y(t_2), \dots, y(t_N)$, one must examine the deviations of the data from the mean value, and determine if these deviations occur with a specific period. If this signal is pure noise, and contains no periodicity, then the signal is expected follow a Gaussian

distribution, with a mean value \bar{y} , and a variance σ^2 , defined as:

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y(t_i) \quad (2.31)$$

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (y(t_i) - \bar{y})^2 \quad (2.32)$$

In the LS method the data are then distributed with a zero mean, so that only data points that are not equal to the mean contribute to the sum in the numerator. For the data to be centered on zero, every value of $y(t_i)$ has the mean value, \bar{y} , subtracted from it. With this substitution the normalized periodogram becomes:

$$P_n(\omega) = \frac{1}{2\sigma^2} \left(\frac{\left[\sum_{i=1}^N (y(t_i) - \bar{y}) \cos \omega(t_i - \tau) \right]^2}{\sum_{i=1}^N \cos^2 \omega(t_i - \tau)} + \frac{\left[\sum_{i=1}^N (y(t_i) - \bar{y}) \sin \omega(t_i - \tau) \right]^2}{\sum_{i=1}^N \sin^2 \omega(t_i - \tau)} \right). \quad (2.33)$$

This is the regular LS periodogram that will be used to perform a basic search for periodic behavior in SNO's solar neutrino data.

For evenly spaced data the maximum frequency that can be sampled is given by the Nyquist frequency, which depends upon the constant time interval between the data points. With unevenly spaced data, the time interval is not constant, so the frequency range and maximum frequency are no longer well defined. The Nyquist frequency was defined in the case of even sampling because periods between sampling intervals can not be determined, with uneven sampling consecutive time intervals can be, and likely are, different, so in theory this allows a wider range of frequencies to be sampled.

For evenly spaced data, the range of frequencies sampled is predefined and depends upon the Nyquist frequency, along with the number of data points, N . For unevenly

spaced data, the choice of how many frequencies and with what frequency interval to sample, is undefined and left to the user to decide. In this case, the number of independent frequencies is unknown, which leaves the FAP, Equation (2.29), with an unknown parameter. However, a Monte Carlo (MC) study can be used to determine the FAP of a given power. Each toy MC is generated with the goal of simulating a constant signal. Then, each MC is examined with the LS technique, and the maximum power found. The distribution of maximum powers can be used to find the FAP at a given power, and can be used to fit for the number of independent frequencies. Further details of this technique can be found in Chapter 5.

2.3.2 Modified Lomb-Scargle Method

Often, a data point $y(t_i)$ has an error associated with its measurement, stemming from statistical and systematic uncertainties. In the regular LS periodogram these errors are not taken into account. An extension of the regular LS method is to incorporate these errors as weights, allowing those data points with small errors to have an increased influence on the periodogram, over other data points with large associated errors. In Appendix D of [51], Scargle briefly outlines a method to incorporate such weights into the periodogram formula,

$$P_n(\omega) = \frac{1}{2\sigma_w^2} \times \left(\frac{\left[\sum_{i=1}^N w'_i (y(t_i) - \bar{y}_w) \cos \omega(t_i - \tau) \right]^2}{\sum_{i=1}^N w'_i \cos^2 \omega(t_i - \tau)} + \frac{\left[\sum_{i=1}^N w'_i (y(t_i) - \bar{y}_w) \sin \omega(t_i - \tau) \right]^2}{\sum_{i=1}^N w'_i \sin^2 \omega(t_i - \tau)} \right). \quad (2.34)$$

The definition of τ is modified to become:

$$\tan(2\omega\tau) = \frac{\sum_{i=1}^N w'_i \sin(2\omega t_i)}{\sum_{i=1}^N w'_i \cos(2\omega t_i)}. \quad (2.35)$$

The weights w'_i are normalized, such that the average weight is equal to unity, and $\sum_{i=1}^N w'_i = N$. For each of the N data points the weight, w_i , is set by the user. Then the mean of the N w_i 's, \bar{w} , is found by applying Equation (2.31). The normalized weight for each data point is then $w'_i = w_i/\bar{w}$.

When each data point has an associated normalized weight, the mean value \bar{y} in Equation (2.34) is taken to be the weighted mean,

$$\bar{y}_w = \frac{1}{N} \sum_{i=1}^N w'_i y(t_i). \quad (2.36)$$

The variance σ_w^2 remains as the variance on the data, however now the spread of the data is also weighted,

$$\sigma_w^2 = \frac{1}{N-1} \sum_{i=1}^N w'_i [y(t_i) - \bar{y}]^2, \quad (2.37)$$

note that the mean value \bar{y} is the regular mean, as defined in Equation (2.31).

This method allows errors on the measured data points to be incorporated into the periodogram calculation through the use of weights, by setting the weights in proportion to the inverse variance. In this thesis analysis, this modified LS periodogram is used in parallel with the non-weighted, regular method to search for a periodic signal in the data.

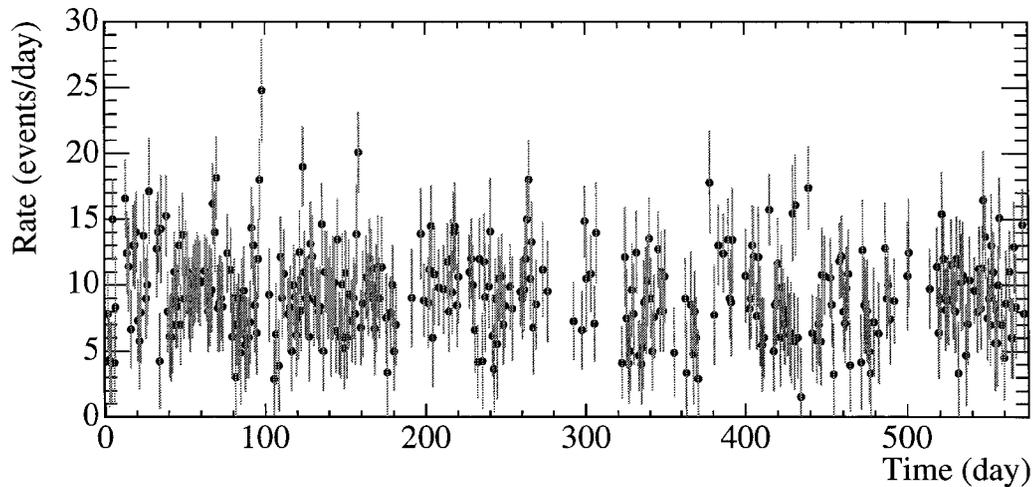


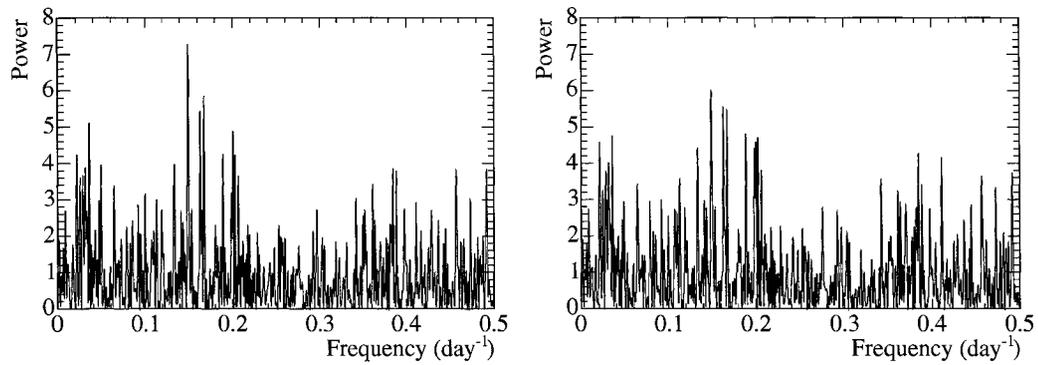
Figure 2.1: A plot of event rate, $y(t_i)$, with associated error bars, versus time, t_i , for a sample data set.

2.3.3 A Monte Carlo Example of the Regular and Modified Lomb-Scargle Methods

To demonstrate the regular and modified LS periodograms, consider the sample distribution of data shown in Figure 2.1. This data spans 575 days and is unevenly distributed among 351 bins, with an average rate of 9.35 events/day. In Figure 2.1 the x-axis shows the time, in units of days, corresponding to the event rates, shown along the y-axis. The error bars represent the expected statistical uncertainty on the data. Both, the regular and modified LS methods will attempt to fit sinusoidal functions with a range of frequencies to the data. The regular LS method ignores the error bars and treats each data point with equal weight.

For this example, the periodogram power is calculated for 7300 frequencies in a range of $1/3650 \text{ day}^{-1}$ to $1/2 \text{ day}^{-1}$. The regular and modified periodograms are shown in Figure 2.2.

The largest power occurs at the same frequency for both the regular and modified



(a) Regular LS.

(b) Modified LS.

Figure 2.2: The (a) regular and (b) modified Lomb-Scargle periodograms (power versus frequency) for the sample data set in Figure 2.1.

methods. However, the value of this power differs, 7.27 and 5.96 for the regular and modified, respectively. Figure 2.3 is a scatter plot showing the modified versus regular power at each of the 7300 frequencies. The two methods appear well correlated. Discrepancies are due to increased, or decreased, weight of the data points.

Determining whether or not a periodic signal has been found in this or any data set is the subject of Chapter 5.

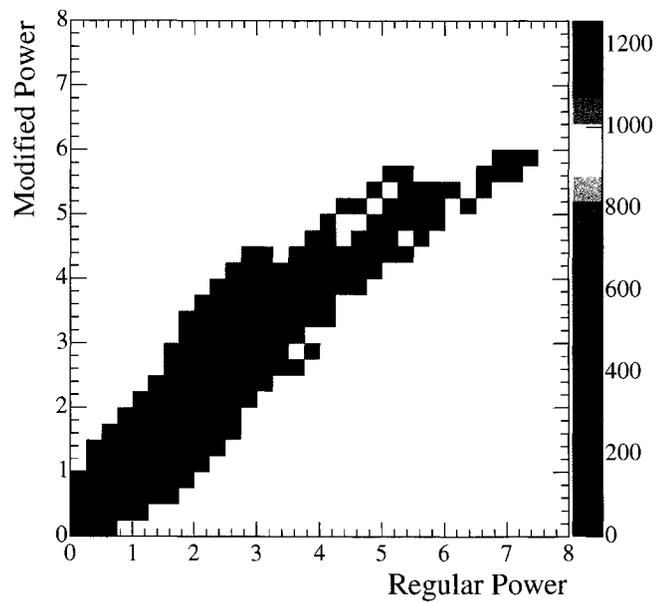


Figure 2.3: A scatter plot of the modified versus regular powers at each of the 7300 frequencies sampled against the data set shown in Figure 2.1. The colour legend to the right indicates the number of frequencies that occur at each power point in the plane.

Chapter 3

Detecting Solar Neutrinos at SNO

Neutrinos produced in the Sun travel approximately 1.5×10^8 km across a vacuum-like space to reach the Earth, where only a tiny fraction are detected by the Sudbury Neutrino Observatory (SNO). This chapter provides a brief description of the SNO detector, along with a summary of the processes used to detect the candidate neutrino events that are used in the time series analysis.

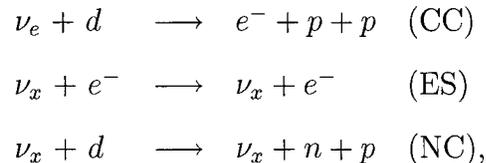
3.1 The Sudbury Neutrino Observatory

The SNO experiment is a real-time, solar neutrino experiment, designed to solve the SNP by providing a measurement of the total flux of all active ^8B neutrinos. The detector is located 2039 m below the surface in the Creighton mine, near Sudbury, Ontario, Canada. The detecting medium is 1000 tons of heavy water, D_2O , which is enclosed in a spherical acrylic vessel (AV), 12 m in diameter. The AV is surrounded by 7000 tons of light water, H_2O , providing support for the inner components and shielding from radioactive background sources. At a radius of 8.9 m an array of

9438 photomultiplier tubes (PMTs) are arranged on a spherical support called the Photomultiplier Support Structure (PSUP). These PMTs sit in the H₂O and detect Čerenkov light generated within the volume. The layout of the detector is shown in Figure 3.1.

Čerenkov radiation is produced by charged particles when they travel faster than the speed of light in that medium. Water has a refractive index of $n = 1.33$, and so the minimum speed a charged particle must reach to produce this radiation is $v = c/n \simeq 2.554 \times 10^8$ m/s. The Čerenkov light of interest is generated when solar neutrinos interact with the D₂O to produce energetic electrons, with kinetic energy greater than 0.26 MeV.

There are three main neutrino reactions that can occur within the D₂O that produce a signal. These three reactions are



where $x = e, \mu, \tau$, and d is a deuteron (composed of a neutron, n , and a proton, p).

The charged current (CC) reaction occurs when an electron-neutrino, ν_e , weakly interacts with the deuteron nucleus via W boson exchange, to produce two protons and an electron. This reaction requires a minimum of 1.44 MeV from the incoming neutrino. The outgoing electron will produce Čerenkov radiation, which will be detected by the PMTs.

The elastic scattering (ES) reaction is the result of any neutrino, ν_x , elastically ‘bouncing’ off an electron in the D₂O. This reaction is sensitive to all flavours of neutrinos, however, the ν_e is approximately six times as likely to interact than ν_μ or ν_τ . This is because the ν_e can elastically interact via both W and Z boson exchange,

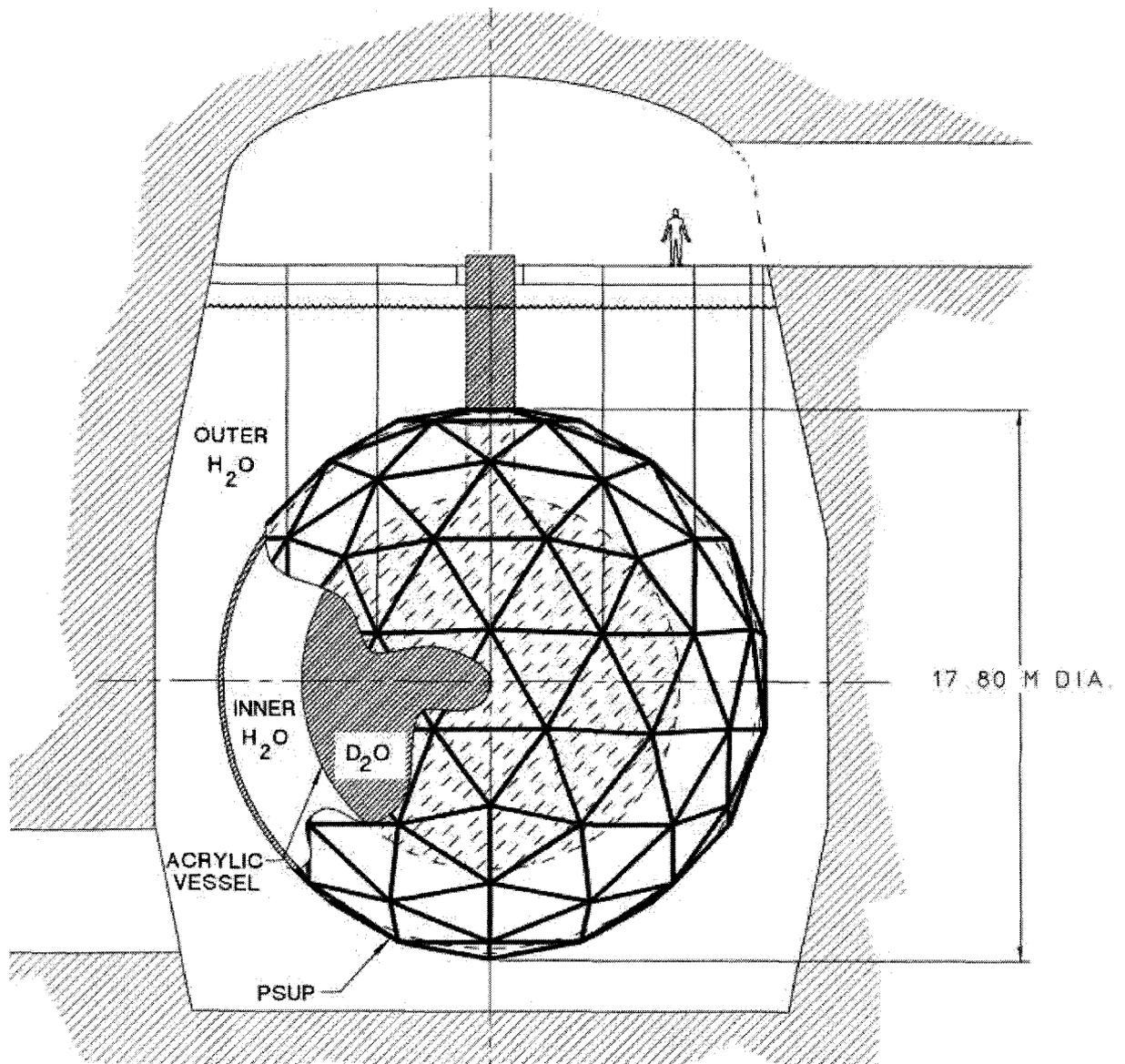


Figure 3.1: Layout of the SNO detector, 2039 m underground [53].

whereas the ν_μ and ν_τ can only interact via the neutral Z boson. Again, the outgoing electron from the interaction will produce Čerenkov light if it has sufficient energy.

The neutral current (NC) reaction occurs when a ν_x with energy greater than 2.2 MeV interacts with the deuteron and separates the proton and neutron pair. The neutron moves around the volume and is eventually captured by the detecting medium. Upon neutron capture, gamma rays are produced that Compton scatter electrons which produce Čerenkov light.

From start to finish, the SNO experiment involves three distinct phases; each phase offering a different method to measure the NC rate. The first phase, which ran from November 2, 1999 to May 28, 2001, employed pure D_2O to provide the measurement. In this phase, the neutron captured on a deuteron to produce an excited state of a tritium, which subsequently decayed to release a single 6.25 MeV γ . This γ Compton scattered electrons to produce the Čerenkov light. With pure D_2O the neutron detection efficiency was 24%. The detailed results of this phase have been published in [39, 40, 42, 54].

The second phase is known as the salt phase, as two tons of NaCl were added to the D_2O to increase the neutron capture efficiency to 83%. In this phase, the liberated neutron captured on ^{35}Cl to form an excited state of chlorine, $^{36}\text{Cl}^*$. Multiple γ rays were emitted with the transition of the chlorine atom into its ground state. These γ rays, with a total energy of 8.6 MeV, Compton scattered many electrons, creating an isotropic distribution of the Čerenkov light. This phase successfully ran from July 26, 2001 to August 28, 2003. The documentation for these results can be found in [55, 56].

At the time of this thesis SNO is currently taking data in the final phase, the neutral current detector (NCD) phase, which began taking data on December 1, 2004. Forty cylindrical proportional counters 5 cm in diameter are strung vertically in the

AV. All but four NCDs are filled with ^3He , which detect neutrons via the reaction:



The other four NCD strings contain ^4He and are used to calibrate the internal NCD backgrounds.

For this thesis the candidate neutrino events are from the D_2O and salt phases.

3.2 Timing System

The timing system is a key electronic component of the data acquisition, and is essential for a periodicity study. This system must ensure accurate event times and run boundaries. There are two independent clocks used to provide this timing information. The two clocks are used to verify each other, and if one fails, the other can provide the necessary timing information.

One clock is a 10 MHz clock that keeps the Coordinated Universal Time (UTC) by synchronizing with the Global Positioning System (GPS). Based on the 53-bit register that tracks the clock ticks, the clock is projected to keep time for 28.54 years, which encompasses the entire duration of the experiment, and is therefore said to keep the absolute time. This clock has an accuracy of 100 ns [53].

For verification of the 10 MHz clock and increased accuracy, the second clock, a 50 MHz clock, is also used to track timing information. This clock determines the global trigger times. Due to the 43-bit register the 50 MHz clock ‘rolls’ over after 2.036 days [57].

There is, in addition, a Pulsed Global Trigger (PGT) that simultaneously fires all of the enabled triggers and provides an unbiased picture of the detector systems. This

runs with a frequency of 5 Hz. Counting the number of PGT events can provide an estimate of the time. This system is not independent as it relies on the 50 MHz clock. However, it can be used as a verification tool [58].

The 10 MHz clock is synchronized to the GPS so that rare neutrino events, or bursts of events, can be linked and verified with other international neutrino experiments, as would be necessary in the case of a suspected supernova. The clock is kept in sync with the GPS through a series of electronic signals. At predetermined times a signal from the GPS is received on the surface and is transmitted along approximately 4 km of optical fiber to the underground laboratory, this is called the sync pulse. At the same predetermined time global triggers are latched onto the underground 10 MHz and 50 MHz counters and a pulse is sent to surface through a separate fiber optic cable, this is called the ping pulse. Once the ping reaches the surface it is reflected back along the cable, and is now referred to as the pong pulse. The ping-pong signal is used to determine the delay time for the original sync pulse, sent from the surface GPS receiver to the underground timing card. A comparison of the global triggered 10 MHz clock tick is made to the time of the GPS sync signal to correct for any time discrepancies between the 10 MHz clock and the GPS [59, 60].

Occasionally, service to the GPS has been lost. In these cases the 10 MHz clock was rebuilt from the 50 MHz clock [57]. These situations are often the result of severed fiber optical cables; a hazard of operating in an active mine.

3.3 Run List Selection

Throughout the operational phases of SNO the detector livetime has been broken up into runs, reflecting the various detector activities. The start and stop times of a run

	10 MHz Raw Livetime (day)	50 MHz Raw Livetime (day)
D ₂ O	312.93	312.92
Salt	398.59	398.58

Table 3.1: Raw livetimes for the 10 MHz and 50 MHz clocks.

are calculated with both the 10 MHz and 50 MHz clocks. The start of a run is tagged at the time of the earliest event, and the stop time by the last event. An event is classified as anything that meets the conditions set by any of the enabled triggers, and is likely not to be a neutrino event. Since the PGT runs at 5 Hz the maximum time between the actual start (end) of a run and the first (last) event is 0.2 seconds.

The task of SNO's Run Selection Committee is to evaluate runs that are suitable for neutrino data analysis. There are obvious cuts made on runs tagged for detector calibration and maintenance. Other runs are cut based on electronic configurations, background levels, environmental conditions, and unusual disruptive activity.

After these cuts were performed 559 and 1212 runs were approved for neutrino analysis in the D₂O and salt phases, respectively. The total raw livetimes for the 10 MHz and 50 MHz clocks are calculated by summing the run lengths; these are shown in Table 3.1. There is a slight difference between the total clock livetimes; this difference is distributed throughout the runs, and typically amounts to individual run lengths that differ by times less than a second. This difference is not expected to effect this periodicity analysis which searches for periods on the order of days.

Within these neutrino runs there are several time cuts imposed to reduce potential backgrounds. These cuts reduce the livetime of a run, hence are referred to as dead time cuts. A dominant source of these dead time cuts is called the muon follower cut, used to reduce spallation products caused by cosmic ray muons that can mimic the NC

	Raw Livetime (day)	Corrected Livetime (day)	% Dead Time Cut
D ₂ O	312.93	306.39	2.1%
Salt	398.59	391.43	1.8%

Table 3.2: Livetime statistics for the approved neutrino runs. The corrected livetime corresponds to the amount of livetime after the dead time cuts. The dead time cut values have been taken from [39, 56].

interaction. Muons passing through the detector can interact with the D₂O to produce neutrons causing a background to the NC reaction, or can create spallation products in the surrounding materials which can also falsely mimic a neutrino interaction. The muon follower cut removes all data 20 seconds after a muon is identified, long enough for the half-life decay of most of the spallation products. Muon identification is based on the amount of light deposited at various positions throughout the detector. There are a handful of other dead time cuts that are applied to SNO's neutrino analysis. Table 3.2 shows the raw livetime and corrected livetime for the 10 MHz clock in both phases.

Candidate neutrino events are selected from the approved runs within the live-times, after dead time cuts. For this analysis the raw livetime for each run is used, which is defined as the run start plus the total length, neglecting dead time cuts. The dead time cuts are ignored as the selected events do not come from these times, and per run the fraction of dead time is very small (see Table 3.2) compared to the periods of interest in this analysis, and is randomly distributed throughout the runs. The 50 MHz clock has been used to define the run lengths. The distribution of run lengths for the D₂O and salt phases are shown in Figure 3.2. There are a large number of runs corresponding to 1 day and 0.292 days (= 7 hours), a direct result of the maximum run lengths programmed into the SNO data acquisition software. The run boundaries

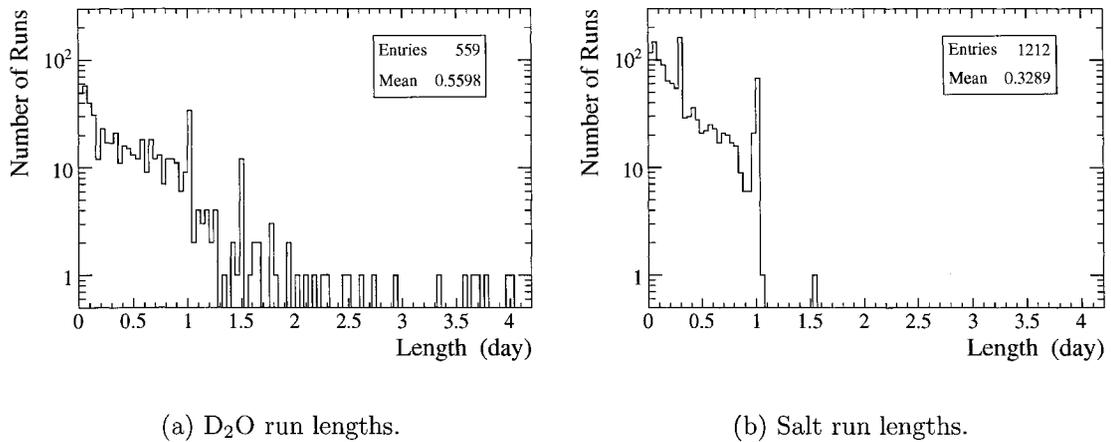


Figure 3.2: Run lengths for the D₂O and salt phases.

are used in this periodicity analysis to define certain criteria for the LS periodogram, explained in Chapter 4, and are also used to generate the Monte Carlo data sets, described in Chapter 5.

3.4 Event Selection

The approved runs are analyzed to find the candidate solar neutrino events. Typical event rates in the detector during a neutrino run can range from 15 to 30 Hz, so over the course of several hours, hundreds of thousands of events can accumulate. An event is classified as anything that meets any of the enabled trigger conditions.

3.4.1 Trigger System

There are 26 different triggers that can initiate the creation of an event file. For neutrino data analysis the triggers of importance are:

NHIT		100	20	
ESUM		HI	LO	
OWL		N	HI	LO

The NHIT triggers are used to determine the number of PMTs that fire within a given time window. This time window is 100 ns for NHIT 100, and 20 ns for NHIT 20. The NHIT signals are discrete pulses of uniform height for each triggered PMT. For NHIT 100 there are three trigger levels, NHIT 100 LO/MED/HI. Standard neutrino data acquisition requires NHIT 100 MED to be 16, implying that within a 100 ns time window 16 PMTs must meet their trigger condition. The ESUM triggers are a sum of the individual PMT pulses. ESUM HI and ESUM LO are triggers with high and low gain, respectively. The OWL triggers are equivalent to the NHIT and ESUM triggers, only these are for the 91 outward looking PMTs used for background reduction. OWL N is equivalent to NHIT 100, and OWL HI and OWL LO are like the ESUM HI and ESUM LO triggers.

When a trigger condition is satisfied the event is tagged with a global trigger. An event file is written that stores information about the event, such as the 10 MHz and 50 MHz times, and PMT charge, timing, and location information. This information is used by the data builder to reconstruct events.

3.4.2 Backgrounds

The average solar neutrino interaction rate in the detector is on the order of ten events per day, so a majority of the events are due to backgrounds. Backgrounds can be broken up into a few categories:

- Neutron backgrounds : There are background neutrons that are generated inside and outside the D₂O volume, typically as a result of radioactive decays.

- Čerenkov backgrounds : These can occur as a result of radioactive decays.
- Instrumental backgrounds : These backgrounds are caused by the detector electronics.
- Other backgrounds : This includes muon backgrounds, AV backgrounds, α reactions on elements, antineutrinos, and atmospheric neutrinos.

Neutron backgrounds can be generated by ^{208}Tl and ^{214}Bi , daughter elements in the decay chain of ^{232}Th and ^{238}U , if they produce a γ with enough energy (> 2.2 MeV) to break up the deuteron nucleus, mimicking the NC signal. This background is minimized by limiting the amount of ^{232}Th and ^{238}U in the detector components. There is a dedicated task to monitor these levels and ensure they remain below the target concentrations of $3.8 \times 10^{-15} \text{gTh/gD}_2\text{O}$ and $30 \times 10^{-15} \text{gU/gD}_2\text{O}$ [56]. For both phases, these radioactive limits were met. Measuring the levels of ^{232}Th and ^{238}U gives an upper limit on the number of these neutron backgrounds contaminating the neutrino NC signal.

These radioactive decays can also generate Čerenkov backgrounds. To determine the number of these backgrounds in the neutrino signal, Th and U sources are placed at various positions outside the D_2O . The known concentrations from these sources, along with the measured Čerenkov events due to these sources, are used to estimate the number of events these backgrounds contribute, given the measured concentrations of ^{232}Th and ^{238}U in the detector components.

The detector electronics can create false events, these are classified as instrumental backgrounds. One type of instrumental background is called a flasher. These are believed to be caused by static discharges of the PMTs. A signature of these backgrounds is the distribution of hit PMTs; there is a small area of hit PMTs near the

flashing PMT, and a larger group of hit PMTs on the opposite side of the detector. Other instrumental backgrounds are caused by electronic pickup and breakdowns, and upon initiation of various water monitoring procedures. Most of these backgrounds are readily cut based on the signature charge and timing distributions of the triggered PMTs.

To reduce the amount of backgrounds from external sources that may creep into the interior volume, a fiducial volume less than the total D₂O volume is chosen for analysis. For the D₂O phase, only events reconstructed within a 500 cm radius from the center of the detector are analyzed. For the salt phase, this value is 550 cm. Backgrounds from radioactive decays are also reduced by examining events with a reconstructed kinetic energy greater than 5.0 MeV for the D₂O phase, and 5.5 MeV for the salt phase.

3.4.3 Candidate Solar Neutrino Events

After these background cuts are applied the remaining events are primarily neutrino events, with a small contribution from backgrounds. As reported in [39, 56], the remaining candidate neutrino events are 2924¹ in the D₂O phase, and 4722 in the salt phase.

As mentioned above, the candidate events are primarily neutrino events, however, the final data set does contain a small number of background events. MC simulations and measured values of radioactive isotopes provide upper limits on the amount of expected backgrounds remaining in the event data set. Table 3.3 shows the percentage of non-neutrino events included this analysis. These are minimal, on the order of five

¹There is a four event difference between that quoted in this thesis and [39], due to an additional background dead time cut that was implemented in the salt phase, and retroactively applied to the event list in the D₂O phase in 2005.

	D ₂ O	Salt
Total Candidate Neutrino Events	2924	4722
Total Estimated Backgrounds	123.0	260.4
Purity of Events	95.8%	94.5%

Table 3.3: Background contributions to the total neutrino event rate. The breakdown of contributing backgrounds can be found in [39, 56] for the D₂O and salt phases, respectively.

percent of the total events.

Using MC simulations of the three neutrino reactions, probability density functions (PDFs) are created for the characteristic energy, radial, and directional distributions for each signal. The CC reaction is distinguishable through its energy distribution. The ES reaction has a distinct directional distribution. When an incoming solar neutrino elastically scatters off an electron in the water the outgoing electron generally scatters in the forward direction, due to conservation of momentum. The NC reaction can be identified by its characteristic energy, radial, and light isotropy distributions. The rate for each of the three neutrino signals is found by fitting the candidate events to these PDFs using an extended maximum likelihood fit.

It is not possible to determine which of the three reactions initiated a single event, so this analysis will only examine the total candidate neutrino data set, and consequently only examine the periodic behavior of the combined signal solar neutrino rates.

Chapter 4

Lomb-Scargle Method Applied to the SNO Data

The information necessary to produce a LS periodogram is rate with a corresponding time, (see, *eg.* Figure 2.1). The SNO data provides run and event time information. The task is to group this data in such a way that meets the LS input requirements, and limits the loss of information. Then the data are analyzed to search for periodic behavior at a number of frequencies. This frequency range should be broad enough to maximize the data span, and limited to reduce potential aliasing effects. This chapter describes how the SNO data were used to produce the LS periodograms.

4.1 Input Variables

4.1.1 Bin size

For this analysis the SNO data are put into 1-day bins. A bin size of 1-day was chosen because it provides sensitivity for periods of a few days, and with the average

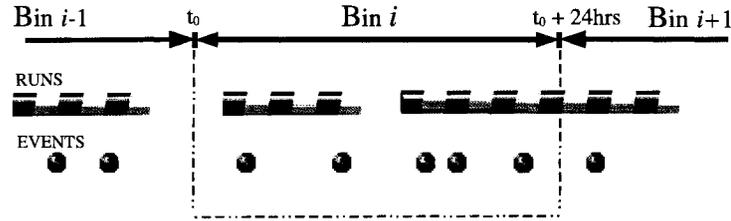


Figure 4.1: A graphical representation of how runs and events are put into bins with predefined time boundaries.

event rate on the order of 10 events per day, most bins have adequate statistics. A systematic test of the bin size choice was performed, where the bin size was varied from 1 to 5 days. The choice of bin size does not affect the outcome. The full analysis is described in Chapter 6.

Figure 4.1 demonstrates the procedure. Bin i is defined as a 24 hour period, spanning some time t_0 to $t_0 + 24$ hours. All candidate neutrino events recorded in this time are grouped, and give a raw number of events to that bin. The sum of all neutrino runs falling within this 24 hour period contribute to the livetime in that bin. In some cases, a run must be split between neighboring bins, depending upon run boundaries.

4.1.2 Time

One input parameter for Equation (2.33) is the time, t_i corresponding to a given rate $y(t_i)$. The rate is proportional to the accumulation of all events occurring within the bin boundaries, and these events can only occur during run livetimes. Therefore, the effective bin time is chosen as a run-lifetime weighted value, defined as,

$$t_i = \left(\frac{1}{\int_{b_s}^{b_f} w dt} \right) \int_{b_s}^{b_f} w t dt, \quad \text{where } w = \begin{cases} 1 & \text{during livetime,} \\ 0 & \text{otherwise.} \end{cases} \quad (4.1)$$

Here b_s and b_f are the bin start and finish times, respectively. The definition of the time variable ensures an accurate representation of the time for binned candidate neutrino events. This choice results in an uneven time spacing of the data, which is why the LS method, described in Chapter 2, is applied to this data.

The first bin starts at t_0 and ends at $t_0 + 24$ hours, the second bin from $t_0 + 24$ hours to $t_0 + 48$ hours, etc. In general, the n^{th} bin has time boundaries of $[t_0 + (n \times 24 \text{ hours})]$ to $[t_0 + ((n + 1) \times 24 \text{ hours})]$. The start bin boundary, t_0 , is chosen as time $t = 0.00$ seconds on the day of the earliest run. To demonstrate the robustness of the method, and prove time invariance of the LS periodogram, this t_0 was incrementally changed. The choice of t_0 does not affect the outcome of the analysis. Further details can be found in Chapter 6.

4.1.3 Rate

A 1-day rate is calculated for every bin using the raw number of events and livetime. The resulting rate is a livetime corrected rate. If for example five events are detected in a livetime of 12 hours, then the livetime corrected rate is 10 events per day.

Unfortunately, binning has the potential to create low statistic bins. It is possible that the range of a given bin only encompasses a small fraction of livetime, leading to a low raw event rate. For example, if a bin has a total of one event in one hour, it is said that 24 events would be seen in one day. Even worse, if zero events are seen in a limited livetime, then the livetime correction would imply a rate of zero events per day. These low statistic bins provide a poor representation of the rate, and may bias the resulting LS power spectrum.

To avoid this potential bias, raw 1-day bins that have an expected number of events of five, or less, are combined with neighboring bins. The expected number of

events is directly related to the amount of livetime, and is defined as,

$$(\text{expected number of events}) = \left[\frac{\text{total number of events}}{\text{total livetime}} \times (\text{bin livetime}) \right]. \quad (4.2)$$

The condition that all bins have an expected number of events greater than five is important for statistical error assignment (discussed below), and necessary to provide a Gaussian-like distribution of data. Combining based on expected versus observed number of events is crucial when using the MC to determine the significance and sensitivity of the data, because the MC should reflect the same time distribution as seen in the data. An analysis was performed to ensure combining bins does not introduce a bias, this is explained in Chapter 6. The limit of five events requires less than half of all bins to be combined, and therefore maintains a good sensitivity in the detection of a periodic signal.

If, and when, bins are combined, a 1-day rate is calculated from the new number of raw events and livetime. Similarly, a new time variable is calculated using Equation (4.1), where b_s is the start of the earliest bin and b_f is the end of the last bin. The final livetime corrected rates are the values for $y(t_i)$.

4.1.4 Weight

The modified LS method requires that each data point be weighted, allowing data points to have greater (or lesser) influence on the power calculation. For this analysis, the unnormalized weight, w_i (refer to section 2.3.2 for details on the (un)normalized weights), in bin i equals the inverse variance,

$$w_i = \frac{1}{\sigma_i^2}. \quad (4.3)$$

The errors, σ_i , are assigned based on statistical fluctuations and livetime uncertainties of the data.

The LS periodogram is a null hypothesis test: it tests the data distribution against the hypothesis that the data are normally distributed, constant noise. The false alarm probability (FAP), and method of MC simulations, tests the significance of peak powers in the data against this null hypothesis. The raw statistical errors, σ_i^{raw} , are chosen to reflect this model, and are taken to be the square root of the expected number of events,

$$\sigma_i^{\text{raw}} = \sqrt{(\text{expected number of events})_i}, \quad (4.4)$$

see Equation (4.2) for the definition of expected number of events.

Similar to the method of livetime correcting the rate, the errors undergo the same livetime correction, such that,

$$\sigma_i = \left[\frac{\text{maximum bin livetime}}{(\text{actual bin livetime})_i} \right] \sigma_i^{\text{raw}}. \quad (4.5)$$

Bins with more livetime (maximum bin livetime / actual bin livetime) $\simeq 1$ have a smaller error, and therefore have a larger weight in the analysis.

4.1.5 Frequency Range

The data under analysis is unevenly spaced in time. As discussed in Chapter 2, there is no clear prescription for choosing the sampled frequency range. The data are put into 1-day bins, so by Equation (4.1), the difference in time between consecutive bins could be as small as a few seconds, or as large as several days. However, the Nyquist frequency defined for evenly spaced data will be applied to this analysis, assuming an even spacing of 1-day. By this logic the maximum frequency to be sampled is 0.5

day⁻¹. Although, it may be possible to obtain information for larger frequencies, the statistics will be limited, and the risk of aliasing greater.

This analysis has chosen to sample to a maximum period of 10 years. Sensitivity to periods much greater than 8 years is not expected, due to the four year calendar span of the combined data sets.

The frequency interval is, again, arbitrary. The choice is limited by computer processing time, realistic expectations of possible periods, and resolution between neighboring frequencies. A minimum frequency of 1/3650 day⁻¹ and a maximum frequency of 1/2 day⁻¹ has been chosen. Sampling 7300 frequencies at an interval of 1/(4 × 3650) day⁻¹ is a convenient choice, and will sample a wide variety of frequencies in the defined range. This range will allow a search of the expected yearly solar neutrino periodicity, and the 20 – 40 day rotational periods of the solar layers. Due to the limited data, a search for the 11 year solar cycle will not be studied.

4.1.6 Summary

Table 4.1 summarizes the parameters that have been defined for the SNO data.

4.2 Lomb-Scargle Analysis

The first two phases of the SNO experiment have been described in Chapter 3. This section now presents the LS analysis that has been performed on these two data sets.

4.2.1 D₂O and Salt Data Separately

A basic analysis is to analyze the distribution of neutrino events from the D₂O and salt phases, separately. Table 4.2 provides a summary of the data from these two

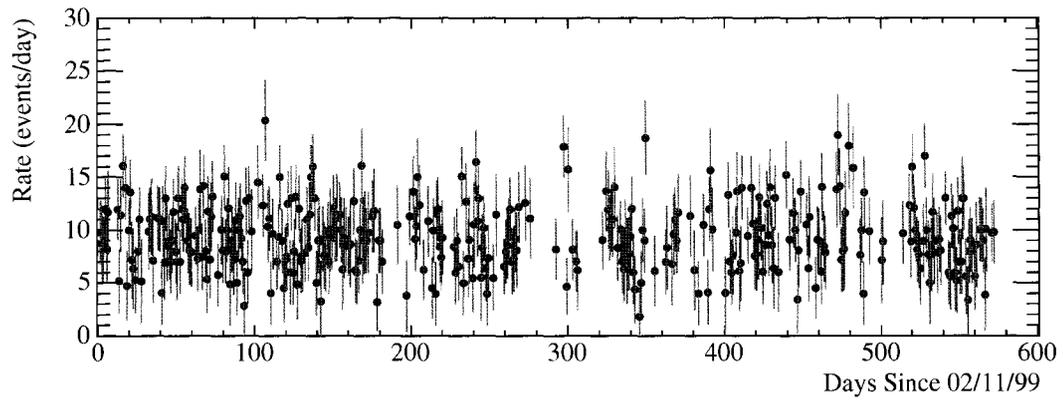
Variable	Definition	Comments
Bin size	24 hours	initial bin boundary, t_0 , $t = 0.00$ s on day of earliest run
Time t_i	$t_i = \left(\frac{1}{\int_{b_s}^{b_f} w dt} \right) \int_{b_s}^{b_f} wt dt$	$w = 1$ during livetime $w = 0$ otherwise
Rate $y(t_i)$	$y(t_i) = (\text{raw number of events}) \times$ $(\text{maximum livetime} / \text{bin livetime})$	each bin has an expected number of events $>$ five
Weight w_i	$w_i = \frac{1}{\sigma_i^2}$	$\sigma_i = \left[\frac{\text{maximum livetime}}{(\text{bin livetime})_i} \right] \sigma_i^{\text{raw}}$ $\sigma_i^{\text{raw}} = \sqrt{\text{expected number of events}}$
Frequency	$[\nu_{min}, \nu_{max}] =$ $[1/3650, 1/2] \text{ day}^{-1}$	$\Delta\nu = \frac{1}{4 \times 3650} \text{ day}^{-1}$ 7300 frequencies sampled

Table 4.1: A summary of the variables to be used in this LS analysis.

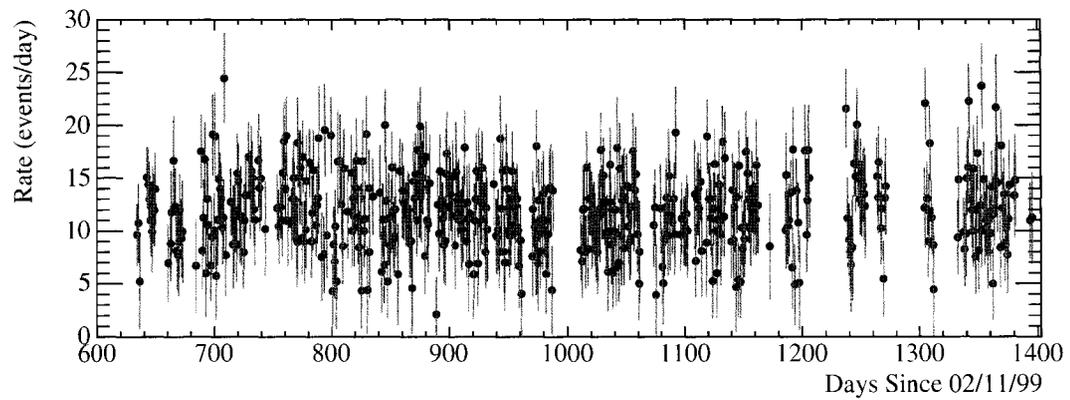
phases. The data have been grouped in 1-day bins as previously described, resulting in 351 data points for the D₂O phase, and 469 data points for the salt phase.

The 1-day livetime corrected rate versus time, with associated errors, is shown in Figure 4.2. The mean number of events for each 1-day bin in the D₂O and salt phases are 9.35 ± 0.17 and 11.85 ± 0.17 , respectively. The increase in the mean rate during the salt phase is a result of the increased neutron capture efficiency, with the introduction of salt. This phase has increased detection of the total number of all active solar neutrinos. There is a 59 day gap in the data between the end of the D₂O phase and beginning of the salt phase, a reflection of the preparation and transition time.

A projection of Figure 4.2 onto the y-axis is shown in Figure 4.3, showing the distribution of rates. Recall that the LS method is testing the null hypothesis that the data have a Gaussian distribution, so a fit to a Gaussian function is shown. The fit probabilities are 38% and 51% for the D₂O and salt phases, respectively. The rates



(a) D₂O Phase, with a mean rate of 9.35 ± 0.17 events per day.



(b) Salt Phase, with a mean rate of 11.85 ± 0.17 events per day.

Figure 4.2: The livetime corrected 1-day rate of events as a function of time for the (a) D₂O and (b) salt phases.

	D ₂ O	Salt
Number of events	2924	4722
Run numbers	10000 – 16013	20684 – 33520
Calendar dates (dd/mm/yy)	02/11/99 – 28/05/01	26/07/01 – 28/08/03
Number of Runs	559	1212
Total Raw Livetime (days)	312.93	398.3
Number of 1-Day Bins	351	469

Table 4.2: Statistics of the SNO data used in this analysis.

shown in this figure have not been weighted by their associated errors.

Figure 4.4 shows the time difference between consecutive bins. Both phases show a peak at $\Delta t \simeq 1$ day. This reaffirms the use of the Nyquist frequency to reduce aliasing effects for frequencies greater than 0.5 day^{-1} . As the time difference between neighboring bins is increased from 1 day, the number of bins with this time difference falls off.

The time spacing of the data points directly affects the number of independent frequencies, by the ability to resolve large frequencies. If the data were binned into less than a 24 hour period the time difference between consecutive bins would be smaller and there would be an increased ability to resolve larger frequencies.

4.2.2 D₂O and Salt Data Combined

A more complex analysis involves combining data from the two phases. This is motivated by increased statistics, leading to better resolution of short periods; and a longer time span of the data, providing a sensitivity to longer periods.

The data sets can not be blindly combined, since the mean rates of both phases are different. The distribution of data under analysis should have a mean value \bar{y}

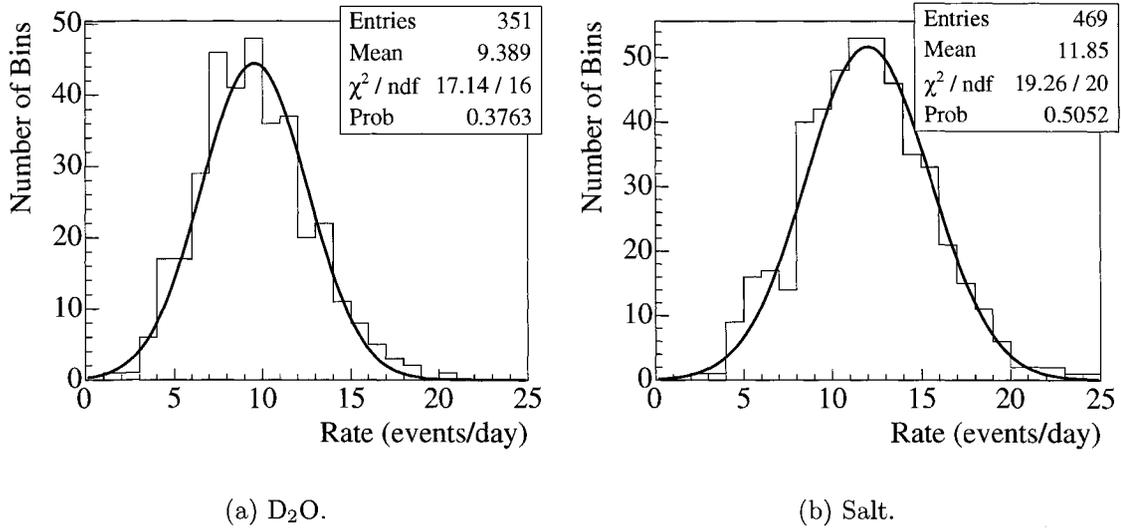


Figure 4.3: Distribution of the neutrino 1-day corrected bin rates for the (a) D₂O and (b) salt phases. Each distribution has been fit to a Gaussian, and the fit results shown.

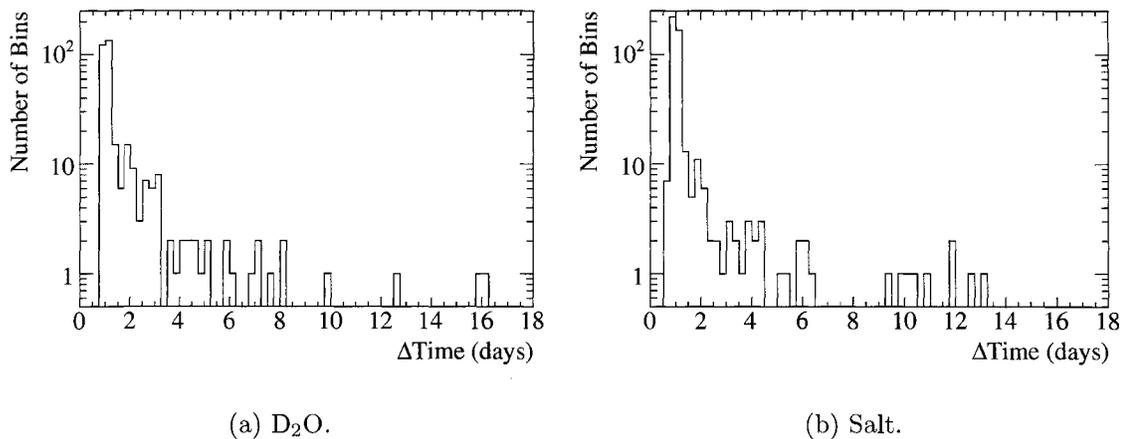


Figure 4.4: The time difference between consecutive bins for the (a) D₂O and (b) salt phases.

and variance σ^2 . Combining the D₂O and salt data sets would resemble a double Gaussian, centered on the respective mean rates. The D₂O and salt data sets are similar in length, so combining these two data sets would yield a step function-like distribution of the the data. Such a distribution would create a false signal at a low frequency.

To avoid this problem, the rates in one phase must be adjusted, such that the two phases have an equal mean rate. This is achieved by multiplying each D₂O bin by a common factor, bringing the mean rate up to match that of the salt rate. The multiplicative factor is the ratio of the mean salt rate to the mean D₂O rate. The statistical uncertainty on this ratio of 2.4% is ignored, and the systematic error is not considered in this analysis. A similar technique can be used to adjust the salt rate down. Both adjustments have been performed, and as expected, both cases yield the same results. Appendix A describes the results of both techniques.

4.2.3 Results

The data from SNO's D₂O and salt phases have been analyzed separately, and then combined for analysis, following the procedures outlined above. The periodograms for the regular and modified LS methods are shown in Figure 4.5. The numerical results are displayed in Table 4.3. The frequency of largest power corresponds to periods of 2.45 days, 2.33 days, and 2.42 days for the D₂O , salt, and combined phases, respectively, using the modified LS method.

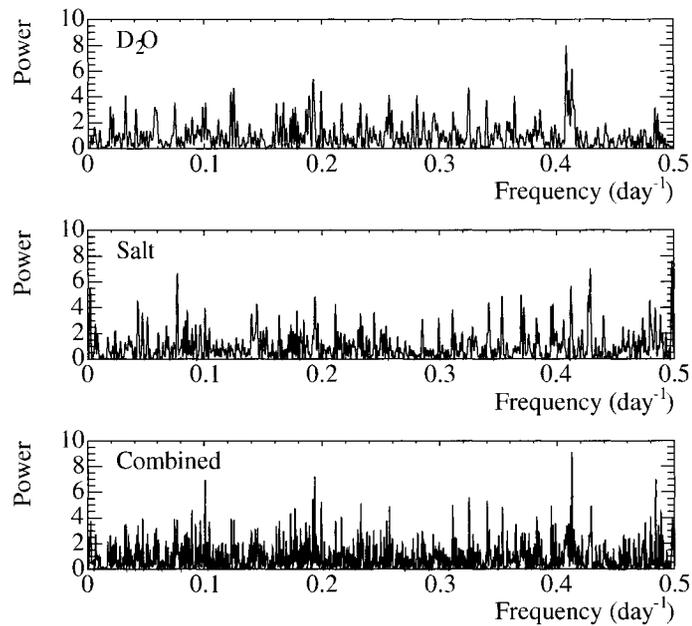
All analyses sample the same 7300 frequencies, but the number of independent frequencies, in other words, the resolution of neighboring frequencies, increases with data. This is seen by the decreasing width of the frequency peaks with the increasing statistics in the D₂O, salt, and combined data sets (see Figure 4.5). By Equation

Phase	Method	Maximum Power	Frequency (day ⁻¹)
D₂O	REGULAR	7.911	0.408
	MODIFIED	7.134	0.408
Salt	REGULAR	7.601	0.498
	MODIFIED	6.769	0.429
Combined	REGULAR	9.093	0.413
	MODIFIED	8.734	0.413

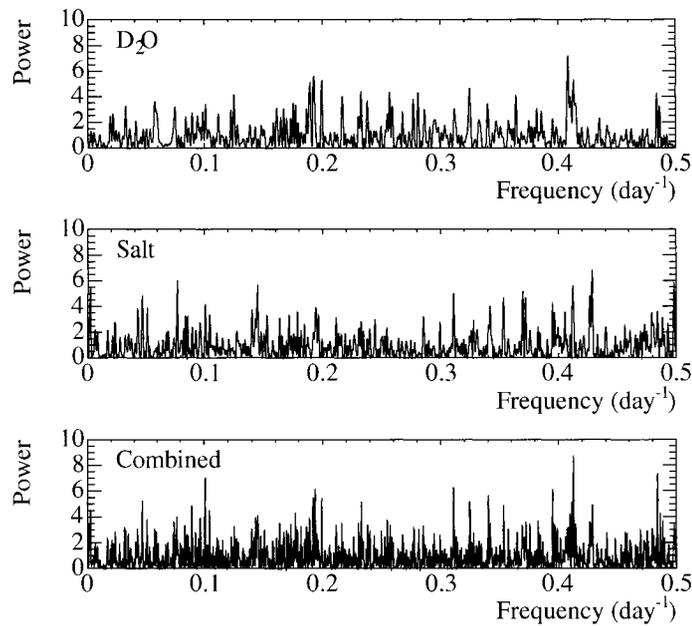
Table 4.3: The regular and modified LS periodogram results for the three data sets. The frequency corresponds to the frequency at the maximum power.

(2.29), as the number of independent frequencies increases, the FAP at a given power decreases. So the larger power in the combined phase of 8.73 does not necessarily represent a more likely periodic signal. The actual number of independent frequencies can not be quantitatively determined from these distributions, so the MC analysis in the following chapter will be used to determine the significance of any peaks.

The modified and regular LS methods produce slightly different periodograms. Figure 4.6 is a scatter plot showing the modified versus regular power at each of the 7300 frequencies sampled in the combined analysis. The figure shows that the powers at each frequency are well correlated, where differences arise from weighting the bins by statistical and livetime uncertainties. Unless otherwise stated, the remainder of this thesis will only present the results of the modified LS method. This method includes more information about the data, hence gives more accurate results.



(a) Regular LS method.



(b) Modified LS method.

Figure 4.5: Lomb-Scargle periodograms for the SNO data, D_2O and salt separately, and the combined phases, using the (a) regular and (b) modified methods.

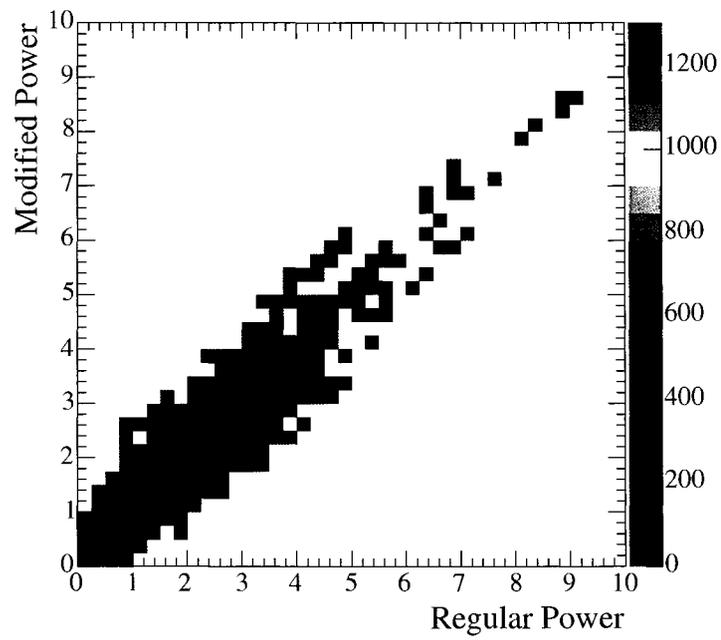


Figure 4.6: The modified versus regular powers at all 7300 frequencies sampled in the combined analysis.

Chapter 5

Monte Carlo Analysis of the Periodogram Results

The LS analysis has been applied to the SNO solar neutrino data, and it is now necessary to determine if a significant periodic signal has been found. As a reminder, the analytic expression used to determine the probability of a power z , or greater, being obtained under the null hypothesis is called the false alarm probability (FAP),

$$\text{Prob}(Z > z) \equiv F(z) = 1 - \text{Prob}(Z < z) = 1 - [1 - e^{-z}]^M. \quad (5.1)$$

The number of independent frequencies, M , is an unknown function of the time distribution of the data, number of data points, and to a lesser extent, the number of frequencies sampled. Since M can not be found from the data alone, this analytic expression can not be used to interpret the maximum power found in the data. Instead, a Monte Carlo (MC) study is used to interpret the data results.

MC simulations are used to replicate the characteristics of an experiment. In this

case, the quantity of interest is the time distribution of the neutrino events. The SNO data set is reproduced 10 000 times with MC techniques, to characterize the typical periodogram if the data have no periodic behavior. Each of these MC data sets will produce a periodogram. The distribution of 10 000 maximum powers obtained from each of the 10 000 MC experiments, will be compared to the maximum power in the data, and then used to determine the probability that the data are constant. The constant rate, null hypothesis is rejected if the chance probability for a power of z , or greater, to be obtained in a constant rate signal is less than, or equal to, 1%.

Also in this chapter, MC studies are used to determine the sensitivity of this analysis to various periodic signals, by examining the probability of rejecting the null hypothesis at the 99% confidence level (CL).

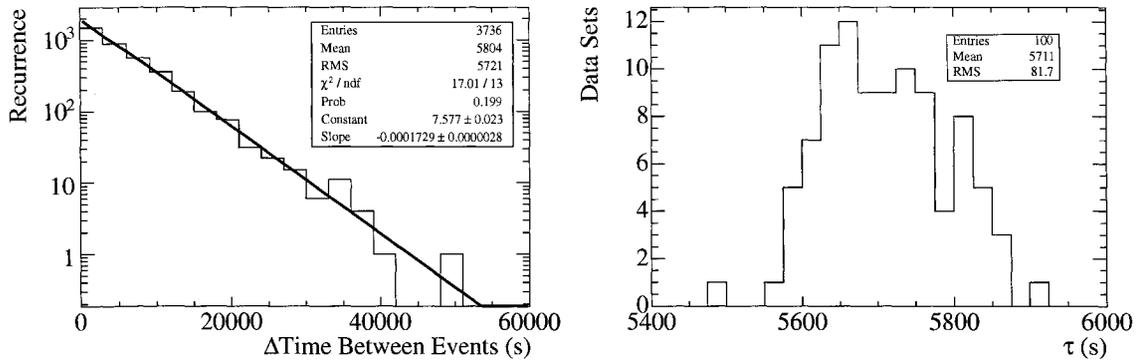
5.1 Monte Carlo Generation

10 000 independent MC data sets were generated in likeness of both the D₂O and salt phases. Each data set is generated with a random number of events, where this random number is normally distributed, centered on the actual number of events in the respective phases. Event times are randomly generated and only accepted if they occur within the actual SNO neutrino run boundaries.

These MC data sets have similar event time signatures as those of the data. The time difference between successive events should follow an exponential function,

$$f(t) = \kappa e^{-\frac{t}{\tau}}, \quad (5.2)$$

where κ is a constant, t is the time difference between neutrino events in the same run, and τ is the mean time difference between successive neutrino events. This τ



(a) Time difference between neutrino events for the salt data.

(b) Distribution of τ values for 100 salt MC data sets.

Figure 5.1: Average time between neutrino events for the salt phase and Monte Carlo data sets.

represents the average time between neutrino events, and is similar for the data and MC. Figure 5.1(a) shows the time difference between successive neutrino events for the salt data set. Fitting to an exponential function indicates that $\tau = (1/0.0001729) = 5784_{-92}^{+95}$ seconds. Then, performing this same fit on 100 of the salt MC data sets, Figure 5.1(b) shows the histogram of these 100 τ values. The mean value from these 100 MC is $\tau = 5711$ seconds, in good agreement with the data distribution.

The MC data sets also share the same mean event rate as the actual SNO data. Figure 5.2 shows the distribution of mean rates for the 10 000 MC in both phases. These mean rates have been calculated using the weighted mean formula defined with Equation (2.36). These agree well with the mean data rates quoted in Chapter 4, of 9.35 events/day and 11.85 events/day for the D_2O and salt phases.

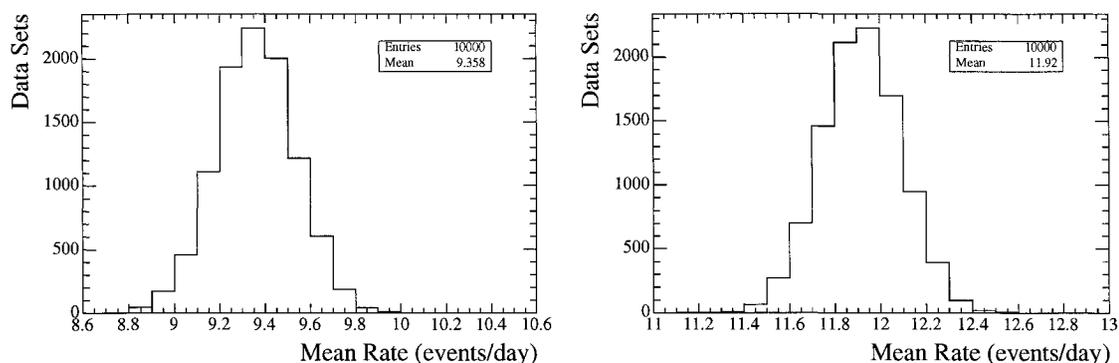


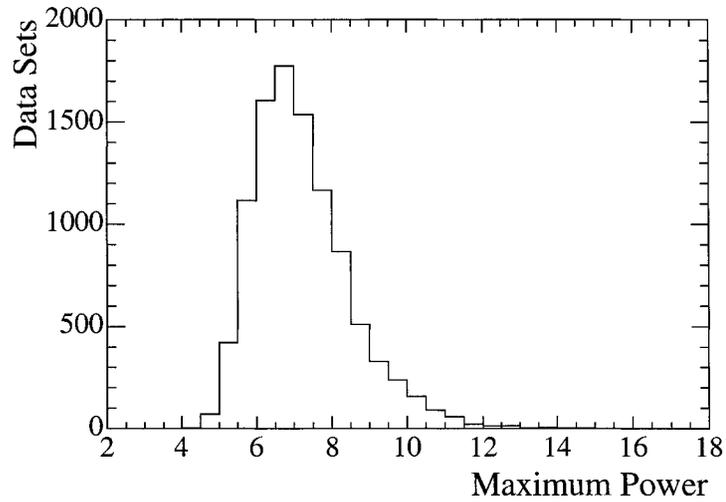
Figure 5.2: Mean event rates for the 10,000 D₂O (left) and salt (right) MC data sets with no periodic behavior.

5.2 Null Hypothesis Monte Carlo

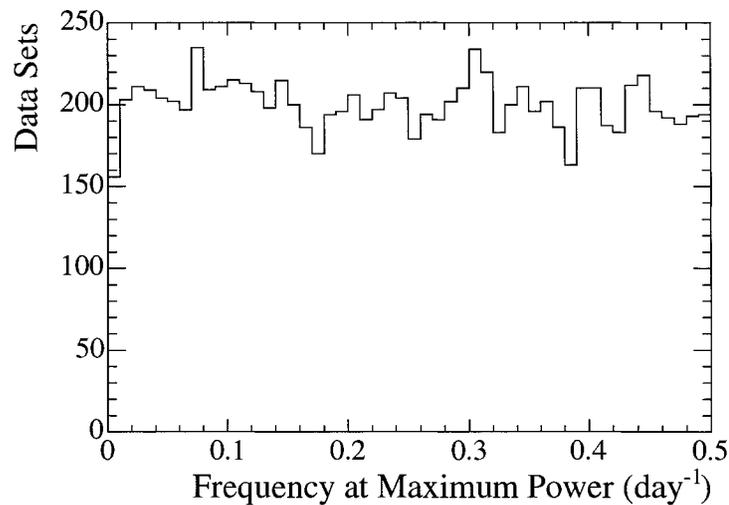
The MC data sets for each phase were analyzed separately, and then combined, using the same techniques as applied to the data, described in Chapter 4. The distribution of maximum powers are shown in Figures 5.3(a), 5.4(a), 5.5(a) for each of the three data sets. These represent the largest powers obtainable in a data set that has no periodic behavior. Since these data sets have no intrinsic periodicity, the distribution of frequencies at these maximum powers should also be uniformly, randomly distributed in the sampled range of $[0, 0.5] \text{ day}^{-1}$. These frequency distributions are shown in Figure 5.3(b), 5.4(b), 5.5(b).

5.2.1 Rejecting the Null Hypothesis

As previously described, the number of independent frequencies, and hence the FAP is primarily sensitive to the number of data points and time distribution of the data. To use the MC to determine the significance of a power, the MC must share the same number of bins and time distribution as the data. This is achieved because low

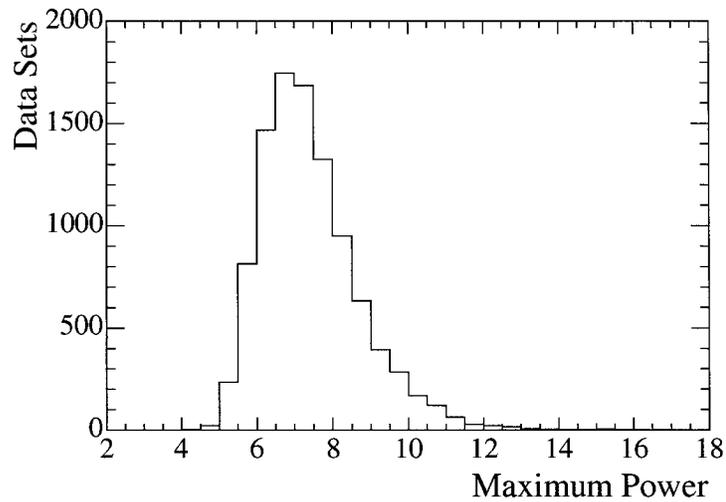


(a) Maximum powers from 10 000 D₂O phase null hypothesis MC data sets.

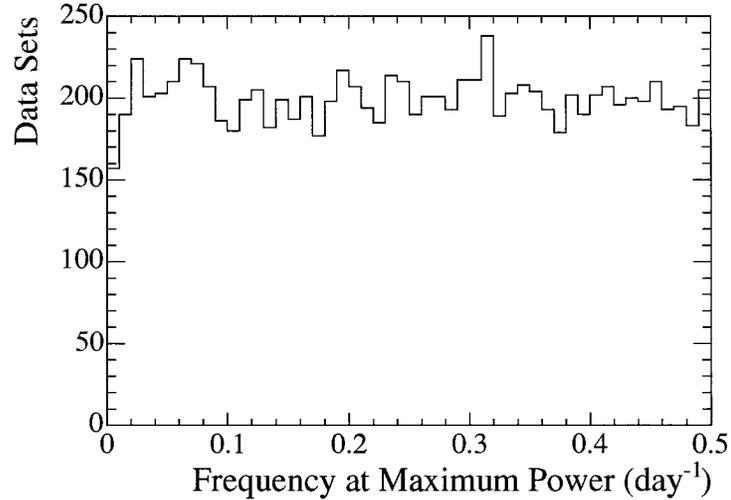


(b) Frequencies at the maximum powers shown in (a).

Figure 5.3: Distribution of (a) maximum powers and (b) frequencies at the maximum powers for the 10 000 D₂O MC data sets. These MC have been generated to simulate the null hypothesis.

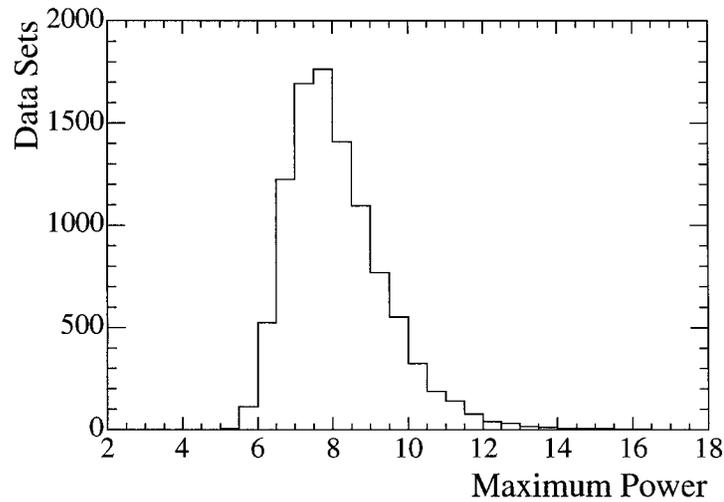


(a) Maximum powers from 10 000 salt phase null hypothesis MC data sets.

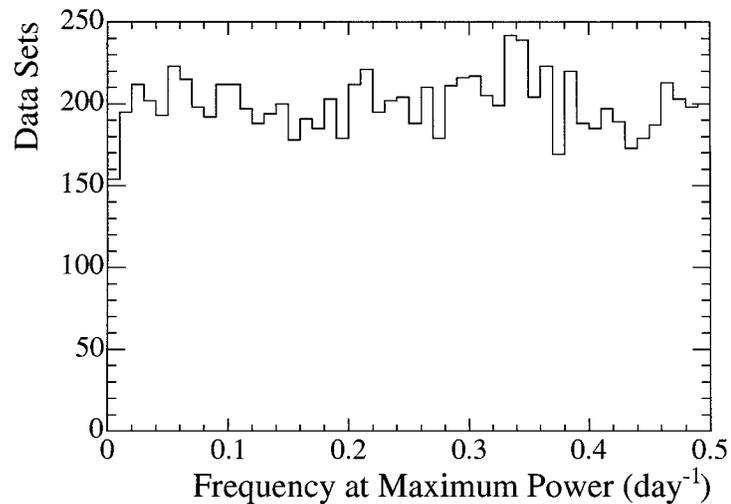


(b) Frequencies at the maximum powers shown in (a).

Figure 5.4: Distribution of (a) maximum powers and (b) frequencies at the maximum powers for the 10 000 **salt** MC data sets. These MC have been generated to simulate the null hypothesis.



(a) Maximum powers from 10 000 combined phase null hypothesis MC data sets.



(b) Frequencies at the maximum powers shown in (a).

Figure 5.5: Distribution of (a) maximum powers and (b) frequencies at the maximum powers for the 10 000 **combined** MC data sets. These MC have been generated to simulate the null hypothesis.

statistic bins are combined when the expected number of events are five or less. The expected number of events is a reflection of the bin livetime and run boundaries, a trait common to the data and all MC¹.

Using the 10 000 MC maximum powers, the probability for a power z , or greater, to be the result of a constant signal is found. For example, in the combined phase maximum power distribution, Figure 5.5(a), 21.7% of the MC data sets had a maximum power of 9.0 or greater. So if a maximum power of 9.0 is obtained in the SNO combined phase data sets, there is a 21.7% probability that this power would be obtained for a data set that is constant in time. This corresponds to a FAP of 21.7% in Equation (5.1).

In a similar way, given a probability, the corresponding power can be found. Table 5.1 shows the powers necessary to obtain 50%, 10%, 5%, and 1% probabilities of the null hypothesis. For the combined phase, if a power of 12.24 is found there is a 1% probability that this data set is constant. Therefore, if a power of 12.24 or greater is found the null hypothesis is rejected at the 99% CL. Otherwise, the null hypothesis can not be rejected. The maximum powers for the null hypothesis combined phase MC have been re-plotted on a logarithmic scale in Figure 5.6 with a smaller bin size. The darkened maximum powers are the 100 MC with maximum powers greater than or equal to 12.24.

5.2.2 Fit for the Number of Independent Frequencies

Equation (5.1) can be used in conjunction with the MC maximum power distributions to provide an estimate for the number of independent frequencies. Equation (5.1)

¹There is slight variation of the number of data points between MC files, since the expected number of events relies on the total number of events, and this varies between MC files. But this variation is small (RMS = 1.577 bins) and is not expected to affect this analysis.

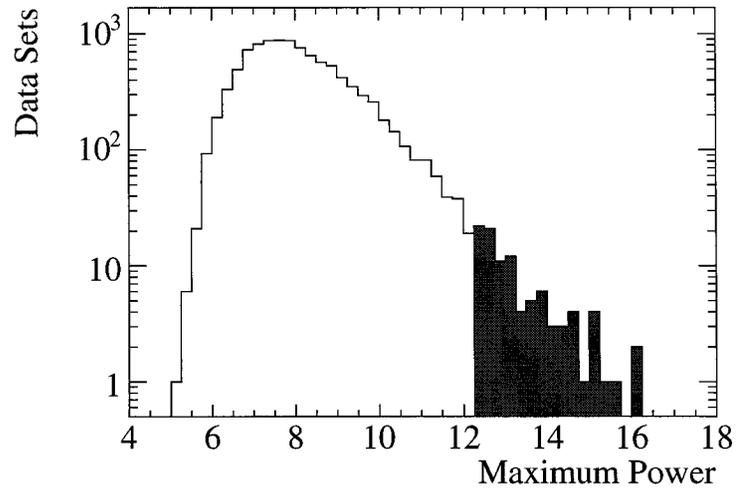


Figure 5.6: The 10 000 combined phase MC maximum powers, where the y-axis is shown with a logarithmic scale. One percent of the MC files with the largest maximum powers are shaded.

	50%	10%	5%	1%
D ₂ O	7.00	8.93	9.67	11.15
Salt	7.20	9.11	9.90	11.43
Combined	7.91	9.84	10.56	12.24

Table 5.1: Power values corresponding to the probability that noise fluctuations would yield this power in the various data sets under analysis. These powers have been found using the MC maximum power distributions.

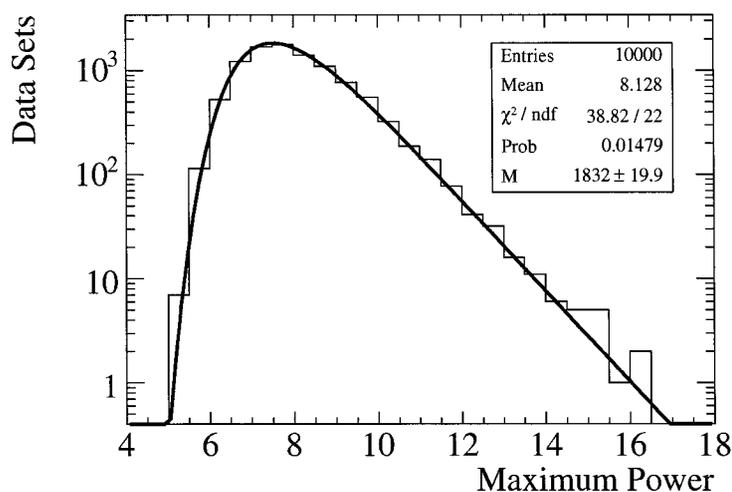


Figure 5.7: The 10 000 combined phase MC maximum powers, fit with Equation (5.4) to find the value for M .

is the reverse cumulative probability distribution of the powers. The cumulative probability distribution is just,

$$1 - F(z) = \text{Prob}(Z < z) = [1 - e^{-z}]^M. \quad (5.3)$$

Taking the derivative of this cumulative probability distribution,

$$f(z) = M e^{-z} (1 - e^{-z})^{(M-1)} \quad (5.4)$$

gives the probability density function (PDF). By fitting this PDF to the maximum power distributions in Figures 5.3(a), 5.4(a), 5.5(a), a value for M can be obtained. To demonstrate this, Figure 5.6 has been fit with Equation (5.4); this fit is shown in Figure 5.7. Table 5.2 shows the general fit results for every phase.

The small fit probabilities are because Equation (5.4) does not completely describe the maximum power distributions. However, they both provide a method of inter-

SNO Phase	Fit for M	χ^2/dof	Probability (%)
D ₂ O	740 ± 8	42.83 / 21	0.3
Salt	902 ± 10	36.59 / 23	3.6
Combined	1832 ± 20	38.82 / 22	1.5

Table 5.2: Using the MC maximum power distributions and Equation (5.4) to fit for the number of independent frequencies.

preting the powers found in the data periodograms. Substituting the values for M into Equation (5.1) one can find the powers necessary to obtain a FAP of 1%. For the D₂O data set this power is 11.23, for the salt data set this is 11.40, and for the combined data sets this power is 12.11. These values are in good agreement with the actual power values obtained using the MC maximum power distributions shown in Table 5.1.

The approximate fits for M prove that as a general trend, the number of independent frequencies increases with data, which impacts the interpretation of a given power in each data set.

5.2.3 Summary

The MC maximum power distributions are used to shed light on the results of the data, and determine the sensitivity of this analysis. The null hypothesis is rejected at the 99% CL in the D₂O data set if a power of 11.15, or greater, is obtained. This power is 11.43 for the salt phase, and 12.24 for the combined phases.

Periods (day)	2.1, 2.5, 5, 10, 20, 50, 100, 500, 1000
Amplitudes (%)	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18, 20, 22, 25

Table 5.3: 1000 MC data sets are generated for each period and amplitude pair.

5.3 Monte Carlo Sensitivity Study

Suppose the SNO data did have a periodic distribution of the neutrino times, given the methods employed in this thesis, would this periodic signal be distinguishable from the null hypothesis? The answer is a function of both the period and amplitude of the periodic signal. To determine the sensitivity to periodic signals in the SNO data, MC data sets are again utilized, as in the previous section, only now these data sets are generated with a sinusoidal distribution of the neutrino event times.

5.3.1 Generating Periodic Monte Carlo Data Sets

For each period and amplitude pair listed in Table 5.3 1000 MC data sets were generated for both the D₂O and salt phases. The MC data sets are generated using the accept/reject method, with the function

$$n(t) = A + B \sin(2\pi\nu t), \quad (5.5)$$

such that $A + B = 1$, A is greater than or equal to B , and B/A is the signal amplitude, ν is the signal frequency, and t is the candidate neutrino event time. The time, t , is randomly generated, and passes the first test if it occurs within the approved runs, as before. However, now event times are accepted only if the value of $n(t)$ is greater than a second random number generated in the range $[0, 1]$.

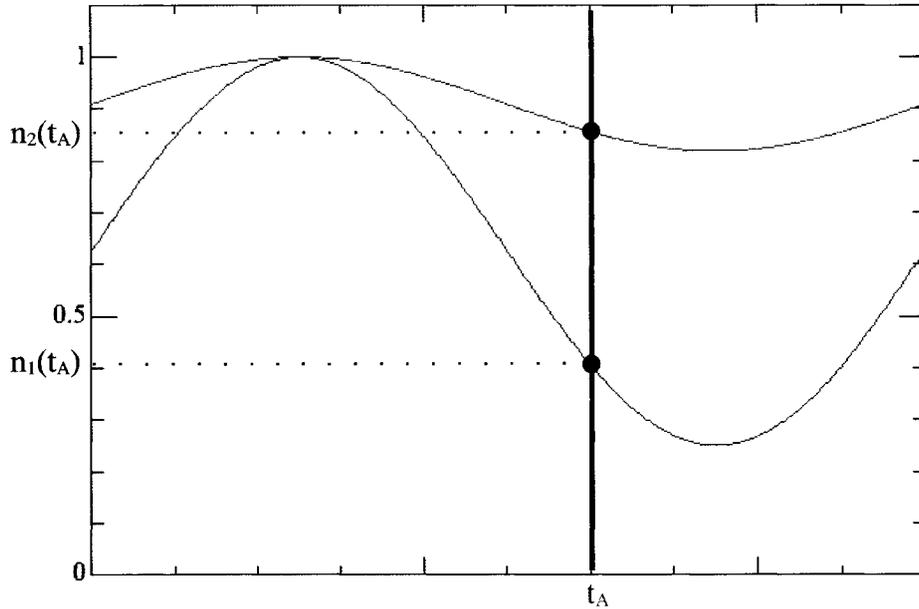


Figure 5.8: Demonstration of sinusoidal functions used in the accept/reject method to generate the periodic MC. Both functions $n_1(t)$ and $n_2(t)$ have the same frequency. The function $n_1(t)$ has a larger amplitude of 60%, the function $n_2(t)$ has an amplitude of 10%.

Figure 5.8 demonstrates this idea. Consider two sinusoidal functions obeying Equation (5.5), each with the same frequency ν_o , but $n_1(t)$ with an amplitude of

$$B_1/A_1 = 0.60 \quad (\text{where } A_1 = 0.625, B_1 = 0.375), \quad (5.6)$$

and $n_2(t)$ with an amplitude

$$B_2/A_2 = 0.10 \quad (\text{where } A_2 = 0.909, B_2 = 0.091). \quad (5.7)$$

Now let the random time value be some value t_A . The corresponding values for $n_1(t_A)$ and $n_2(t_A)$ are marked on the y-axis. Now a second random number is generated in the range $[0, 1]$. If this second number is less than the value of $n_i(t_A)$ (and the event is

within the run boundaries) then the time t_A is accepted for the periodic MC following $n_i(t)$. For all $n(t)$, the event time has the greatest probability of being accepted if the time t sets $\sin(2\pi\nu t) = +1$, since $n(t)$ will be equal to 1, $(A + B)$. The probability of accepting a time is at a minimum when $\sin(2\pi\nu t) = -1$ because $n(t)$ will have the lowest value $(A - B)$. This implies the event distribution will be periodic in time with the form,

$$x(t) = A + B \sin(2\pi\nu t) . \quad (5.8)$$

This method creates periodic behavior in the neutrino rates. The periodic MC listed in Table 5.3 are generated according to this method and are used to determine if the analysis is sensitive to signals with frequencies in the search range.

5.3.2 Sensitivity Analysis

Selected periodic MC from Table 5.3 are analyzed using the same LS analysis as all previous data and null hypothesis MC data sets. For the 1000 periodic MC data sets of one periodic behavior, defined by a period and amplitude, the maximum powers are found. These maximum powers are used to calculate the probability of rejecting the null hypothesis, for this periodic behavior.

Sample distributions of the maximum powers obtained for a 20 day period with amplitudes of 10%, 15%, 20%, and 25% in the combined phase are shown in Figure 5.9, with the null hypothesis shown in black at the far left. As the amplitude of the periodic signal is increased, the average mean maximum power also increases, increasing the ability to resolve the signal from the null hypothesis.

For the same 20 day period, Figure 5.10 separately examines the maximum power distributions at amplitudes of 8%, 9%, 10%, and 11%. With an amplitude of 8%,

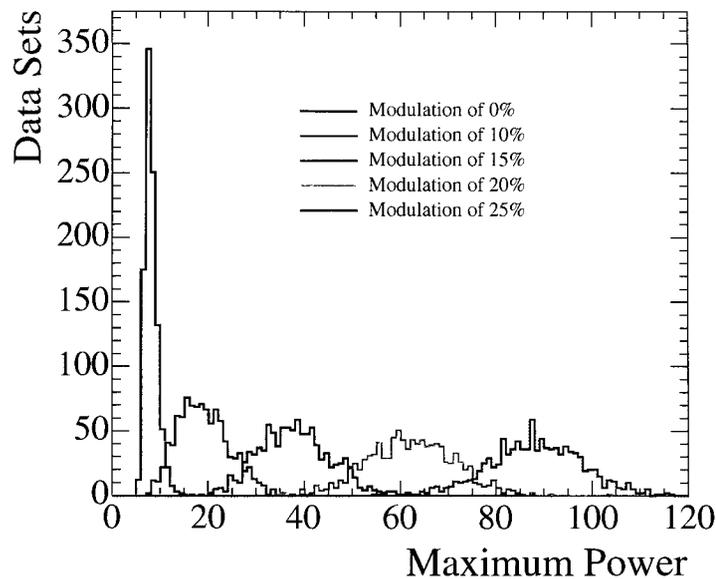


Figure 5.9: Maximum powers obtained for the combined phase periodic MC data sets with periods of 20 days, and amplitudes of 10%, 15%, 20%, and 25%. Each period and amplitude pair has 1000 entries.

shown in Figure 5.10(a), 48.8% of the 1000 periodic MC data sets had a maximum power greater than 12.24. If the combined SNO data set has this periodic behavior, the analysis has a 48.8% probability of rejecting the null hypothesis. Figure 5.10 again shows that as the amplitude increases, the probability of detecting the signal and rejecting the null hypothesis also increases. This trend is true for all periods.

The sensitivity of the method to periodic signals is also seen by plotting the frequency at these maximum powers. Again, using the 20 day period as the example, Figure 5.11 shows the distribution of frequencies at maximum power at four different amplitudes. The frequency of the maximum power is often found at the exact signal frequency, or close neighboring frequencies. Neighboring frequencies are often found with large powers because the number of independent frequencies is less than the number of sampled frequencies, due to the limited data set. This idea is examined

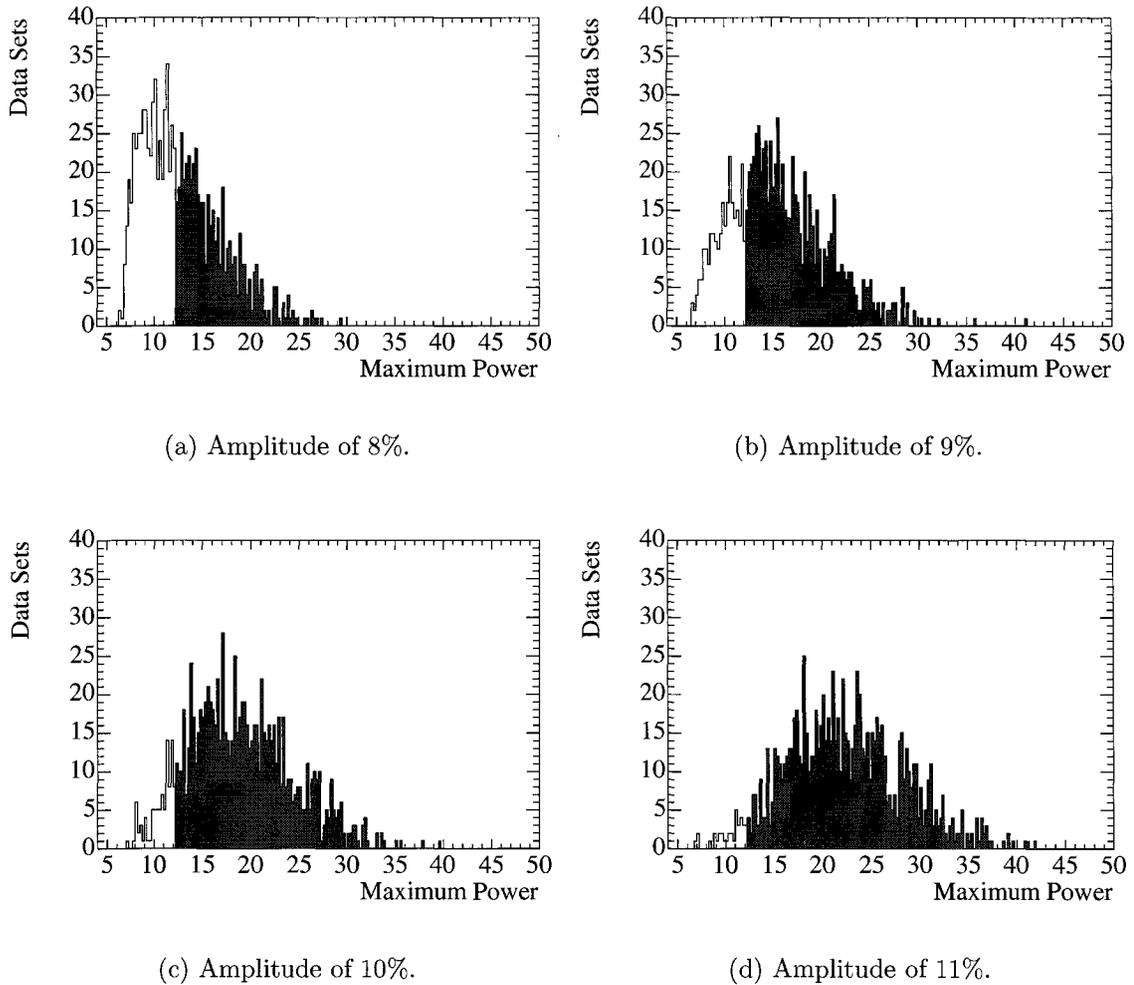


Figure 5.10: Distribution of maximum powers for a 20 day period with amplitudes 8%, 9%, 10%, and 11%, for the combined data sets. The shaded portions of the distributions show those MC data sets with maximum powers greater than 12.24, and represent the probability of rejecting the null hypothesis.

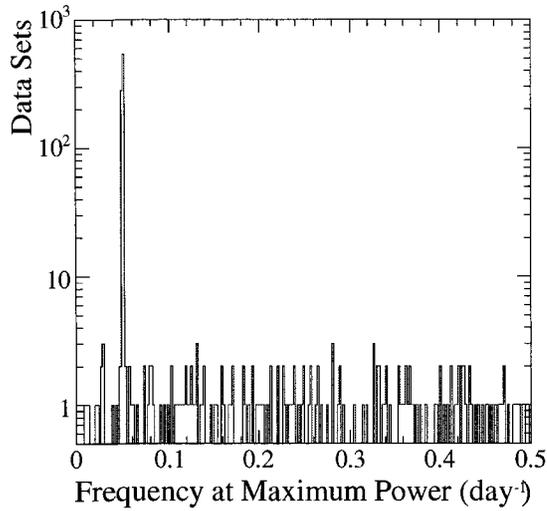
in Chapter 6. With increasing amplitude the probability of finding the correct signal frequency also increases.

If the method does not find the correct signal frequency, the maximum powers are smaller. In this case, the frequency at maximum power is likely to be random within the sampled frequency range. A scatter plot of the 1000 maximum powers versus frequency at maximum power is shown in Figure 5.12 for a 20 day period with an amplitude of 10%. The largest maximum powers correspond to the MC data sets where the signal frequency is correctly identified. The ability of the analysis to find the correct signal frequency is another tool to identify the sensitivity.

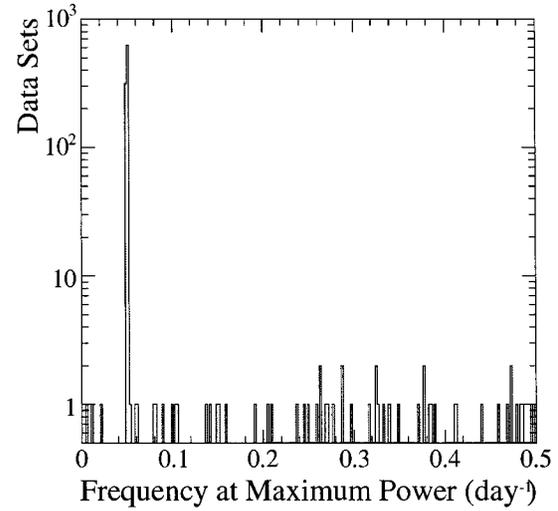
The LS analysis rejects the null hypothesis if the power in the combined data set is greater than, or equal to, 12.24. Hence, the maximum powers in the periodic MC are used to determine the sensitivity of the analysis, by finding the percentage of data sets with maximum powers greater than this value.

Table 5.4 shows the percentage of data sets with a maximum power greater than, or equal to, the power necessary to reject the null hypothesis, for the different periodic MC. This table shows a limited range of the period-amplitude space that was analyzed for each phase. The full range of results is provided in Appendix B.

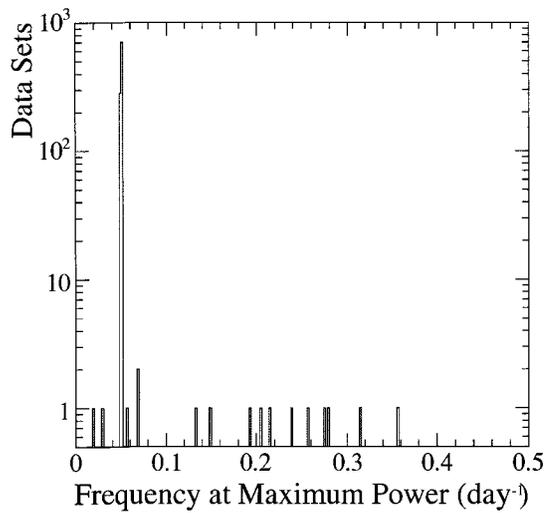
The percentage of MC data sets with a power greater than, or equal to, the 99% CL power reflects the probability of detecting this signal from the null hypothesis. As the quantity and time interval of data increases, the sensitivity to periodic signals increases. For example, a signal with a period of 50 days and an amplitude of 10% the D₂O phase has a 21.6% probability of detecting this signal. This probability increases to 59.4% in the salt phase, and 87.8% in the combined data sets.



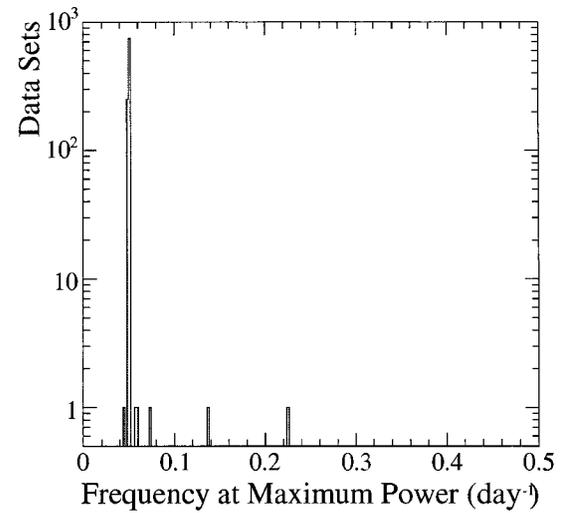
(a) Amplitude of 8%.



(b) Amplitude of 9%.



(c) Amplitude of 10%.



(d) Amplitude of 11%.

Figure 5.11: Frequencies at maximum power for a 20 day period at amplitudes of 8%, 9%, 10%, and 11% for the combined data sets. A 20 day period corresponds to a frequency of 0.05 day^{-1} .

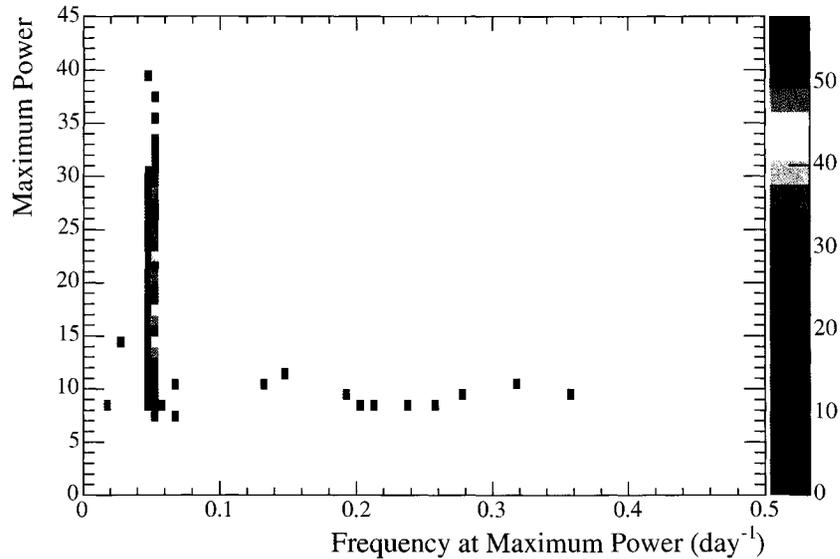


Figure 5.12: Maximum power versus frequency at maximum power for 1 000 combined MC data sets with a period of 20 days and amplitude of 10%.

5.3.3 Sensitivity Region

For each period, the amplitude at which there is a 50% and 90% probability of rejecting the null hypothesis is found by interpolating the known values, from Appendix B. The relationship between the amplitude of a signal and the probability of detecting the signal is not linear, but it is taken to be linear in this interpolation, since the upper and lower points lie close together.

The 50% and 90% probabilities of rejecting the null hypothesis for the combined phase are shown as a sensitivity contour in Figure 5.13. The x-axis shows a logarithmic scale of period, and the y-axis shows the amplitude. The space above the 50% and 90% lines corresponds to the increasing ability to detect signals with periods and amplitudes in this region. For example, if the SNO data are periodic with a 10 day, 11% amplitude signal, there is a 90% probability of detecting this signal.

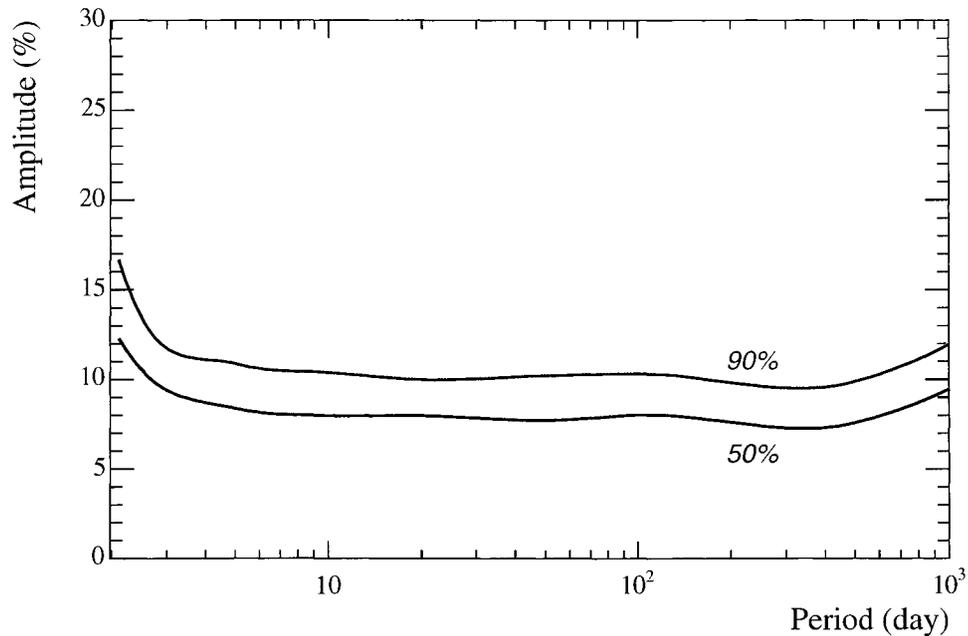


Figure 5.13: The sensitivity contour for the combined data set. The region above the contour lines represents increasing sensitivity to periodic behavior.

At periods smaller than 10 days the detection sensitivity decreases. This is expected since the data are assembled into unevenly spaced bins, and the time separation between these bins is on the order of days. As the period increases towards 1000 days, the limited calendar span of the data reduces the sensitivity of detecting these periods.

5.3.4 Summary

The sensitivity analysis shows that the LS method would be able to detect a periodic signal if it were present in the SNO data set. In general, there is a 50% probability of detecting a signal, and rejecting the null hypothesis at the 99% CL in the combined SNO data sets, if the data are periodic with a period of 5 to 500 days, with an amplitude greater than 8%.

D₂O		Amplitude (%)					
		10	11	12	13	14	15
Period (day)	2.5	4.1	7.2	9.9	12.8	18.9	28.6
	5.0	13.7	21.2	28.9	40.1	52.7	65.2
	20.0	19.8	29.0	40.6	54.6	69.7	78.4
	50.0	21.6	33.3	44.8	61.4	72.9	83.7
	100.0	18.6	28.9	42.2	56.1	69.7	80.7
	1000.0	1.1	1.9	2.2	3.6	4.0	5.1
Salt		Amplitude (%)					
		10	11	12	13	14	15
Period (day)	2.5	15.3	25.9	36.6	52.5	65.6	75.0
	5.0	43.9	63.4	78.9	87.9	94.5	98.8
	20.0	56.0	68.6	82.4	91.8	97.8	99.3
	50.0	59.4	75.1	88.4	95.8	98.6	99.6
	100.0	56.2	75.4	86.0	93.9	98.0	99.5
	1000.0	31.6	42.6	60.3	75.4	87.6	93.0
Combined		Amplitude (%)					
		10	11	12	13	14	15
Period (day)	2.5	43.4	57.4	75.8	87.6	94.1	97.3
	5.0	80.9	91.2	96.8	99.3	99.7	100.0
	20.0	90.4	96.5	99.3	99.7	100.0	100.0
	50.0	87.8	96.7	99.3	99.8	100.0	100.0
	100.0	87.6	96.4	99.2	99.5	100.0	100.0
	1000.0	58.9	76.7	90.5	97.5	99.3	99.9

Table 5.4: The percentage of periodic MC data sets with maximum powers greater than, or equal to, the 99% CL power. The 99% CL powers have been taken from the MC null hypothesis analysis, and correspond to 11.15 for the D₂O phase, 11.43 for the salt phase, and 12.24 for the combined data.

Chapter 6

Systematic Checks

The LS method, described in Chapter 2, was applied to the SNO data. There were imprecise parameters in this method, such as bin size and number of sampled frequencies. These parameters were defined in Chapter 4, and were based upon information in the data, such as the time span of the data, daily rates, and run boundaries. To ensure that these choices have not imposed a bias in the results, systematic checks of these variables have been performed, and are discussed in this chapter. This chapter also examines potential systematic effects of the SNO experiment, specifically the variability of the detector acceptance, and time varying backgrounds, that could potentially impact the results of this periodicity analysis.

Unless otherwise stated, all of these checks have been performed on the combined data sets using the modified LS method.

6.1 Alternative Bin Size

To test the robustness of the LS periodogram, the SNO data from the combined phase have been reanalyzed using 101 different bin sizes in the range of 1 to 5 days, with a step size of 3456 seconds. To reduce effects of aliasing, the maximum frequency sampled is modified for each bin size to reflect that predicted by the Nyquist frequency assuming even spacing of the data. The frequency interval of $\frac{1}{(4 \times 3650)} \text{ day}^{-1}$ is maintained throughout, and only the number of frequencies sampled, N_{freq} , is changed,

$$N_{\text{freq}} = (4 \times 3650) / (2 \times \text{bin size}), \quad (6.1)$$

where the bin size is in units of days. For example, a 3-day bin samples 2433 frequencies, up to a maximum frequency of $1/6 \text{ day}^{-1}$, and a 5-day bin samples 1460 frequencies, to a maximum of $1/10 \text{ day}^{-1}$.

6.1.1 Data Results

The maximum power for each of the 101 bin sizes is shown in Figure 6.1. The largest maximum power is 11.05, and occurs when the data are put into a bin size of 3 days. The period of this maximum power is found to be 10 days. The distribution of the maximum power is wide, and reflects the change in the number of data points and sampled frequencies with each trial.

The frequency at each maximum power is shown in Figure 6.2. The frequency at maximum power is not the same for all bin sizes. The number of times a given frequency appears with the largest power appears to fall off with increasing frequency. This is because with increasing bin size the analysis is sampled to a smaller maximum frequency.

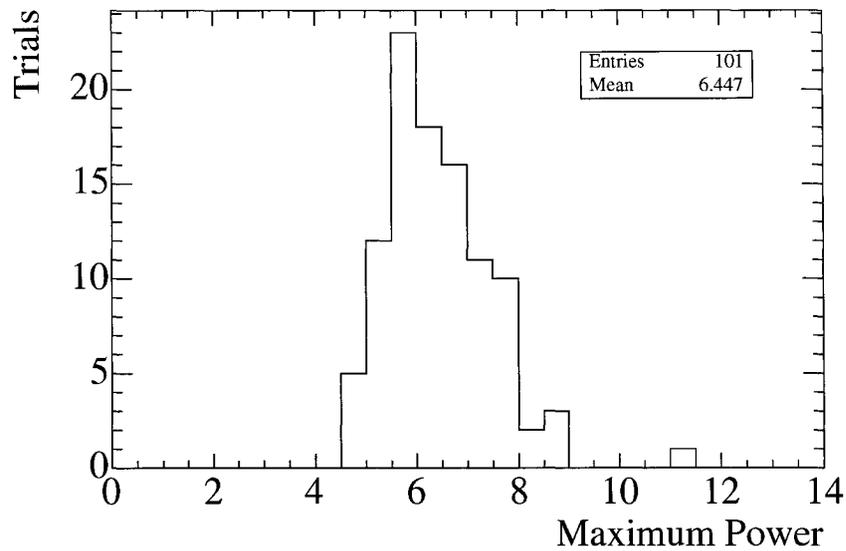


Figure 6.1: The maximum power found at each of the 101 different bin sizes in the range of 1 day to 5 days for the combined phase data. The largest maximum power is 11.05 at a frequency of 0.10089 day^{-1} when the data are binned in 3 days.

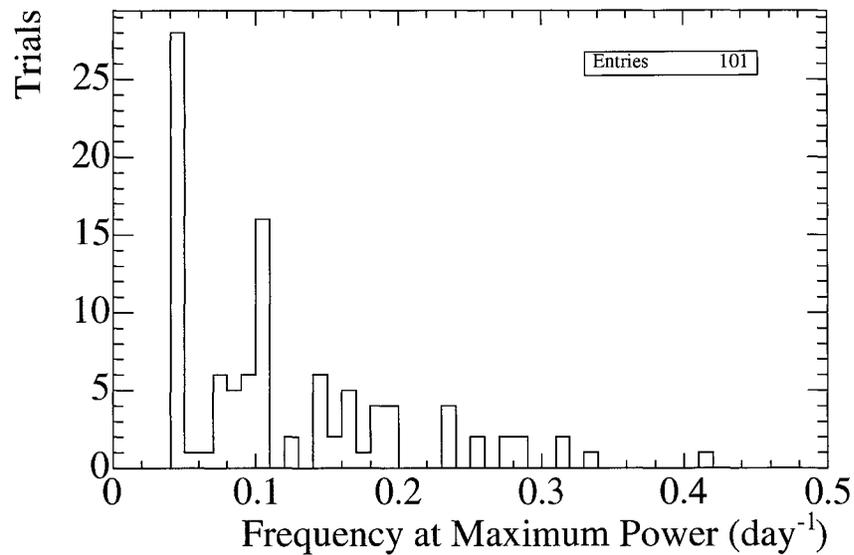


Figure 6.2: The frequency at maximum power for 101 different bin sizes in the range of 1 day to 5 days, for the combined phase data.

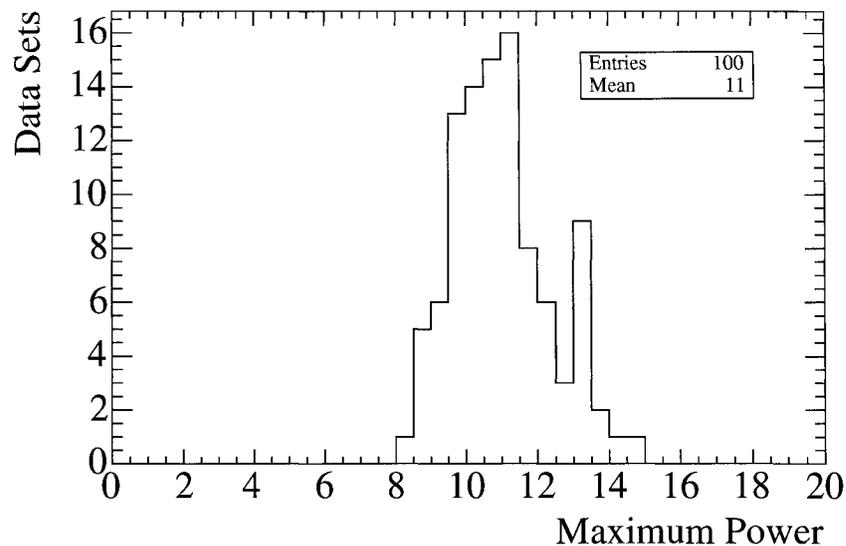


Figure 6.3: The maximum power in 100 MC files that have been sampled with 101 different bin sizes.

6.1.2 Monte Carlo Study

MC data sets are used to determine the FAP of the largest maximum power of 11.05, found when the data are put into 3-day bins. There is a trials penalty that must be considered when reanalyzing the same data in multiple ways. Joint FAP functions must be examined to determine the likelihood of a power when multiple methods are simultaneously applied to the same data set. This is equivalent to sampling the same number of frequencies in the MC, as in the data, to determine the FAP of periodogram powers in the data.

A total of 100 MC event files with a constant rate, were analyzed with the same 101 bin sizes. Figure 6.3 shows the maximum power found in each of the 100 MC when they are analyzed with the same 101 bin sizes. The maximum power in the data, of 11.05, falls approximately in the middle of this MC distribution. Hence, the largest maximum power in the data falls within the expected range of a data set

that resembles a constant rate. In other words, there is a 50% chance that this data resembles a constant rate when analyzed with 101 different bin sizes.

Thus, the SNO data remains consistent with pure noise, even when analyzed with a multitude of bin sizes.

6.2 Vary Bin Start Boundary

In the primary analysis, the first bin begins at $t_0 = 0.00$ seconds on the day of the earliest run. All bin boundaries are then defined with reference to this starting value, t_0 , so all bins begin at $t = 0.00$ seconds on their respective days.

To show that the results of the LS analysis do not change under this choice, the bin start boundary is allowed to vary in increments of 4320 seconds (an arbitrary choice). On a given day, there are 20 choices for t_0 . Many of the individual bin rates, $y(t_i)$, and times, t_i , will change with each definition of t_0 , since the new bin boundaries may distribute the events and livetime differently among neighboring bins.

The default bin start time is $t_0 = 0.00$ seconds on November 2, 1999. If instead it is defined as $t_0 = 0.00$ seconds on November 1, 1999, there is no change in the periodogram, since a shift by exactly 24 hours when putting the data into 1-day bins should not effect the result, as long as all of the data are still included. Figure 6.4 shows the periodograms when $t_0 = 0.00$ seconds on November 1 and November 2. As expected, both periodograms are identical.

This is true for all times shifted by 24 hours, before the start of the first run. Using this logic, it is only necessary to examine the power spectrum for the 20 different t_0 values within a 24 hour period. Figure 6.5 shows the maximum powers and frequencies at each maximum power obtained for each of the 20 different t_0 values. The

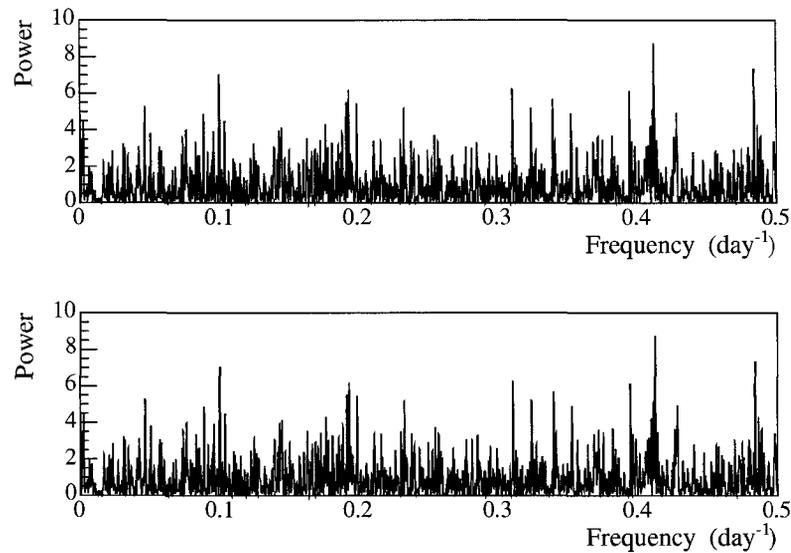


Figure 6.4: Periodograms when $t_0 = 0.00$ seconds on November 1 (top) and November 2 (bottom), 1999.

frequency of maximum power at 0.4 day^{-1} occurs in 40% of the trials. The deviation in the maximum powers is not large, with the largest maximum power being 8.98 and smallest 6.08. This small deviation is expected, since it is the same data set, only with events and livetime shuffled between neighboring bins.

The 1% FAP power for the combined data set analyzed with one value of t_0 is 12.24, as explained in Chapter 5. By increasing the number of trials applied to the same data set, the 1% FAP power will increase to a value greater than 12.24. Therefore, this analysis concludes that there is no significant peak in the data when the value of t_0 is varied, and the choice of $t_0 = 0.00$ seconds on the earliest run day does not mask, or falsely produce, a signal.

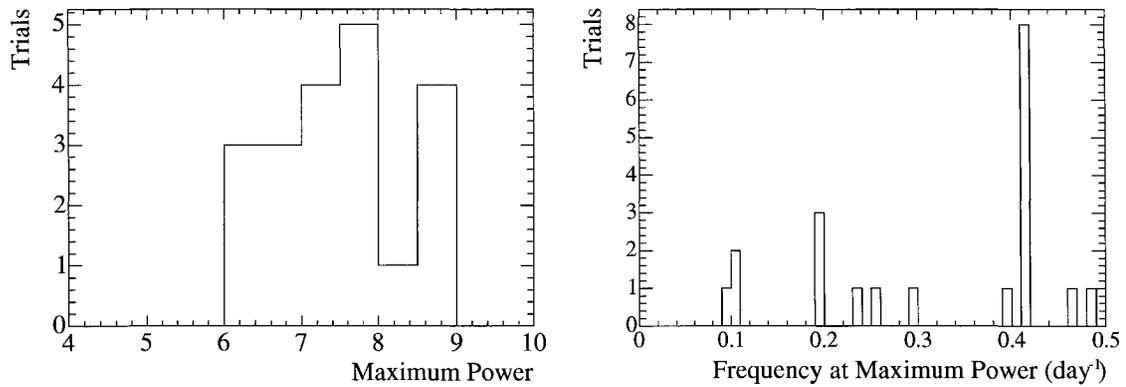


Figure 6.5: Distribution of (a) maximum powers and (b) frequencies at the maximum powers when the data are analyzed with 20 different values for the start bin boundary, t_0 .

6.3 Sampled Frequencies

The SNO data are fit with 7300 frequencies in the range of $1/3650 \text{ day}^{-1}$ to $1/2 \text{ day}^{-1}$, at a constant interval of $1/(4 \times 3650) \text{ day}^{-1}$. The large number and range of frequencies was chosen because it was not known *a priori* the frequency of a periodic solar neutrino rate. To maximize the range of frequencies, the full calendar span of the data set and bin size were used as limiting parameters in choosing the minimum and maximum sampled frequencies. This section examines the frequencies that are sampled to determine if they are sufficient to find a signal within the sampled range, even if the exact signal frequency is not explicitly sampled. This relates to the ability of resolving neighboring frequencies, given the finite data. If the sampling width between consecutive sampled frequencies is sufficient then for a signal that is not one of the 7300 sampled, the next closest frequency, and possibly its neighbors, should have powers reflecting the power obtained if this signal had been sampled.

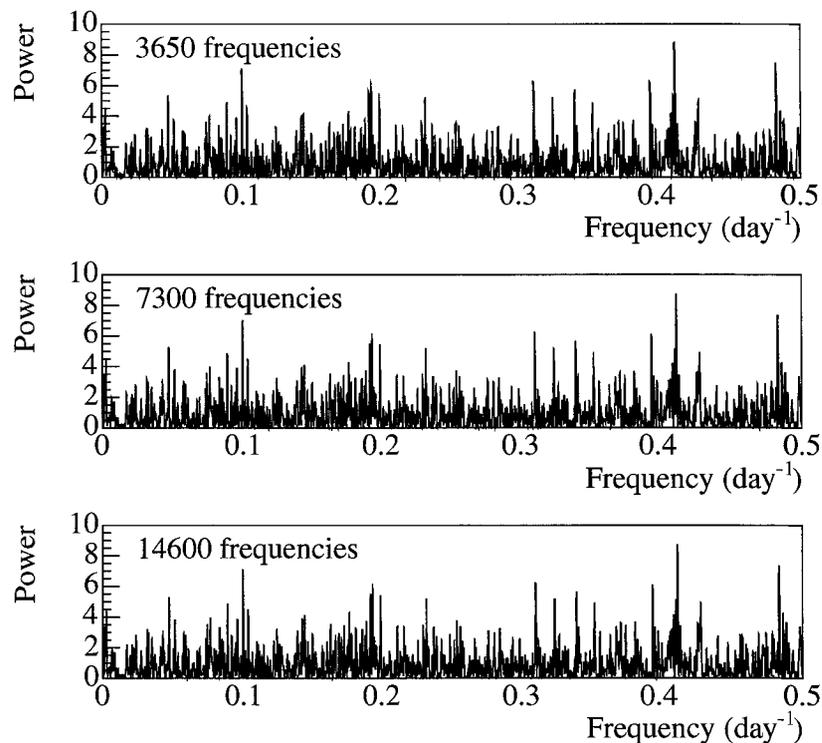


Figure 6.6: Periodograms for the combined phases when the data are fit to 3650, 7300, and 14 600 frequencies in the range of $1/3650 \text{ day}^{-1}$ to $1/2 \text{ day}^{-1}$.

6.3.1 Sampling the Data Against Alternative Frequencies

As mentioned above, the original analysis sampled 7300 frequencies. The data have been reanalyzed with half the number of frequencies, 3650, at a frequency interval twice as large, $1/(2 \times 3650) \text{ day}^{-1}$. And then re-analyzed with twice as many frequencies, 14 600, at an interval of $1/(8 \times 3650)$. The same frequency range is maintained in all cases. Figure 6.6 shows the combined phase periodograms for these three cases, with increasing number of sampled frequencies. These periodograms are essentially identical, and again no significant periodic behavior is found in any case.

The width of a frequency peak is a reflection of the dependence between neighboring frequencies. These periodograms can be used to determine if the number of

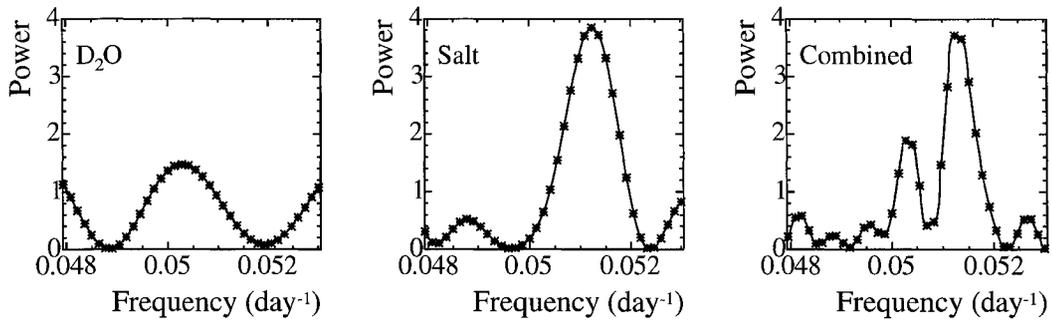
independent frequencies changes with the number of sampled frequencies, by examining the frequency widths. Figure 6.7 shows the data periodograms for the D₂O, salt, and combined phases, zoomed into the frequency range 0.048 – 0.053 day⁻¹, for the 3650, 7300, and 14 600 sampled frequencies. The amount of data increases with each phase, moving from left to right across a row. For this case, as the amount of data increases, the width of the frequency peak decreases, reflecting an increase in the number of independent frequencies.

By looking from top to bottom along one column it is seen that increasing the number of sampled frequencies from 3650 does not alter the frequency peak widths. In other words, the independence of neighboring frequencies does not change when more frequencies are sampled. Indicating that sampling 7300 frequencies does not miss a signal that might otherwise be sampled in a coarser sampling interval.

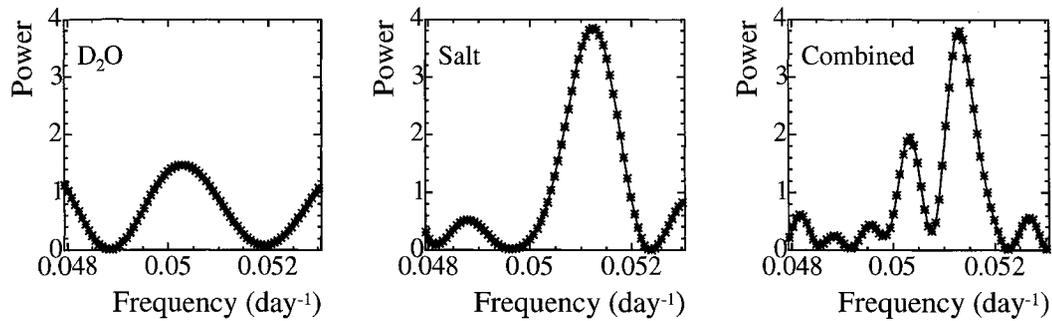
6.3.2 Monte Carlo Event Files

The width of the peaks shown in Figure 6.7 is directly related to the dependence between neighboring frequencies. To determine if this width, and dependence on neighboring frequencies is constant for all frequencies, MC data sets with periodic behavior are used. The periodic MC, described in Chapter 5, are generated with specific known frequencies. The dependence of powers between neighboring frequencies implies that for sufficient amplitudes, the signal frequency will have a large LS power, along with the neighboring frequencies.

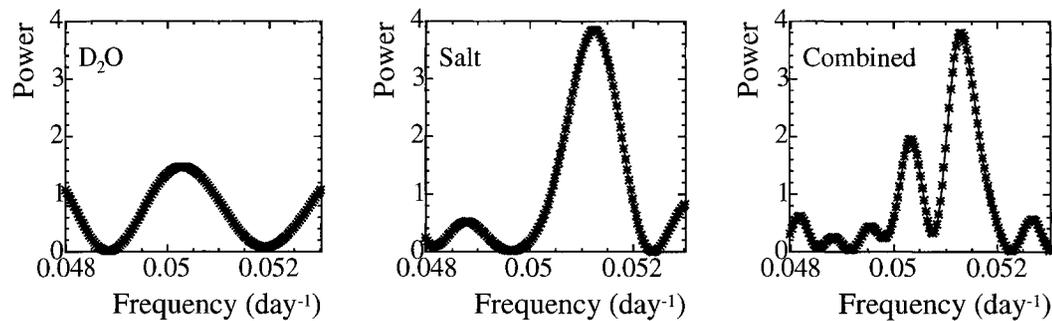
Consider periodic MC with signal frequencies of 0.4, 0.2, 0.1, 0.02, 0.01, and 0.02 day⁻¹, each with an amplitude of 15%. Figure 6.8 shows a portion of the periodograms for each of these six types of MC, centered on the signal frequency, shown within ± 0.00025 day⁻¹. Each of these distributions has been fit with a Gaussian function



(a) 3650 frequencies sampled.



(b) 7300 frequencies sampled.



(c) 14 600 frequencies sampled.

Figure 6.7: Data periodograms when (a) 3650, (b) 7300, and (c) 14 600 frequencies are sampled, zoomed into the frequency range of 0.048 – 0.053 day⁻¹. In the figure the points represent frequencies that are actually sampled, the smooth contour line joins these points.

to determine the peak width. For the signal frequency of 0.4 day^{-1} the Gaussian fit probability is 99.9% and fits the mean frequency as 0.4 day^{-1} and standard deviation as $0.000289 \text{ day}^{-1}$. This fit is done for all of these signal frequencies, within $\pm 0.00025 \text{ day}^{-1}$ of the imposed signal frequency. For all six signals the fit to a Gaussian is good, with the smallest fit probability of 7% at a frequency of 0.002 day^{-1} . Fitting the standard deviation of the Gaussian distribution is a good measure of the width, and hence the dependence between neighboring frequencies. Note that in all cases the sampled neighboring frequencies also have large powers, greater than the 99% CL power.

This exercise has been repeated for a total of 100 MC data sets at each signal frequency listed above. The Gaussian fits for sigma are shown in Figure 6.9. The general width for all cases is on the order of 0.0003 day^{-1} . A scatter plot of the signal width and frequency is shown in Figure 6.10. It appears from this limited test that the dependence upon neighboring frequencies is constant across the frequency domain. When 7300 frequencies are sampled in the combined data sets, powers at signal frequencies tend to influence several neighboring sampled frequencies on either side of this signal frequency.

6.3.3 Summary

It has been shown that sampling 7300 frequencies is sufficient to find a significant periodic signal within the sampled frequency range. The number of independent frequencies is less than the number sampled frequencies, which results in a correlation between neighboring frequencies. The degree of dependence appears to be constant across the frequency domain.

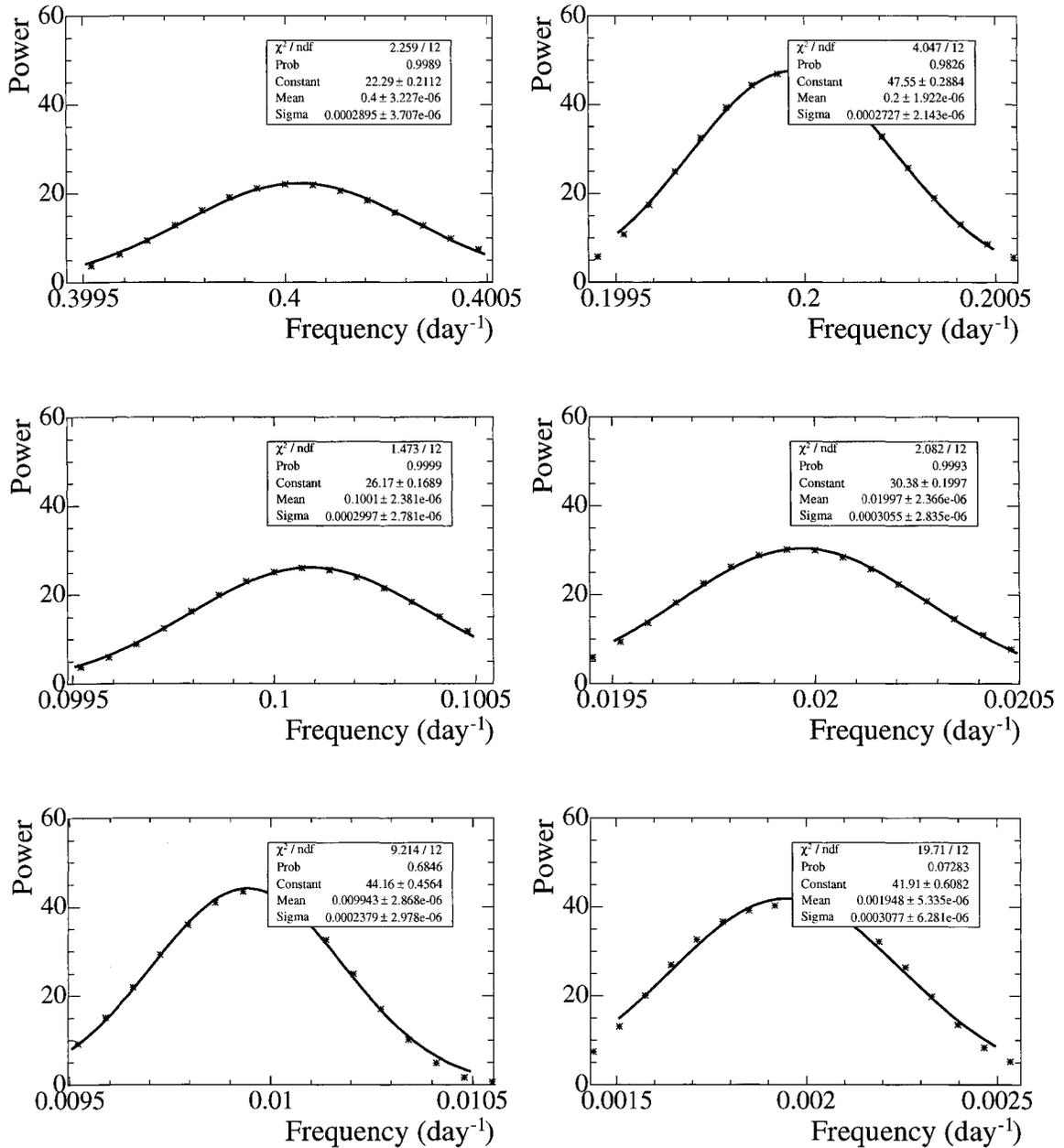
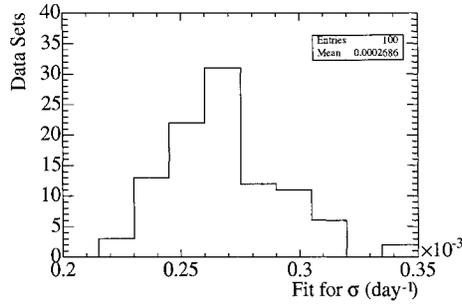
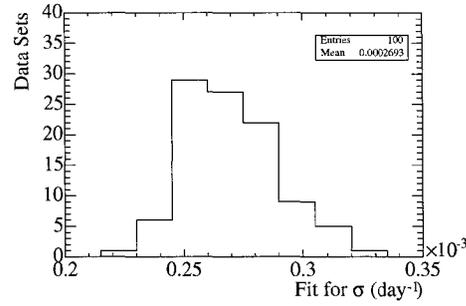


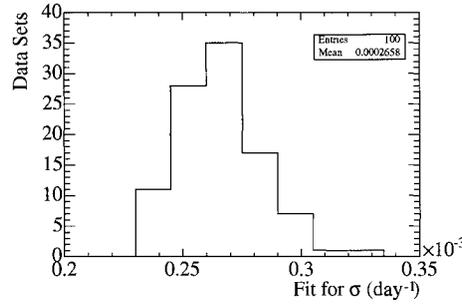
Figure 6.8: Periodograms of periodic MC zoomed into their respective signal frequencies. A Gaussian function is fit to these distributions.



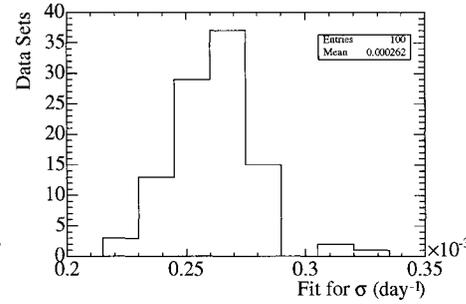
(a) $f_s = 0.4 \text{ day}^{-1}$, mean σ fit = $0.000269 \text{ day}^{-1}$.



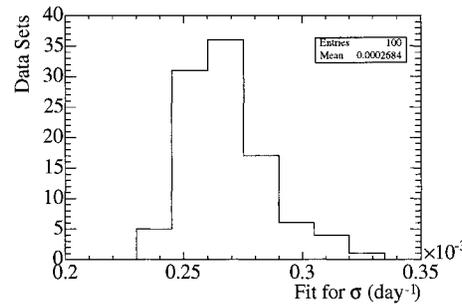
(b) $f_s = 0.2 \text{ day}^{-1}$, mean σ fit = $0.000269 \text{ day}^{-1}$.



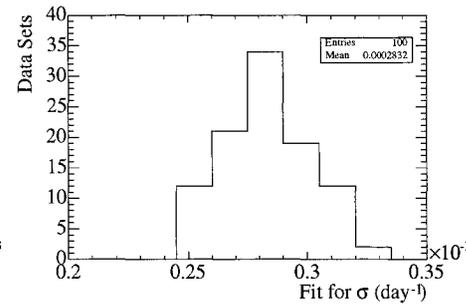
(c) $f_s = 0.1 \text{ day}^{-1}$, mean σ fit = $0.000266 \text{ day}^{-1}$.



(d) $f_s = 0.02 \text{ day}^{-1}$, mean σ fit = $0.000262 \text{ day}^{-1}$.



(e) $f_s = 0.01 \text{ day}^{-1}$, mean σ fit = $0.000268 \text{ day}^{-1}$.



(f) $f_s = 0.002 \text{ day}^{-1}$, mean σ fit = $0.000283 \text{ day}^{-1}$.

Figure 6.9: Fits for the frequency width surrounding periodograms for periodic MC. The variable f_s indicates the signal frequency. The mean fit for σ from the Gaussian fits is shown under each plot.

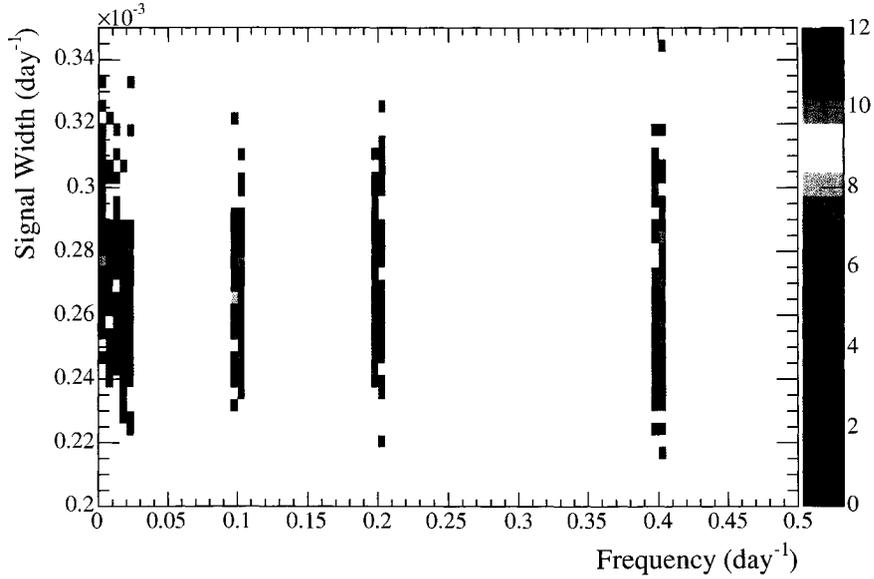


Figure 6.10: Width of a frequency peak versus the mean frequency surrounding the signal frequency for periodic MC. 100 MC data sets have been analyzed for each of the six signal frequencies.

6.4 Combining Condition

In this LS periodicity analysis low statistic 1-day bins are combined with neighboring bins. This is done to ensure an approximate Gaussian distribution of the rate, and avoid possible biasing effects that low statistic bins might create. Low statistic bins are characterized by the amount of livetime in a one day bin, and are defined as any bin that has an expected number of events less than, or equal to, five. The expected number of events in bin i was defined in Equation (4.2) as:

$$(\text{expected number of events}) = \left[\frac{\text{total number of events}}{\text{total livetime}} \times (\text{bin livetime}) \right]. \quad (6.2)$$

where the total number of events and livetime is across the respective phases.

This section compares aspects of the modified LS analysis under two scenarios:

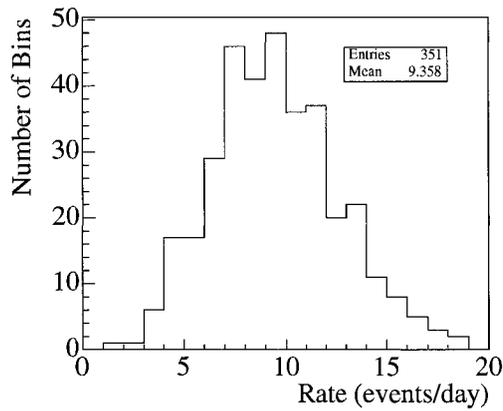
1. When low-statistic bins are combined if the number of expected events in a bin are less than, or equal to, five.
2. When no bins are combined.

For the D₂O phase, there are a total of 435 bins that have livetime, upon combining low statistic bins, 351 bins remain. For the salt phase, there are a total of 518 bins that have livetime, and after combining low statistic bins, 469 bins remain. Figure 6.11 shows the distribution of 1-day rates for the D₂O and salt phases when low statistic bins are, and are not, combined. When livetime correcting the raw number of events to get a 1-day rate, the raw number of events in a bin is multiplied by a factor proportional to the livetime. For bins with a small amount of livetime, it is possible that the raw number of events will be zero. Thus, when multiplied by any value, bins with zero events will have a 1-day rate of zero events, which may not be an accurate representation of the true rate. This explains why Figures 6.11(b) and 6.11(d) show a large, anomalous peak in the zero events/day histogram bin.

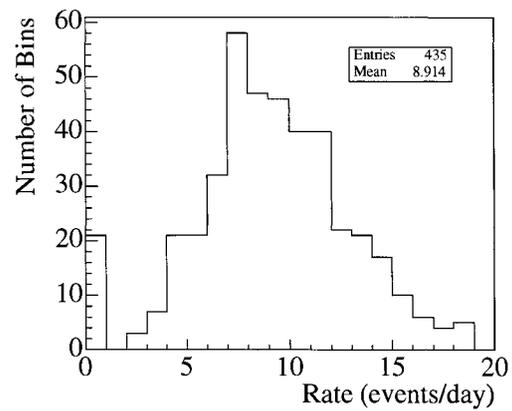
When these low statistic bins are combined, there is a decrease in the number of data points that enter into the LS periodogram calculation. For the D₂O and salt phases this represents a loss of 84 and 49 bins, respectively. Does this loss of information affect the resulting LS periodogram? And, to what extent is the sensitivity of the method compromised?

6.4.1 Analysis of Combined Data Set

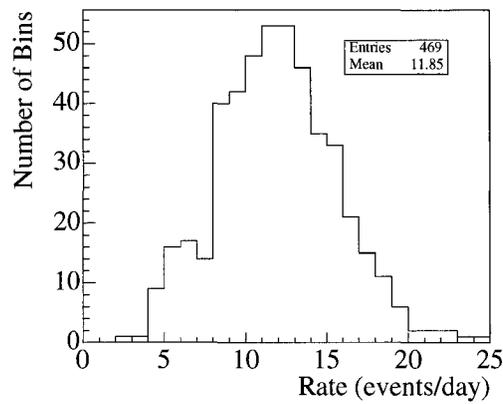
To address the first question, the LS analysis has been performed on the combined data sets using both scenarios. Figure 6.12 shows a scatter plot of the power obtained at each frequency when bins are not combined, versus when bins are combined. The



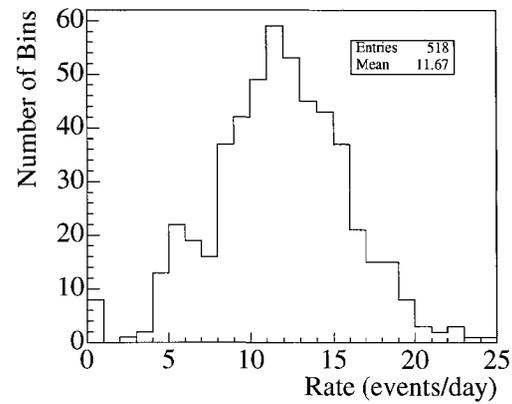
(a) D₂O phase, when low statistic bins are combined.



(b) D₂O phase, when low statistic bins are *not* combined.



(c) Salt phase, when low statistic bins are combined.



(d) Salt phase, when low statistic bins are *not* combined.

Figure 6.11: One day bin rates for the D₂O phase when bins (a) are, and (b) are not combined. Similarly, for the salt phase when bins (c) are, and (d) are not combined.

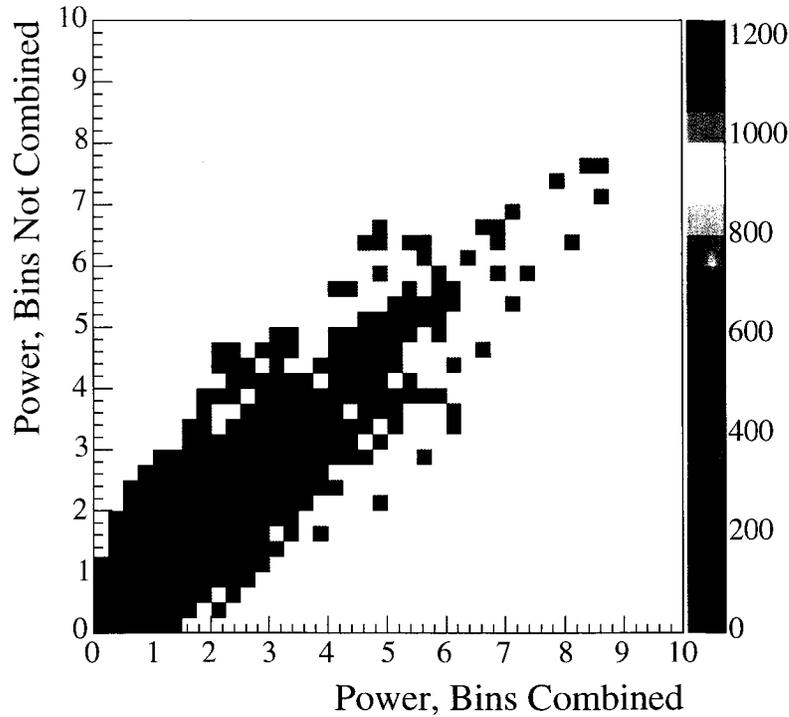


Figure 6.12: Powers for the combined phases when low statistic bins are, and are not, combined with neighboring bins.

powers are clearly correlated, indicating that the loss of 133 bins does not change the outcome of the analysis. There is no evidence that a periodic signal is lost in the data when combining the low statistic bins.

6.4.2 Monte Carlo Sensitivity Study

To answer the second question, the sensitivity analysis described in Chapter 5 has been redone for the D_2O phase, in the case where no bins have been combined. This will be compared with the sensitivity of the original D_2O phase, when bins are combined.

First, it is necessary to determine the power at which the null hypothesis can

be rejected at the 99% CL. This will be different for both scenarios, since the total number of bins are different. In the original analysis, the power for the 99% CL in the D₂O phase was 11.15. When no bins are combined, the power for the 99% confidence level is found to be 11.59. This value has been obtained by examining the maximum LS power distribution of 1000 MC event files with no intrinsic periodicity.

To determine the sensitivity, MC files with predefined periodicity have been analyzed using both methods. A range of periods with various amplitudes have been sampled, with each period and amplitude pair having 1000 MC event files. If 50% of the maximum powers in the periodic MC have a maximum power greater than the 99% CL power, it can be said that at this period and amplitude there is a 50% chance of rejecting the null hypothesis at the 99% confidence level.

The 50% contour line is shown on a sensitivity contour for the D₂O phase in Figure 6.13, for both scenarios. For periods greater than 10 days, both methods have the same sensitivity. As the period decreases from 10 days, the method of combining bins becomes less sensitive. When bins are combined there are a minimum of two calendar days that ‘merging’. Instead of having a time and rate associated with two separate days, there is now only one data point to represent the two days, obviously decreasing the sensitivity at these low periods.

6.4.3 Summary

There is a loss of information when data are put into bins of any size. The LS analysis has been performed with data put into 1-day bins. When these 1-day bins have a small amount of livetime, they are combined, to provide a better representation of the true solar neutrino rate. The sensitivity analyses have shown that the method is most sensitive to periods in the range of 10 days to several hundreds of days, which is the

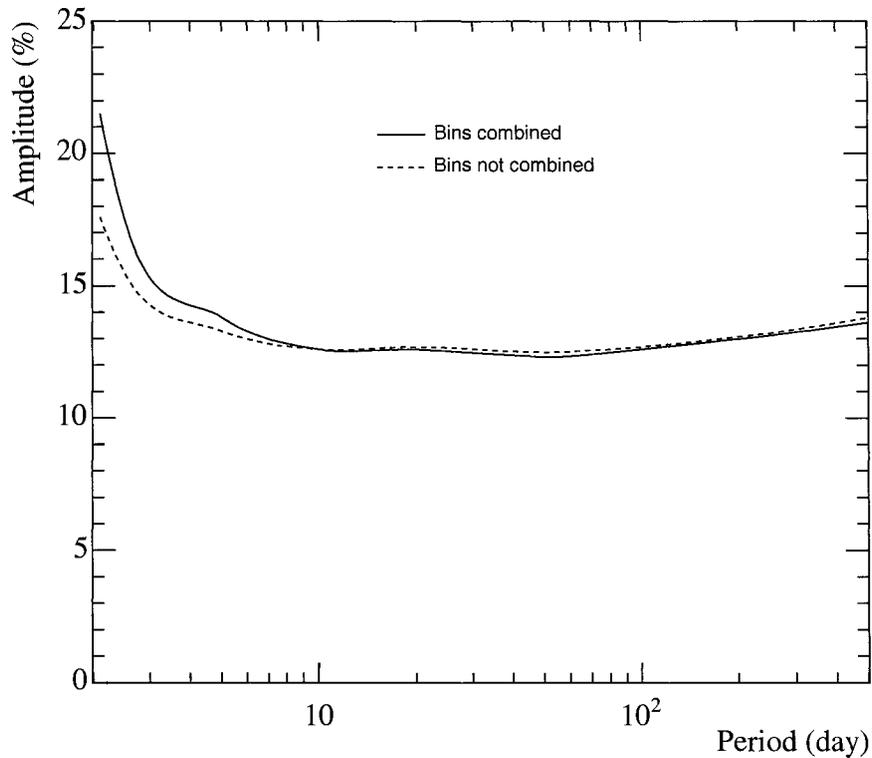


Figure 6.13: D_2O 50% contour sensitivities in the case where low statistic bins are and are not combined with neighboring bins.

primary range of interest, based on results from previous solar neutrino periodicity searches. For periods less than 10 days, combining bins slightly reduces the sensitivity, but this loss is small, and is not expected to impact the results.

6.5 Experimental Systematics

In a periodicity analysis one must ensure that the experimental conditions do not falsely produce an apparent signal, or hide the true signal in the data. For this analysis on the SNO data, the two experimental conditions that are of concern are effects of detector performance, and the small number of background events in the

data sample.

Admittedly, these issues are of key importance when a periodic signal is found. However, based on the MC analysis presented in Chapter 5, the SNO data appears not to have any significant periodic behavior. Nevertheless, a small discussion on these topics is presented in this section.

6.5.1 Detector Acceptance

A colleague in the SNO collaboration, Oser, examined the stability of the detector as a function of time using two calibrations sources [61]. The response of the detector is characterized by the ability of the detector to detect a neutrino event. This includes all aspects of the detector, from the hardware in the D₂O volume, to the exterior electronic components.

If the detector response varies with time, with some periodic behavior, this could affect the neutrino rate as measured by the detector, and result in a false periodic signal in the neutrino rate. Since the presence of a signal is not observed in the data, this topic is not investigated further.

An effect of the detector response that could affect this analysis, is if the detector response randomly creates a background potentially making a true signal difficult to detect. This would imply that at a time t the detector has some detection efficiency ϵ , and a later time t' a detection efficiency of ϵ' . With the assumption of random noise, ϵ at t is independent of ϵ' at t' . Generally speaking in this model, there is no relationship between the ability of a detector to accept events at different times. This model of random variations in the detector acceptance was used by Oser to determine the maximum systematic uncertainty of the detector at any given time.

The SNO ¹⁶N calibration source emits 6.13 MeV γ s, and is used to calibrate the

energy response of the detector. These ^{16}N calibration runs have been performed frequently throughout the D_2O and salt phases, and typically last from several hours to one day. Oser found the scatter of reconstructed energy in the D_2O phase to be 0.25% per day, and 0.18% per day in the salt phase. Assuming the worse variation of 0.25% per day on the energy scale, it was necessary to determine the variation this may cause in the number of observed neutrino events. As described in Chapter 3, in the D_2O and salt phases the 2924 and 4722 candidate neutrino events have a kinetic energy greater than, or equal to, 5.0 MeV and 5.5 MeV, respectively. Oser increased the analysis energy threshold by 1% (keeping all other thresholds the same), and there was a change in the total number of events by 1.8% in the D_2O phase, and 1.4% in the salt phase. When the energy scale is varied by 1% the number of events is altered by a maximum of 1.8%. Therefore, a variation of 0.25% in one day, would vary the number of events seen in one day by 0.45%.

The second calibration source used in the analysis was a Cf source. This source produces neutrons at a known intensity and so the ability of the detector to measure these neutrons can be found. Using the calibration data available for this source, Oser found the variation of the neutron detection to be 1% per day (see Figure 4.2).

These two calibration sources were used because they are performed frequently, and therefore provide a good representation of the detector acceptance across the full data set under analysis.

With the method described, the largest detector variation was found to be on the order of 1% per day. This uncertainty can be taken as a systematic uncertainty, and added in quadrature to the statistical uncertainty, such that each data point in the LS analysis is weighted by the usual statistical uncertainty and the new systematic

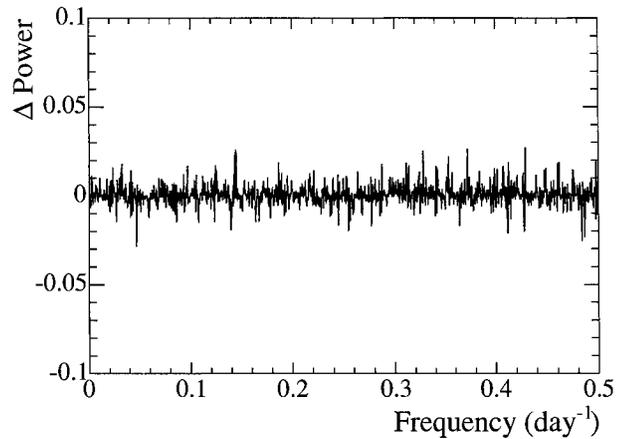


Figure 6.14: Difference in combined data sets powers at each of the 7300 frequencies, between the primary analysis in which the statistical error is used, and combination of the statistical and detector systematic errors.

uncertainty,

$$w_i = \frac{1}{\sigma_i^2}, \quad \text{where now } \sigma_i^2 = \sigma_{\text{stat}}^2 + \sigma_{\text{sys}}^2. \quad (6.3)$$

The statistical error has been discussed in Chapter 4. And the systematic variance is taken as 1% of the 1-day rate in each bin. The LS periodogram for the combined data set was computed with this additional uncertainty. The difference in powers at each frequency between the statistics only, and statistical plus systematic analysis is shown in Figure 6.14. There is virtually no change in the periodogram when this small detector variability is included as a systematic error. The 1% effect is expected to be small since the statistical error alone is far greater, on the order of 30%.

6.5.2 Background Systematics

In the 7646 candidate neutrino events, there are expected to be 383.4 non-solar neutrino background events. The breakdown of these is shown in Table 6.1. These back-

ground events come from a variety of sources, and are assumed to occur randomly in time. It is unlikely that these backgrounds are periodic, and even less likely that all 383.4 events would be periodic with the same period. The presence of randomly distributed non-solar neutrino events could create a background noise to the neutrino rate, if the neutrino rate is periodic.

If a periodic signal was found in the SNO data with some periodicity, a MC analysis could be employed to determine if this signal could have been caused by these background events. With that technique, MC event files with 383 events would be generated, using the same MC techniques as described in Chapter 5. These 383 events would be periodic, with periods at the found signal frequency, for various amplitudes. By binning these events, and subtracting these bin totals from the original data bins, the data could be reanalyzed (minus the assumed periodic backgrounds), to determine if the periodic signal is still found. If it is not found, then it is *possible* that these 383 events spread across four years could create a false signal.

The primary source of background events is neutrons originating exterior and interior to the D₂O. It has been suggested to perform a periodicity analysis on a data set with an increased radial cut or decreased energy cut, allowing a larger fraction of these background events to contaminate the data set, to see if a periodic signal is found. This is possible, although not pursued in this thesis. With such an extension, the previous MC analysis of the powers would not apply, and new MC would need to be generated and analyzed reflecting the larger data set, to determine the significance of any peak found in the periodogram. If a significant periodic signal had been found in the SNO data set, this path would have been explored.

	D ₂ O	Salt	Sub-Total
Internal neutrons*	51	125.1	176.1
Internal γ s	0	3.2	3.2
Čerenkov events inside the D ₂ O	23.6	3.6	23.6
External backgrounds	52	128.5	180.5
TOTAL	123.0	260.4	383.4

Table 6.1: A breakdown of the background events that contaminate the solar neutrino data for the combined D₂O and salt phases, [39, 56]. *A large fraction of the internal neutrons are the result of the photodisintegration of D₂O due to the radioactive decays of ²³²Th and ²³⁸U, and ²⁴Na (in the salt phase).

6.5.3 Summary

The effect of detector performance has been discussed in this section, and is treated as nothing more than noise that could impact the sensitivity of detecting a signal. Including the systematic variation of 1% on a 1-day event rate proved to have little effect on the periodogram. A more stringent approach to dealing with these systematics would be necessary if a significant periodic signal had been found in the data set. If a signal were found with a specific frequency, the detector acceptance and background time variations would be analyzed in greater detail at the suspected signal frequency. Without detecting a periodic signal in the data, these potential systematics could have an infinite combination of periods and amplitudes. At present, there is no experimental or theoretical motivation to further pursue these sources of systematic uncertainty.

Chapter 7

Summary and Conclusion

7.1 Summary

This thesis opened with a review of solar neutrino physics, and presented motivations to search for periodic behavior in the neutrino rate. This was followed by a description of the Lomb-Scargle (LS) periodogram, a mathematical tool used to search for sinusoidal periodic behavior in unevenly sampled data. The remainder of this thesis focussed on applying this LS technique to the solar neutrino data taken by the Sudbury Neutrino Observatory (SNO), in the first two phases of the experiment. In the D₂O phase 2924 candidate neutrino events were detected, and in the salt phase 4722 candidate neutrino events were found. The events from the D₂O and salt phases were analyzed as separate data sets, and then combined. In going from the D₂O, to the salt, and to the combined data sets the amount and calendar span of the data increased, thereby increasing the overall ability to detect a periodic signal in the ⁸B neutrino flux.

The largest LS power in SNO's D₂O data set is 7.1 at a frequency of 0.4082 day⁻¹.

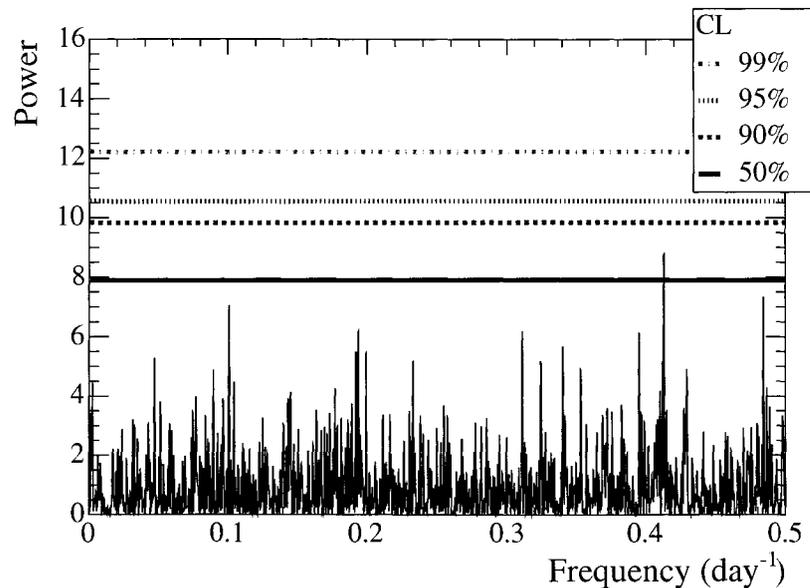


Figure 7.1: The combined D₂O and salt phase LS periodogram. Superimposed are the 50%, 90%, 95%, and 99% CL of rejecting the null hypothesis, based on the MC analysis presented in Chapter 5.

At this power, 46% of the constant rate, null hypothesis Monte Carlo (MC) data sets had a maximum power of this value, or greater. Implying that there is a 54% probability of rejecting the null hypothesis. For the salt data set, the largest power is 6.8, at a frequency of 0.4292 day⁻¹, with a 35% probability of rejecting the null hypothesis. And finally, when these data sets are combined, the largest LS power is 8.7 at a frequency of 0.4132 day⁻¹. In this case, there is a 73% probability of rejecting the null hypothesis.

Figure 7.1 shows the LS periodogram for the combined data sets. Superimposed are the probabilities of rejecting the null hypothesis at the 50%, 90%, 95%, and 99% confidence level (CL), obtained from the MC analysis. The null hypothesis is rejected at the 99% CL if the combined data set shows a power of 12.24, or greater. For the D₂O and salt data sets, this 99% CL power is 11.15 and 11.43, respectively.

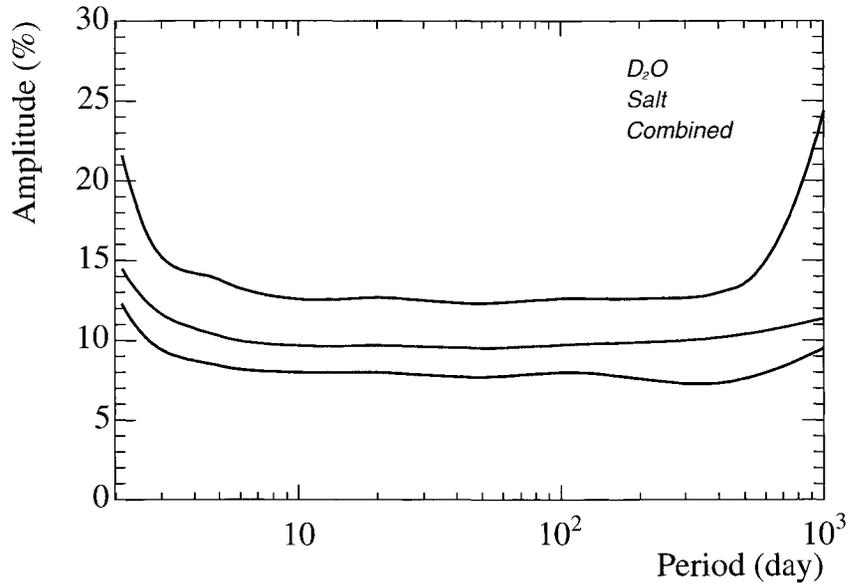


Figure 7.2: Sensitivity contours for a 50% probability of detecting a periodic signal for the separate D_2O and salt phases, and when these two phases are combined.

Figure 7.2 shows the characteristic period and amplitudes a periodic signal would need to have in each data set, to have a 50% probability of being detected. As the amount of data are increased, so too is the ability to detect a wider range of periods with decreasing amplitudes.

7.2 Comments on Two Specific Frequencies

In this section two frequencies are investigated in more detail. There is motivation to believe that the neutrino rate should, or could, be periodic at these frequencies.

The first is the modulation of the solar neutrino flux by the Earth's orbital eccentricity. This would occur at a frequency of $1/365.25 = 0.00273785 \text{ day}^{-1}$. The eccentricity of the Earth's path around the Sun is 1.7%, which should cause a yearly modulation of the flux with an amplitude of approximately 3.5%. This is a small

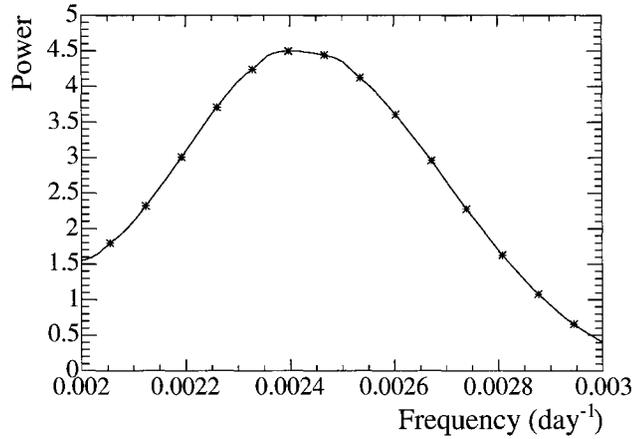


Figure 7.3: Periodogram from the combined data sets, zoomed into the yearly frequency of $1/365.25 = 0.00273785 \text{ day}^{-1}$.

effect, and based on the sensitivity analysis, the probability of detecting a signal with this amplitude in this analysis is less than 50%. Taking the periodogram from the combined D₂O and salt phases, and focussing on this frequency, the closest frequency sampled was $0.002739 \text{ day}^{-1}$ and was found with a power of 2.27. The portion of this periodogram is shown in Figure 7.3.

The second frequency of interest, is a frequency of 9.43 yr^{-1} . Some authors [35–38] claim the Super-Kamiokande solar neutrino data displays periodic behavior at this frequency. A frequency of 9.43 yr^{-1} corresponds to a frequency of 0.02583 day^{-1} . This portion of the combined phase periodogram is shown in Figure 7.4. The closest frequency that is sampled is 0.02582 day^{-1} , and is found with a power of 0.38.

At both of these specific frequencies there is no evidence of any significant periodic behavior. The yearly modulation has a small amplitude, and is therefore difficult to detect; in parallel with the fact that the data only spans a four year period, limiting the statistics at this period. The claimed solar neutrino signal in the Super-Kamiokande data, at a frequency of 9.43 yr^{-1} , has a reported amplitude of 7%. Based on the

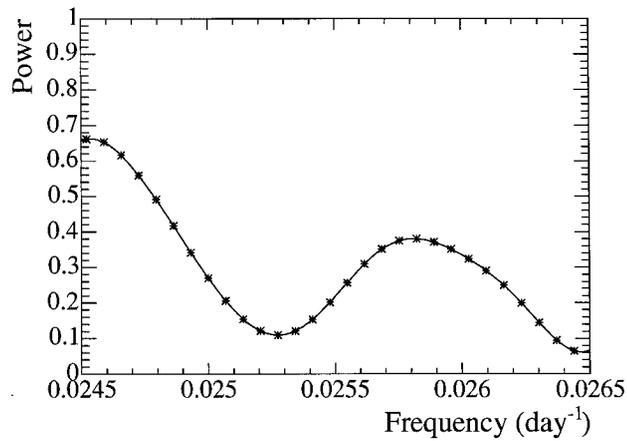


Figure 7.4: Periodogram from the combined data sets, zoomed into a frequency of 9.43 yr^{-1} or 0.02583 day^{-1} .

sensitivity analysis, there would be approximately a 50% probability of detecting a signal with this periodic behavior. It is not seen in the SNO data.

7.3 Future Prospects

Although no evidence for a periodic solar neutrino rate was found in the SNO data sets, there are additional analyses that can be performed.

The SNO data presented in this thesis can be, in principle, joined with the data from the Super-Kamiokande experiment. This would increase statistics, and possibly remove systematic effects in the experimental setups.

SNO has the unique ability to be able to measure the rate of the total number of solar neutrinos using the neutral-current reaction. And in a separate charged-current interaction, measure just the rate of the solar electron-neutrinos. The third interaction measured by SNO, the elastic scattering interaction, is sensitive to all neutrino flavours, but with a larger sensitivity to the electron-neutrino flavour. In the

D₂O and salt phases it is almost impossible to determine on an event-by-event basis which of the three interactions generated the triggering Čerenkov light, so a flavour sensitive periodicity analysis would be very difficult to pursue.

Chapter 1 outlined methods by which only the rate of solar electron-neutrinos would be periodic, but the total rate remain constant. If this were true, the large number of neutral-current events in the data set could wash out this signal, by adding a noisy background to a signal that would be present only in the electron neutrino rate. An alternative possibility is that the total solar neutrino rate is periodic, this would effect all three signals in the SNO data set. This is best probed by examining the events from all three interactions, as has been done in this thesis.

SNO is currently taking data for the third, and final phase of the experiment. The neutral-current detectors are essentially an independent experiment that measure only the total rate of solar neutrinos. The neutrino events detected in these counters can be used to perform an event-by-event periodicity analysis on the total rate.

In this third phase, the PMT-D₂O detection system is still active and recording neutrino events. By combining the data from this phase with the previous phase, the total calendar span of data will be seven years, and may allow a search for longer periods, including the 11 year period related to the solar magnetic activity. Including additional data will also increase the sensitivity of seeing periodic behavior, and lower the sensitivity contours.

7.4 The Last Word

Based on the Monte Carlo studies performed in this thesis, the solar neutrino rate, as measured by SNO in the D₂O and salt phases, can not reject the null hypothesis of a

constant rate distribution of the data. Despite the lack of evidence of a periodic solar neutrino rate, it is still possible that one exists, outside the frequency and amplitude regions of sensitivity. It is important to continue to test neutrino properties, since with each passing decade this elusive particle continues to reveal new properties.

Appendix A

Comparison of Adjusting Rates for the Combined Analysis

This appendix compares the results of

1. adjusting the D₂O phase rates up, such that the mean D₂O phase event rate equals the mean salt phase event rate (D₂O up),
2. adjusting the salt phase rates down, such that the mean salt phase event rate equals the mean D₂O phase event rate (salt down),

for the combined analysis. For the first option, each D₂O bin is multiplied by the ratio of

$$\frac{\text{mean salt phase rate}}{\text{mean D}_2\text{O phase rate}},$$

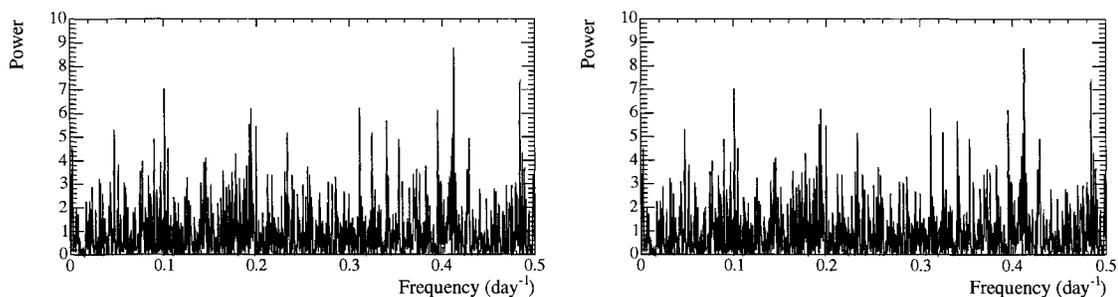
and the second option multiplies each salt bin by the ratio

$$\frac{\text{mean D}_2\text{O phase rate}}{\text{mean salt phase rate}}.$$

Method	Mean D ₂ O Rate	Mean Salt Rate	Ratio (D ₂ O up)	Ratio (salt down)
Regular	9.389 ± 0.177	11.847 ± 0.177	1.262 ± 0.030	0.793 ± 0.0190
Modified	9.345 ± 0.173	11.848 ± 0.172	1.268 ± 0.030	0.789 ± 0.0190

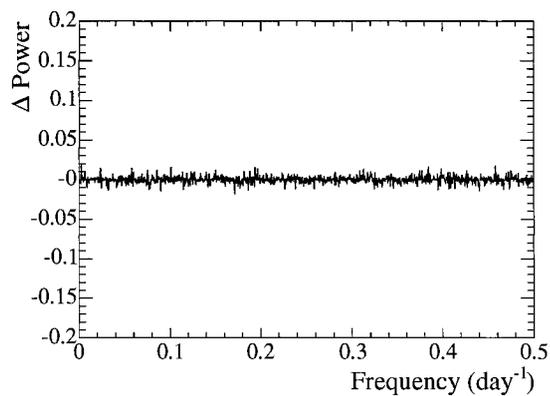
Table A.1: Mean values and ratios used to adjust the rate of one phase to match the other when the two phases are combined.

It must be noted that the regular LS method uses the standard mean definition, and the modified LS method uses the weighted mean definition. The difference is subtle, but reflects the subsequent mean value calculations that follow in the respective LS periodograms. Table A.1 summarizes the (weighted) mean rates and ratios that were used in the combined data analysis. The periodograms for the modified method, for both adjustment options are shown in Figure A.1. As can be seen, the periodograms are identical.



(a) D₂O phase rate adjusted up to the salt phase rate.

(b) Salt phase rate adjusted down to the D₂O phase rate.



(c) Difference in powers between the two methods.

Figure A.1: Periodogram for the combined data sets when the rates of the (a) D₂O phase are brought up to match the salt phase rates, (b) salt phase are brought down to match the D₂O rate. The difference between the powers at each frequency is shown in (c).

Appendix B

Detailed Sensitivity Analysis

Please refer to the following pages for the detailed sensitivity analysis for the D₂O, salt, and combined phases.

Note that for each period and amplitude pair 1000 MC data sets have been analyzed. The table entries with a dash (-) have not been analyzed.

B.1 D₂O

		Amplitude (%)												
		10	11	12	13	14	15	17	18	20	22	25	30	35
Period (day)	2.1	2.6	2.7	3.6	5.2	8.0	11.0	20.3	–	38.0	52.6	78.4	96.2	–
	2.5	4.1	7.2	9.9	12.8	18.9	28.6	43.8	57.3	72.7	87.4	97.3	–	–
	5.0	13.7	21.2	28.9	40.1	52.7	65.2	84.0	91.0	97.9	–	–	–	–
	10.0	18.9	32.0	41.4	56.9	69.3	79.9	94.1	97.6	99.7	–	–	–	–
	20.0	19.8	29.0	40.6	54.6	69.7	78.4	94.5	97.0	99.6	–	–	–	–
	50.0	21.6	33.3	44.8	61.4	72.9	83.7	96.1	98.3	99.6	–	–	–	–
	100.0	18.6	28.9	42.2	56.1	69.7	80.7	93.7	97.2	99.7	–	–	–	–
	500.0	11.1	18.1	30.2	43.7	54.8	68.3	89.6	94.1	98.7	–	–	–	–
	1000.0	1.1	1.9	2.2	3.6	4.0	5.1	10.9	12.8	23.4	–	53.3	79.2	95.8

Table B.1: The percentage of D₂O phase MC data sets that had a maximum power greater than 11.15, the power necessary to reject the null hypothesis at the 99% CL.

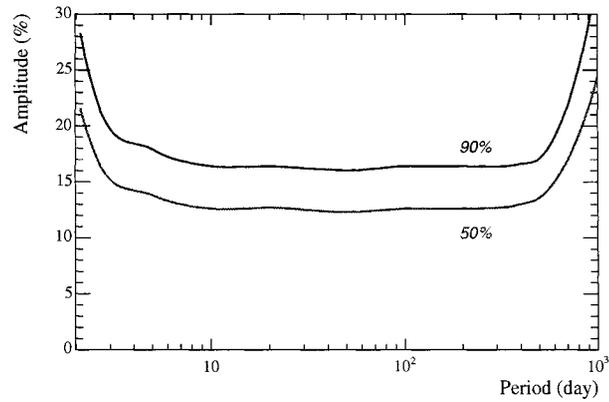


Figure B.1: 50% and 90% probability of rejecting the null hypothesis in the D₂O phase. These values have been interpolated from the data in Table B.1.

B.2 Salt

		Amplitude (%)										
		8	9	10	11	12	13	14	15	17	18	20
Period (day)	2.1	2.7	5.3	8.9	13.2	22.0	31.8	44.0	57.4	79.9	86.7	96.3
	2.5	–	–	15.3	25.9	36.6	52.5	65.6	75.0	92.4	–	99.3
	5.0	–	–	43.9	63.4	78.9	87.9	94.5	98.8	99.7	99.9	100.0
	10.0	22.6	37.7	56.7	74.7	84.5	94.2	98.3	99.2	99.9	100.0	100.0
	20.0	20.0	37.2	56.0	68.6	82.4	91.8	97.8	99.3	100.0	100.0	100.0
	50.0	24.4	40.3	59.4	75.1	88.4	95.8	98.6	99.6	99.9	100.0	100.0
	100.0	26.3	37.5	56.2	75.4	86.0	93.9	98.0	99.5	100.0	100.0	100.0
	500.0	15.2	24.9	41.2	61.0	75.3	87.6	93.2	96.6	100.0	100.0	100.0
	1000.0	–	–	31.6	42.6	60.3	75.4	87.6	93.0	99.0	99.8	100.0

Table B.2: The percentage of salt phase MC data sets that had a maximum power greater than 11.43, the power necessary to reject the null hypothesis at the 99% CL.

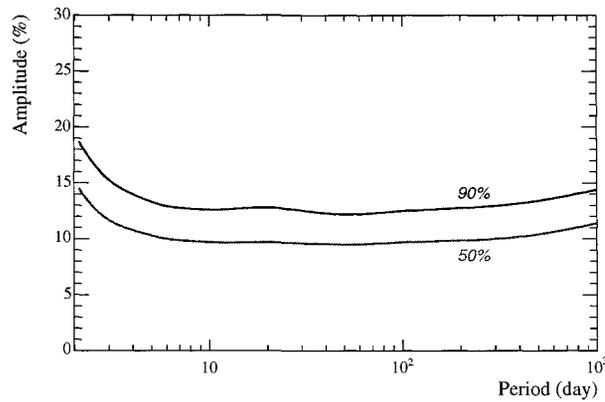


Figure B.2: 50% and 90% probability of rejecting the null hypothesis in the salt phase. These values have been interpolated from the data in Table B.2.

B.3 Combined

Period (day)	Amplitude (%)									
	7	8	9	10	11	12	13	14	15	17
2.1	–	6.4	10.4	18.6	29.9	46.6	56.6	70.2	84.8	96.6
2.5	–	–	–	43.4	57.4	75.8	87.6	94.1	97.3	–
5.0	–	41.2	62.7	80.9	91.2	96.8	99.3	99.7	100.0	–
10.0	–	48.7	70.7	86.2	95.7	99.6	99.7	100.0	100.0	–
20.0	–	48.8	73.7	90.4	96.5	99.3	99.7	100.0	100.0	–
50.0	30.4	55.7	76.2	87.8	96.7	99.3	99.8	100.0	100.0	–
100.0	29.0	50.7	73.2	87.6	96.4	99.2	99.5	100.0	100.0	–
500.0	32.1	56.8	75.9	91.2	96.5	99.7	99.7	100.0	100.0	–
1000.0	–	–	39.4	58.9	76.7	90.5	97.5	99.3	99.9	–

Table B.3: The percentage of combined phase MC data sets that had a maximum power greater than 12.24, the power necessary to reject the null hypothesis at the 99% CL.

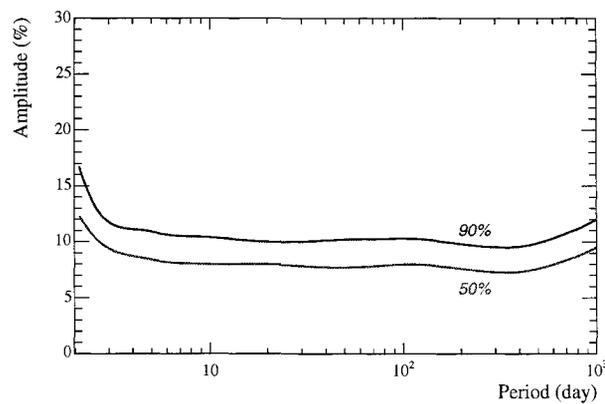


Figure B.3: 50% and 90% probability of rejecting the null hypothesis in the combined phases. These values have been interpolated from the data in Table B.3.

Bibliography

- [1] W.H. Press *et al.*, *Numerical Recipes in C*, Second ed. (Cambridge University Press, 1992).
- [2] E. Inrig, *A Search for Periodicity in the Solar Neutrino Flux Data from the Sudbury Neutrino Observatory*, SNO Collaboration internal report, 2004.
- [3] SNO Collaboration, *A Search for Periodicities in the 8B Solar Neutrino Flux Measured by the Sudbury Neutrino Observatory*, submitted to Phys. Rev. D , [arXiv: hep-ex/0507079].
- [4] J. Bahcall, *Neutrino Astrophysics* (Cambridge University Press, 1990).
- [5] E. Nesme-Ribes, S. Baliunas, and D. Sokoloff, *The Stellar Dynamo*, Scientific American: Special Edition **14**, 35 (2004).
- [6] P. Wilson, *Solar and Stellar Activity* (Cambridge University Press, 1994).
- [7] D. Hathaway, Website: <http://science.msfc.nasa.gov/ssl/pad/solar/sunspots.htm>, Last updated: January 4, 2005.
- [8] J. Christensen-Dalsgaard, *Heliosiesmology*, Rev. Mod. Phys. **74**, 1073 (2003), [arXiv: astro-ph/0207403].
- [9] J. Bahcall, M. Pinsonneault, and S. Basu, *Solar Models: current epoch and time dependences, neutrinos, and helioseimological properties*, Astroph. J. **555**, 990 (2001), [arXiv: astro-ph/0010346].
- [10] J. Bahcall and C. Pena-Garay, *Solar models and solar neutrino oscillations*, New J. Phys. **6**, 63 (2004), [arXiv: hep-ph/0404061].

- [11] S. Mikheyev and A. Smirnov, *Resonance enhancement of oscillations in matter and solar neutrino spectroscopy*, Sov. Jour. Nucl. Phys. **42**, 913 (1985).
- [12] L. Wolfenstein, *Neutrino Oscillations in Matter*, Phys. Rev. D **17**, 2369 (1978).
- [13] B. Pontecorvo, Zh.Eskp.Teor.Fiz. **33**, 549 (1957).
- [14] A. Cisneros, *Effect of Neutrino Magnetic Moment on Solar Neutrino Observations*, Astroph. Space Sci. **10**, 87 (1971).
- [15] J. Schechter and J. Valle, *Majorana neutrinos and magnetic fields*, Phys. Rev. D **24**, 1883 (1981).
- [16] M. Voloshin and M. Vysotsky, *Neutrino Magnetic Moment and Time Variation of Solar Neutrino Flux*, Sov. J. Nucl. Phys. **44**, 845 (1986).
- [17] M. Voloshin, M. Vysotsky, and L. Okum, *Neutrino electrodynamics and possible effects for solar neutrinos*, Sov. Phy. JETP **64**, 446 (1986).
- [18] R. Barbieri and G. Fiorentini, *The solar neutrino puzzle and the $\nu_L \rightarrow \nu_R$ conversion hypothesis*, Nucl. Phys. B **304**, 909 (1988).
- [19] E. Akhmedov, *The neutrino magnetic moment and time variations of the solar neutrino flux*, Talk given at 4th International Solar Neutrino Conference (April 8-11, 1997), [arXiv: hep-ph/9705451].
- [20] C.-S. Lim and W. Marciano, *Resonant spin-flavor precession of solar and supernova neutrinos*, Phys. Rev. D **37**, 1368 (1988).
- [21] E. Akhmedov, *Resonant amplification of neutrino spin rotation in matter and the solar-neutrino problem*, Phys. Lett. B **213**, 64 (1988).
- [22] G. Bazilevskaya *et al.*, *Cosmic rays, solar activity, and neutrino flux from the Sun*, JETP Letters **35**, 341 (1982).
- [23] D. Basu, *Solar neutrino and solar particles*, Solar Phys. **81**, 363 (1982).
- [24] V. Garvin, A. Kopylov, and N. Makeev, *Study of the Sun's neutrino brightness curve with the help of a chlorine-argon neutrino detector*, JETP Letters **35**, 608 (1982).

- [25] L. Dorman and A. Wolfendale, *The correlation of the solar neutrino rate with solar activity*, J. Phys. G. Nucl. Part. Phys. **17**, 769 (1991).
- [26] R. Ehrlich, *Possible time variations in ^{37}Cl solar neutrino data and neutrino oscillations*, Phys. Rev. D **25**, 2282 (1982).
- [27] J. Bahcall, G. Field, and W. Press, *Is solar neutrino capture rate correlated with sunspot number*, Astroph. J. **320**, L69 (1987).
- [28] J. Bahcall and W. Press, *Solar-cycle modulation of event rates in the chlorine solar neutrino experiment*, Astroph. J. **370**, 730 (1991).
- [29] G. Walther, *Absence of correlation between the solar neutrino flux and the sunspot number*, Phys. Rev. Lett. **79**, 4522 (1997), [arXiv: astrop-ph/9710031].
- [30] J. Boger, R. Hahn, and J. Cummings, *Do statistically significant correlations exist between the Homestake solar neutrino data and sunspots?*, Astroph. J. **537**, 1080 (2000), [arXiv: astro-ph/0001482].
- [31] R. Wilson, *Correlative aspects of the solar electron neutrino flux and solar activity*, Astroph. J. **545**, 532 (2000).
- [32] L. Pandola, *Search for time modulations in the Gallex/GNO solar neutrino data*, Astropart. Phys. **22**, 219 (2004), [arXiv: hep-ph/0406248].
- [33] P. Sturrock and M. Weber, *Comparative analysis of Gallex-GNO solar neutrino data and SOHO/MDI helioseismological data; Further evidence for rotational modulation of the solar neutrino flux*, Astrophys. J. **565**, 1366 (2002), [arXiv: astro-ph/0103154].
- [34] J. Yoo *et al.*, *A search for periodic modulations of the solar neutrino flux in Super-Kamiokande-I*, Phys. Rev. D **68**, 092002 (2003), [arXiv: hep-ex/0307070].
- [35] P. Sturrock, *Time-series analysis of Super-Kamiokande measurements of the solar neutrino flux*, Astrophys. J. **594**, 1102 (2003), [arXiv: hep-ph/0304073].
- [36] P. Sturrock, *Analysis of Super-Kamiokande 5-day measurements of the solar neutrino flux*, Astrophys. J. **605**, 568 (2004), [arXiv: hep-ph/0309239].

- [37] P. Sturrock, *Power spectrum analysis of Super-Kamiokande solar neutrino data: variability and its implications for solar physics and neutrino physics*, [arXiv: hep-ph/0501205].
- [38] G. Ranucci, *Time series analysis methods applied to the Super-Kamiokande I data*, [arXiv: hep-ph/0505062].
- [39] SNO Collaboration, *Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002), 011301.
- [40] SNO Collaboration, *Electron Antineutrino Search at the Sudbury Neutrino Observatory*, Phys. Rev. D **70** (2004), 093014.
- [41] J. Barranco, O. Miranda, T. Rashba, V. Semikoz, and J. Valle, *Confronting SFP solutions with current and future solar neutrino data*, Phys. Rev. D **66**, 093009 (2002), [arXiv: hep-ph/0207326].
- [42] SNO Collaboration, *Measurement of Day and Night Neutrino Energy Spectra at SNO and Constraints on Neutrino Mixing Parameters*, Phys. Rev. Lett. **89** (2002), 011302.
- [43] K. Eguchi *et al.*, *First results from KamLAND: Evidence for reactor anti-neutrino disappearance*, Phys. Rev. Lett. **90**, 021802 (2003), [arXiv: hep-ex/0212021].
- [44] H.O. Back *et al.*, *Study of neutrino electromagnetic properties with the prototype of the Borexino detector*, Phys. Lett. B **563**, 35 (2003).
- [45] E. Akhmedov and J. Pulido, *Solar neutrino oscillations and bounds on neutrino magnetic moment and solar magnetic field*, Phys. Lett. B **553**, 7 (2003), [arXiv: hep-ph/0209192].
- [46] J. Bahcall and A. Ulmer, *Temperature dependence of solar neutrino fluxes*, Phys. Rev. D **53**, 4202 (1996), [arXiv: astro-ph/9602012].
- [47] C. Burgess *et al.*, *Large mixing angle oscillations as a probe of the deep solar interior*, Astrophys. J. **588**, L65 (2003), [arXiv: hep-ph/0209094].

- [48] N. Morrison, *Introduction to Fourier Analysis* (John Wiley and Sons, Inc, 1994).
- [49] N. Lomb, *Least-squares frequency analysis of unequally spaced data*, *Astroph. and Space Science* **39**, 447 (1976).
- [50] J. Scargle, *Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data*, *Astroph. J.* **263**, 835 (1982).
- [51] J. Scargle, *Studies in astronomical time series analysis. III. Fourier transforms, autocorrelation functions, and cross-correlation functions of unevenly spaced data*, *Astroph. J.* **343**, 874 (1989).
- [52] J. Horne and S. Baliunas, *A prescription for period analysis of unevenly sampled time series*, *Astroph. J.* **302**, 757 (1986).
- [53] SNO Collaboration, *The Sudbury Neutrino Observatory*, *Nucl. Instr. and Meth. A* **449**, 172 (2000).
- [54] SNO Collaboration, *Measurement of charge current interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **87** (2001), 071301.
- [55] SNO Collaboration, *Measurement of the Total Active ^8B Solar Neutrino Flux at the Sudbury Neutrino Observatory with Enhanced Neutral Current Sensitivity*, *Phys. Rev. Lett.* **92** (2004), 181301.
- [56] SNO Collaboration, *Measurements of Electron Energy Spectra, Flux, and Day-Night Asymmetries of ^8B Solar Neutrinos from the 391-Day Salt Phase SNO Data Set*, submitted to *Phys. Rev. C* (2005), [arXiv: nucl-ex/0502021].
- [57] SNO Collaboration, *The SNOMAN User Manual*, (2004), SNO Collaboration internal reference manual.
- [58] K. Minknaitis and S. Oser, *Lifetime for the Full Salt Data Set*, SNO Collaboration internal report, 2003.
- [59] R. Van Berg *et al.*, *Design of SNO Timing System*, (1995), SNO Collaboration internal technical report, SNO-STR-95-008.

-
- [60] J. Klein *et al.*, *The SNO Trigger System*, SNO Collaboration internal technical report, SNO-STR-97-035, 1997.
- [61] S. Oser, *Limits on Detector Acceptance Variations for a Periodicity Analysis*, SNO Collaboration internal report, 2004.