MODERN

APPLIED

SCIENCE

CCSE

Canadian Center of Science and Education

Editorial Board

Contents

Microbial Composition in Cadmium Contaminated Soils around Zinc Mining Area, Thailand

Uratchwee Unhalekhaka (Corresponding author)

Environmental Science (Interdisciplinary program)

Graduate School, Chulalongkorn University

Phyathai Road, Pathumwan, Bangkok, Thailand 10330

Tel: 66-2-218-7666   E-mail: aummiee@yahoo.com

Charnwit Kositanont

Department of Microbiology, Faculty of Science, Chulalongkorn University

Phyathai Road, Pathumwan, Bangkok, Thailand 10330

Tel: 66-2-218-7667   E-mail: charnwit_k@yahoo.com

Abstract

According to the problems of cadmium contamination in soil at Mae Sot district, Tak province, this study attempts to examine the shift of active microbial population according to cadmium concentration variation around the zinc mine by using RT-PCR of the bacterial 16S rRNA gene and DGGE profiles analysis. The results showed that the amount of total cadmium and zinc was in the range from 1.1 to 34.95 mg/kg soil and from 27.05 to 1998 mg/kg soil, respectively. At the highest cadmium concentration (34.95 mg/kg soil) the highest diversity index was obtained. Cluster analysis from DGGE pattern indicated two different clusters in the samples. The two soil samples with the cadmium amounts were similar to the EU Maximum Permissible level (3.0 mg/kg soil) belonged to one cluster and the leftovers belonged to another. The Pearson correlation between the diversity index and cadmium values was significantly positive.

Keywords: Cadmium, Zinc mine, 16S rRNA, DGGE, Microbial population

1. Introduction

Cadmium (Cd) usually occurs in association with zinc ore (Zn) and accumulated in water and soil as by-product of Zn mining. Cd was absorbed into human body through respiration and eating. It would be accumulated into liver and kidney. Above body's threshold, Cd causes nausea, abdominal pain, kidney dysfunction and osteomalacia (Simmons et al., 2005).

The first case study was observed in 1950 at Jinzu river basin, Toyama prefecture, Japan. As a result of Kamioka Zn mine released liquid wastes into the river that normally used within the community and paddy field irrigation causing Itai-itai disease and greater than hundred lives were lost. Thereafter, they found Cd in contaminated soil around the river bank at 4.85 mg/kg soil or around 14 times of unaffected soil (0.34 mg/kg soil) (ICETT, 1998).

In Thailand, the biggest Zn deposit situates in Mae Sot district, Tak province and the estimated mine production capacity is 50 000 metric tons (Padaeng Industry Public Co., Ltd., 2008). Mining actions for example drilling, several explosions, material transfer, mine tailings disposal and drainage may cause the Cd distribution throughout the area, as mentioned in the research by Soil Analysis Division, Land Development Department, Thailand. The average Cd level in the sediment of cinder stacks was 228.5 mg/kg soil (Tunmanee and Thongmarg, 2007).

High Cd concentration was reported as the negative factor for soil microorganism metabolism (Renella et al., 2005). Therefore, the microbial community changing would be unavoidable. On the other point of view, changing of microbial community could be the indicator for Cd contamination in soil.

Formerly, the estimation of soil bacteria are culture-dependent methods which many biases from culturing media and laboratory conditions are always found. Less than 10% of the total microbial population can grow under these artificial status (Torsvik et al., 1998). Therefore, the culture-independent approaches are widely used in the present to eliminate these limitations.

In this study, the impact of Cd contamination on the shift of the active portions of soil bacterial diversity around Zn mining area were examined through 16S rRNA RT-PCR and denaturing gradient gel electrophoresis (DGGE) that show functioning soil bacteria genetic fingerprints (Muyzer, 1999).

2. Materials and Methods

2.1 Sampling sites

Soil samples were collected from 4 different creeks around Zn mine, located in Mae Sot district of Tak province, which are Huai Mae Tao Ngae Sai, Huai Mae Tao, Huai Mae Ku and Huai Nong Khiao (Figure 1).

2.2 Soil sampling

The eight collecting points were specified in the Royal Thai Survey Department topographical paper map scale 1:50 000. For each soil sampling site, three 10 cm depth cores were taken, tightly sealed in plastic bag and kept at -20^0C until the experiment began.

2.2.1 Sample preparation

Three soil cores from each sampling point were mixed and air dried. To removed plant debris and gravel dried soil was sieved through a 2 mm wire mesh.

2.2.2 Soil characteristics determination

Soil pH was investigated on 1:1 soil:water solution. Soil texture was estimated using Hydrometer. Organic matter (OM) was determined according to Walkley and Black method. The cation exchange capacity (CEC) was measured by ammonium saturation method (Department of Land Development, 2004). The total Cd and Zn were extracted with boiling *aqua regia* and detected by an inductively coupled plasma-optical emission spectrometry (ICP-OES), Optima 2001 DV (Perkin Elmer, USA) (Amacher, 1996).

2.3 RNA extraction

RNA from soil samples were extracted based on Griffiths et al.'s method (2004) and treated with RQ1 RNase-free DNase (Promega) as the manufacturer's instructions.

2.4 Reverse transcription PCR (RT-PCR)

The amplification of RNA templates was performed by a *Bca*Best[TM] RNA PCR Kit Ver.1.1 (Takara, Japan) according to the manufacturer's instructions (Hoshino and Matsumoto, 2007). The cDNA was amplified further with the primer pair: GC-341F and 907R (Muyzer et al., 2004) by using a Takara Ex Taq and performed in a GeneQ thermal cycler TC-24/H (Bioer Technology, China). The touch-down temperature cycling condition was used, lowering the annealing temperature 1^0C every cycle from 64^0C to 55^0C and then run further 10 cycles at 55^0C. The RT-PCR was verified for the absence of DNA contamination by subjected the same reaction without Reverse transcriptase.

The product were analyzed after electrophoresis (100V, 40 min) on 1.5% agarose gel then stained with ethidium bromide (EtBr).

2.5 DGGE determination

The mix PCR product from each soil sample (30µl) was loaded to 6% polyacrylamide in 1x TAE buffer at 60^0C then run at 100V for 18 h with gradient between 35% to 70% using the DCode system (Bio-Rad, USA). Gel was stained with EtBr and photographed by GelDoc XR (Bio-Rad, USA).

2.6 Microbial pattern analysis

DGGE image was analyzed with Quantity One version 4.6.1 software (Bio-Rad, USA), the total number of bands within a gel lane was determined as species richness (*R*) (Ampe et al., 2000) and the Shannon-Weaver index (*H*) of soil bacterial diversity was calculated based on the band intensity and number of bands as follows:

$H = -\Sigma(P_i)(\log P_i)$

where $P_i = n_i/N$, n_i is the peak intensity of a band and N is the sum of peak intensities within each lane (Briones et al., 2008).

Correlation between the Shannon index and other parameters was performed by program R for Windows version R-2.9.0 (R Development Core Team, 2009).

3. Results and Discussions

3.1 Sampling points

The 4 creeks (Huai, in Thai) of sampling sites are dendritic pattern and flow from east to west into the Moei river. Huai Mae Tao Ngae Sai places in the north-eastern highland of Zn mining while Huai Mae Tao situates on north-western part, flows through the mine and discharges to alluvial plain at Ban Mae Tao Mai. Huai Nong Khiao, a creek in the south-western highland of Zn mine, does not related to mining since it is on the other side of the hill. On the other hand, Huai Mae Ku locates along the south boundary of the mine and flows through Ban Mae Ku Nuea.

3.2 Soil properties

As presented in Table 1, soil pH were almost neutral while the OM varied from 4.03% to 7.98%. Therefore, they are categorized as high organic level soil (Department of Land Development, 2004). Every sample was coarse (sandy loam) except MT2 which was medium-texture soil (loam). For CEC, three of them (MT4, MT5 and R2) can be classified into rather low exchange rating and the rest five of them were moderate.

From the analysis results, the soils in this area are rich top soil. The neutral pH and high organic matter make them as a suitable site for agriculture.

3.3 Concentrations of total Cd and Zn

Total Cd and Zn contents in eight soil samples are shown in Table 1. Focus on Cd, the sample that was collected from Huai Mae Tao Ngae Sai (MT1), the creek which not pass through the mining area, had Cd quantity near the EU Maximum Permissible (MP) level (3.0 mg/kg soil) (Simmons et al., 2005) while the sample at Huai Nong Khiao (R1) was found Cd level only 1.1 mg/kg soil. Huai Mae Ku had high Cd concentrations especially the R2 site which is a dike for the community irrigation. It might suggest that the water retention in the area causes the Cd accumulation. At the downstream of Huai Mae Ku (R3) found Cd amount 5 times lower than the upstream (R2), implied that Cd in soils are higher than water and water is an intermediate which carries Cd to accumulate at the dike.

Huai Mae Tao which flows through the Zn mine, found the lower level of Cd in the upriver (MT2) and increased at the downstream (MT3-MT5). The accumulation range of Cd between MT2 and MT3 is estimated as 2 times whereas the expansion rate at MT3 to MT5 is 13% per site. From these data, it might assume that the water usage within the community lead to the Cd accumulation in soils into the hazardous level.

3.4 Soil microbial profiles

The DGGE patterns generated from the universal primers for general bacteria (GC-341F and 907R) were exposed the active bacterial composition in Cd-contaminated soil samples (Figure 2).

The low Cd soil ($\leq$ 3.05 mg/kg soil) which is MT1 and R1 samples gave banding patterns 8 bands while the highest (R2) showed 18 bands with distinctive intensity were observed, indicating the presence of different kinds of active bacteria.

It is obviously to notice that the intensity of bands in sample with 3.05 mg/kg soil (MT1) is greater than the lowest Cd soil (R1). Band ① was found when Cd amounts are 3.05-7.55 mg/kg soil (MT1 and R3) in range and faded away in higher Cd concentrations while band ②, the 2 closest bands, appeared deeply sharp only in the highest Cd quantity sample (R2). This might believed that band ① should be a good indicator of soil that has the Cd exceed the standard level whereas band ② should be an indicator when soil carries much higher amount.

The similarity of each lane were analyzed by using UPGAMA clustering method and illustrated into a dendrogram (Figure 3). From the dendrogram, there were two major clusters showed about 30% similarity which are MT1 and R1. Cd levels of both sites were similar to the EU standard, while another group had much higher Cd amounts.

According to the *H* index (Table 3), the sample R2 had the highest microbial diversity. Furthermore, when considered the correlation of these results and the soil properties with Pearson correlation (*r*) (Table 4) it was found that only Cd values of soil samples had rather strong positive coefficient (*r* = 0.768) at the 0.05 level where there are no correlation occurred with other variables. These may be concluded that only Cd concentrations had the direct effect on the active soil bacterial communities.

When consider in the Cd levels above 7.55 or at 8.45 to 34.95 mg/kg soil in range (MT2-MT5, and R2) found that the relationship between Cd quantities and diversity index is clearly corresponded as shown in Figure 4. It might suggested that the active bacterial diversity is higher when the Cd amounts are greater than or approximate to 9 mg/kg soil as similar to the evidence that stated by Wang et al. (2007). In case of the lower levels, there could be other factors influenced to the bacterial functions especially nutrients and conditions of each sample apart from Cd concentrations (Wang et al., 2007).

The results from this study suggest that Cd contamination impact on the active soil bacteria provide the information in the real field situation and further research is needed to discover the role of these functioning bacteria and explain their influence within the contaminated environment.

References

Amacher, M. C. (1996). Methods of soil analysis: nickel, cadmium and lead. In D. L. Sparks. (Ed.), *Methods of soil analysis. Part3: Chemical methods.* Wisconsin: American Society of Agronomy and Soil Science Society of America. pp. 739-768.

Ampe, F., & Miambi, E. (2000). Cluster analysis, richness and biodiversity indexed derived from denaturing gradient gel electrophoresis fingerprints of bacterial communities demonstrate that traditional maize fermentations are driven by the transformation process. *International Journal of Food Microbiology,* 60, 91-97.

Briones, A. M., Shililu, J., Githure, J., Novak, R., & Raskin, L. (2008). *Thorsellia anophelis* is the dominant bacterium in a Kenyan population of adult *Anopheles gambiae* mosquitoes. *The ISME Journal,* 2(1), 74-82.

Department of Land Development. (2004). *Manual for analysis of soil, water, fertilizer, plant, soil amendment and standard quality control of merchandise.* Bangkok: WJ Properties, (Chapter 2-3).

Griffiths, R. I., Manefield, M., Whiteley, A. S., & Bailey, M. J. (2004). DNA and RNA extraction from soil. In G. A. Kowalchuk, F. J. de Bruijin, I. M. Head, A. D. Akkermans, & J. D. van Elsas. (Eds.), *Molecular microbial ecology manual.* (2nd ed.). Netherlands: Kluwer Academic Publishers. pp. 1.11/149-158.

Hoshino, Y. T., & Matsumoto, N. (2007). DNA- versus RNA-based denaturing gradient gel electrophoresis profiles of a bacterial community during replenishment after soil fumigation. *Soil Biology & Biochemistry,* 39(2), 434-444.

International Center for Environmental Technology Transfer (ICETT). (1998). Preventative measures against water pollution Jinzu river, Toyama prefecture. [Online] Available: http://www.icett.or.jp/lpca_jp.nsf/a21a0d8b94740fbd492567ca000d5879?OpenView (March 23, 2007)

Muyzer, G. (1999). DGGE/TGGE a method for identifying genes from natural ecosystems. *Current Opinion in Microbiology,* 2(3), 317-322.

Muyzer, G., Brinkhoff, T., Nübel, U., Santegoeds, C., Schäfer, H., & Wawer, C. (2004). Denaturing gradient gel electrophoresis (DGGE) in microbial ecology. In G. A. Kowalchuk, F. J. de Bruijin, I. M. Head, A. D. Akkermans, & J. D. van Elsas. (Eds.), *Molecular microbial ecology manual.* (2nd ed.). Netherlands: Kluwer Academic Publishers. pp. 3.13/743-770.

Padaeng Industry Public Company Limited. (2008). Annual report 2008. [Online] Available: http://www.padaeng.com/pdf/Annual/2008_AnnaulReport.pdf (June 11, 2009).

R Development Core Team. (2009). A language and environment for statistical computing. R Foundation for Statistical Computing. [Online] Available: http://www.R-project.org (April 22, 2009)

Renella, G., Mench, M., Landi, L., & Nannipieri, P. (2005). Microbial diversity and hydrolase synthesis in long-term Cd- contaminated soils. *Soil Biology & Biochemistry,* 37(1), 133-139.

Simmons, R. W., Pongsakul, P., Saiyasitpanich, D., & Klinphoklap, S. (2005). Elevated levels of cadmium and zinc in paddy soils and elevated levels of cadmium in rice grain downstream of zinc mineralized area in Thailand: Implications for public health. *Environmental Geochemistry & Health*, 27(5-6), 501-511.

Torsvik, V., Daae, F. L., Sandaa, R. A., & Øvreås, L. (1998). Novel techniques for analysing microbial diversity in natural and perturbed environments. *Journal of Biotechnology,* 64(1), 53-62.

Tunmanee, N., & Thongmarg, J. (1994), Contamination of cadmium in soils in some area of Thailand. [Online] Available: http://www.ldd.go.th/pldweb/tech/RISabst/analyst37c.htm (April 15, 2007).

Wang, Y., Shi, J., Wang, H., Lin, Q., Chen, X., & Chen, Y. (2007). The influence of soil heavy metals pollution on soil microbial biomass, enzyme activity, and community composition near a copper smelter. *Ecotoxicology & Environmental Safety,* 67(1), 75-81.

Wang, J., Lu, Y., & Shen, G. (2007). Combined effects of cadmium and butachlor on soil enzyme activities and microbial community structure. *Environmental Geology,* 51, 1221-1228.

Table 1. Soil samples characteristics with total Cd and Zn amounts

Creek[1]	Sample	pH	OM (%)	Texture[2]	CEC (cmol/kg)	Cd (mg/kg soil)	Zn (mg/kg soil)
A	MT1	6.3	4.89	SL	10.10	3.05	97.45
B	MT2	6.5	7.98	L	13.41	8.45	193.05
B	MT3	6.5	5.81	SL	12.91	16.85	1998
B	MT4	6.8	5.97	SL	8.60	18.80	954
B	MT5	6.8	4.03	SL	6.99	22.50	1163
C	R1	6.7	6.32	SL	14.71	1.10	27.05
D	R2	6.7	4.31	SL	7.39	34.95	789
D	R3	6.8	5.39	SL	14.57	7.55	316.05

† [1] A: Huai Mae Tao Ngae Sai; B: Huai Mae Tao; C: Huai Nong Khiao; D: Huai Mae Ku.

[2] SL: Sandy Loam; L: Loam.

Table 2. Biodiversity index (*H*) calculated from the DGGE pattern and species richness (*R*)

Sample	*H*	*R*
MT1	0.68	8
MT2	0.60	5
MT3	0.93	14
MT4	0.78	8
MT5	0.87	9
R1	0.64	8
R2	1.18	18
R3	0.99	16

Table 3. Pearson correlation coefficients between microbial diversity and soil parameters ($P < 0.05$)

Parameters	Correlation with diversity index
Cd	0.768*
Zn	0.471
pH	0.403
OM	-0.646
CEC	-0.377
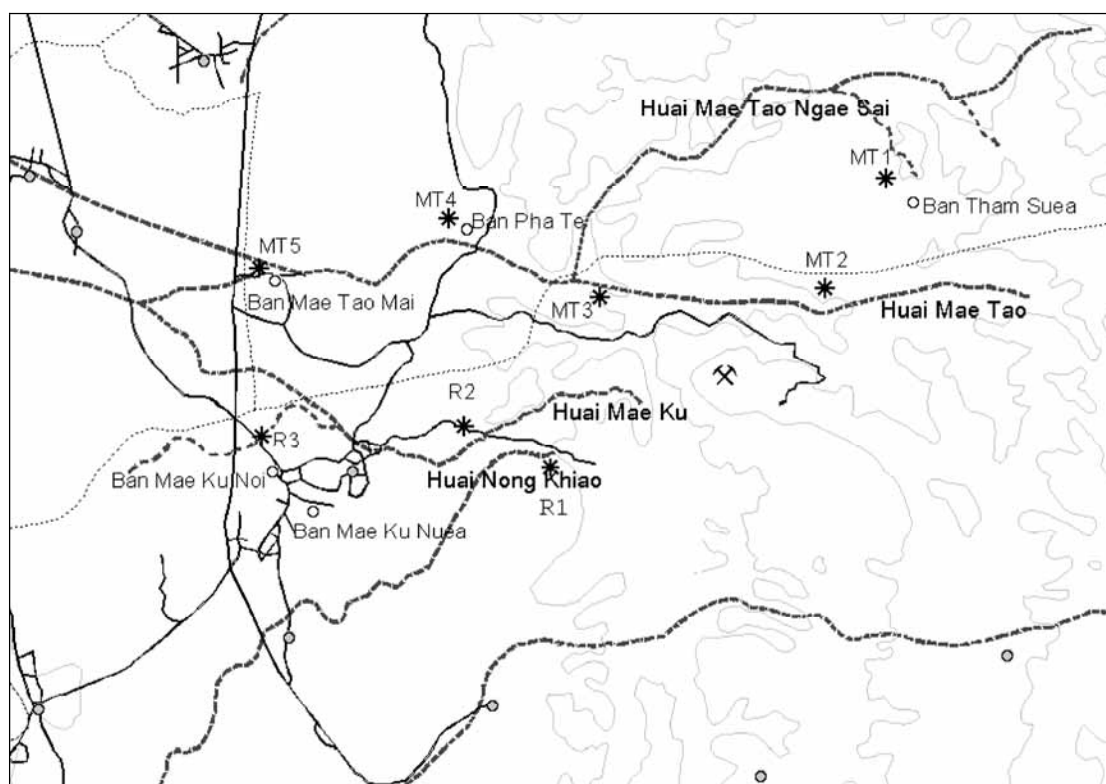
† o = villages in Mae Sot district   ∗ = sampling points   ⚒ = Zn mine

──── , ... = roads   ‒ ‒ ‒ ‒ = creeks

Figure 1. Map of soil sampling points
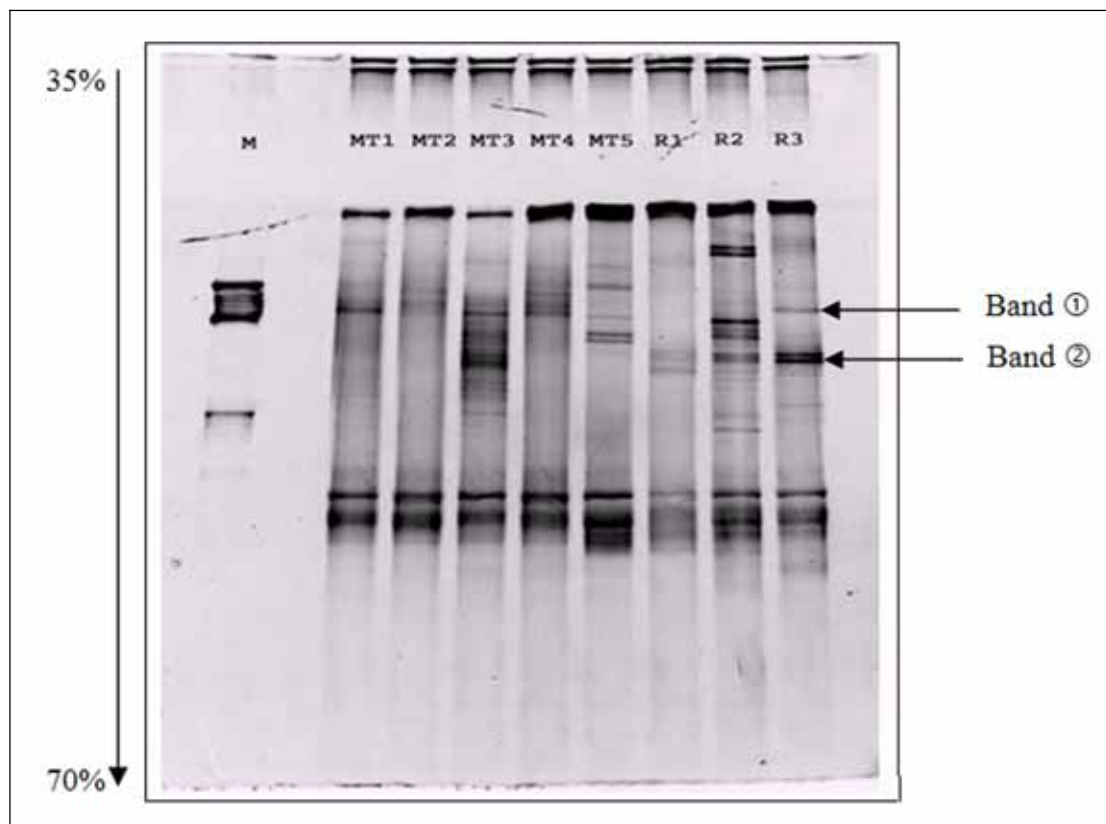
Figure 2. DGGE profiles of functioning bacterial communities based on 16S rRNA from eight Cd-polluted soil samples
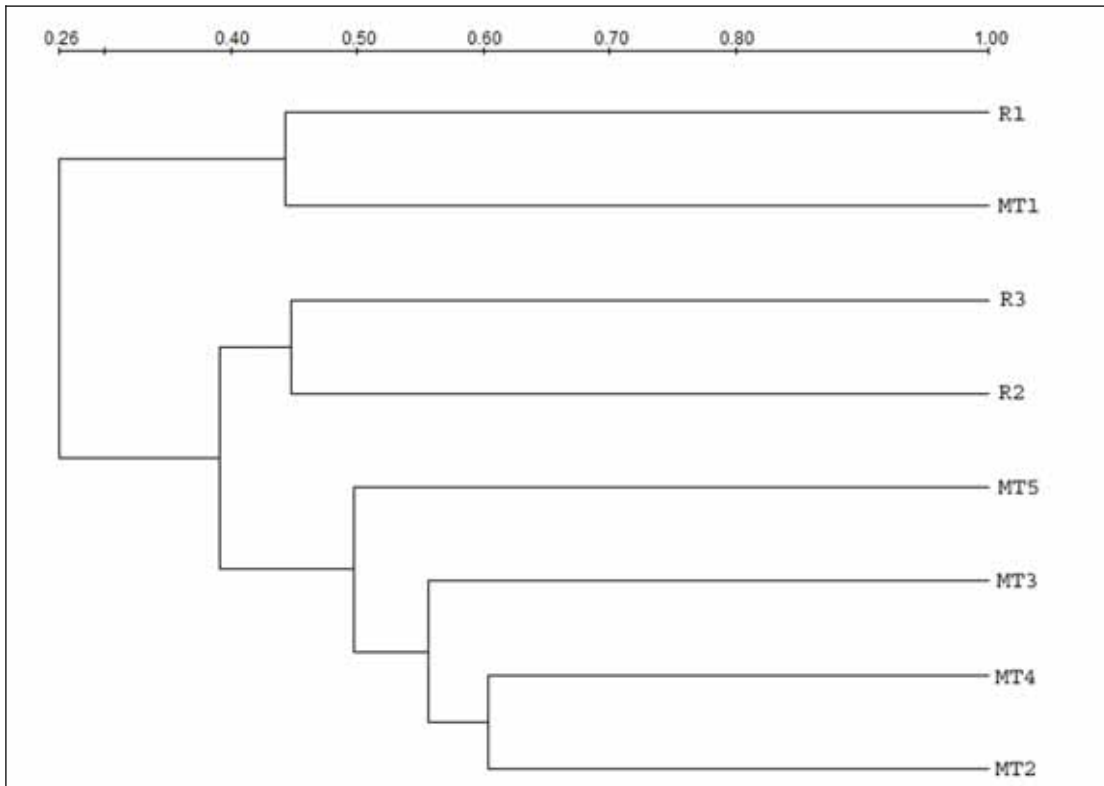
Figure 3. Similarity of the DGGE bacterial communities in soil samples MT1-R3
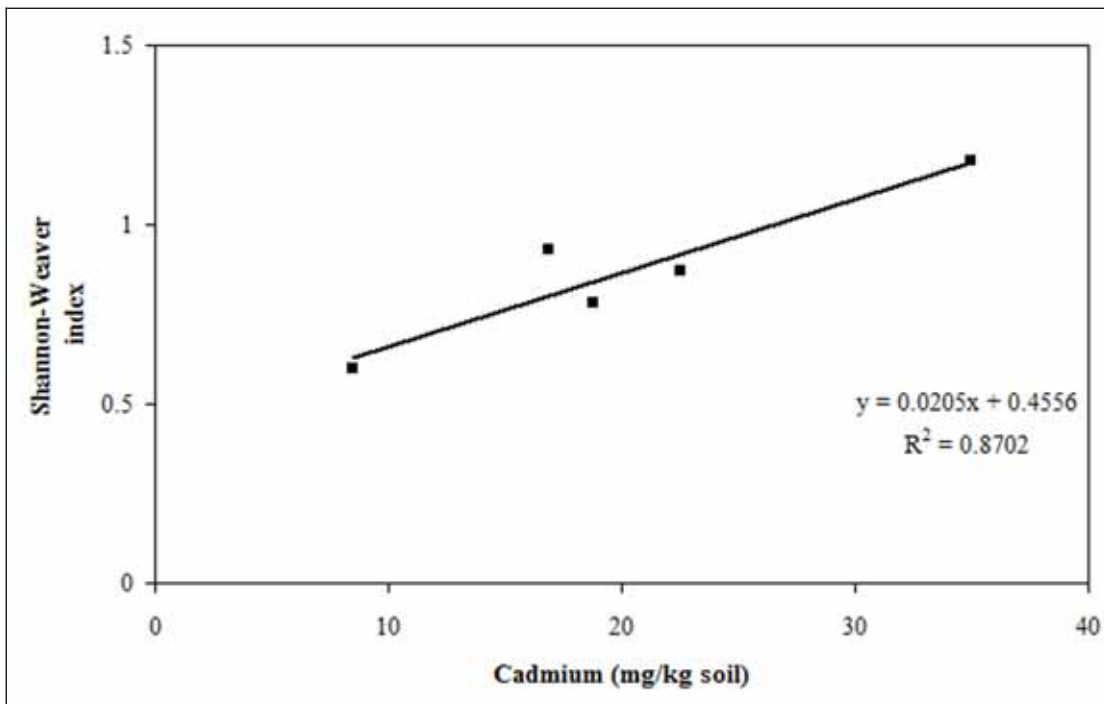
Figure 4. Trend of the active bacterial diversity index at Cd levels 8.45 to 34.95 mg/kg soil

Study on the Workover Fluid Formula and Performance

of the Prevention Reservoir

Fengying Zhang, Jienian Yan & Zhiyong Li

MOE Key Laboratory of Petroleum Engineering

China University of Petroleum Beijing

Beijing 102249, China

Tel: 86-10-8973-3893   E-mail: glyh2000@163.com

Abstract

In the workover treatment of Yakela condensate gas field, the workover fluids are easily leaked, and many problems such as clay expansion, rate sensitivity, migration of fines, and change of wetting property will occur, which will induce serious reservoir harms. In the article, the performances of the existing workover fluid of the Yakela condensate gas field with high temperature and high pressure are evaluated, and the influencing factors harming the reservoir are analyzed and obtained, and the effective measures and methods preventing the reservoir harm are confirmed. The technical principle, material composing, performance index, and the evaluation of stratum liquid compatibility are introduced in the article. The result of performance evaluation indicates that the core return permeability of the workover liquid system can achieve above 80%, and it is an ideal low-harm workover fluid system.

Keywords: Workover fluid, Stratum harm prevention, Permeability

Because of many years' exploitation, the reservoir conditions of the Yakela condensate gas field have changed a lot. The pollutions of the wells which still adopt the original workover fluid system occur in succession, and the reservoirs were harmed to different extents. According to existing materials of tested wells, the average skin coefficient after pollutions is about 35, and the upper gas layer skin coefficient of the Well 15 achieves 131.4, which seriously influences the production energy design and normal production of single well. To make clear the cause of reservoir harm, it is necessary to analyze and evaluate the harm mechanism and the harm degree of the original workover fluid, and based on room researches, the good workover fluid which can fulfill the reservoir of the Yakela condensate gas field can be selected, which can reduce the harms of the oil gas field in the workover process as more as possibly.

1. Mechanism analysis of reservoir harm

1.1 Performance analysis of existing workover fluid

At present, the workover fluid used in the Yakela condensate gas field mianly is the stratum water. Only to comprehensively analyze and study the performance of the existing workover fluid and completion fluid, the reservoir harm induced by the workover fluid and the completion fluid and the main harm mechanism can be scientifically diagnosed and analyzed. The analysis result is seen in Table 1.

From the experiment result in Table 1, the stratum water belongs to the calcium chloride of super high salinity, and in the workover process, the stratum scaling will be easily induced because of the invasion of exterior fluids, and in the production process, the potential problem of the production string (oil/ casing string) erosion may exist, and corresponding protective measures such as corrosion mitigation and scale inhibition should be adopted to effectively prevent the invasion of exterior fluids. The total iron content in the stratum water has achieved the standard, so the corrosion resistance of the production string should be considered.

1.2 Analysis of reservoir harm factors

Yakela condensate gas field mainly is the low permeability reservoir, and combining with the geology of the oil reservoir and the character of the liquid, the reservoir harm factors mainly include following four aspects from the harm

mechanism.

(1) The clay expansion plugging induced by the water sensitivity

In Yakela condensate gas field, some parts conclude 20%~35% mixed-layer minerals of illite and smectite, and the highest content can achieve 50%. When the workover fluid filters into the stratum, the smectites will be expanded quickly to reduce the hole-tunnels on the oil layer. In addition, the dispersion and transfer of clay after hydrating expansion will prick up, which will make the plugging more seriously.

(2) Plugging of water locking

The experiment of the water locking plugging shows that the harm of water locking to the reservoir is serious and the harm rate achieves 50%~70%.

(3) Organic scaling and inorganic scaling plugging

Because the mineralization of stratum water is quite high and contains large number of Ca^{2+}, so the inorganic scaling ($CaCO_3$ and $CaSO_4 \cdot 2H_2O$ scaling) is one of main harm mechanisms. The organic scaling mainly occurs in the oil extraction. Giving priority to gas reservoirs, the condensate oils contain certain paraffin, so the possibility of paraffin deposit (organic scaling) exists, which should be emphasized in the middle and post terms in the oil and gas field development.

(4) Emulsion plugging

In the operation process of low-pressure well, the workover fluid and the stratum crude oil could easily form high-viscosity emulsion fluid under the function of surface active materials to plug the filtering pass of the oil and gas reservoir, which will induce the increase of filtration-resistance.

2. Study of new workover fluid technology

According to the survey of materials, the basic idea of developing the workover fluid of Yakela condensate gas field is to take the reservoir protection as the premise and fulfill the requirements of the workover works simultaneously, i.e. new workover fluid system should possess many characters such as good compatibility of stratum liquid, little corrosion and low stratum harm. At the same time, comparing with the existing workover fluid, the new workover fluid has many advantages such as good expansion resistance, small filtration loss, easily flowing-back, obvious reservoir protection, and high corrosion inhibition rate.

2.1 Technical principle

New low-harm workover fluid system is mainly composed by viscosity increaser, fluid loss additive, inhibitor, and special particles with temporary plugging function on the pore throat of reservoir (Yan, 2001). The technical principle is to add oil temporary plugging agent matching with the pore throat of reservoir, and under the function of certain pressure difference, the temporary plugging agent will form thin layer of low filtration shielding at the entrance of stratum hole to effectively prevent the further invasion of workover fluid to enter into the reservoir. When the workover is completed and the oil well begins to produce oils normally, under the reverse pressures, part of oil temporary plugging agents will be washed out from the holes, and the other part will be dissolved by the produced condensate oil to recover the filtration rate of the reservoir (Lin, 1999 & Li, 2005, P.24-28 & Luo, 2006, P.16-20) and protect the oil and gas reservoir.

2.2 Material composing and main performance indexes

(1) Oil temporary plugging agent is new product, and its oil solution ratio exceeds 90%, and the softening point is 140℃.

(2) HTB mutual solvent is water-soluble non-ionic surface active agent, and when it is added, the oil-wetted surface of the oil temporary plugging agent will be converted into hydrophilic surface, which can help the oil temporary plugging agent to be decentralized into new low-harm workover fluid system.

(3) For the thickening agent, XC with good salt resistance which still has thickening ability in high-concentration salt solution can rapidly dispersed and increase viscosity in water without the support of other solutions (Zhao, 2002, P.75-76), and the "fish eye" and "block mass" don't exist in the solution.

(4) Main performance indexes. The adjustable range of density is 1.00~1.30kg/L, and the proper temperature is in 100~140℃.

(5) Selected workover fluid formula. Stratum water + 1%DG-HS1 (inhibitor) + 0.3%HTB (discharge aiding agent) + 1%DG-NW1 (clay stabilizer) + 1%SMP-2+ 0.3%XC + 2%NaCOOH + 1%PAC141

2.3 Experiment result and discussion

2.3.1 Experiment method

According to the petroleum gas industry standard SY/T 6540-2002, the JHCF-1 dynamic core pollution harm evaluation laboratory instrument (seen in Figure 1) is adopted to implement the room experiment. (1) Vacuumize the experimental core, saturate it for 24 hours by the stratum water, and measure the initial permeability of the core positively drove by the coal oil, (2) under the pressure difference pressure of 3.5 MPa, simulate the borehole and circulate the drilling fluid for 12 min, and measure the dynamic filtration loss of drilling fluid, (3) stop the circulation of drilling fluid, and measure the permeability of the harmed core positively drove by the coal oil (Fan, 1996).

2.3.2 Rheology and lubricity and formation protection ability

Measure the rheology of the low-harm workover fluid system by the six-speed rotating viscometer, and the result is seen in Table 2. From Table 2, the workover fluid has good rheology which can fulfill the requirement of the locale construction.

New workover fluid is the low solid fluid which can not form compact mud cake, and the temporary plugging method is used to reduce the filtration loss, and in the research, the fluid loss additive and temporary plugging agent are added to reduce the filtration loss, which can obtain good effect, and before and after hot aging, the filtration loss all can be controlled in 8ml.

2.3.3 Return permeability

By the core flow test, the reservoir protection of the workover fluid is evaluated in the laboratory, and the permeability of the workover fluid before and after pollution in the experiment are respectively measured, and the loop-pressure high-temperature and high-pressure core laboratory instrument is used to reversely pollute the core sample, and the result is seen in Table 3. From Table 3, for the reservoir core sample of YK10 in Yakela condensate gas field, respectively use the stratum water and selected workover fluid to pollute, when the mud cake (without solid fluid pollution) can not be formed basically, the return permeability of core could be enhanced form 65.21% to above 80%, which indicates that the selective workover fluid harms the reservoir little, and it can protect the reservoir well.

2.3.4 Study of workover fluid and reservoir rock compatibility

Use the selected new workover fluid formula and the reservoir core sample of YK10 well to make the experiment of compatibility. Take certain core for drying and weighing, mix them with the workover fluid after being mashed, filter, clear, drying and weigh after 16h' hot aging in the temperature of 150℃, compare the qualities before and after the experiment. The result of the experiment is seen in Table 4.

From Table 4, after mixing with the workover fluid, the change of core quality is not obvious, which indicates that the dispersion degree is not high, and three kinds of workover fluids all have good compatibility with the reservoir rocks.

2.3.5 Study of workover fluid and reservoir liquid compatibility

By the turbidity experiment, the compatibility between the workover fluid and the stratum water is evaluated. Mix the selected new workover fluid filtrate with the stratum water according to two different proportions, and the data measured by the turbidity meter is seen in Table 5.

From Table 5, after mixing the workover fluid filtrate with the stratum water, the turbidity of the fluid can not be changed basically, and in eyes, there are not any depositions, which indicate that the above selected formula has good compatibility with the stratum water.

3. Conclusions

(1) The present workover fluid in Yakela condensate gas field is analyzed, and the possible reservoir harm mechanism is also analyzed.

(2) Aiming at the characters of Yakela condensate gas field, the formula of the low-harm workover fluid system can be confirmed as follows.

Stratum water + 1%DG-HS1 (inhibitor) + 0.3%HTB (discharge aiding agent) + 1%DG-NW1 (clay stabilizer) + 1%SMP-2+ 0.3%XC + 2%NaCOOH + 1%PAC141

(3) New low-harm workover fluid system has many characters such as oil temporary plugging function, good plugging effect, good flow-back and plugging solving, small harm permeability, and extended applied temperature range. The laboratory return permeability recovery of the workover fluid can achieve above 80%.

(4) New low-harm workover fluid technology has wide application prospect in the workover of Yakela condensate gas filed.

References

Fan, Shizhong, Yan, Jienian & Zhou, Dachen. (1996). *Drilling Fluid, Competition Fluid and Oil-gas Reservoir Protection Technology.* Beijing: Petroleum University Press.

Lin, Yongxue, Yan, Xunbiao & Duan, Qingrong. (1999). *Low-harm Workover Fluid Study and Application Technology.* Dezhou: Petroleum Drilling Institute of Sinopec New Star Petroleum Co., Ltd.

Li, Weiping, Xiang, Quanjin & Shu, Fuchang. (2005). Laboratory Evaluation on a New Oil-soluble Temporary Plu0gging Agent DFA for Workover Fluid. *Drilling Fluid & Completion Fluid.* No.22(2). P.24-28.

Luo, Jiansheng, Yan, Jienian & Fang, Dake. (2006). Low Pressure Reservoirs Protection in the Middle Block of WZ12-1 Oilfield. *Drilling Fluid & Completion Fluid.* No.23(4). P.16-20.

Yan, Jienian. (2001). *Process Engineering of Drilling Fluid.* Dongying: University of Petroleum Press.

Zhao, Xinqing. (2002). Applications of TB-O Low-harm Workover Fluid in Yakela-Dalaoba Condensate Field. *Petroleum Drilling Techniques.* No.30(3). P.75-76.

Table 1. Performance analysis of workover fluid (Unit: mg/l)

Water type	Total mineralization	Cation						Anion			
		Total Fe	Zn^{2+}	Si^+	Ca^{2+}	Mg^{2+}	Ba^{2+}	Cl^-	SO_4^{2-}	S^{2-}	HCO_3^-
$CaCl_2$	255000	106.3	68.3	29.4	2790	574.1	8.8	153000	468	Null	3.45

Table 2. Rheology of the system

Formula	Experiment condition	Density (g/cm3)	AV mPa·s	PV mPa·s	YP Pa	API Vf ml	pH
New workover fluid	Before hot aging	1.17	20	10	10	6.5	8.0
	Hot aging (120℃×16h)	1.17	13	9	4	8	8.0

Table 3. Evaluation test result of reservoir core permeability recovery

Core No.	Perm-plug method (10-3μm2)	Brine permeability (10-3μm2)	Oleic permeability before pollution (10-3μm2)	Oleic permeability after pollution (10-3μm2)	Return permeability (%)	Pollution medium
13	18.7	14.1	10.717	6.989	65.21	Stratum water
25	35.2	26.4	15.312	12.676	82.78	New workover fluid

Table 4. Workover fluid and reservoir rock compatibility study

New workover fluid	Rock sample quality before mixing (g)	37.26
	Rock sample quality after mixing (g)	35.15
	Difference (g)	2.11

Table 5. Turbidity test result of workover fluid filtrate

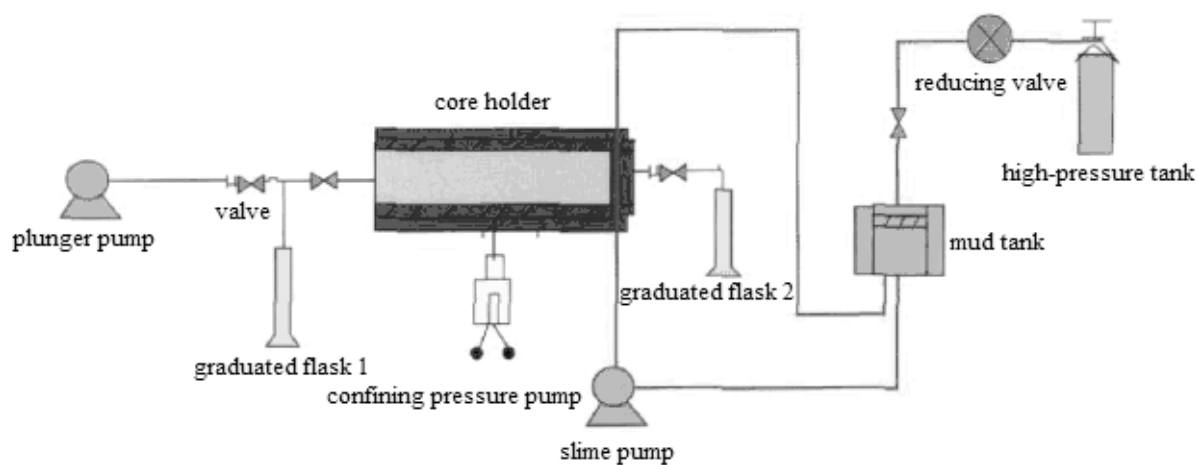| Formula	Liquid	Turbidity	Phenomenon
New workover fluid	Workover fluid filtrate	12.8	Without depositions
	Mixing of workover fluid filtrate and stratum water by 1:1	13.1	Without depositions
	Mixing of workover fluid filtrate and stratum water by 2:1	13.0	Without depositions

Figure 1. Laboratory Instrument of Dynamic Core Pollution Harm Evaluation

Drying Shrinkage of Heat-Cured Fly Ash-Based Geopolymer Concrete

Steenie Edward Wallah

Department of Civil Engineering, Faculty of Engineering

Sam Ratulangi University, Manado 95115, Indonesia

E-mail: wsteenie@yahoo.com

Abstract

Fly ash-based geopolymer concrete is manufactured without using portland cement at all. This type of concrete has environmental benefit as it has very low greenhouse gas emission compared to that resulted from the production of portland cement. In addition, it also utilizes waste or by-product material that makes it more environmentally friendly. This paper presents the study of drying shrinkage of heat-cured fly ash-based geopolymer concrete. Geopolymer concrete in this study used low-calcium fly ash as its source material, alkaline solution and aggregates normally used for ordinary portland cement concrete. Four series of test specimens with different compressive strength were prepared to study the drying shrinkage of this concrete. The test results were then compared with the calculated results of drying shrinkage as predicted by Gilbert Method which is normally used for ordinary portland cement concrete. Test results show that the heat-cured fly ash-based geopolymer concrete undergoes very low drying shrinkage. The drying shrinkage strain at one year as calculated using Gilbert Method is much higher, about five to seven times, compared to the measured drying shrinkage strain from the tests.

Keywords: Drying shrinkage, Fly ash, Geopolymer concrete, Heat-cured

1. Introduction

Environmental issue has become a crucial issue in concrete industry. This is mostly because of the emission of greenhouse gasses from the production of portland cement as the primary binder in making concrete in the meantime. Many efforts have been trying in order to reduce the use of portland cement in concrete that in turn will also reduce the greenhouse gas emission. Those efforts include the use of supplementary cementing materials and finding alternatives for portland cement. In this regard, fly ash-based geopolymer concrete is a good alternative since it does not use portland cement at all. This type of concrete becomes more environmentally friendly because it also uses waste or by-product material which is fly ash. As a relatively new material, extensive studies are still needed to explore this type of concrete as a construction material. One area that needs to be studied is its shrinkage behaviour, which is an important long-term property of concrete. This paper reports the study of drying shrinkage of fly ash-based geopolymer concrete by conducting relevant testing of drying shrinkage and comparing the test results with the developed formula or equation normally used for predicting drying shrinkage of ordinary portland cement (OPC) concrete.

2. Shrinkage of Concrete

Shrinkage is the decrease in volume of concrete with time. Unlike creep, another long-term property of concrete, shrinkage is independent of the external actions to the concrete. There are some types of shrinkage in the concrete which should be distinguished. Gilbert (2002) divided them into plastic shrinkage, chemical shrinkage, thermal shrinkage and drying shrinkage.

Plastic shrinkage occurs in wet concrete or when the concrete is still in plastic state due to loss of water by evaporation or suction by the underlying concrete or soil. This could lead to significant cracking during setting. The magnitude of plastic shrinkage is affected by temperature, ambient relative humidity and wind velocity (Neville, 2000). Plastic shrinkage also depends on the cement content of the mix and the water cement ratio, it is greater for greater cement content and low water cement ratio.

Chemical shrinkage is caused by various chemical reactions within the cement paste, including the hydration shrinkage. While thermal shrinkage is related to the liberation of the heat of hydration as Portland cement reacts with water.

Drying shrinkage is the reduction in volume which is primarily caused by the loss of water during the drying process. Drying shrinkage normally accounts for the biggest proportion of the total long-term shrinkage. Factors which affect the

drying of concrete also affect the magnitude and rate of development of drying shrinkage. Those factors include the type and content of cement or binder, water content and water to cement ratio, type of aggregate, maximum size and its proportion in the concrete, relative humidity and the size and shape of the member.

The aggregates plays a significant role in affecting the shrinkage of concrete (de Larrard et. al., 1994; Neville, 2000). This is related to the restraining effect of the aggregate on shrinkage. The higher aggregate content results in smaller shrinkage and also concrete with aggregates of higher modulus or rougher surfaces is more resistance to the shrinkage process.

The higher water to cement ratio normally results in higher shrinkage due to interrelated effects. As water to cement ratio increases, paste strength and stiffness decrease and as the water content increases, shrinkage potential increases because it also reduces the volume of restraining aggregates.

The relative humidity affects the magnitude of shrinkage as the rate of shrinkage is lower at higher values of relative humidity. The rate and magnitude of shrinkage decrease with an increase in the volume of concrete member, but the duration of shrinkage is longer for large members since more time is needed for shrinkage effects to reach the interior regions (de Larrard et. al., 1994).

The shrinkage strain of concrete which is usually considered to be the sum of drying, chemical and thermal shrinkage components, continues to increase with time at a decreasing rate (Gilbert, 2002). And as for creep, shrinkage is also assumed to approach a final value as time approaches infinity. Shrinkage is a concern in concrete structures since it is probably the most common cause of cracking. Shrinkage also causes axial deformation and warping which could lead to significant deflection and the shrinkage induces tension and resulting cracks, if not controlled, can lead to serviceability, durability and even shear strength failure (Gilbert, 1988; Rusch et. al., 1983).

There are some formulas or equations that have been developed to estimate the shrinkage of concrete. One of them is as proposed by Gilbert (2002) that will be used in this study as a comparison to the experimental results.

The method proposed by Gilbert divides the total shrinkage strain (ε_{cs}) into endogenous shrinkage (ε_{cse}) and drying shrinkage (ε_{csd}). Endogenous shrinkage is taken to be the sum of chemical and thermal shrinkage. The total shrinkage strain is given by Equation 1 and the endogenous shrinkage at any time t (in days) after concrete placement is given by Equation 2.

$$\varepsilon_{cs} = \varepsilon_{cse} + \varepsilon_{csd} \tag{1}$$

$$\varepsilon_{cse} = \varepsilon_{cse}{}^* (1.0 - e^{-0.1t}) \tag{2}$$

Where $\varepsilon_{cse}{}^*$ is the final endogenous shrinkage and may be taken as

$$\varepsilon_{cse}{}^* = (0.06 f'_c - 1.0) \times 50 \times 10^{-6} \tag{3}$$

in which f'$_c$ is in MPa.

The drying shrinkage at time t (in days) after the commencement of drying may be taken as

$$\varepsilon_{csd} = k_1 k_4 \varepsilon_{csd.b} \tag{4}$$

where $\varepsilon_{csd.b}$ is given by Equation 5. In Equation 5, $\varepsilon_{csd.b}{}^*$ depends on the quality of the local aggregates and may be taken as 800 x 10^{-6} for Sydney and Brisbane, 900 x 10^{-6} for Melbourne and 1000 x 10^{-6} elsewhere.

$$\varepsilon_{csd.b} = (1.0 - 0.008 f'_c) \times \varepsilon_{csd.b}{}^* \tag{5}$$

The factor k_1 in Equation 4 is given by Equation 6, and the factor k_4 is taken equal to 0.7 for an arid environment, 0.65 for an interior environment, 0.6 for a temperate inland environment and 0.5 for a tropical or near-coastal environment.

$$k_1 = \frac{\alpha_1 t^{0.8}}{t^{0.8} + 0.15 t_h} \tag{6}$$

where

$$\alpha_1 = 0.8 + 1.2 e^{-0.005 t_h} \tag{7}$$

and the hypothetical thickness, t_h is given by Equation 8, where A is the cross-sectional area of the member and u_e is that part of the perimeter of the member cross-section which is exposed to the atmosphere.

$$t_h = \frac{2A}{u_e} \tag{8}$$

3. Fly Ash-Based Geopolymer Concrete

Fly ash-based geopolymer concrete is geopolymer concrete utilizes fly ash as its source material. The source material for making geopolymer concrete should be rich in silica and alumina. Unlike the cement-based concretes that utilise the formation of calcium-silica hydrates (CSHs) for matrix formation and strength, geopolymers involve the chemical reaction of alumino-silicate oxides with alkali polysilicates yielding polymeric $Si - O - Al$ bonds (Davidovits, 1991; van Jaarsveld, et. al., 2002). In geopolymer concrete, the silica and the alumina present in the source materials are first induced by alkaline activators to form a gel. This geopolymer gel binds the loose aggregates and other unreacted materials in the mixture to form the geopolymer concrete. In this experimental work, fly ash is used as the source material to make geopolymer paste as the binder, instead of cement paste, to produce concrete. The geopolymer paste binds the loose coarse aggregates, fine aggregates and other un-reacted materials together to form the fly ash-based geopolymer concrete. The manufacture of geopolymer concrete is carried out using the usual concrete technology methods. As in the Portland cement concrete, the aggregates occupy the largest volume, i.e. about 75-80 % by mass, in fly ash-based geopolymer concrete. The silicon and the aluminium in the fly ash are activated by a combination of sodium hydroxide and sodium silicate solutions to form the geopolymer paste that binds the aggregates and other un-reacted materials.

Previous research has been reported on the studies of fly ash-based geopolymer concrete as in Hardjito et. al. (2004) who studied the development of this concrete including the effects of various parameters. Previous studies also indicates that fly ash-based geopolymer concrete possesses good long-term properties and durability (Wallah, et al., 2004; Wallah and Rangan, 2006). Moreover, fly ash-based geopolymer concrete is a potential material for structural application (Sumajouw and Rangan, 2006; Chang, et. al., 2007). The use of geopolymer technology in making concrete has environmental benefit as it could reduce the CO_2 emission to the atmosphere up to 80% compared to OPC concrete (Davidovits, 1994).

4. Experimental Work

Geopolymer concrete in this study utilized the low calcium (class F) fly ash from Collie Power Station, Western Australia as the source material. Aggregates, comprising 20 mm, 14 mm and 7 mm coarse aggregates and fine aggregate in saturated surface dry conditions, were used. The coarse aggregates were crushed granite-type aggregates and the fine aggregate was fine sand. The alkaline activator was a combination of analytical grade sodium hydroxide (NaOH) in flake form with 98% purity dissolved in water and sodium silicate (Na_2O= 14.7%, SiO_2=29.4%, and water=55.9% by mass solution). A high range water-reducing admixture with a dosage of 1.5% by mass of the fly ash was added to the mixture. The specimens were cured at 60^oC for 24 hours. Two types of curing were applied, dry curing or steam curing. For dry curing, the specimens were cured in an oven and for steam curing the specimens were cured in the steam curing chamber. After curing, the specimens were left to air-dry in the laboratory until testing.

This study focused on the drying shrinkage of fly ash-based geopolymer concrete, where the procedure as in Australian standard, AS 1012.13 (1992) was used as the basis to determine the drying shrinkage through an experimental or laboratory testing.

Test specimens for drying shrinkage test were 75x75x285 mm prisms with the gauge studs as shown in Figure 1. Three specimens were prepared for each type of test. In addition, for each type of test, four 100x200 mm cylindrical specimens were also prepared for compressive strength test.

Four series of concrete specimens (Table 1) were used for drying shrinkage test designated with 1DS, 2DS, 3DS and 4DS. The first two comes from one mixture proportion designed for higher strength while the other two comes from one mixture proportion with lower strength. Two types of heat curing which is dry curing and steam curing, were applied with one type for one test series in each mixture.

The shrinkage strain measurements started on the third day after casting the concrete. On the third day after casting, the specimens were demoulded and the first measurement was taken. Horizontal length comparator (Figure 2) was used for length measurements. The next measurement was on the fourth day of casting, considered as Day 1 for the drying shrinkage measurements. The measurements then continued every day in the first week, once a week until the fourth week, once in two weeks until the twelfth week, and then once in four weeks until one year.

During the drying shrinkage tests, the specimens were kept in a laboratory room where the temperature was maintained at approximately at 23^oC. The relative humidity of the room varied between 40% and 60%.

5. Results and Discussions

5.1 Shrinkage Test Results

Table 2 presents the resulted 7th day compressive strength of each test category. For concrete from mixture-1, the compressive strength is 65 MPa and 57 MPa for 1DS and 2DS respectively, while for mixture-2, those values are 50 MPa and 41 MPa for 3DS and 4DS respectively.

Figures 3 and Figure 4 show the plots of drying shrinkage strain versus age in days for the heat-cured test specimens. It can be seen from these Figures that heat-cured fly ash-based geopolymer concrete undergoes very low drying shrinkage. For all test specimens, from both mixtures and curing types and different compressive strength, the final value of drying shrinkage strain after a one-year period was only around 100 microstrain. Although each type of mixture and curing type results in different compressive strength, the drying shrinkage strain does not have significant difference among those four series of tests.

The test data plotted in Figures 3 and 4 show that the drying shrinkage strains fluctuated slightly over the period of measurement. This could be attributed to the moisture movement from the environment to the concrete or vice versa which causes reversible shrinkage or swelling of the concrete. Also, there were some minor differences in the measured values of drying shrinkage strains between dry and steam cured specimens. However, these variations are considered to be insignificant in the context of the very low drying shrinkage experienced by the heat-cured geopolymer concrete specimens.

5.2 Comparison of Test Results and Prediction

The drying shrinkage of heat-cured fly ash-based geopolymer concrete is generally very low compared to that of ordinary portland cement concrete. This can also be seen if the test results are compared with the values predicted by using one of the available prediction formula for OPC concrete.

In this study, the measured drying shrinkage strains are compared with the values predicted by a method proposed by Gilbert (2002). which is also included in the Australian Standard for Concrete Structures AS3600 as described in Section 2.

The measured shrinkage strains are compared with the predictions by Gilbert method in Figure 5 to Figure 8. In these calculations, the factor k_4 was taken as equal to 0.65 as the test specimens were exposed to an interior environment and the value of f'$_c$ was taken as the 7th day compressive strength of the test specimens as given in Table 2.

It can be seen from Figures 5 to Figure 8 that the measured drying shrinkage strains of heat-cured fly ash-based geopolymer concrete specimens are significantly smaller than the predicted values.

In heat-cured fly ash-based geopolymer concrete, most of the water released during the chemical reaction may evaporate during the curing process (Davidovits, 1999; Hardjito & Rangan, 2005). Because the remaining water contained in the micro-pores of the hardened concrete is small, the induced drying shrinkage is also very low.

Davidovits (Personal communication) suggested that the smaller drying shrinkage strain of heat-cured fly ash-based geopolymer concrete may be explained by the 'block-polymerisation' concept. According to this concept, the silicon and aluminium atoms in the fly ash are not entirely dissolved by the alkaline liquid. The 'polymerisation' that takes place only on the surface of the atoms is sufficient to form the 'blocks' necessary to produce the geopolymer binder. Therefore, the insides of the atoms are not destroyed and remain stable, so that they can act as 'micro-aggregates' in the system and this could 'increase' the aggregate content in concrete. As for OPC concrete, aggregate content will influence the magnitude of shrinkage as the shrinkage of concrete will decrease with the increase in the quantity of aggregates. The proportion of aggregates in the mixtures of fly ash-based geopolymer concrete used in this work is approximately similar to that used in OPC concrete. However, the presence of the 'micro-aggregates' due to the 'block-polymerisation' gives the effect of increasing the aggregate content in the concrete. In other words, the presence of the 'micro-aggregates' increases the restraining effect of the aggregates on drying shrinkage.

7. Conclusion

Heat-cured fly-ash based geopolymer concrete undergoes very low drying shrinkage. The drying shrinkage strains fluctuated slightly over the period of measurement and the value at one year measurement is only around 100 microstrain. The test measurement at one year for all test series of specimens with different compressive strength, which were produced from different mixtures and curing types, does not have significant difference. The values of drying shrinkage strain predicted using Gilbert Method is much higher, about five to seven times of the measured drying shrinkage strain..

References

Chang, E. H., Sarker, P., Lloyd, N., & Rangan, B. V. (2007). Shear behaviour of reinforced fly ash-based geopolymer concrete beams. Paper presented at the The 23rd Biennial Conference of the Concrete Institute of Australia, Adelaide, Australia.

Davidovits, J. (1999, 30 June - 2 July 1999). Chemistry of Geopolymeric Systems, Terminology. Paper presented at the Geopolymere '99 International Conference, Saint-Quentin, France.

Davidovits, J. (1994). Global Warming Impact on the Cement and Aggregates Industries. *World Resource Review*, 6(2), 263-278.

Davidovits, J. (1991). Geopolymers: Inorganic Polymeric New Materials. *Journal of Thermal Analysis*, 37, 1633-1656.

de Larrard, F., Acker, P., & Roy, R. L. (1994). Shrinkage creep and thermal properties. In S. P. Shah & S. H. Ahmad (Eds.), *High Performance Concretes and Applications* (pp. 65-114). London: Edward Arnold.

Gilbert, R. I. (2002). Creep and shrinkage models for high strength concrete - proposal for inclusion in AS3600. *Australian Journal of Structural Engineering*, 4(2), 95-106.

Gilbert, R. I. (1988). *Time Effects in Concrete Structures*. Amsterdam: Elsevier.

Hardjito, D., Wallah, S. E., Sumajouw, D. M. J., & Rangan, B. V. (2004). On the Development of Fly Ash-Based Geopolymer Concrete. *ACI Materials Journal*, 101(6), 467-472.

Hardjito, D., & Rangan, B. V. (2005). Development and Properties of Low-Calcium Fly Ash-Based Geopolymer Concrete. Perth, Australia: Faculty of Engineering, Curtin University of Technology.

Neville, A. M. (2000). *Properties of Concrete* (Fourth and Final ed.). Essex, England: Pearson Education, Longman Group.

Rusch, H., Jungwirth, D., & hilsdorf, H. K. (1983). Creep and Shrinkage, Their Effect on the Behaviour of Concrete Structures. New York: Springer-Verlag.

Standards-Australia. (1992). Methods of testing concrete - Determination of the drying shrinkage of concrete for samples prepared in the field or in the laboratory. AS 1012.13 - 1992.

Sumajouw, M. D. J., & Rangan, B. V. (2006). Low-Calcium Fly Ash-Based Geopolymer Concrete: Reinforced Beams and Columns. Perth, Australia: Faculty of Engineering, Curtin University of Technology.

van Jaarsveld, J. G. S., van Deventer, J. S. J., & Lukey, G. C. (2002). The effect of composition and temperature on the properties of fly ash- and kaolinite-based geopolymers. *Chemical Engineering Journal*, 89(1-3), 63-73.

Wallah, S. E., Hardjito, D., Sumajouw, D. M. J., & Rangan, B. V. (2004). Geopolymer Concrete: A Key for Better Long-Term Performance and Durability. Paper presented at the ICFRC International Conference on Fibre Composites, High Performance Concretes and Smart Materials, Chennai, India.

Wallah, S. E., & Rangan, B. V. (2006). Low-Calcium Fly Ash-Based Geopolymer Concrete: Long-Term Properties. Perth, Australia: Faculty of Engineering, Curtin University of Technology.

Table 1. Test Parameters for Drying Shrinkage Test

Test Designation	Mixture	Curing type
1DS	Mixture-1	Dry
2DS	Mixture-1	Steam
3DS	Mixture-2	Dry
4DS	Mixture-2	Steam

Table 2. Compressive Strength of Heat-Cured Fly Ash-based Geopolymer Concrete Shrinkage Specimens

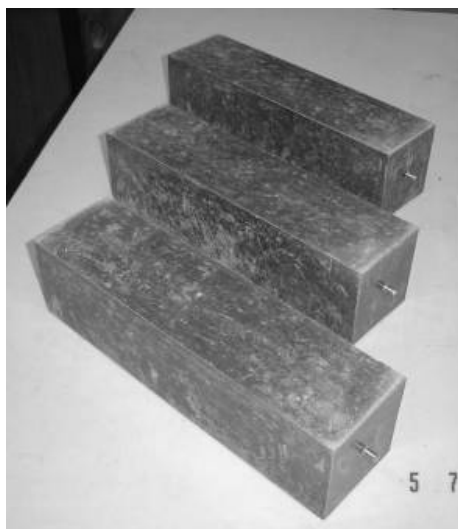| Test Designation	Type of mixture	Curing type	7th Day compressive strength (MPa)
1DS	Mixture-1	dry	65
2DS	Mixture-1	steam	57
3DS	Mixture-2	dry	50
4DS	Mixture-2	steam	41

Figure 1. Specimens for Drying Shrinkage Test

Figure 2. Horizontal Length Comparator with a Drying Shrinkage Test Specimen

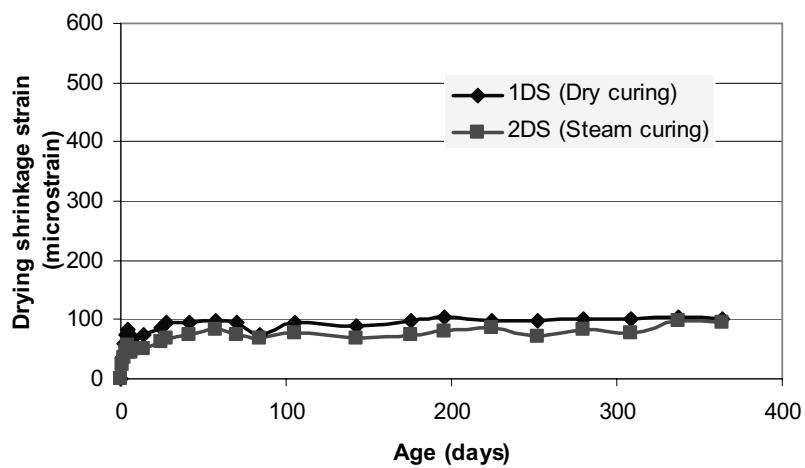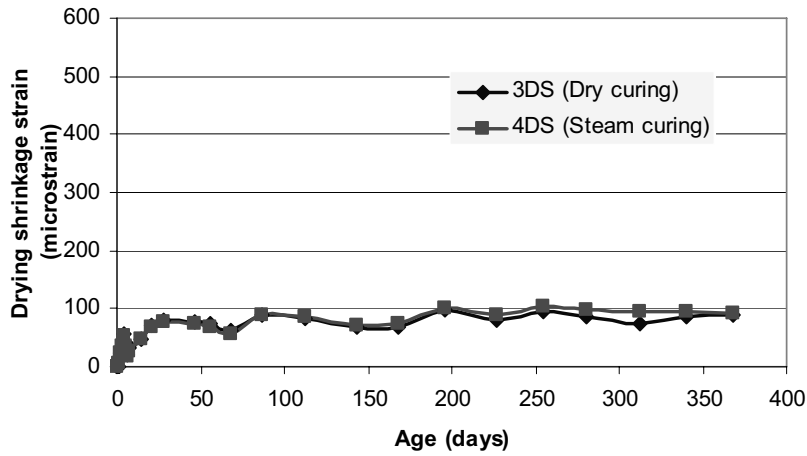Figure 3. Drying Shrinkage of Heat-Cured Mixture-1 Specimens

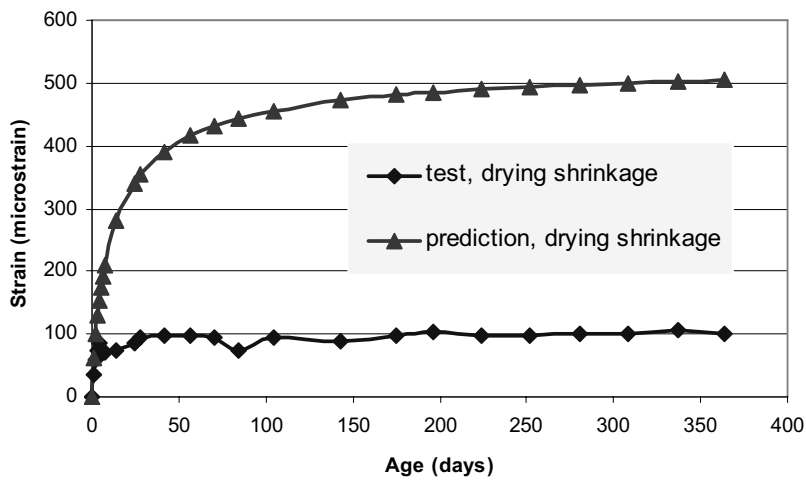Figure 4. Drying Shrinkage of Heat-Cured Mixture-2 Specimens

Figure 5. Comparison of Test and Predicted Shrinkage Strains for 1DS
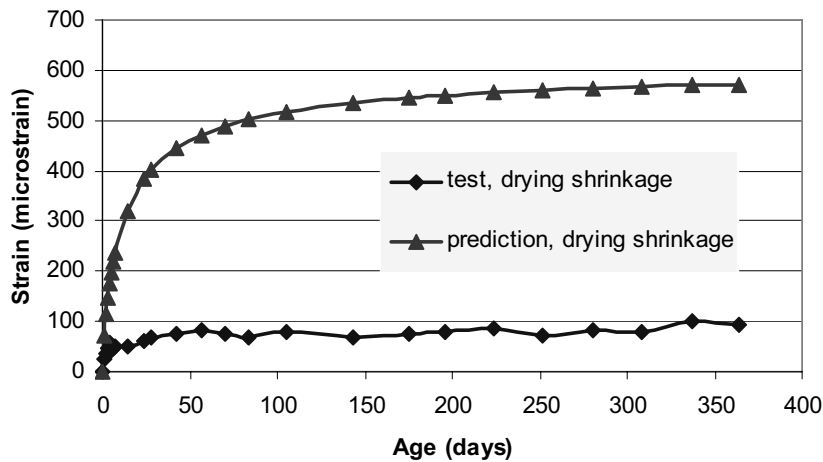

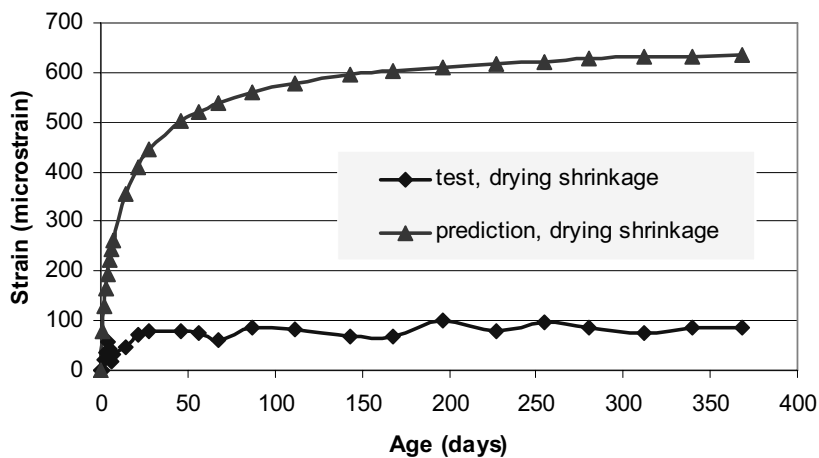Figure 6. Comparison of Test and Predicted Shrinkage Strains for 2DS

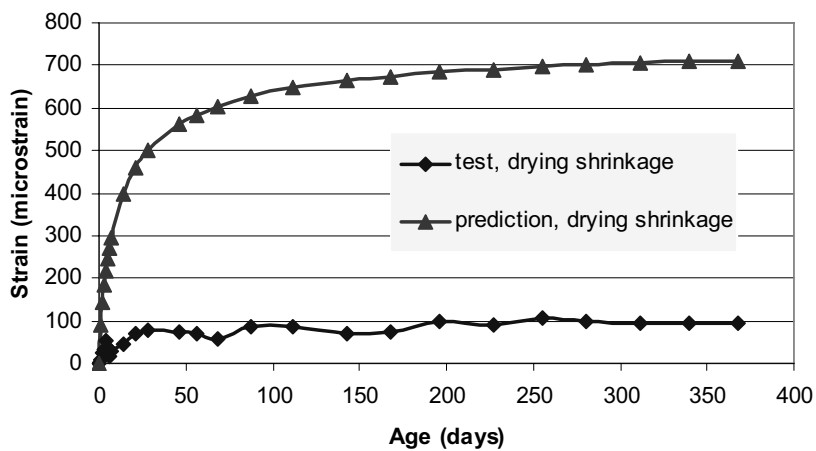Figure 7. Comparison of Test and Predicted Shrinkage Strains for 3DS

Figure 8. Comparison of Test and Predicted Shrinkage Strains for 4DS

Finite Element Analysis of the Stability of Tunnel Surrounding Rock with Weak Rock Layer

Yangsong Zhang

Nanjing University of Science and Technology, Nanjing 210094, China

Tel: 86-25-8431-5773   E-mail: zysnj318@126.com

Xiaozhao Li

Nanjing University, Nanjing 210093, China

Tel: 86-25-8359-7196   E-mail: lixz@nju.edu.cn

Abstract

The evaluation of local stability of the tunnel surrounding rock mass with weak rock layer is very important for the economy and security of the tunnel construction. Based on the rock tunnel geological conditions in the construction, the finite element model of the tunnel surrounding rock mass stability analysis with weak rock layer is established in the article by selecting the ABAQUS program linear D-P materials model as the gneiss yield criterion. The computation result indicates that the explicit dynamic analysis program could simulate the whole process of the instabilization and failure of the weak rock layer, and avoid the convergence existing in the general finite element computation, and the analysis model and program could offer effective measure for the analysis of local stability of the tunnel surrounding rock mass with weak rock layer.

Keywords: Rock tunnel, Weak rock layer, Explicit finite element analysis, Stability

1. Introduction

The empirical method and the numerical analysis method are two main methods to analyze the stability of the tunnel surrounding rock mass at present. The empirical method is to mainly qualitatively analyze the stability of tunnel surrounding rock mass according to the quality classification result of rocks, and the quality classification methods extensively used in the world include the Geomechanics Classification System (RMR, Bieniawski, 1989), the Q-system (Barton et al., 1974 & Barton, 2002), the Geological Strength Index (GSI, Hoek and Brown, 1997), and the Chinese national standard BQ classification method which is extensively used in China (The Ministry of Water Resources People' Republic of China, 1995). These classification methods are the results of researches in many tunnel cases, and have been applied in many evaluations about the stability of practical tunnel surrounding rock mass. But these methods could not quantitatively compute the stress redistribution aroused in tunnelling works (Basarir et al., 2005), and could not reasonably consider the local stability of unequal tunnel surrounding rocks with weak rock layer. On the other hand, the finite element method could offer the displacement of the surrounding rock mass after tunnel is excavated, the stress redistribution, the plastic zone distribution, and other useful information, which can help to quantitatively evaluate the stability of the surrounding rock mass (Zhu Weishen et al., 2003 & Basarir et al., 2005). However, similarly, the reasonability of the results computed by many numerical methods such as finite element method depends on whether the space distribution of rock medium in the established numerical analysis models is close to the actual situation to large extent, and the reasonability of input parameters in various rock mass computations.

For the surrounding rock mass that the rock mass have good quality and high strength, when analyzing the stability of tunnel surrounding rock mass by the finite element method, the security coefficient is often used to denote the factor of safety of stability, but to solve the factor of safety of the stability of tunnel surrounding rock mass, the strength reduction method is mainly adopted at present (Zheng Yingren and Zhao Shangyi, 2005). But when computing the local stability of surrounding rock mass with weak rock layer, and the strength parameter of rock mass is not reduced, the weak rock layer may be failed locally, and it is very important to mainly study when the weak rock layer begins to become into instabilization and failure, and the influences of the whole stability of tunnel surrounding rock mass after

the weak rock layer is failured, and the local supporting opportunity, and the supporting effect. There are not systematic research results about these problems at present. In the article, taking geological conditions of Yuntaishan Road Tunnel in Lianyungang of Jiangsu, China, as the example, the finite element computation analysis method about the local stability of the tunnel surrounding rock with weak rock layer is discussed as follows.

2. Explicit finite element analysis of the whole process of instabilization and failure of tunnel surrounding rock mass

2.1 Instabilization criteria of tunnel surrounding rock mass

At present, the criterion that the finite element method is adopted to compute extreme status of instabilization and collapsed of rock-soil mass includes following kinds.

(1) The distribution of plastic zone. The size of the plastic zone of surrounding rock mass in excavaiing is one standard to evaluate the stability of surrounding rock mass (Ferrero Anna Maria, Migliazza Maria, Giani Gian Paolo, 2004). But for tunnels, the concept of the size of the plastic zone of surrounding rock mass or whether it is connected or not is not specific like the slope and the specific borderlines are difficult to be confirmed.

(2) The computating that the finite element is not convergent indicates that the rock-soil mass has been destroyed (Dawson E.M., Roth W. H., 1999 & Griffiths D.V., Lane, P.A., 1999). According to this standard, different finite element programs and different convergence standards will produce different computation results. And in the finite element simulation analysis of tunnelling, the result of the general finite element program is not convergent often, so the further change of the surrounding rock mass status after local destroying can not be known.

(3) The sign that the rock-soil mass is collapsed should be that the slide moves infinitely, and the slide surface of rock-soil mass and the displacement change and development infinitely. If the numerical method can compute the deformation of surrounding rock mass, and follow the whole process of the deformation and instability of surrounding rock mass, so this rule is very good. And it can monitor various points of the surrounding rock mass of different classes in different parts of the tunnel, which can not only know the local instability, but know the total instability induced by the local instability. It specially suits for the analysis and judgment of the local stability of tunnel surrounding rock mass with weak rock layer.

2.2 Constitutive model of rock mass

Mohr-Coulomb strength rule and the Drucker-Prager strength rule are two kinds of usual rules in the rock-soil mass computation. The strength parameter of the Mohr-Coulomb is easily obtained usually, but in the extreme analysis of rock-soil mass, its computation result often is not convergent at the sharp corner. In the article, the linear Drucker-Prager model is adopted. The linear Drucker-Prager yielding function and the plastic potential energy respectively are (ABAQUS Inc, 2006)

$$F = t - p \tan \beta - d = 0 \tag{1}$$

$$G = t - p \tan \psi \tag{2}$$

Here, $t = \dfrac{1}{2} q \left[1 + \dfrac{1}{K} - \left(1 - \dfrac{1}{K} \right) \left(\dfrac{\gamma}{q} \right)^3 \right]$, $r = (\dfrac{9}{2} S \cdot S : S)^3$, $S = \sigma + p\mathrm{I}$

$\beta(\theta, f_i)$ is the slope of the linear yielding plate on the p–t stress plane, and it is usually called as the friction angle of materials.

d is the cohesive strength of materials, and $d = \left(1 - \dfrac{1}{3} \tan \beta \right) \sigma_c^0$, σ_c^0 is the uniaxial compressive yielding stress, and

$K(\theta, f_i)$ is the ratio between the uniaxial extension yielding stress and the uniaxial compressive yielding stress, so it control the dependence of the middle main stress of the yielding plane.

$\psi(\theta, f_i)$ is the dilation angle in the p–t plane.

The Mohr-Coulomb rule can be denoted as (ABAQUS Inc, 2006)

$$\sigma_1 - \sigma_3 + (\sigma_1 + \sigma_3) \sin \varphi - 2c \cdot \cos \varphi = 0 \tag{3}$$

For the triaxial compression, the formula (1) can be rewrote as

$$\sigma_1 - \sigma_3 + \frac{\tan \beta}{2 + \dfrac{1}{3} \tan \beta} (\sigma_1 + \sigma_3) - \frac{1 - \dfrac{1}{3} \tan \beta}{1 + \dfrac{1}{6} \tan \beta} \sigma_c^0 = 0 \tag{4}$$

By Comparing the formula (3) with the formula (4), β and σ_c^0 corresponding with the Drucker-Prager compression cone and the Mohr-Coulomb rules are

$$\tan \beta = \frac{6 \sin \varphi}{3 - \sin \varphi} \tag{5}$$

$$\sigma_c^0 = 2c \frac{\cos \varphi}{1 - \sin \varphi} \tag{6}$$

In the same way, β and σ_c^0 corresponding with the Drucker-Prager tensile cone and the Mohr-Coulomb rules respectively are

$$\tan \beta = \frac{6 \sin \varphi}{(3 + \sin \varphi)} \tag{7}$$

$$\sigma_c^0 = c \frac{\cos \varphi (6 - \tan \beta)}{3 - \tan \beta} \tag{8}$$

$$K = \frac{1}{1 + \frac{1}{3} \tan \beta} \tag{9}$$

Fit the average value of the compression cone and the tensile cone of the Drucker-Prager yielding rule with the Mohr-Coulomb rule, and take the average $\tan \beta$ value of the formula (5) and the formula (7) to compute $K(\theta, f_i)$ according to the formula (9), and $0.778 \leq K(\theta, f_i) \leq 1$, and in the same way, $K(\theta, f_i)$ controls the dependence of the middle main stress of the yielding plane. Large numerous of comparisons and computations indicated that the parameters computed by the linear Drucker-Prager yielding rule and above methods and Mohr-Coulomb had good fitting precision.

2.3 Analysis approaches

In general finite element analysis, there are three methods to simulate the tunnelling, i.e. (1) gradually reducing the elasticity modulus of the excavating rock mass, (2) adopting the killing element or removing element method, (3) substituting the rock mass in excavating tunnels by the fixed boundary conditions, and first computing the initial stress to balance the reverse stress of above borderlines, and in subsequent computation, putting the borderline reverse function on the responding nodes, and removing corresponding fixed borderline conditions on the tunnelling plane, and computing the ground stress balance, and in subsequent computation, gradually reduce the reverse stress on the tunnelling face to simulate the stress release in tunnelling according to certain proportion. In ABAQUS/Explicit analysis, above two methods can not obtain satisfactory excavating modeling effect. Though the third method is complex, but it fits for the ABAQUS/Explicit analysis, and it is convenient to accurately control the stress release.

According to the surrounding rock mass condition, each step to excavate tunnel is 3.0m, and the detailed simulation process includes following steps.

(1) Compute the reverse forces of various nodes on the walls of the excavating tunnel. It only needs to implement the fixed borderline conditions on the tunnelling face in the step of the initial geostatic stress balance, so the reverse forces of various nodes can be obtained.

(2) Implement the reverse forces of various nodes in the computation of first step on the corresponding nodes to compute the initial geostatic stress balance again.

(3) Gradually reduce the reverse forces of various nodes on the tunnel surfaces to zero step by step, to simulate the stress release of rock mass in each time. The final simulation computation includes the release of the reverse forces of the nodes on the surface of surrounding rock mass in front of the tunnel.

The first step is implemented in ABAQUS/Standard, and the second step and the third step are implemented continually in ABAQUS/Explicit.

3. Example analysis

3.1 Geological conditions and physical mechanical parameters of surrounding rock masses

The computation antitype is the surrounding rock mass section of the tunnel, mainly includes the solid metamorphic rocks, and the saturated uniaxial compressive strength of rocks is R_C=50~100MPa. For two main joints strikes are NE and NW in the rock mass, the joint plane is planar and coarse, and presents close-parting opening status without filling. The schistose structural plane occurrence average dip is 155^0, and the average obliquity angle is 28^0. Some parts with

the green schist weak rock layer unveil, and the saturated uniaxial comprehensive strength of rocks is R_C=5~17MPa, and the schistose structural plane of rock mass grows. The weak rock distribute as the intercalated status along the schistose structural plane direction. The field survey shows that when the weak rock layer is distributed on the top of the tunnel, if the local support measure is not adopted, the local instabilization and failure of surrounding rock mass easily occur. According to the field survey, the geometric relationship model between the weak rock distribution and the tunnel is seen in Figure 1.

The physical mechanical parameters confirmed by the modified basic quality index [BQ] of surrounding rock masses are seen in Table 1, and the Table 1 also shows the internal friction angle β and the compressive initial yielding stress σ_c^0 of the linear Drucker-Prager rule.

For the linear Drucker-Prager rule in the ABAQUS program, to ensure that the yielding plane is out-extrusive, the value of *K* should equal or exceed 0.778, so when value of *k* confirmed by the formula (6) is less than 0.778, *K*=0.778. Take nonassociated flow rules, suppose the expansion angle is $\psi(\theta, f_i)$ =β/4 (Hock and Brown, 1997).

3.2 Finite element analysis model

In the computation, for the model, the length is 200m (x axis), the height is 120m (z axis), the depth is 45m (y axis), the span of the tunnel is 16m, and the height of the tunnel is 9.05m, and the simulated excavating depth of the tunnel is 15m, and whole sections are dug. The length from the top of the tunnel to the top of the model is 41m, and the rock mass of 200m covering on it adopts the average press on the surface of the model to be simulated. As seen in Figure 2, the model is divided into 121714 4-node linear tetrahedron elements and 21815 nodes.

3.3 Results of computation and analysis

According to the numerical simulation results, through observing the vertical displacement of the tunnel top and the distribution of the plastic zone of surrounding rock masses, following rules can be summarized.

3.3.1 Vertical displacement of tunnel top

The distribution of the final vertical displacement (U3) of tunnel surrounding rock mass obtained in the finite element computation is seen in Figure 3. From Figure 3, the vertical displacement value from the top of the tunnel in the weak rock mass ([BQ]=240) distribution zone obviously exceeds the corresponding values surrounding the solid rock zone ([BQ]=500). The maximum of the vertical displacement of the tunnel top is at the node A, and it is 17.23mm in the weak rock layer, and the minus denotes the displacement direction downwards, which is observe of the direction of Z axis. But the displacements in the solid rock zone of the tunnel top are small, for example, the vertical displacement of the node C far from the weak rock distribution is 6.38mm.

In the finite simulation computation of tunnelling, the load release proportion factor (LPD) and the computation time have same meanings. With the increase of the excavating step computation time, the node force on the excavating borderline gradually reduces to zero. So from the result of the finite element computation, the relationship between the node displacement on the excavating plane and the load release proportion factor can be easily obtained.

Figure 4 shows the relationship curve between the vertical displacement and the load release proportion factor of three places on the top of tunnel. Obvious inflexion occurs in the curve of the node A (weak rock layer) with the increase of the load release proportion factor when LPD is close to 0.9, i.e. the displacement increases quickly, so according to the result of the finite element computation, comparing with the extreme status rule 2, if any support measure is adopted, the weak rock layer on the top of tunnel will be collapsed. That accords with actual situation, in the tunnelling construction of Yuntaishan in Lianyungang, in the solid II-class surrounding rock mass, because the quality of rock mass was good, the systematic primary support had not been adopted according to relative standards, and the stability was good when the weak rock layer didn't exist, and the deformation of surrounding rock mass was small, but for the same solid II-class surrounding rock mass, when the top of tunnel had the weak rock layer with big depth (1~3m) and effective local support had not been adopted, several collapses occurred in the weak rock surrounding layer on the top of tunnel in certain time after the tunnel was excavated.

From Figure 4, the vertical displacements of the node B and the node C in the solid rock zone are obviously less than the vertical displacement of the node A, and because the node B is close to the edge of the weak rock layer, the vertical displacement is larger than the node C far from the weak rock mass because of the influence of the weak rock plastic failure. The vertical displacements of the node C and the node B increase slowly, which indicates that they have not achieved the extreme status.

3.3.2 Plastic zone of surrounding rock mass

Figure 5 shows the distribution of the plastic zone of surrounding rock mass. From Figure 5, the plastic zone occurs in the top of tunnel and the upper weak rock layer on the excavating plane. The solid rock distribution zones all have not achieved the plastic failure. From the development process of the plastic zone, when the load proportion factor is 0.9, obvious plastic variance begins in the weak rock layer, which accords with the time when the inflexion of the vertical

displacement-load proportion factor curve of the node A, and that indicates that the quick increase of displacement is the induced by the obvious plastic variance in surrounding rock mass.

4. Conclusions

Combining with the linear Drucker-Prager yielding rule, the ABAQUA/Explicit program is adopted to analyze the stability of the tunnel surrounding rock mass with the weak rock layer. By the research, following conclusions can be obtained.

(1) The numerical analysis program in the article can be used to compute the whole process of the deformation and failure of the tunnel surrounding rock mass, and it can avoid the problem that the results of the finite element computation is not convergent, so it specially suits for the computation of the stability of the tunnel surrounding rock mass with weak rock layer.

(2) It is reasonable and feasible to apply the ABAQUS/Explicit program to compute and adopt the rule that the deformation increases quickly as the extreme status of the instabilization and failure of tunnel surrounding rock mass. On the one hand, the computation program can ensure to obtain the inflexion that the displacement quickly increases, and on the other hand, the results of the computation and judgment are coherent with the actual situation.

(3) In the article, the quality classification standard of rock mass is adopted to evaluate the strength parameters of Mohr-Coulomb rule, and these parameters are converted into the strength parameters of the linear Drucker-Prager yielding rule by the method in the article. Then they are used to compute the stability and the displacement of tunnel surrounding rock mass, and the computation results is coherent with the actual survey result.

References

ABAQUS Inc. (2006). ABAQUS User's Manual, Version 6.6.

Barton, N., Lien, R., Lunde, J. (1974). Engineering classification of rock masses for the design of tunnel support. *Rock Mech.* No.6(4). P.189-236.

Basarir, H., Ozsan, A., Karakus, M. (2005). Analysis of support requirements for a shallow diversion tunnel at Guledar dam site, Turkey. *Engineering Geology.* No.81. P.131-145.

Bieniawski, Z.T. (1989). Engineering Rock Mass Classification: A Complete Manual for Engineers and Geologists in Mining, Civil, and Petroleum Engineering. *Wiley, New York.* 251 pp.

Dawson E.M. and Roth W. H. (1999). Drescher A. Slope stability analysis by strength reduction. *Geotechnique.* No.49 (6). P.835-840.

Ferrero Anna Maria, Migliazza Maria, Giani Gian Paolo. (2004). Analysis of tunnel stability: comparison between continuous and discontinuous approaches. *Int J Rock Mech. Min. Sci.* No.41(3). CD-ROM. P.1-6.

Hoek E and Brown ET. (1997). Practical estimates of rock mass strength. *Int J Rock Mech Min Sci Geomech Absrr.* No.34(8). P.1165-1186.

The Ministry of Water Resources People' Republic of China. (1995). *The Chinese Standards-Standard for engineering classification of rock masses (GB50128-94).* Beijing: China Planning Press (in Chinese).

Zheng, Yingren and Zhao, Shangyi. (2005). Limit state finite element method for geotechnical engineering analysis and its applications. *China civil engineering journal.* No.38(1). P.91-104 (in Chinese).

Zhu, Weishen, Li, Shucai, Li, Shuchen, et al. (2003). Systematic numerical simulation of rock tunnel stability considering different rock conditions and construction effects. *Tunnelling and Underground Space Technology.* No.18. P.531-536.

Table 1. Physical mechanic parameters based on the modification of the basic quality index [BQ] of surrounding rock mass

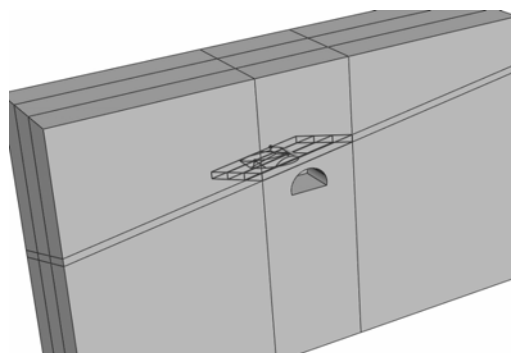| Type of rock	[BQ]	Class of rock mass	Deformation modulus E_{mass} (GPa)	Passion ratio v	Density of gravity (kN/m^3)	Mohr-Coulomb criterion		Drucker-Prager criterion	
						$\varphi(°)$	c (MPa)	$\beta(°)$	σ_c^0 (MPa)
Solid rock	500	II	26.5	0.23	26.0	55	1.8	59.1	7.13
Weak rock	240	V	1.25	0.36	22.0	26	0.19	41.59	0.504

Figure 1. Weak Rock layer on the Top of the Tunnel (the lines in the model present the range of weak rock layer)

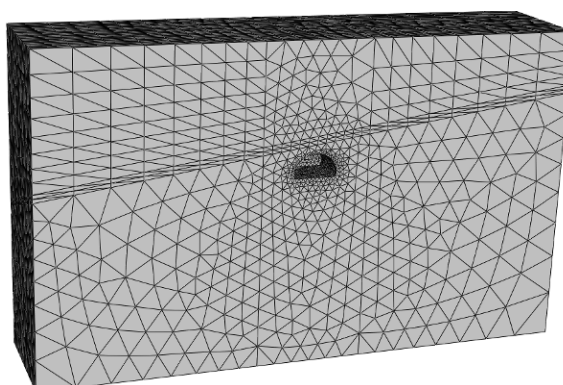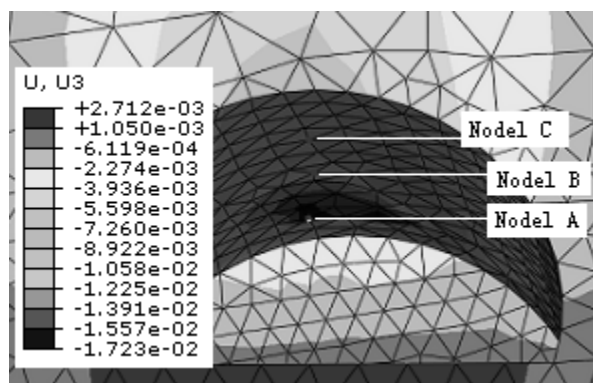Figure 2. Finite Mesh of the Tunnel Computation Model

Figure 3. Vertical Displacement U3 of Tunnel Surrounding Rock mass (Unit: m) (Nodel A: weak rock layer,
Nodel B: solid rocks near to weak rock layer, Nodel C: solid rocks far from weak rock layer)
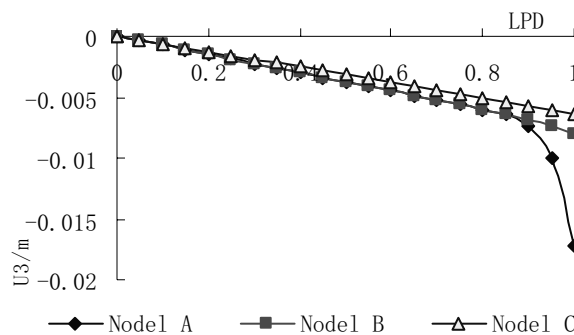

Figure 4. Relationship Curve between Vertical Displacement of Surrounding Rock Mass
on the Top of Tunnel and the Load Proportion factor (LPD)

Prediction of Stock Market Index Movement by Ten Data Mining Techniques

Phichhang Ou (Corresponding author)

School of Business, University of Shanghai for Science and Technology

Rm 101, International Exchange Center, No. 516, Jun Gong Road, Shanghai 200093, China

Tel: 86-136-617-515-547, Fax: +86-21-55271502   E-mail: phichhang@gmail.com

Hengshan Wang

School of Business, University of Shanghai for Science and Technology

Box 461, No. 516, Jun Gong Road, Shanghai 200093, China

Tel: 86-21-5527-5971   E-mail: wanghs@usst.edu.cn

Abstract

Ability to predict direction of stock/index price accurately is crucial for market dealers or investors to maximize their profits. Data mining techniques have been successfully shown to generate high forecasting accuracy of stock price movement. Nowadays, in stead of a single method, traders need to use various forecasting techniques to gain multiple signals and more information about the future of the markets. In this paper, ten different techniques of data mining are discussed and applied to predict price movement of Hang Seng index of Hong Kong stock market. The approaches include Linear discriminant analysis (LDA), Quadratic discriminant analysis (QDA), K-nearest neighbor classification, Naïve Bayes based on kernel estimation, Logit model, Tree based classification, neural network, Bayesian classification with Gaussian process, Support vector machine (SVM) and Least squares support vector machine (LS-SVM). Experimental results show that the SVM and LS-SVM generate superior predictive performances among the other models. Specifically, SVM is better than LS-SVM for in-sample prediction but LS-SVM is, in turn, better than the SVM for the out-of-sample forecasts in term of hit rate and error rate criteria.

Keywords: Data mining, Stock price movement prediction, SVM, LS-SVM, NN, Bayesian classification with Gaussian processes

1. Introduction

Financial market is a complex, nonstationary, noisy, chaotic, nonlinear and dynamic system but it does not follow random walk process, (Lo & Mackinlay, 1988; Deng, 2006). There are many factors that may cause the fluctuation of financial market movement. The main factors include economic condition, political situation, traders' expectations, catastrophes and other unexpected events. Therefore, predictions of stock market price and its direction are quite difficult. In response to such difficulty, data mining (or machine learning) techniques have been introduced and applied for this financial prediction. Most of the studies have focused on the accurate forecasting of the value of stock price. However, different investors adopt different trading strategies; therefore, the forecasting model based on minimizing the error between the actual values and the forecasts may not be suitable for them. In stead, accurate prediction of movement direction of stock index is crucial for them to make effective market trading strategies. Some recent studies have suggested that trading strategies illustrated by the forecasts based on the direction of stock price change may be more effective and generate higher profit. Specifically, investors could effectively hedge against potential market risk and speculators as well as arbitrageurs could have opportunity of making profit by trading stock index whenever they could obtain the accurate prediction of stock price direction. That is why there have been a number of studies looking at direction or trend of movement of various kinds of financial instruments (such as Wu & Zhang, 1997; O'connor et al, 1997). But, these studies do not use data mining based classification techniques. Data mining techniques have been

introduced for prediction of movement sign of stock market index since the results of Leung et al (2000) and Chen et al (2001), where LDA, Logit and Probit and Neural network were proposed and compared with parametric models, GMM-Kalman filter. Kim (2003) applied newly and powerful techniques of data mining, SVM and Neural network, to forecast the direction of stock index price based on economic indicators. To obtain more profits from the stock market, more and more "best" forecasting techniques are used by different traders. In stead of a single method, the traders need to use various forecasting techniques to gain multiple signals and more information about the future of the markets. Kumar & Thenmozhi (2006) collected five different approaches including SVM, Random forecast, Neural network, Logit and LDA to predict Indian stock index movement based on economic variable indicators. From the comparison, the SVM outperformed the others in forecasting S&P CNX NIFTY index direction as the model does not require any priori assumptions on data property and its algorithm results global optimal solution which is unique. Huang et al (2005) also forecasted the movement direction of Japanese stock market (NIKKEI 225 index) by various techniques such as SVM, LDA, QDA, NN and the all-in-one combined approach. The SVM approach also gives better predictive capability than other models: LDA, QDA and NN, following the out-performance of the combined model. In the study, they defined the movement of the NIKKEI 225 index based on two main factors including American stock market, S&P 500 index, which is the most influence on the world stock markets including Japanese market, and the currency exchange rate between Japanese Yen and US dollar.

In our study, ten different techniques of data mining are discussed and applied to predict price movement of Hang Seng index of Hong Kong stock market. The approaches include Linear discriminant analysis (LDA), Quadratic discriminant analysis (QDA), K-nearest neighbor classification, Naïve Bayes based on kernel estimation, Logit model, Tree based classification, neural network, Bayesian classification with Gaussian process, Support vector machine (SVM) and Least squares support vector machine (LS-SVM). The main goal is to explore the predictive ability of the ten data-mining techniques in forecasting movement direction of Hang Seng Index based on five factors, including its open price, high price, low price, S&P 500 index, and currency exchange rate between HK dollar and US dollar. The general model of stock price movement is defined as

$$D_t = f(O_{t-1}, H_{t-1}, L_{t-1}, S\&P500_{t-1}, FX_{t-1}),$$

where D_t is the direction of HSI movement at time t and is defined as a categorical value "1" if the closing price at time t is greater than the closing price at time $t-1$ and as "0", otherwise. The function $f(.)$ can be linear or nonlinear and it is estimated by the ten data-mining algorithms.

O_{t-1} denotes the open price of HSI at time $t-1$;

H_{t-1} is the high price of HSI at time $t-1$;

L_{t-1} is the low price in a day of HSI at time $t-1$;

$S\&P500_{t-1}$ is the closing price of S&P 500 index at time $t-1$.

FX_{t-1} is the currency exchange rate between HK dollar and US dollar.

All the inputs are transformed into log return to remove any trend pattern.

The remaining of the paper is organized as follow. Next sections describe the data and prediction evaluation. Section 3 briefly discusses the ten different algorithms. The final section is for conclusion.

2. Data description

We examine the daily change of closing prices of Hang Seng index based on five predictors, Open price, High price, Low price, S&P 500 index price, and Exchange rate USD against HKD. The stock prices are downloaded from the Yahoo finance and the foreign exchange rate is taken from website of Federal Reserve Bank of ST. Louis. The sample period is from Jan 03 2000 to Dec. 29 2006 so that the whole sample is of 1732 trading days. The data is divided into two sub-samples where the in-sample or training data spans from Jan 03 2000 to Dec 30 2005 with 1481 trading days. The whole year 2006 from Jan 1 2006 to Dec 29 2006 of size 250 trading days are reserved for out-of-sample or test data. Figure1 displays the actual movement of HSI closing prices for the whole sample. Figure 2 plots S&P 500 price and its log return and Figure 3 shows the plots of price and log return of exchange rate of HKD against USD. To measure the predictive performances by different models, Hit rate and Error rate are employed and defined as Hit rate = $\frac{1}{m} \sum_{i=1}^{m} I_{[A_i = P_i]}$ and Error rate = $\frac{1}{m} \sum_{i=1}^{m} I_{[A_i \neq P_i]}$ where A_i is the actual output for ith trading day and P_i is the predicted value for ith trading day, obtained from each model. Here m is the number of the out-of-sample. R software with related packages is used to conduct the whole experiment; for example Karatzoglou (2004, 2006) illustrates the R commands for SVM, LSSVM, and Bayesian classification with Gaussian Processes.

3. Data-mining methods

Let the training data be $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ where $X = (X_1, \ldots, X_p)$ denote real-valued random input vector and the response Y is categorical, i.e. $Y \in \{1, 2, \ldots, K\}$. The goal is to form a predictor $G(x)$ to predict Y based on X. So $G(x)$ divides the input space (or feature vector space) into a collection of regions where each labeled by one class. For binary class problem, $Y \in \{1, 2\}$, the decision boundary between the two classes is a hyperplane in the feature space. A hyperplane in the p- dimensional input space is the set: $\{x : \beta_0 + \sum_{i=1}^{P} \beta_i x_i = 0\}$. (1)

The two regions separated by the hyperplane are $\{x : \beta_0 + \sum_{i=1}^{P} \beta_i x_i > 0\}$ and $\{x : \beta_0 + \sum_{i=1}^{P} \beta_i x_i < 0\}$.

Now we define Bayes classification rule. Suppose the training data $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ are independent samples from the joint distribution of X and Y, $f_{X,Y}(x, y) = p_Y(y) f_{X/Y}(x / Y = y)$ and the loss function of classifying Y as $G(X) = \hat{Y}$ is $L(\hat{Y}, Y)$, where the marginal distribution of Y is specified by the pmf $p_Y(y)$ and $f_{X/Y}(x / Y = y)$ is the conditional distribution of X given $Y = y$. The goal of classification is to minimize the expected loss defined as $E_{X,Y} L(G(X), Y) = E_X [E_{Y/X} L(G(X), Y)]$. To minimize the left hand side of the expected loss, it suffices to minimize $E_{Y/X} L(G(X), Y)$ for each X. Hence the optimal classifier is $G(X) = \arg \min_y E_{Y/X=x} L(y, Y)$. For 0-1 loss function $L(y, y') = 0$ for $y = y'$ and 0 otherwise, we have $E_{Y/X=x} L(y, Y) = 1 - \Pr(Y = y / X = x)$. Therefore, the classification rule, called Bayes rule, becomes the rule of maximum a posterior probability: $G(x) = \arg \min_y E_{Y/X=x} L(y, Y) = \arg \max_y \Pr(Y = y / X = x)$.

We consider ten algorithms for classification where some of them attempt to estimate $\Pr(Y = y / X = x)$ and then apply the Bayes rule $G(x) = \arg \max \Pr(Y = y / X = x)$. The algorithms include linear discriminant analysis (LDA), Quadratic discriminant analysis (QDA), Naïve Bayes based on kernel estimation, and Logit model. Other types of data mining techniques are K-nearest neighbor classification, Tree based classification, neural network, Bayesian classification with Gaussian process, Support vector machine (SVM) and Least squares support vector machine (LS-SVM). The last three models take more advantage via kernel based methods.

3.1 Linear Discriminant Analysis

The goal here is to obtain class posteriors $\Pr(Y / X)$ for optimal classification. Suppose $f_k(x)$ is the class-conditional density of X in class $Y = K$ and let π_k be the prior probability of class K with $\sum_{k=1}^{K} \pi_k = 1$. By Bayes theorem, $\Pr(Y = k / X = x) = f_k(x) \pi_k / \sum_{i=1}^{K} f_i(x) \pi_i$. Suppose each class density is from Gaussian distribution defined as $f_k(x) = (2\pi)^{-p/2} |\Sigma_k|^{-1/2} \exp(-\frac{1}{2}(x - \mu_k)^T \Sigma_k^{-1}(x - \mu_k))$ and the classes are assumed to have a common covariance matrix $\Sigma_k = \Sigma \ \forall k$. Considering the log ratio of the two classes k and l posteriors

$\log [\Pr(Y = k / X = x) / \Pr(Y = l / X = x)]$

$= \log [f_k(x) / f_l(x) + \log [\pi_k / \pi_l] = \log [\pi_k / \pi_l] - \frac{1}{2}(\mu_k + \mu_l)^T \Sigma^{-1}(\mu_k - \mu_l) + x^T \Sigma^{-1}(\mu_k - \mu_l)$ which is a linear equation in x in p dimensional hyperplane defined in (1), the linear discriminant functions are obtained $\delta_k(x) = x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k^T \Sigma^{-1} \mu_k + \log \pi_k$ and the LDA classifier is $G(x) = \arg \max_k \delta_k(x)$. From the training data, we can estimate the Gaussian distribution parameters as

$\hat{\pi}_k = N_k / N$ where N_k is the number of class k observations; $\hat{\mu}_k = \sum_{y_i=k} x_i / N_k$ and $\hat{\Sigma} = \sum_{k=1}^{K} \sum_{y_i=k} (x_i - \hat{\mu}_k)(x_i - \hat{\mu}_k)^T / (N - K)$. For two classes {1, 2}, the LDA rule classifies to class 2 if $x^T \hat{\Sigma}^{-1}(\hat{\mu}_2 - \hat{\mu}_1) > \frac{1}{2} \hat{\mu}_2^T \hat{\Sigma}^{-1} \hat{\mu}_2 - \frac{1}{2} \hat{\mu}_1^T \hat{\Sigma}^{-1} \hat{\mu}_1 + \log(N_1 / N) - \log(N_2 / N)$ and class 1 otherwise. (2)

From the experimental result, we have $X = (X_1, \ldots X_5)$ with dimension of p =5 so that five coefficients of the linear discriminant in (2) are obtained as the following: V1 = -0.79761825, V2 = 0.91216594, V3 = 0.74655315, V4 = -0.02241239, V5 = -1.85401191. Prior probabilities of groups: 0 and 1 are 0.5037137 and 0.4962863 respectively. For the in-sample data, the hit rate is 0.8393 and the error rate is 0.1607, while in the out-of-sample data, the hit rate and error rate are 0.8440 and 0.1560 respectively.

3.2 Quadratic discriminant analysis

For the QDA, the above Σ_k are not assumed equal for each k, then $\delta_k(x) = -\frac{1}{2}\log|\Sigma_k| - \frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1}(x-\mu_k) + \log \pi_k$. The decision boundary between two classes k and l is the quadratic equation $\{x : \delta_k(x) = \delta_l(x)\}$. The estimates for QDA are similar to those LDA except that separate covariance matrix must be estimated for each class. Here QDA needs $(K-1)\{p(p+3)/2+1\}$ parameters. See McLachlan (1992) and Duda et al (2000) for comprehensive discussion on discriminant analysis. From the prediction results, the hit rate and error rate for training data by QDA are 0.8305 and 0.1695 respectively. For the test data, the hit rate and error rate are 0.8480 and 0.1520 respectively.

3.3 K-nearest neighborhood method

K-nearest neighbor method is one of the simplest machine learning algorithms used for classifying objects based on closest training examples in the feature space. An object is classified by a majority being assigned to the class most common amongst its k nearest neighbors. Formally, the k-nearest neighbor approach uses the training data set $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ closest in input space to x to form $\hat{Y}$. Specifically, the k-nearest neighbor fit for $\hat{Y}$ is defined as $\hat{y}(x) = \frac{1}{k}\sum_{x_i \in N_k(x)} y_i$ where $N_k(x)$ is the neighborhood of x defined by k closest points x_i in the training example. That is, we find the k observations with x_i closest to x in input space and average their responses. k is estimated by cross-validation technique. The algorithm starts with the determination of the optimal k based on RMSE done by cross validation technique, then calculate the distance between the query distance and all the training samples. After sorting the distance and determination of the nearest neighbors based on the kth minimum distance, gather the Y of the nearest neighbors. Finally, use simple majority of the category Y of nearest neighbors as the prediction value of the query distance. Noticeably, the k-nearest neighbor approach does not rely on prior probabilities like LDA and QDA.

Results: Table 1 displays the process of choosing k by cross-validation technique in the experiment. We consider k = 30 as an initial range and then select optimal k in the range. The best k = 10 is obtained corresponding to smallest error (0.1708). [Insert Table 1 around here]. The performance results are given as the following. For the training data, the hit rate is 0.8312 and the error rate is 0.1688 and for the test data, the hit rate is 0.7960 and the error rate is 0.2040.

3.4 Naïve Bayes classification method

This is a well established Bayesian method primarily formulated for performing classification tasks. Given its simplicity, i.e., the assumption that the independent variables are statistically independent, Naive Bayes models are effective classification tools that are easy to use and interpret. Naive Bayes is particularly appropriate when the dimensionality of the feature space (i.e., number of input variables) is high (a problem known as the curse of dimensionality). Mathemcally, Naïve Bayes model requires an assumption that given a class $Y = j$, the features X_k are independent so that $f_j(X) = \prod_{k=1}^{p} f_{jk}(X_k)$. The estimates of $f_{jk}(.)$ is from the training data via kernel smoothing. Naïve Bayes classification is $G(x) = \arg\max_j f_j(X)\pi_j$ and π_j is estimated by the sample proportions. See Mitchell (2005) for precise explanation. From the result obtained in the experiment, the hit rate and error rate for in-sample data are 0.8386 and 0.1614 respectively. For the out-of-sample, the hit rate is 0.8280 and the error rate is 0.1720.

3.5 Logit model

Logistic regression refers to methods for describing the relationship between a categorical response variable and a set of predictor variables. It can be used to predict a dependent variable on the basis of independents and to determine the percent of variance in the dependent variable explained by the independents. The logistic regression applies maximum likelihood estimation after transforming the dependent into a logit variable. In this way, logistic regression estimates the probability of a certain event occurring. Note that logistic regression calculates changes in the log odds of the dependent, not changes in the dependent itself. From the Friedman et al (2008), the model for logistic regression is given as: $\pi(x) = \Pr(Y = 1/X = x) = \dfrac{\exp(\beta_0 + \sum_{i=1}^{p}\beta_i X_i)}{1 + \exp(\beta_0 + \sum_{i=1}^{p}\beta_i X_i)}$ for two classes of output Y. We obtain β's using the maximum likelihood approach.

Logit is given by: $G(x) = \log\left(\dfrac{\pi(x)}{1+\pi(x)}\right) = \log\dfrac{\Pr(Y=1/X=x)}{\Pr(Y=0/X=x)} = \beta_0 + \sum_{i=1}^{p}\beta_j X_j$. The curve of $\pi(x)$ are called sigmoid because they are S-shape and therefore nonlinear. Statisticians have chosen the logistic distribution to model binary data because of its flexibility and interpretability. The minimum for $\pi(x)$ is attained at $\lim_{a \to -\infty} e^a/(1+e^a) = 0$, and the maximum for $\pi(x)$ is obtained at $\lim_{a \to \infty} e^a/(1+e^a) = 1$.

From the experiment, the coefficients of β's estimated by maximum likelihood are obtained as $\beta_0 = -0.0031$, $\beta_1 = -2.3285$, $\beta_2 = 2.6697$, $\beta_3 = 2.4129$, $\beta_4 = 0.0736$ and $\beta_5 = -3.1932$. The performance results show that the hit rate and error rate for in-sample data is 0.8474 and 0.1526 respectively while the hit rate and error rate for out-of-sample data are 0.8560 and 0.1440 respectively.

3.6 Tree based classification

Classification tree is one of the main techniques used in Data Mining. Classification trees are used to predict membership of objects in the classes of a categorical dependent variable from their measurements on one or more predictor variables. The goal of classification trees is to predict or explain responses on a categorical dependent variable, and as such, the available techniques have much in common with the techniques used in the more traditional methods of Discriminant Analysis, Cluster Analysis, Nonparametric Statistics, and Nonlinear Estimation. The flexibility of classification trees makes them a very attractive analysis option as it does not require any assumption on the distribution like traditional statistical methods. Technically, the Tree-based methods partition the feature space into a set of rectangles, and then fit a simple model in each one. Starting from the root of a tree, the feature space X containing all examples is split recursively into subsets usually two at a time. Each split depends on the value of only a unique variable of input x. If x is categorical, the split is of the form $x \in A$ or $x \notin A$ where A is subset of X. The goodness of split is measured by an impurity function defined for each node. The basic idea is to choose a split such that the child nodes are purer than their parent node. The split continues till the end subsets (leaf nodes) are 'pure'; that is till one class dominates. For an impurity function ϕ, define the impurity measure by $i(t) = \phi(\Pr(1/t), \Pr(2/t), \ldots, \Pr(K/t))$ where $\Pr(j/t)$ is the estimated probability of class j within node t. The goodness of a split s for node t is

$\phi(s,t) = \Delta I(s,t) = i(t) - p_R i(t_R) - p_L i(t_L)$ where p_R and p_L are the proportions of the samples in node t that go to the right node and left node respectively. Possible impurity functions:

a. Entropy: $i(t) = \sum_{j=1}^{K} \Pr(j/t) \log(1/\Pr(j/t))$

b. Gini index: $i(t) = \sum_{j=1}^{K} \Pr(j/t)(1 - \Pr(j/t))$ where $\Pr(j/t) = \dfrac{1}{N(t)} \sum_{x \in t} I_{[j_x = j]}$, $I_{[j_x = j]} = 1$ if $j_x = j$ and 0 otherwise,

and $N(t)$ is the total number of samples in node t. The criteria is to stop splitting a node t when $\max_{s \in S} \Delta I(s,t) < \beta$, β is the chosen threshold. It is not trivial to choose β as it leads to overfitting or underfitting problems for a new data prediction. To solve this problem, we go for pruning approach. The idea is to obtain the subtree from the initial large tree. One of the most popular techniques is the Cost-complexity pruning. Let $T_{\max}$ be initial large tree and let the pruned subtree by $T \leq T_{\max}$. Then the cost complexity measure $C_\alpha(T)$ is defined as $C_\alpha(T) = R(T) + \alpha \left| \tilde{T} \right|$, $\left| \tilde{T} \right|$ denotes the number of leaf nodes. $R(T) = \sum_{t \in \tilde{T}} \Pr(t) \, r(t)$ is the error measure of T, $r(t) = 1 - \max_{j} \Pr(j/t)$ and $\alpha \geq 0$ is the complexity parameter. Here $C_\alpha(T)$ represents the tradeoff between the cost of a tree and its complexity. The goal of cost-complexity pruning is for each α choose a tree $T(\alpha) \leq T_{\max}$ such that $C_\alpha(T)$ is minimized. The estimation of α is achieved by crossvalidation. We choose $\hat{\alpha}$ that minimizes the cross-validation sum of squares. Thus the final tree is $T(\hat{\alpha})$. Tree based method is considered one of top ten algorithm of data mining technique (Wu et al, 2008). We refer to Quinlan (1986) and Friedman et al (2008) for detailed discussion on the tree based classification. For illustrative application of Tree method in predicting stock price behavior is referred to Pearson (2004).

Experimental results:

Denote variables actually used in tree construction by V1 = Open price, V2 = Low price, V3=High price, V4= S&P500, V5 = FX, V6 = Close price. The class "1" is when the next price is larger than the previous price, while the class "0" is refers to when the next price is smaller than the previous price. Table 2 show the process of pruning tree.

From the Table 2, we choose CP = 0.0036281 corresponding to 17 splits based on the smallest value of the x_error: 0.40272. The initial large tree is not displayed to reduce the space but pruned tree is illustrated in the appendix section. The performance result shows that the hit rate is 0.8717 and error rate is 0.1283 in the training data; while the test data, the hit rate is 0.8 and the error rate is 0.2.

3.7 Neural network for classification

Haykin (1994) defines neural network as a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects: (1) Knowledge is acquired by the network through a learning process, and (2) Interneuron connection strengths known as synaptic weights are used to store the knowledge. The literature on neural network is enormous, and its application spreads over

many scientific areas see Bishop (1995) and Ripley (1996) for detailed. Recently, the neural network has been well known for good capability in forecasting stock market movement. Let's go directly to its formulation.

Let X be input vector and Y be output taking value as categorical. Following notations in Hastie (1996), the neural network model can be represented as

$z_j = \sigma(\alpha_{0j} + \alpha_j^T x)$, $\quad j = 1, \ldots, m$

$\hat{y}_k = f_k(\beta_{0k} + \beta_k^T z)$, $\quad k = 1, \ldots, q$

where $\sigma(z) = 1/(1 + e^{-z})$ activation function called sigmoid. The parameters α_{jl} and β_{kj} are known as weights and α_{0j} and β_{0k} are bias. Here we use $f_k(v) = 1/(1 + e^{-v})$ the inverse logit for binary classification.

To learn the neural network, back propagation is used.

Suppose we use least squares on a sample of training data to learn the weights:

$R(\alpha, \beta) = \sum_{i=1}^{N} \sum_{j=1}^{k} (y_k^i - \hat{y}_k^i)^2$,

$\dfrac{\partial R^i}{\partial \beta_{kj}} = -2(y_k^i - \hat{y}_k^i) f_k'(\beta_k^T z^i) z_j^i$,

$\dfrac{\partial R^i}{\partial \alpha_{jl}} = -\sum_{k=1}^{K} 2(y_k^i - \hat{y}_k^i) f_k'(\beta_k^T z^i) \beta_{kj} \sigma'(\alpha_j^T x^i) x_l^i$.

The ith component is denoted by superscript i.

Gradient update at the (r+1)st iteration,

$\beta_{kj}^{(k+1)} \quad \leftarrow \quad \beta_{kj}^{(r)} - \gamma_r \sum_{i} \dfrac{\partial R^i}{\partial \beta_{kj}^{(r)}}$

$\alpha_{jl}^{(r+1)} \quad \leftarrow \quad \alpha_{jl}^{(r)} - \gamma_r \sum_{i} \dfrac{\partial R^i}{\partial \alpha_{jl}^{(r)}}$.

Here γ_r is the learning rate.

In the experimental analysis, $l = 1, \ldots, 5$; $j = 1, 2$; $k = 1, 2$ so that a 5-2-2 network with 18 weights are obtained with the following estimated weights:

$\alpha_{01} = -9.33 \quad \alpha_{11} = 2.57 \quad \alpha_{12} = -6.14 \quad \alpha_{13} = -3.06 \quad \alpha_{14} = 2.40 \quad \alpha_{15} = 2.70$

$\alpha_{02} = -0.19 \quad \alpha_{21} = -0.84 \quad \alpha_{22} = 0.86 \quad \alpha_{23} = 0.87 \quad \alpha_{24} = 0.09 \quad \alpha_{25} = -1.15$

$\beta_{01} = 2.59 \quad \beta_{11} = 3.99 \quad \beta_{21} = -5.82$

$\beta_{02} = -2.57 \quad \beta_{12} = -3.99 \quad \beta_{22} = 5.80$

From the prediction performances, the hit rate is 0.8481 and error rate is 0.1519 in the in-sample data. For the out-of-sample, 0.8520 and 0.1480 are hit rate and error rate respectively.

3.8 Bayesian classification for Gaussian process

Gaussian processes are based on the prior assumption that adjacent observations should convey information about each other. Particularly, it is assumed that the observed variables follow normal distribution and that the coupling between them takes place by covariance matrix of a normal distribution. Using the kernel matrix as the covariance matrix is a convenient way of extending Bayesian modeling of linear estimators to nonlinear situations.

In regression problem, the goal is to predict a real valued output based on a set of input variables. It is possible to carry out nonparametric regression using Gaussian process. With Gaussian prior and Gaussian noise model, the solution of the regression problem can be obtained via Kernel function placed on each training data; the coefficients are determined by solving a linear system. If the parameter θ indexed in the Gaussian process are unknown, Bayesian inference can be carried out for them. Gaussian process can be extended to classification problems by defining a Gaussian process over y, the input to the sigmoid function. The goal is to predict $\Pr(Y = j / X = x)$, $j = 1, \ldots, K$. For binary case, $j = 0, 1$, $\Pr(Y = 1 / X = x)$ is estimated by $\sigma(y(x))$ where $\sigma(y) = 1/(1 + e^{-y})$. The idea is to place the Gaussian prior on $y(x)$ and combine it with the training data $D = (x_i, t_i)$, $i = 1, \ldots, n$ to obtain predictions for new x points. Bayesian treatment is imposed by integrating over uncertainty in y and in the parameters that control the Gaussian prior. Then

the Laplace's approximation is employed to obtain the results of the integration over y. Specifically, let $\Pr(y)$ be the prior of $y = (y(x_1), y(x_2), \ldots, y(x_n))$ so that $\Pr(y_*, y)$ is the joint distribution including y_*. Given new input x_*, we want to predict $y_* = y(x_*)$ based on the $D = (x_i, t_i)$, $i = 1, \ldots, n$. Let $\Pr(t/y)$ be the probability of observing the particular values $t = (t_1, t_2, \ldots, t_n)^T$ given actual values y (i.e., noise model). Then we have

$$\Pr(y_*/t) = \int \Pr(y_*, y/t) dy \quad = \frac{1}{\Pr(t)} \int \Pr(y_*, y) \Pr(y) \Pr(t/y) dy \quad = \int \Pr(y_*, y) \Pr(y/t) dy \;.$$ Hence the predictive distribution for y_* is found from the marginalization of the product of the prior and the noise model.

The integral terms are estimated by Laplace's approximation. William et al (1999) and Rasmussen et al (2006) provided a comprehensive and detailed discussion on the Bayesian classification with Gaussian processes.

The following shortly describes the experimental results obtained from training and forecasting the movement of HSI by Bayesian classification with Gaussian processes.

Problem type: classification

Gaussian Radial Basis kernel function hyperparameter: sigma = 0.414967184016261

Number of training instances learned: 1481

Train error: 0.238758065

Cross validation error: 0.1647560

In-sample data: Hit rate is 0.8595 and error rate is 0.1405.

Out-of-sample data: Hit rate is 0.8520 and error rate is 0.1480.

3.9 Support vector machine for classification

A popular machine learning algorithm with neural network type is SVM, support vector machine, developed by Vapnik (1995). The SVM is a kernel based learning approach like the above Gaussian processes for classification. However, the SVM does not require any assumptions on the data property like in Gaussian processes. The SVM has been successfully applied for various areas of predictions; for instance, in financial time series forecasting (Mukherjee et al, 1997; Tay and Cao, 2001), marketing (Bend-David and Lindenbaum, 1997), estimating manufacturing yields (Stoneking, 1999), text categorization (Joachims, 2002), face detection using image (Osuna et al, 1997), handwritten digit recognition (Burges and Schokopf, 1997); Cortes and Vapnik, 1995), medical diagnosis (Tarassenko et al, 1995).

The SVM formulation can be started as the following.

Given a training set (x_i, y_i), $i = 1, 2, \ldots, N$ with input data $x_i \in R^n$ and corresponding binary class label $y_i \in \{-1, 1\}$, SVM algorithm seeks the separating hyperplane with largest margin. The problem can be formulated as follow:

$$\min_{w, b, \xi} \frac{1}{2} w^T w \tag{3}$$

subject to $\; y_i(w^T x_i + b) \geq 1 \quad i = 1, \ldots, N$. $\tag{4}$

Standard method to solve the problem (3)-(4) is convex programming, where Lagrange method is applied to transfer primal to dual problems of optimization. Specifically, we first construct Lagrangian

$$L_P = \frac{1}{2} w^T w - \sum_{i=1}^{N} \lambda_i [y_i(w^T x_i + b) - 1] \tag{5}$$

λ_i are nonnegative Lagrange multipliers corresponding to (4). The solution is achieved by a saddle point of the Lagrangian which has to be minimized with respect to w and b and maximized with respect to λ_i.

Differentiating (5) and set the results equal to zero,

$$\frac{\partial L_P}{\partial w} = w - \sum_{i=1}^{N} \lambda_i y_i x_i = 0 \tag{6}$$

$$\frac{\partial L_P}{\partial b} = \sum_{i=1}^{N} \lambda_i y_i = 0 \tag{7}$$

The optimal solution is obtained from (6) as $\; w^* = \sum_{i=1}^{N} \lambda_i^* y_i x_i$ $\tag{8}$

where * denotes optimal values. Now substituting (8) and (7) into (5),

$$L_D = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} w^T w = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j y_i y_j x_i^T x_j \qquad (9)$$

The dual problem is posed as quadratic programming:

maximize $\quad L_D$

subject to $\quad \sum_{i=1}^{N} \lambda_i y_i = 0$

$\quad 0 \le \lambda_i \le C, \quad i = 1, \dots, N$.

The conditions $\quad \lambda_i^*[y_i(w^{*T} x_i + b^*) - 1] = 0, \quad i = 1, \dots, N \qquad (10)$

implies that $\lambda_i > 0$ only when constraint (4) active. The vectors for which $\lambda_i > 0$ are called support vectors.

By (10), we obtain $b^* = y_i - w^{*T} x_i$ for any support vector x_i. By linearity of the inner product and (8), the decision function for the linear separable case is

$$f(x) = sign\left(w^T x + b\right) = sign\left(\sum_{i=1}^{N} y_i \lambda_i^* (x^T x_i) + b^*\right).$$

For linearly non-separable case, we introduce a new set of variables $\{\xi_i\}, i = 1, \dots, N$ that measure the amount of violation of the constraints.

Thus (3) and (4) are modified as

$$\min_{w,b,\xi_i} \frac{1}{2} w^T w + C\left(\sum_{i=1}^{N} \xi_i\right)^k \qquad (11)$$

subject to $\quad y_i\left(w^T x_i + b\right) \ge 1 - \xi_i, \quad i = 1, \dots, N \qquad (12)$

$\quad \xi_i \ge 0, \qquad i = 1, \cdots, N \qquad (13)$

where C and k are predetermined parameters defines the cost of constraints.

The Lagrangian is constructed as

$$L_P = \frac{1}{2} w^T w - \sum_{i=1}^{N} \lambda_i [y_i(w^T x_i + b) - 1 + \xi_i] - \sum_{i=1}^{N} \gamma_i \xi_i + C\left(\sum_{i=1}^{N} \xi_i\right)^k \qquad (14)$$

where λ_i and γ_i are Lagrange multipliers which are associated with constraints (12) and (13) respectively. The solution to this problem is determined by minimizing L_P with respect to w, k and b and maximizing with respect to λ_i and γ_i.

Differentiating (14) and setting equal to zero,

$$\frac{\partial L_P}{\partial w} = w - \sum_{i=1}^{N} \lambda_i y_i x_i = 0 \qquad (15)$$

$$\frac{\partial L_P}{\partial b} = \sum_{i=1}^{N} \lambda_i y_i = 0 \qquad (16)$$

$$\frac{\partial L_P}{\partial k} = \begin{cases} kC\left(\sum_{i=1}^{N} \xi_i\right)^{k-1} - \lambda_i - \gamma_i = 0, & k > 1 \\ C - \lambda_i - \gamma_i = 0, & k = 1 \end{cases}$$

or $\delta - \lambda_i - \gamma_i = 0$ by denoting $\sum_{i=1}^{N} \xi_i = (\delta / Ck)^{1/k-1}$ for $k > 1$. $\qquad (17)$

From (15), $w^* = \sum_{i=1}^{N} \lambda_i^* y_i x_i$. $\qquad (18)$

Substituting (18),(17),(16) into (14), we obtain

$$L_D = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j y_i y_j x_i^T x_j - \frac{\delta^{k/k-1}}{(kC)^{1/k-1}}\left(1 - \frac{1}{k}\right).$$

This leads to the dual problem as

maximize $\quad L_D = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j y_i y_j x_i^T x_j - \frac{\delta^{k/k-1}}{(kC)^{1/k-1}}\left(1 - \frac{1}{k}\right)$

subject to $\sum_{i=1}^{N} \lambda_i y_i = 0$

$0 \leq \lambda_i \leq \delta$, $i = 1, \ldots, N$.

When $k = 1$,

$$\max L_D = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j y_i y_j x_i^T x_j$$

subject to $\sum_{i=1}^{N} \lambda_i y_i = 0$

$0 \leq \lambda_i \leq C$, $i = 1, \ldots, N$.

Hence the classifier is

$$f(x) = sign \left(\sum_{i=1}^{N} y_i \lambda_i^* (x^T x_i) + b^* \right)$$

where $b^* = y_i - w^{*T} x_i$ for any support vector x_i such that $0 < \lambda_i < C$ following from $\lambda_i^* [y_i (w^{*T} x_i + b^*) - 1 + \xi_i] = 0$, $i = 1, \ldots, N$.

For nonlinear classifier, we need to map the input variable x into a higher dimensional feature space and work with linear classification in that space, i.e., $x \rightarrow \phi(x) = (\alpha_1 \phi_1(x), \ldots, \alpha_n \phi_n(x), \ldots)$ where $\{\alpha_n\}_{n=1}^{\infty}$ are some real numbers and $\{\phi_n\}_{n=1}^{\infty}$ are some real functions.

The solution of the SVM has the form

$$f(x) = sign(\phi(x)^T w^* + b^*) = sign \left(\sum_{i=1}^{N} y_i \lambda_i^* \phi(x)^T \phi(x_i) + b^* \right)$$

To avoid complex calculation of scalar product $\phi(x)^T \phi(x_i)$, we introduce kernel function:

$K(x, y) \equiv \phi(x)^T \phi(x) = \sum_{n=1}^{\infty} \alpha_n^2 \phi_n(x) \phi_n(y)$ which satisfies Mercer's conditions.

Hence, $f(x) = sign \left(\sum_{i=1}^{N} y_i \lambda_i^* K(x, x_i) + b^* \right)$. In this work, Gaussian kernel or RBF(radial basis function) is used as it tends to give good performance under general smoothing assumptions. The kernel is defined as

$K(x_1, x_2) = \exp(-\gamma \|x_1 - x_2\|^2)$. The kernel and regularized parameters (C, γ) with range $[2^{-5}, 2^5]$ are tuned by gridsearch technique to avoid overfitting problem.

By applying the above algorithm of SVM to train the data under study, training results are obtained. Table 3 illustrates the cross-validation error corresponding to the tuning parameters (C , γ). From the Table 3, we obtain the best hyperparameters C = 2^4 and $\gamma = 2^4$ (the Gaussian kernel function parameter) with the smallest error of ten fold cross-validation = 0.000006. Considering the in-sample data, the hit rate is 1.000 and error rate is 0.000 but for the out-of-sample, the hit rate is 0.860 and the error rate is 0.140.

3.10 Least square support vector machine for classification

LSSVM is a new version of SVM modified by Suykens et al (1999). LSSVM uses least square loss function to obtain a set of linear equations in dual space so that learning rate is faster and the complexity of calculation in convex programming (in SVM) is relaxed. In addition, the LSSVM avoids the drawback faced by SVM such as trade-off parameters (C, σ^2, ε) selection; instead the LSSVM requires only two hyper-parameters (γ, σ^2) to train the model. According to Suykens et al (2001), the equality constraints of LSSVM can act as recurrent neural network and nonlinear optimal control. Due to these nice properties, LSSVM has been successfully applied for classification and regression problems, including time series forecasting. Further application can be found in Van Gestel et al (2004) for detailed discussion on classification performance of LSSVM and Van Gestel et al (2001) and Ye et al (2004) for predictive capability of LSSVM in chaotic time series prediction.

Formally, given a training set $(x_i, y_i), i = 1, 2, \ldots, N$ with input data $x_i \in R^n$ and corresponding binary class label $y_i \in \{-1, 1\}$,

LSSVM formulation is represented as follow:

$\min_{w,b,e} J(w,e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{i=1}^{N} e_i^2$ subject to the equality constraints $y_i[w^T\phi(x_i)+b] = 1 - e_i$, $i = 1,\dots,N$. This formulation consists of equality instead of inequality constraints and takes into account a squared error with regularization term similar to ridge regression.

The solution is obtained after constructing the Lagrangian:

$L(w,b,e\,;\alpha) = J(w,b,e) - \sum_{i=1}^{N} \alpha_i \{ y_i[w^T\phi(x_i)+b] - 1 + e_i \}$,

where α_i are Lagrange multipliers that can be positive or negative in the LSSVM formulation. From the conditions for optimality, one obtains the Karush-Kuhn-Tucker (KKT)

$$
\begin{cases}
\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^{N} \alpha_i y_i \phi(x_i) \\[2mm]
\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{N} \alpha_i y_i = 0 \\[2mm]
\frac{\partial L}{\partial e_i} = 0 \rightarrow \alpha_i = \gamma_i e_i \quad i = 1,\dots,N \\[2mm]
\frac{\partial L}{\partial \alpha_i} = 0 \rightarrow y_i[w^T\phi(x_i)+b] - 1 + e_i = 0, \quad i = 1,\dots,N.
\end{cases}
$$

By eliminating w and e,

$$
\begin{bmatrix} 0 & y^T \\ \hline y & \Omega + \gamma^{-1}I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ 1_v \end{bmatrix}
$$

with $y = [y_1,\dots,y_N]$, $1_v = [1,\dots,1]$ $e = [e_1,\dots,e_N]$, $\alpha = [\alpha_1,\dots,\alpha_N]$.

Matrix $\Omega_{ij} = y_i y_j \phi(x_i)^T \phi(x_j) = y_i y_j K(x_i,x_j)$ for $i,j = 1,\dots,N$ satisfies Mercer's condition and the LS-SVM model for estimating classifier is obtained as

$y(x) = sign\left(\sum_{i=1}^{N} \alpha_i y_i K(x,x_i) + b \right)$.

In this work, Gaussian kernel or RBF(radial basis function) is used as it tends to give good performance under general smoothing assumptions. The kernel is defined as $K(x_1,x_2) = \exp(-\frac{1}{\sigma^2}\|x_1 - x_2\|^2)$. The kernel and regularized parameters (γ, σ^2) are tuned by gridsearch technique to avoid overfiting problem.

The experimental results show that the obtained hyper-parameters: gamma is 0.50691 chosen from the range [0.04978707, 148.4132] and sigma square is 8.7544 selected from the range [0.082085, 12.1825]. The cost of ten fold cross-validation is 0.033291. from the forecasting performance, the in-sample data hit rate is 0.8528 and error rate is 0.1472. For the out-of-sample data, the hit rate is 0.8640 and error rate is 0.1360.

To sum it up, Table 4 below gives a summary result of prediction performances by the ten different approaches. From the Table 4, we can see that almost all the algorithms generate high hit rate (more than 80%) and low error rate (less than 20%). By comparison, LSSVM ranks first one as it outperforms the other model though it is not better than SVM for in-sample prediction. The superior performances of the SVM models in this study also support the results in the literature. Bayesian classification for Gaussian process produces good prediction as neural network, following the SVM and LS-SVM. K-nearest neighbor approach gives the worst predictive ability and then Tree classification is the next one.

4. Conclusion

In this paper, we apply ten different techniques of data mining to forecast movement direction of Hang Seng index from Hong Kong stock market. All algorithms produce good prediction with hit rate more than 80%. The LS-SVM and SVM outperform the other models since theoretically they don't require any priori assumption on data property and their algorithms guarantee to efficiently obtain global optimal solution which is unique. The other models may be reliable for other markets, especially when the data fall into each of their properties. As can be seen in the figure 1-3 in appendix section, different stock prices behave differently. Therefore, all the approaches are recommended for forecasters of stock index movement and the better models, SVM and LS-SVM, are more preferred.

References

Ben-David, S., & Lindenbaum, M. (1997). Learning distributions by their density levels: A paradigm for learning without a teacher, *Journal of Computer and System Sciences* 55, 171-182.

Bishop, C. (1995). *Neural Networks for Pattern Recognition,* Clarendon Press, Oxford.

Burges, C.J.C., & Schokopf, B. (1997). Improving the Accuracy and Speed of Support Vector Machines, Advances in Neural Information Processing Systems, *MIT Press*, Cambridge, MA. pp. 475-481.

Cao, L.J., & Tay, F. (2001). Application of support vector machines in financial time series forecasting, *International Journal of Management Science*, pp. 309-317.

Chen, An-Sing, Daouk, Hazem & Leung, Mark T. Application of Neural Networks to an Emerging Financial Market: Forecasting and Trading the Taiwan Stock Index. [Online] Available: http://ssrn.com/abstract=237038 or DOI: 10.2139/ssrn.237038 (July 2001)

Cortes, C., & Vapnik, V.N. (1995). Support Vector Networks, *Machine Learning* 20, 273-297.

Deng, Min. Pattern and Forecast of Movement in Stock Price. [Online] Available: http://ssrn.com/abstract =217048 or DOI: 10.2139/ssrn.217048 (November 8, 2006)

Duda, R., Hart, P., & Stork, D. (2000). *Pattern Classification* (Second Edition), Wiley, New York.

McLachlan, G. J. (1992), *Discriminant Analysis and Statistical Pattern Recognition*, Wiley, New York.

Friedman, J., Hastie, T., & Tibshirani, R. (2008). *The Elements of Statistical Learning:Data Mining, Inference and Prediction*, Second Edition Springer.

Hastie, T. (1996). Neural Network, *Encyclopedia of Biostatistics*, John Wiley.

Haykin, S. (1994). *Neural Networks: A Comprehensive Foundation*, Macmillan, New York.

Huang, W., Nakamori, Y. & Wang, S.Y. (2005). Forecasting stock market movement direction with support vector machine, *Journal of Computers & Operational Research,* pp. 2513-2522.

Joachims, T. (2002). Learning to Classify Text using Support Vector Machines, *Kluwer Academic Publshers*, London.

Karatzoglou, A., Mayer, D., & Hornik, A. (2006). Support Vector Machines in R, *Journal of Statistical Software.*

Karatzoglou, A., Smola, A., Hornik, A., & Zeileis, A.. (2004). "An S4 Package for Kernel Methods in R", *Journal of Statistical Software*.

Kim, K.J. (2003). Financial time series forecasting using support vector machines, *Neuralcomputing*, 55, 307-319.

Kumar, Manish and Thenmozhi, M. Forecasting Stock Index Movement: A Comparison of Support Vector Machines and Random Forest, Indian Institute of Capital Markets 9th Capital Markets Conference Paper. [Online] Available: http://ssrn.com/abstract=876544(February 06, 2006)

Leung, Mark T., Daouk, Hazem, & Chen, An-Sing. Forecasting Stock Indices: A Comparison of Classification and Level Estimation Models. [Online] Available: http://ssrn.com/abstract=200429 (March 1999)

Lo, A.W., & MacKinlay, A.C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies* 1, 41-66.

Mitchell, T. M. (2007). *Machine learning*, McGraw Hill, USA.

Mukherjee, S., Osuna, E., & Girosi, F. (1997). Nonlinear prediction of chaotic time series using support vector machine. *In: Proceedings of the IEEE workshop on Neural Networks for Signal Processing*, Amelia Island, FL., pp. 511-520.

O'Connor, M., Remus, W., & Griggs, K. (1997). Going up-going down: How good are people at forecasting trends and changes in trends? *Journal of Forecasting* 16, 165-176.

Osuna, E.E, Freund, R., & Girosi, F. (1997). Support Vector Machines: Training and Applications, *MIT press*, USA.

Pearson, R.A.(2004). Tree structures for predicting stock price behaviour, *ANZIAM J.*, Austral. Mathematical Soc. 45, pp C950-C963

Quinlan, J.R. (1986). Induction of Decision Trees, *Journal of Machine Learning*, 1:81-106, Kluwer Academic Publishers.

Rasmussen, C. E. & Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*, the MIT Press, ISBN 026218253X.

Ripley, B.D. (1996). *Pattern recognition and neural networks,* Cambridge University Press.

Stoneking, D. (1999). Improving the manufacturability of electronic designs, *IEEE Spectrum* 36, 70-76.

Suykens, J.A.K, & Vandewalle, J., (1999). Least squares support vector machine classifiers. *Neural Processing Letters* (9) 293-300.

Suykens, J.A.K., Vandewalle, J., & De Moor, B. (2001). Optimal control by least squares support vector machines. *Neural Networks* (14)   23-35.

Tarassenko, L., Hayton, P., Cerneaz, N., & Brady, M. (1995). Novelty detection for the identification of masses in mammograms. *In: Proceedings fourth IEE International Conference on Artificial Neural Networks*, Cambridge, pp. 442-447.

Van Gestel, T., Suykens, J.A.K, Baesens, B., Viaene, S., Vanthienen, J., Dedene, G., Moor B.D., & Vandewalle, G. (2004). Benchmarking Least Squares Support Vector Machine Classifiers, *Machine Learning*, (54) 5-32.

Van Gestel, T.V., Suykens, J.A.K, Baestaens, D.E., Lambrechts, A., Lanckriet G., Vandaele B., De Moor B., & Vandewalle, J.(2001). Financial Time Series Prediction using Least Squares Support Vector Machines within the Evidence Framework, *IEEE Transactions on Neural Networks,* (12) 809-821.

Vapnik, V.N. (1995). *The nature of statistical learning theory*, Springer-Verlag, New York.

Williams, C. K. I. & Barber, D. (1999). Bayesian classification with Gaussian Processes, *Pattern Analysis and Machine Intelligence, 20(12) 1342-1351*, IEEE.

Wu, Y., & Zhang, H. (1997). Forward premiums as unbiased predictors of future currency depreciation: A non-parametric analysis", *Journal of International Money and Finance* 16, 609-623.

Wu, X.D., Kumar V., Quinlan J.R., Ghosh J., Yang Q., Motoda, H., McLachlan G.J., Ng A., Liu B., Yu P.S., Zhou Z.H., Steinbach, M., Hand, D.J., & Steinberg, D. (2008). Top 10 algorithms in data mining, *Knowledge Information System*, Springer-Verlag London.

Ye, M.Y., & Wang, X.D. (2004). Chaotic time series prediction using least squares support vector machine, *J. Chinese Physics*, IOP Publishing Ltd. IP address: 130.120.82.1000.

Appendices:

Table 1. Selection of k by cross-validation technique

k = 1	2	3	4	5	6	7	8	9	10
0.2302	0.2275	0.1964	0.2032	0.1897	0.1917	0.1789	0.1796	0.1735	**0.1708**
k =11	12	13	14	15	16	17	18	19	20
0.1729	0.1729	0.1742	0.1735	0.1769	0.1816	0.1762	0.1823	0.1769	0.1789
k = 21	22	23	24	25	26	27	28	29	30
0.1749	0.1776	0.1729	0.1769	0.1729	0.1769	0.1776	0.1762	0.1749	0.1749

The table displays the mean square error and corresponding value of k. The optimal k is 10 with smallest MSE 0.1708.

Table 2. Process of pruning tree

	CP	n.split	rel error	x_error	x_std
1	0.5387755	0	1.00000	1.01769	0.026178
2	0.0285714	1	0.46122	0.47211	0.022177
3	0.0149660	3	0.40408	0.43673	0.021573
4	0.0117914	5	0.37415	0.41905	0.021250
5	0.0088435	9	0.32653	0.42041	0.021275
6	0.0081633	11	0.30884	0.41905	0.021250
7	0.0068027	13	0.29252	0.41361	0.021148
8	0.0047619	15	0.27891	0.40816	0.021044
9	*0.0036281*	*17*	*0.26939*	*0.40272*	*0.020938*
10	0.0027211	20	0.25850	0.40408	0.020965
11	0.0020408	21	0.25578	0.40680	0.021017
12	0.0017007	25	0.24762	0.41497	0.021173
13	0.0013605	29	0.24082	0.41497	0.021173
14	0.0010000	36	0.23129	0.42857	0.021426

Denote variables actually used in tree construction by V1 = Open price, V2 = Low price, V3=High price, V4= S&p500, V5 = FX, V6 = Close price. The class "1" is when the next price is larger than the previous price, while the class "0" is refers to when the next price is smaller than the previous price.

We choose CP = 0.0036281 corresponding to 17 splits based on the smallest value of the x_error : 0.40272 to obtain the following pruned tree:

Pruned Tree
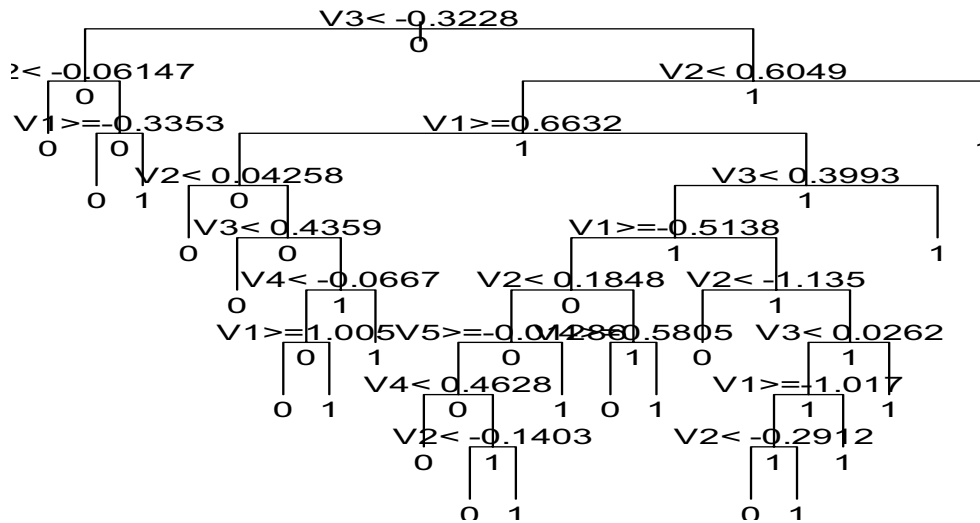
Table 3. Hyperparamters selection based on 10 fold cross-validation for SVM

C/γ	2^{-5}	2^{-4}	2^{-3}	2^{-2}	2^{-1}	2^{0}	2^{1}	2^{2}	2^{3}	2^{4}	2^{5}
2^{-5}	0.1679	0.1596	0.1532	0.1489	0.1455	0.1487	0.1619	0.2479	0.3967	0.4647	0.484945
2^{-4}	0.1576	0.1511	0.1449	0.1424	0.1369	0.1358	0.1426	0.1643	0.2962	0.4112	0.449012
2^{-3}	0.1498	0.1441	0.1400	0.1353	0.1304	0.1246	0.1247	0.1345	0.1699	0.3159	0.381449
2^{-2}	0.1439	0.1391	0.1347	0.1305	0.1233	0.1150	0.1084	0.1073	0.1134	0.1723	0.263409
2^{-1}	0.1405	0.1363	0.1329	0.1269	0.1183	0.1069	0.0930	0.0806	0.0676	0.0618	0.095686
2^{0}	0.1384	0.1341	0.1306	0.1241	0.1141	0.1006	0.0830	0.0626	0.0373	0.0152	0.004014
2^{1}	0.1364	0.1332	0.1285	0.1200	0.1110	0.0926	0.0746	0.0519	0.0238	0.0061	0.000574
2^{2}	0.1351	0.1314	0.1271	0.1170	0.1066	0.0874	0.0680	0.0409	0.0143	0.0018	0.000006
2^{3}	0.1339	0.1296	0.1252	0.1152	0.1007	0.0823	0.0623	0.0306	0.0082	0.00009	0.000006
2^{4}	0.1330	0.1283	0.1226	0.1136	0.0951	0.0776	0.0536	0.0210	0.0030	**0.000006**	0.000006
2^{5}	0.1311	0.1279	0.1206	0.1105	0.0896	0.0719	0.0443	0.0162	0.0004	0.000006	0.000006

The table illustrates the cross-validation error corresponding to the tuning parameters (C, γ). The training result shows that $(C, \gamma) = (2^4, 2^4)$ which corresponds to the smallest training error = 0.000 006.

Table 4. Summary result of prediction performances by ten different approaches

	In-sample		Out-of-sample		
Predictors	Hit Rate	Error Rate	Hit Rate	Error Rate	Rank
LDA	0.8393	0.1607	0.8440	0.1560	6
QDA	0.8305	0.1695	0.8480	0.1520	5
K-NN	0.8312	0.1688	0.7960	0.2040	9
Naïve Bayes	0.8386	0.1614	0.8280	0.1720	7
Logit model	0.8474	0.1526	0.8560	0.1440	3
Tree classification	0.8717	0.1283	0.8000	0.2000	8
Neural network	0.8481	0.1519	0.8520	0.1480	4
Gaussian process	0.8595	0.1405	0.8520	0.1480	4
SVM	1.0000	0.0000	0.8600	0.1400	2
LS-SVM	0.8528	0.1472	0.8640	0.1360	1

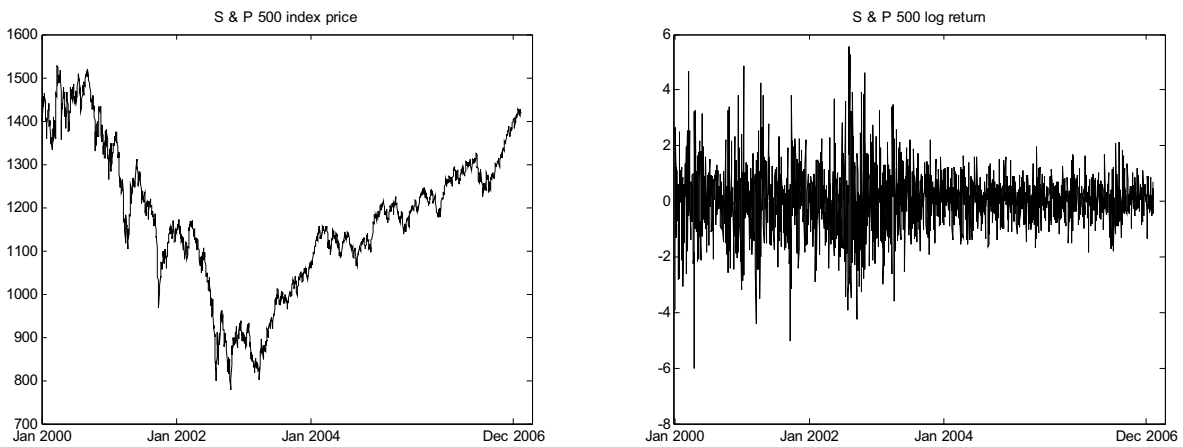Figure 1. Daily closing prices of Hang Seng Index

Figure 2. Daily closing prices (left) and Log return (right) of S & P 500 index
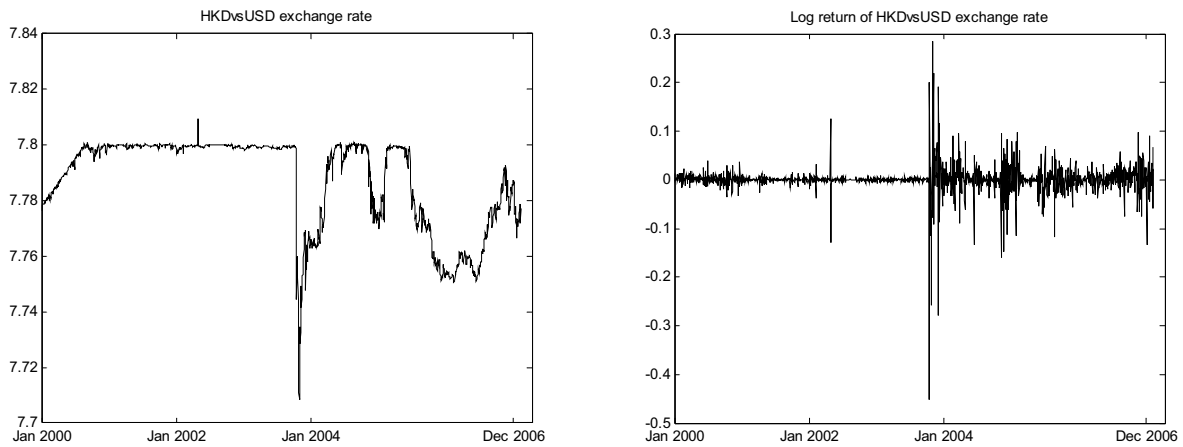

Figure 3. Daily prices (left) and Log return (right) of currency exchange rate between HKD and USD

The Development and Operation of Thermostatic Oscillator

with Network Monitoring

Junyuan Zheng

School of Electronic and information Engineering, Main campus of Soochow University

1 ShiZi Street, Suzhou 215006, China

E-mail: christina1229@sina.com

Xiaoping Chen (Corresponding author)

School of Electronic and information Engineering, Main campus of Soochow University

1 ShiZi Street, Suzhou 215006, China

Tel: 86-512-6716-0719   E-mail: xpchen@suda.edu.cn

Abstract

This paper introduces a method of thermostatic oscillator data transmission system which is based on embedded TCP/IP technology, while most thermostatic oscillator products don't have such a system of network data communication. The TCP/IP technology is used to achieve remote monitoring. It also uses the algorithm of PID and PWM for temperature and speed control, involving hardware and software design.

Keywords: The embedded system, Oscillator, TCP/IP, PWM, PID

1. Introduction

A thermostatic oscillator (also called a shaker) is a thermostatic biochemical instrument whose temperature and speed can be controlled. It is an indispensable precision training preparation laboratory equipment for plant, animal, microbial, genetic, environmental protection, medical viruses, scientific research, education system, and production department. TCP/IP which becomes a fact of popular network communication protocols for the international industrial standards, use a hierarchical structure to provide user with a rich application service. Meanwhile, the embedded system's internet network research capacity and application of the embedded technology are hot research topic.

This paper sets out to explain how the embedded system works, it is based on a core of ARM7 and communicates with the TCP/IP network, carrying out real-time monitoring of temperature and speed for the thermostatic oscillator.

2. The overall design and system structure

This system uses the MCU of LPC2368, which is the kernel of ARM7. The main frequency can reach 72M/Hz. Whether for high-speed, real-time processing or large capacity data transmission, it can all meet all the requirements. The main chip controls temperature and the speed module, etc. It has a built-in Ethernet module: this module contains a 10Mbps or 100Mbps Ethernet MAC (media access controller). The system also uses the National Semiconductor company DP83848 single road 10/100Mb/s Ethernet transceiver and Ethernet interface RJ45 which support 10MBit and 100MBit adaptive network speed to connect to the communication module. The whole design system is robust. The system frame is shown in figure1.

3. Network communication module design

a) Ethernet hardware implementation

The CPU of the system uses LPC2368 by Philips. The main frequency can reach 72M/Hz with an internal Ethernet integration module which supports 10M or 100Mbps PHY devices.

This system's Network Interface Card(NIC) is a 10 Mb/s or 100Mb/s single physical transceiver, which contains an Ethernet intelligent power and has low power usage. It works in a 50MHz crystal frequency and has the characteristics of energy testing mode. It has an intelligent, energy-saving mode.

Finally, connecting the NIC to the Ethernet interface RJ45, the interface supports 10MBit and 100MBit adaptive speed of network connection.

b) System communications protocol and communication process

i. Ethernet data frame transmission and reception

The data transmission in the Ethernet is MAC addressed rather than IP addressed. The IP address and the MAC address have a mapping relationship. An Ethernet packet includes: an introduction to the area, starting frame delimiter and an Ethernet frame. The Ethernet frame consists of a target address, the source address, an optional VLAN area, length/dry-farming, load and frame calibration sequence, as shown in figure 2. Each Ethernet frame is composed of one or more pieces, each corresponding to a descriptor. The DMA Ethernet module of Ethernet is able to manage an Ethernet frame multiple fragments scattered (for receiving) and concentrated (used to send).

ii. The TCP/IP protocol

The TCP/IP protocol is a connection-oriented, reliable end-to-end communication protocol. It is divided into four layers, namely: network interface layer, network layer, transport layer and network, as shown in figure 3.

iii. The system TCP/IP communication process

The data based on the TCP/IP protocol can be divided into three stages: connection, data transmission and disconnect. Its implementation process can be described by the state of the machine. There are two methods to establish a connection, namely the active open and passive open. The server, open passively, detects a connection request. The client, open actively, and transmits the connection request. There are also two ways to disconnect: one is active disconnection, the other is passive disconnection. To achieve active disconnection, a FIN packet is sent. On receiving confirmation of receipt FIN packets, a packet of reset is sent to achieve active disconnection.

This system's communication process is divided into 4 main steps:

The first part is a network which is composed of DP83848 chip and RJ45 hardware interface.

The second part is the TCP/IP protocol stack. Using this protocol stack to realize data communications, the network chip is solidified from the MAC layer and network layer to the transport layer protocol completely, thus there is no need to understand method or code.

The third part is the sending and receiving buffer. The main chip LPC2368 communicates with other hosts on the net via the buffer.

The fourth part is the Ethernet physical interface. DP83848 achieves 10 / 100 M data transmission rate.

In the master machine talks to slave machine whose control chip is a LPC2368. The host shows the temperature and the speed of real-time data under the control of LPC2368.

Specific processes are initialized in the first layer which includes physical equipment and register. Then regular program continues to process the network, TCP/IP stack and users. After that, the listening socket is set to decide the frame processing. Finally, is the dynamic HTTP server. Here the HTML code is stored in a custom array: the array includes the real-time data of temperature and speed. If an application-layer processes calls, the array can visit the real-time data and the real-time data can be displayed to web pages.

Network communication processes as shown in figure 4:

4. Speed and temperature control

4.1 Temperature module

4.1.1 Hardware component

This system produces and controls PWM with the different duty factor in main chip LPC2368 and it adopts full-bridge drive IGBT inverter. When the control requires, the inverter commences work. Inverter's IGBT module is composed of a FSBS10CH60.

4.1.2 PWM control technology

In accordance to certain rules, the main chip LPC2368 controls the inverter switch components, conduction and shutoff. The inverter produces a group of rectangular pulses with the same amplitude but different width, approximately equivalent voltage wave in sine. The pulse width corresponds proportionally to the pulse of area under the sine wave, as shown in figure 5.

4.1.3 PWM control method

PWM control methods have many rules. The sampling method is one of these methods is used widely in engineering. It uses a triangle wave as carrier generally. The principle is to get ladder wave by the sampling sine wave with the triangular wave. The control switch device operates at the intersection of the ladder and triangular wave. And then bring into play the speed control function. The method is used to determine the switch time, as shown in figure 6, a1, a2,... a18;, b1, b2,... b18;, c1, c2,... c18 at the intersection of   A, B and C phase sine and the triangular signals, respectively.

The phase output voltage is high when the value of sine is higher than the triangle wave at the same moment; but if lower, the phase output voltage is 0. Through the determination of value, each switch point can be obtained by A, B and C phase modulation of output voltage waveform. Then the switch pattern with different time can be coded by three binary putting all this data into a table. In this way PWM modulation and speed control are maintained..

4.2 Temperature module

This module achieves the temperature control. First the temperature parameter is set in the running state. Then the oscillator cabinet environmental temperature is collected by the A/D conversion. According to the deviation between cabinet and setting temperature, the PID algorithm is used to do the area control. This oscillator divides into three interval situations. And automatically adjust for temperature fluctuation.

4.2.1 Software workflow

The temperature control module mainly executes the process flowcharts, as shown in figure 7: (1)initialize (2) input the preset temperature by keyboard (3) collect and calculate the temperature in the box (4) call integral PID subroutines, heating power or compressor power to adjust the temperature (5) display the Settings of the temperature inside the cabinet.

4.2.2 Control method

This system uses PID control method for temperature control. The existing like product is under constant temperature control and the needs compressor to be at the full power operation which is because the compressor is not allowed to start frequently. Frequent stopping and starting can cause many problems. Such things as energy, noise and life span of the system. This system fully utilizes the advantages of the embedded ARM function, using the capacity of PD frequency adjustment to realize the compressor control of temperature. This overcomes the deficiencies found in the similar product, improves product quality. And importantly conserves energy while protecting the environment.

This system uses integral-separation PID method can eliminate the big overshoots what is produced by the integral effect and keep shaking by instrument in the start-stop or sharp drops.

The control algorithm process which as shown in figure 8.

The temperature module has high precision and effective control methods. These are the main indexes: Temperature range: - 30--+ 80℃; Temperature measurement precision≤+0.1℃; The precision of temperature control≤ +0.4℃; and more than 20 percent energy-saving which compare with similar products.

5. Conclusion

This paper had introduced the design principle of the thermostatic oscillator, based on TCP/IP, which used embedded ARM LPC2368 as the main control chip. It controlled the realization of the function of all modules. This paper provided the details of the system's remote monitoring function. While most of the current market similar products do not have such a function. The paper also outlined the working and operating principle for speed and temperature modules. The combination of hardware and software ensures steady working state and constant temperature and speed. The design features of the thermostatic oscillator ensure it is both practical and stable. It is able to meet user's demands very well which have been proved already.

References

Arnold Berger. (2002). *Embedded system design*. LvJunYi translation. Beijing: publishing house of electronics industry.

Chen,QingZhang, Zhao, XiaoMin. (2006). *TCP/IP network technology and principle*. Beijing: higher education press.

http://www.zlgmcu.com/philips/arm/lpc23xx/lpc23xx_us_en.pdf, Rev.01-6

Liang, Jinze, once Yuena, Zhang, Xuequn, etc. (2008). Phase voltage source PWM rectifier modeling and simulation [J]. *journal of communication power supply*, 25, technology (1) : 26-31.

Taoyonghua. (2003). *New PID control and its application*. Beijing: mechanical industry press.

Wrgiht G.R, Stevens W.R. (2008). *TCP/IP steps 2*. LiuXueYing translation. Beijing: science press.

Xujianping. (2006). *DXP Protel novice road*. shanghai: Shanghai popular science press.

Zhou, Bing, Wu, Wendou, Tubenwei. (2007). Based on embedded TCP/IP real-time data transmission [J]. *Journal of the micro computer information (SOC) embedded with* 23, the first 3-2 52-54.

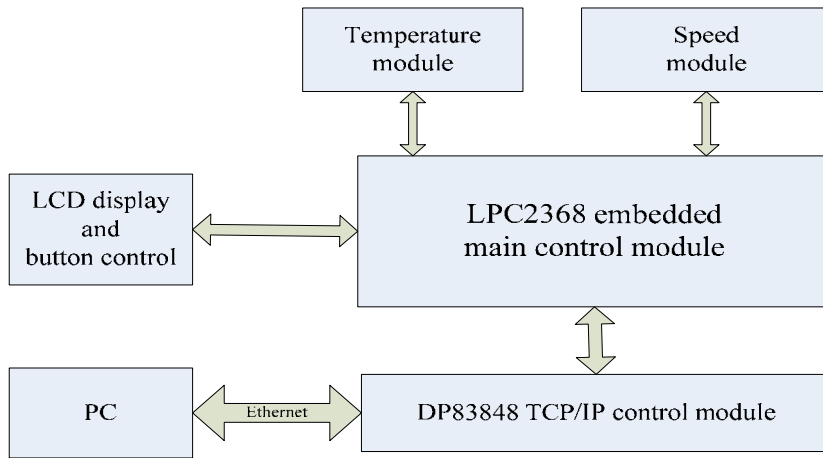Zhou, Ligong company. (2007). LPC23XX micro-controller user manual [EB/OL].
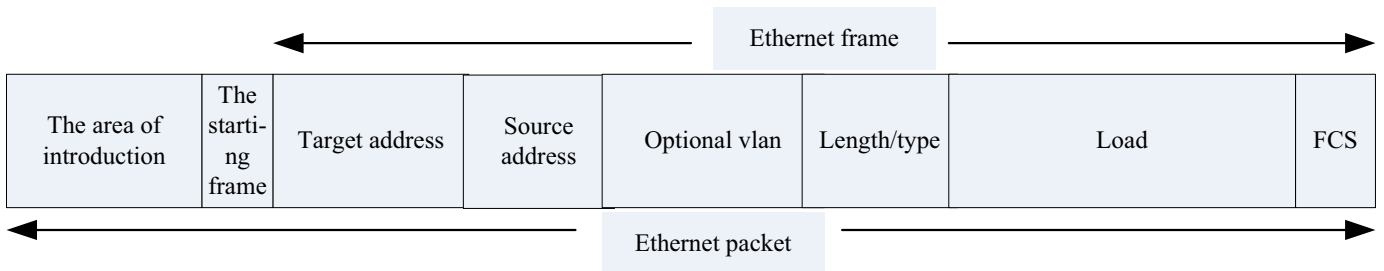
Figure 1. Thermostatic oscillator based on TCP/IP

Figure 2. Ethernet packet structure and data frame structure

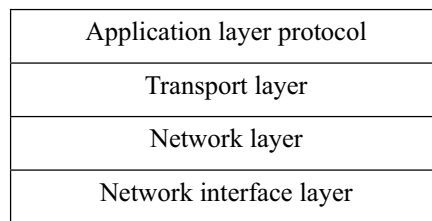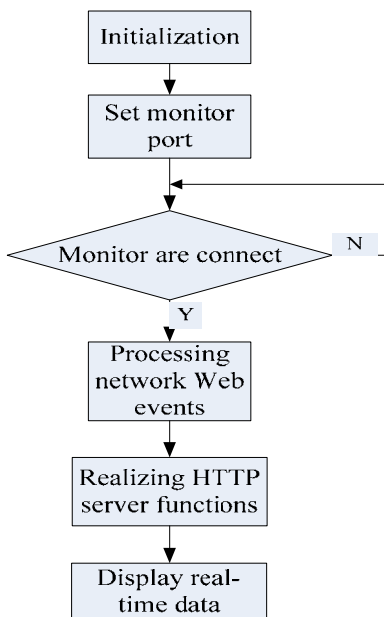| Application layer protocol |
| --- |
| Transport layer |
| Network layer |
| Network interface layer |

Figure 3. TCP/IP protocol model

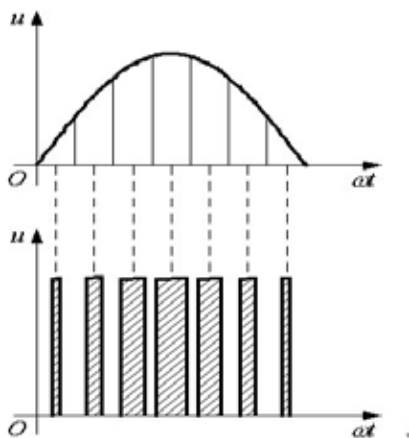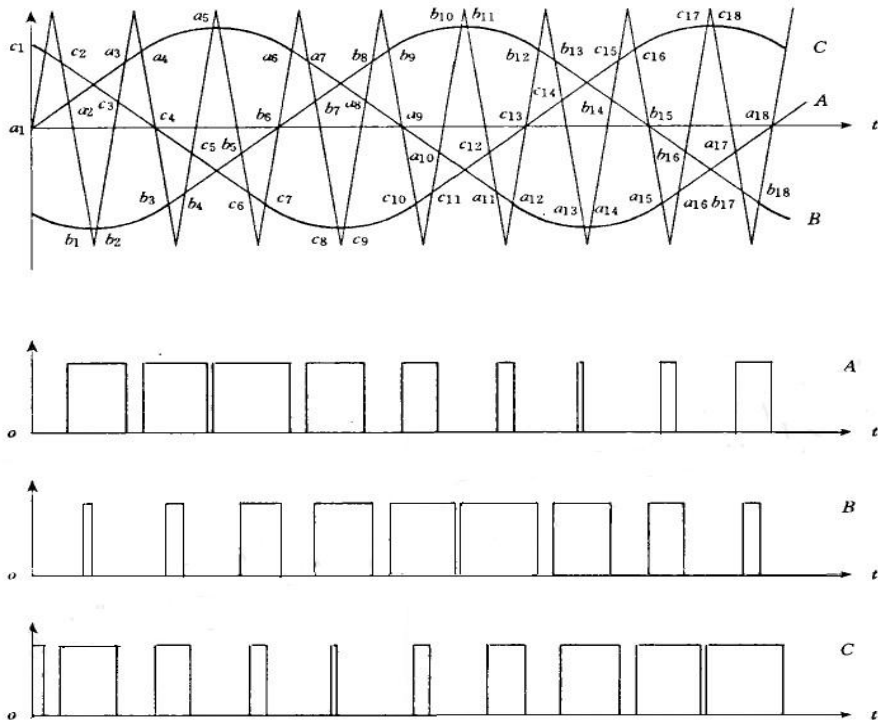Figure 4. Network communication flowcharts

Figure 5. PWM pulse width modulation
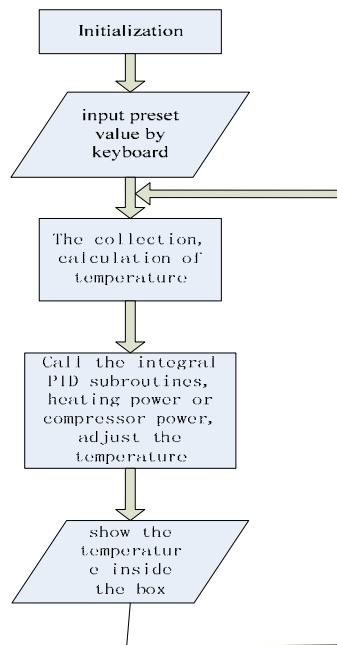
Figure 6. PWM modulation method in this system
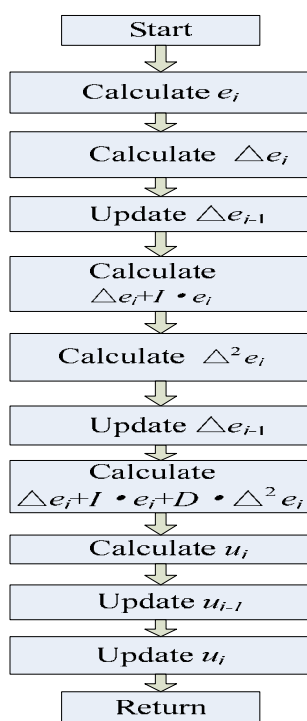

Figure 7. Temperature control chart

Figure 8. PID algorithm process

e_i is the deviation for the given and the actual value, u_i is the output value, I and D are the coefficients of integral and differential separately.

The Quality Test of Wire Bonding

Caiyuan Wang

School of Mechanical & Electronic Engineering

Tianjin Polytechnic University

Tianjin 300160, China

E-mail: llwinnie@126.com.cn

Ronglu Sun

School of Mechanical & Electronic Engineering

Advanced Mechatronics Equipment Technology Tianjin Area Major Laboratory

Tianjin 300160, China

E-mail: rlsun@tjpu.edu.cn

Abstract

The reliability of the IC chip during performance of its function in any application is very much dependant on the quality of the wire bond interconnection. If the quality of the wire bond interconnection is poor and not consistent, it has a significant impact on the reliability and dependability of the device. The quality of a wire bond is determined by the strength of the interfaces between the bond wire and the bond pad as well as the bond wire and the lead finger. In this article, we discuss in detail established methodologies such as bond pull test, bond shear test, visual inspection and some special test of quality.

Keywords: Wire bonding, Quality test, Failure modes

1. Introduction

Today, most of the critical applications like high performance computers and industrial systems are controlled by microchips which have wire bond interconnections. The reliability of the IC chip during performance of its function in any application, be it in an automobile, space shuttle or a personal computer, is very much dependant on the quality of the wire bond interconnection. If the quality of the wire bond interconnection is poor and not consistent, it has a significant impact not only on the reliability and dependability of the device, but also has a huge economic impact. Therefore, quality assurance of the wire bond is of utmost importance in the IC assembly work flow.

2. The Bond Pull Test

The bond pull test is the most widely used technique for the evaluation and control of wire bond quality. This technique is also used for wire bonding process control and process optimization during assembly manufacturing. The bond pull test involves placing a hook under the wire and applying a normal upward force and the wire is pulled till it fails, as is shown in figure1.

During bond pull test, if a ball lift failure or weld lift failure occurs even at very high pull value, still it is unacceptable and the bond is considered of poor bond quality. In case of such failures the cause of the non-stick must be analyzed thoroughly and necessary remedial action needs to be taken. Examples of ball lift and weld lift failures are shown in figure2.

2.1 Different Modes of Pull Test Failures and their Interpretation

The major failure modes that one observes during a gold ball bond pull test are, ball lift (A type),ball neck break (B type), mid span break (C type) , heel break (D type) and weld lift (E type), as shown in figure3. In special cases, the failure may also occur at other interfaces such as lift off of bond pad metallization, caterings under the bond pad and peel off of lead finger metallization. Let us review and discuss the five types of failure modes in detail.

2.2 Ball Lift Failure

The occurrence of ball lift failure indicates that the ball to bond pad interface is very weak. The ball lift failure indicates either the bond has not formed or the bond has become weak. Ball lift can be due to a variety of factors. Poor wire

bonder set-up and bond pad surface contamination are primary causes of ball lifting. Poor set-up includes improper parameter settings, unstable work piece holders, and worn-out tools. These result in poor initial welding and inadequate intermetallic formation between the bond pad and the ball.

2.3 Ball Neck Failure

The failure of the wire just above the ball neck is shown in figure4. This is the predominant failure mode that we observe during bond pull test. Neck breaking is the severing of the wire from its ball bond due to a fracture in the neck. The neck is the portion where the wire meets the ball bond. Neck breaking is commonly due to poor wire bonder set-up such as improper bonding parameter settings, boneheads movement settings, and worn-out or contaminated tools. Incorrect bonding parameters can deform the bond excessively, resulting in a thin, weak, or cracked neck which can easily fracture. Crack propagation can also be caused by defects in the wires due to worn-out and contaminated tools.

2.4 Wire Break at Mid Span

Bond wire failure at the midspan is shown in figure4. This failure mode indicates that the wire failed by plastic deformation and ductile fracture, rather than by brittle fracture at the heel or at the ball neck. Usually the failure at mid span gives the highest bond pull value, close to the ultimate tensile strength (UTS) of the wire.

2.5 Heel Break

Heel breaking is the severing of the wire from its wedge or crescent bond due to a fracture in the heel. The heel is the portion where the wire tapers off into the bond. When we perform bond pull test, the heel break failure may occur as shown in figure4 and this usually indicates damaged heel area. Improper bonehead's movements and low loop settings may subject the wires to excessive stresses that tend to pull them backward and away from the bonds, resulting in gross heel cracks which may propagate to fracture. The major cause of heel break is the presence of a micro crack at the heel. This micro crack may be formed during bonding. The causes of micro crack may be attributed to bonding parameters such as high bonding force and excessive ultrasonic energy. The other probable causes are damaged capillary, clogged capillary or over used capillary.

2.6 Weld Lift

Weld lifting is the detachment of a wedge bond from the bond pad or bonding post, or the crescent bond from the lead frame bonding finger as shown in figure 5. Weld lift during bond pull test, is an indication of improper process optimization. Similar to ball lift failure, weld lift failures, and even at very high bond pull values is a mark of bad quality and therefore, should not be accepted. The weld lifting could be also caused due to low bond parameters, contamination on lead frame, hard metallization, or improper ultrasonic coupling between the capillary and the lead finger. The cause of this weak coupling between the capillary and the lead finger can be due to lead frame fragility or improper clamping during bonding.

3. Ball-Bond Shear Test

The destructive wire bond pull test has limitations, in that it provides very little information on the strength and overall quality of the first bond. A simple wire pull test does not fully explore the characteristics of a ball bond. A quality ball bond can withstand up to ten times the wire pull destruct force but a low quality bond will still take more force to pull off than a wire pull test. Thus pull tests fail in determining the true strength of ball bonds. This factor has led to the development of a complementary wire bond test methodology, the ball-bond shear test.

Ball shear testing is used to assess the integrity of the gold ball-to-bonding pad interface in the gold ball bonding process. It is also a destructive test. Ball shear data reflects the intermetallic formation and its coverage of bonds. It is measured by gram force over the area of the ball formation.

3.1 Description of the Test

A shear tool is used to perform this test. During testing this tool pushes the ball bond with a sufficient force. Positioning of the tool is very important during shear testing. The shear ram is positioned just above the bond pad such that the bottom of the shear tool is close to the centre of the ball. The shear tool moves parallel to the bonded surface and shears the ball bond. Figure6 shows the placement of the shear tool.

The improper positioning of the tool is one of most common problem in shear test. During the ball shear test the shearing ram needs to be positioned exactly according to criteria described in the ball shear test specification. For a normally deformed ball, the tool should approach from 2.5 to above the substrate and for large, high balls; it must be no higher than

Otherwise the tool could drag on the substrate. Figure7 shows the step by step movement of Shear Ram and the failure modes of the ball bond.

The failure modes generally observed during a ball shear test are: ball lift / partial metallization lift off, ball shear, bond pad lift and cratering. Let us review and discuss each of the above mentioned failure modes in further detail.

3.1.1 Ball lift / Partial Metallization Lift Off

A ball lift is defined as a separation at the bonding pad interface with little or no intermetallic formation present or remaining. Figure7 (a) shows part of the Al metallization sheared with the ball, leaving partial metallization on the bond pad during shear test.

3.1.2 Ball Shear (Weld Interface Separation)

During ball shear test, if the intermetallic formation is pronounced, then the ram may shear the whole ball leaving gold or an intermetallic on the bond pad as shown in the figure7 (b).

3.1.3 Bond Pad Lift (Substrate Metallization Removal)

Another failure mode during shear test is bond pad lift. In this failure one can see the separation between the bond pad and the underlying substrate. The aluminum metallization is peeled off leaving exposed. This may occur due to improper adhesion of aluminum metallization. Here the interface between the ball bond and the residual pad attached to the ball remains intact as shown in the figure7(c).

3.1.4 Cratering

Cratering is a general term used to describe one type of bonding failure that occurs as the result of mechanical damage to the bond pad or underlying material. This failure occurs on the underside of a wire bond during bond pull or shear test. Cratering occurs as a result of cracks generated under the bond pad during bonding. Figure7 (d) shows cratering effect during ball shear test.

4. Visual Inspection of Bond Quality

Inspection of wire bonding had traditionally been done visually, because the human ability to discern faults had been much better than the available machines. This technique is used to ensure the proper ball and wedge bonds. Signs of cracked heels, tearing at the wedge, misplaced wires, inconsistent wire placement, "golf-clubbing" at the gold ball, excessive neck down etc. are visual indications that the wire bonding process is not in control. Visual inspection also verifies whether the bonds are properly placed with respect to the bond pads and bond fingers of the lead frame/substrate. Beside that the visual inspection also screens possible bond defects that may results in open or short, based on a specified defined criterion of wire clearance and close proximity of each bond to the other.

As discussed earlier, bond pull and bond shear test check the quality of bonded interfaces, both on the chip and on the lead finger. However, there are many other failure modes that can only be detected by visual inspection, for example deformed bond, s-ing of wire and probe marks to name a few. Therefore visual inspection is very much essential in today's wire bonding process. Visual inspection of bond quality includes both "Pre-Bonding Inspection" and "Post Bonding Inspection". Pre-Bonding Inspection is necessary for assuring process consistency and reliability. Post Bonding Inspection is necessary to assure package reliability.

4.1 Pre-Bonding Inspection

One of the final steps in the manufacture of integrated circuits is the inspection of the probe marks on bonding pads caused during electrical testing of the integrated circuit wafers. Pre-bonding inspection is useful in isolating potential problem areas caused when: the probe mark extends beyond the pad boundary, the mark is too large in size, or extraneous scratch marks have been introduced during electrical testing.

Inspection of the bond pads is facilitated by evaluating parameters which can be used to quantify these defects. These include: size of the pad, location of the pad centre, size of marks on the pad and percentage coverage of pad, location and size of the largest mark.

Other problems on the bonding area that have an impact on the quality of the bond are: discoloration of bond pad (due to corrosion) and photo resist residue. These can be visually inspected and problems resolved. Photo resist residue is not visible under normal lighting. However using polarized UV lighting under a microscope it is possible to identify passivation and photo resist residue problems.

4.2 Post Bonding Inspection.

The following list gives a range of problems that one may encounter after the bonding process, and these problems can only be identified visually. The post bonding inspection includes bond quality problems such as: smashed bonds, skinny bonds, misplaced and deformed bonds and wire conditions such as loop height etc.

The visual inspection can be done manually with the assistance of an operator, or with automatic visual inspection systems.

5. Special Tests for Quality

The industry uses a variety of techniques for measuring the quality and of bonds. The most frequently used techniques are: visual inspection, bond Pull, bond shear test, bond etching, electrical testing, bake test, thermal cycle/thermal shock

testing, electro magnetic resonance, surface analysis.

Some of these tests are usable as production monitoring techniques, where as some tests can only be used in the laboratory for evaluation and development of wire bond technology.

5.1 Bond Etching

Bond etching measures the condition of layers under the bond pads after bonding. With this technique, the ball (wire) and pad metallization is removed, and the condition of underplayed materials is examined for defects. Wet etching method not only allows detection of subsurface damage caused by equipment parameters, but can also serve as a tool for optimization of the bonding parameters. Typically, NaOH, or KOH are the chemical solutions of choice used to remove the interconnection material and expose the underlying structure. For example to check if cratering has occurred under the bond pad, we remove the bond pad metallization using potassium hydroxide (KOH) solution.

5.2 Electrical Testing

Electrical testing measures the quality of the bond. A non-sticking bond will result in an electrical open, while two adjacent bonds in contact will result in an electrical short. There are special techniques such as "Four Point Probe Testing" to measure the electrical resistance of the Au ball-Al bond pad interface. The resistance increases as the amount of intermetallic compound increases. And an open circuit failure is an indication of extensive voiding at the interface.

5.3 Bake Test

Bake test measures the degree of intermetallic compound formation of gold-aluminum (Au/Al). Over diffusion of aluminum atoms into gold can lead to Kirkendall voiding at Au/Al interface resulting in bond lifting. By baking the bonded devices at high temperature (175°, 240° or 300°) for different lengths of time, one can determine the long term reliability of the

bond.

5.4 Surface Analysis

Surface analysis techniques such as energy-dispersive X-ray analysis (EDX), wave-length dispersive X-ray analysis (WDX), auger, and scanning electron microscopy (SEM) are the failure analysis tools one can use to identify the presence of contamination, extent of intermetallic compound formation, and bond irregularities.

6. Conclusion

The quality of a wire bond is determined by the strength of the interfaces between the bond wire and the bond pad as well as the bond wire and the lead finger. In addition, some of the other quality factors that one needs to check for during the wire bonding process are wire sweep, wire sway, wire sagging, heel cracking, bond deformation and many other defects. The quality of the wire bond is controlled by using established methodologies such as bond pull test, bond shear test, visual inspection and special test for quality.

References

Antony Jiju. (1999). Improving the wire bonding process quality using statistically designed experiments. *Microelectronics Journal*, 30, 161-168.

Brokelmann Michael, Wallaschek J & Hesse Hans J. (2005). Model based development of an integrated sensor-actuator system for online quality monitoring in ultrasonic wire bonding, *American Society of Mechanical Engineers*, 70, 219-226.

Chen J. (2002) Real-time ultrasonic testing of quality of wire bonding. *British institute of non-destructive test*, 44, 443-445.

Lin T Y, Leonq W S, Chua K H & Miao Y. (2002). The impact of copper contamination on the quality of the second wire bonding process using X-ray photoelectron spectroscopy method. *Microelectronics reliability*, 42, 375-380.

LV Lei. (2008). The Process Introduction and quality inspection of wire bonding. *Equipment for electronic products manufacturing, 158, 53-60.*

Srikanth N. (2005). Effect on ultrasonic bond quality of wire bonds. *Proceedings of 7th electronics packaging technology conference, 2, 756-759.*
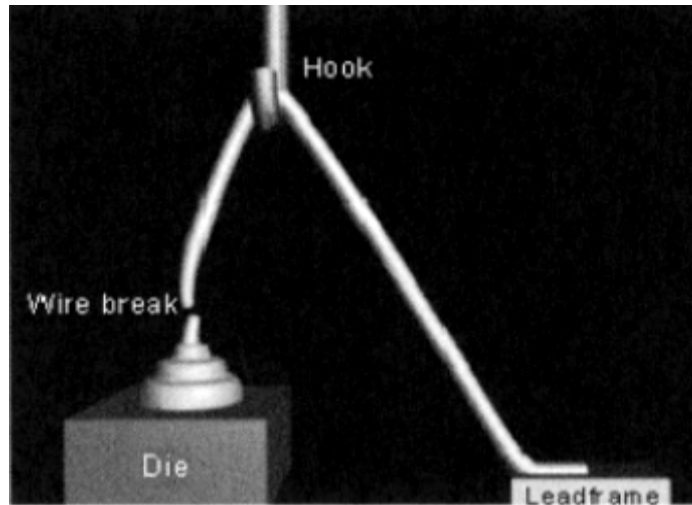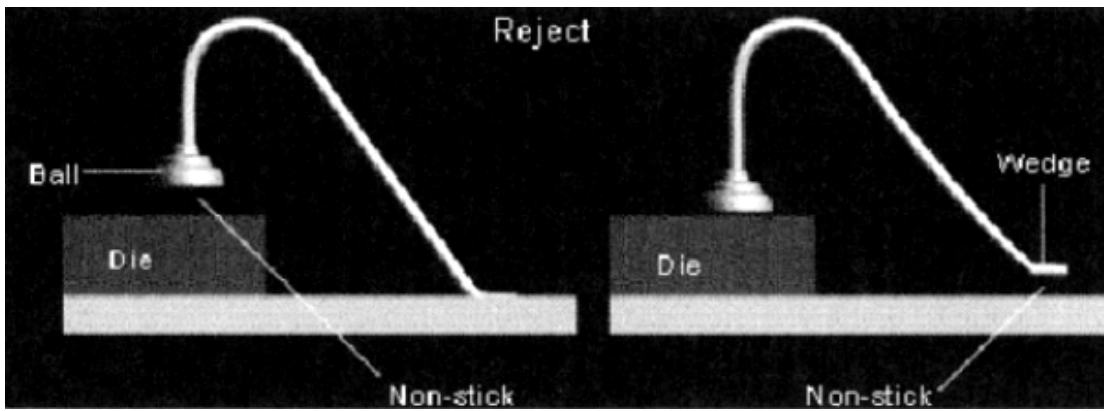
Figure 1. Bond pull test

Figure 2. Ball lift and wedge (weld) lift failure
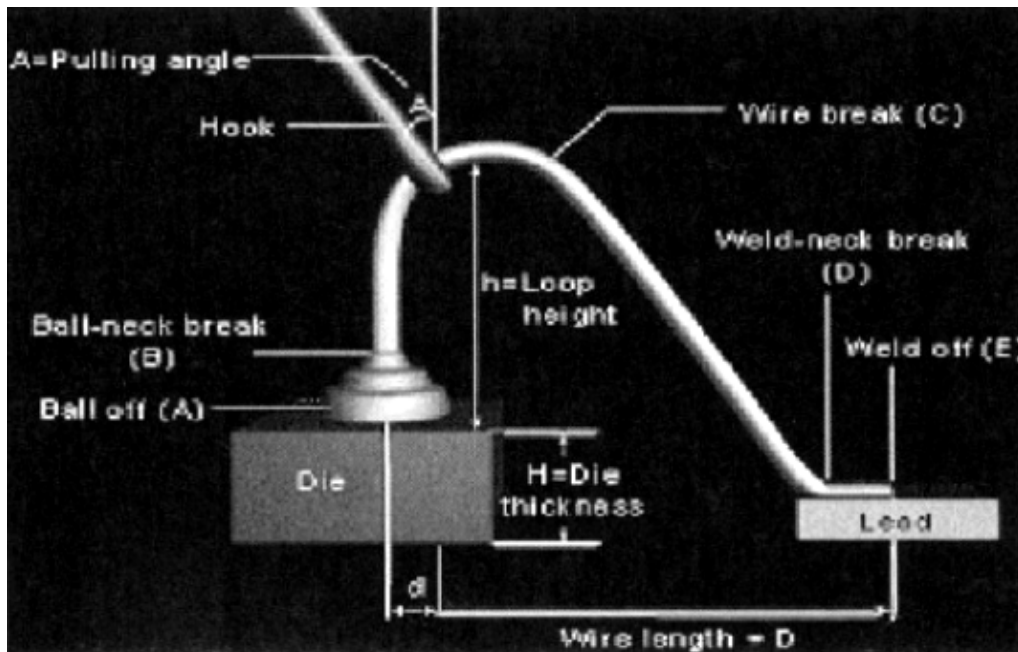

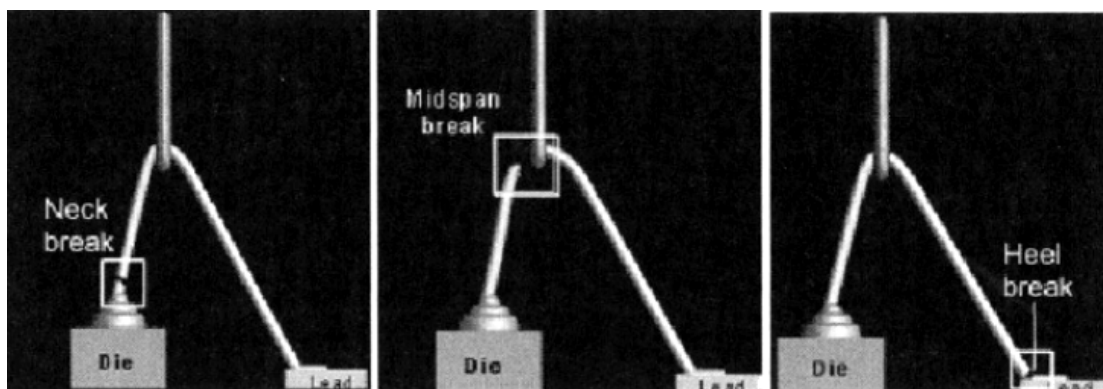Figure 3. Pull test failure modes

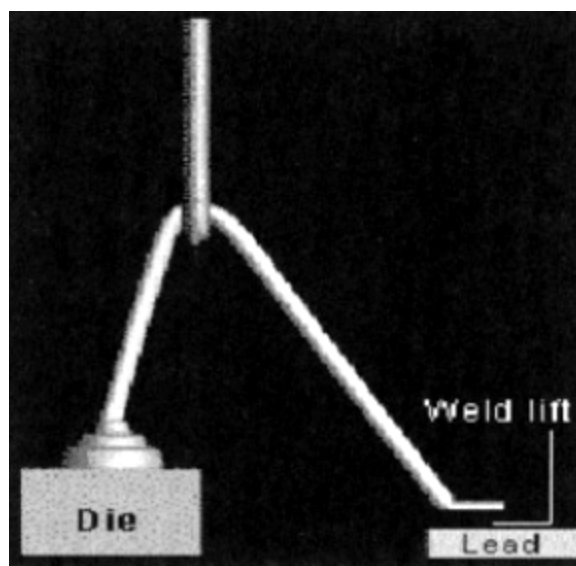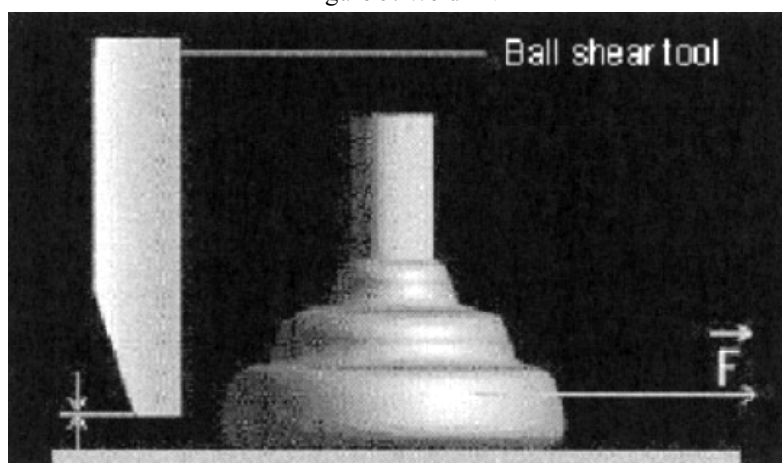Figure 4. Shows neck break, midspan break and heel break

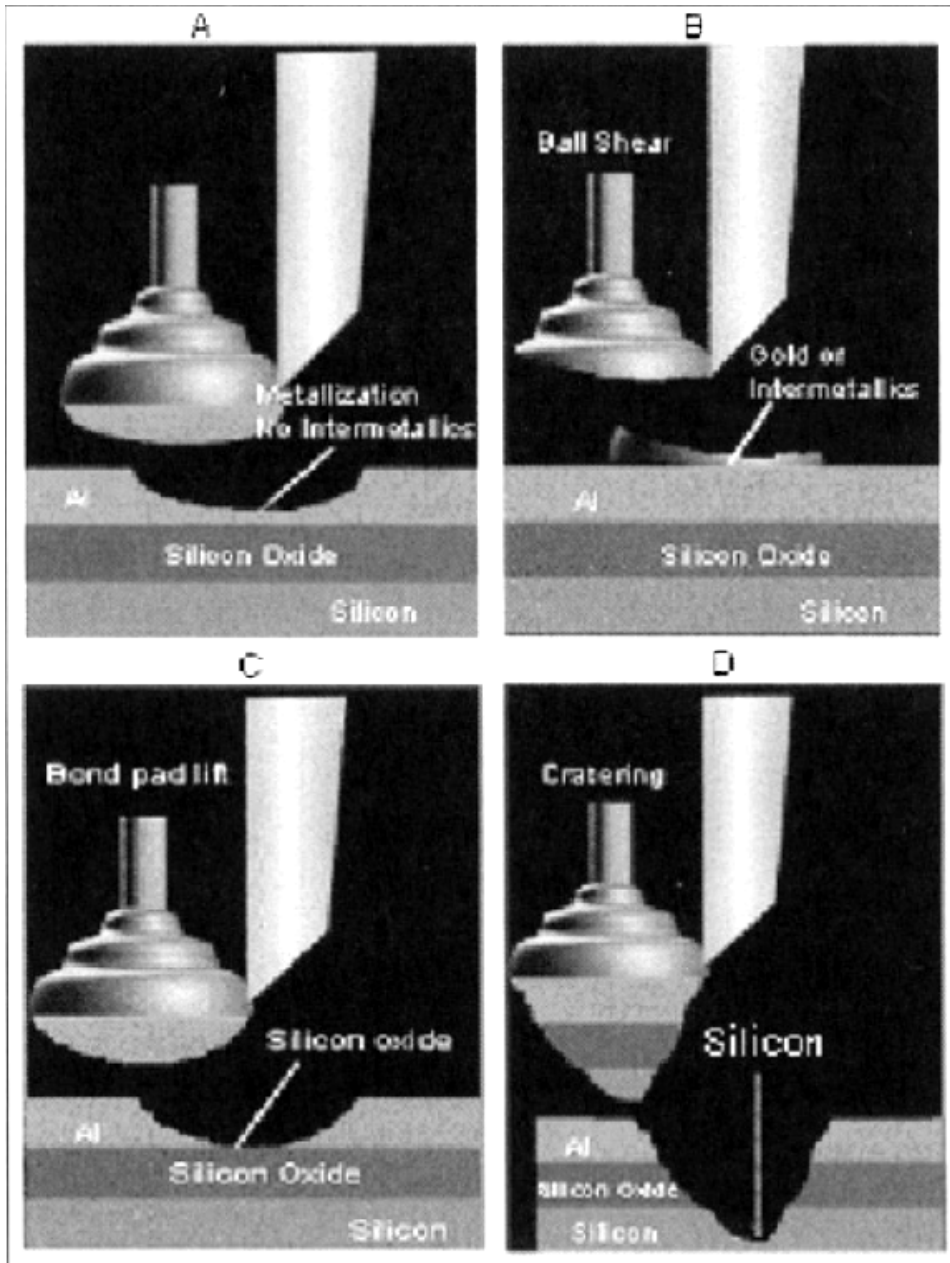

Figure 5. Weld lift

Figure 6. Ball shear test

Figure 7. Shear test failure modes (a) Ball lift (b) Ball shear (c) Bond pad lift (d) Cratering

Studies on the Growth-development Law and Suitable Period for Harvesting of *Pinellia ternata* (Thunb)Breit

Shuhong Wei (Corresponding author)

Life Science College, China West Normal University

Nanchong 637009, Sichuan, China

Tel: 86-817-256-8125   E-mail:weishuhong453@sohu.com

Abstract

Objective: Find out the growth-development law of *P. ternata*, and offer references for suitable period for harvesting and standardized cultural technique. Methods: Periodic sampling, investigation and measurement; Results: Dynamic change of each growth index all show a similar rule, namely that with the elongation of growth-development course, it was characterized an "S" type curve, and peak values are observed during the first "inverted seedlings" period. Especially, average drying rate of tubers during the first "inverted seedlings" period is significantly higher than that during the second period. Conclusions: Initial stage of the first "inverted seedlings" is quite suitable period for harvesting of *P. ternata*, and thus its utilization value is further efficaciously enhanced.

Keywords: *Pinellia ternata (*Thunb) Breit, Growth-development law, Drying rate, Suitable period for harvesting

1. Introduction

Pinellia ternata (Thunb) Breit is perennial herb, and belongs to species of *Pinellia*, the family of *Araceae*. Such species was the recorded toxic grass and plant in Chinese Plant Atlas Database, and its toxicity exists in all individual plant with more in tubers. Meanwhile it also possesses certain medicinal values, viz. transform phlegm and stop coughing, stop nausea and vomiting, address goiters or scrofula, ease pain, anti tumor and so on(2005, PP.78-79). With the increasing demand from market, studies on the cultivation technology for high produce of *P. ternata* have been strengthened. Proclamation of regulations titled "Chinese medicine materials production quality and management regulations" required that the production of medicinal materials should make relative planting regulations and cultivation measures according to their biological characteristics (Ren, 2003, PP. 12-17). Currently, correlated researches on the growth-development characteristics of *P. ternata* cultivation were not yet available. Therefore, studies on the growth-development law of *P. ternata* would be of great significance to ascertain its optimum period for harvesting, implement standardized cultivation and thus enhance its quality and yield.

2. Materials and methods

2.1 Materials

Tested *P. ternata* were sampled from Huohua, Nanchong of Sichuan Province.

2.2 Methods

2.2.1 Cultivation management of *P. ternata*

P. ternata plants were planted in Huohua trail base of Nanchong, Sichuan on March 2rd, 2003. Drill was undertaken in the experimental field with plots of 10m^2(2m×5m). Protection rows were made around the field. Tubers with no disease and the diameter of 1cm or so were selected as seeds, and planting manner was hill seedling. Experimental field was fertilized according to the unified standard, and fertilization level was 2000kg local decomposed manure from pig and cattle yard per acre with 250kg gingko leaf residues and 50kg decomposed oil cake. Cultivation management throughout the trial was identical, and timing irrigation and weed removal by manual work were undertaken.

2.2.2 Cardinal biological characteristics observation of *P. ternata*

During the different growth and development periods such as seedling, full seedling, centralized emergence of bulbil, centralized emergence of spathe, and inverted seedling, 30 samples were collected at random, and divided into 3 groups, each group 10 samples. Indices during different periods such as plant height, compound leaf number, length and width of main leaf, bulbil number and diameter and so on, were measured indoor. Fresh weight of total plant and tubers were

assayed, and then dried at 60℃ for 48h. Dry weight of them was assayed subsequently. According to that analysis, growth-development law of *P. ternata* was ascertained.

3. Results and analysis

3.1 Dynamic change of plant height

Plant height increased with the breeding course after the first "inverted seedlings" period, rapidly grew from seedling to centralized emergence of spathe, attained top value during the period of centralized emergence of spathe, and then kept stable. During the second "inverted seedling" period, dynamic changes of plant height showed a growth trend "slow-fast-slow", and attained peak value during the period of centralized emergence of spathe, higher than those during the second "inverted seedling" period, and then decreased gradually(Figure 1).

3.2 Dynamic change of leaf

Leaf of *P. ternata* was ternate compound leaf, and its compound leaf number increased with the elongation of growth and development course, attaining top values during the period of centralized emergence of spathe. The number of leaf is 3.63 in the earlier half year growth-development period and 2.27 in the later half year, and then decreased gradually. Finally, leaves turned yellow and withered to death during the period of inverted seedlings (Figure 2).

As seen from Figure 3, main leaf length of *P. ternata* showed a similar changing trend during the period of growth-development in earlier and later half year, namely that it increased gradually along with the course of growth-development. It attained top value during the period of centralized emergence of spathe and decreased slightly afterwards.

As seen from Figure 4, main leaf width showed an approximately similar changing trend during the second inverted seedlings period and earlier half year, namely that it increased gradually along with the course of growth-development. It attained top value during the period of centralized emergence of bulbil and decreased slightly afterwards.

3.3 Dynamic change of bulbil number and diameter

As seem from Figure 5, average bulbil number per *P. ternata* showed an identical changing law during the period of the second inverted seedlings and in earlier half year, and attained the rapidest increment from the period of centralized emergence of bulbil to spathe, achieving peak value during the period of centralized emergence of spathe. It was 2.47 during the first inverted seedlings period, 2.50 in the later half year and then remained stable afterwards. Bulbil diameter kept the increasing trend during the first inverted seedlings period, increased slowly from the period of centralized emergence of bulbil to spathe, and increased promptly from the period of centralized emergence of spathe to inverted seedlings (Figure 6).

3.4 Dynamic change of tuber fresh weight, dry weight and drying rate

Tuber fresh weight varied greatly during the first growth-development period, showed an obvious ascending trend from the period of seedling to inverted seedling, decreased slightly during the period of total seedling, then increased promptly subsequently, attained peak value during the period of centralized emergence of spathe and decreased slightly during the period of inverted seedling(Figure 7). Tuber dry weight also decreased slightly during the period of total seedling, attained peak value during the period of centralized emergence of spathe, and then kept stable afterwards(Figure 8). During the second period of growth-development, tuber fresh weight decreased slightly in total seedling period, increased slowly and kept stable subsequently. Compared to the initial seedling, the increment was not significant (Figure 7). Tuber dry weight showed a decreasing trend as a whole during the second period of growth-development (Figure 8). As seen from calculation results, tuber drying rate attained the top value, namely 34% during the period of centralized emergence of bulbil in the second growth-development. It attained the valley value of 23% during the period of centralized emergence of spathe, and then increased to 29% during the period of inverted seedling. As for the second growth-development period, the highest drying rate occurred during the period of total seedling and attained 33%; while the lowest occurred during the period of inverted seedling and attained 18%. In general, average drying rate in the first inverted seedling was larger than that in the second inverted seedling.

4. Conclusions and discussions

Among all growth indices, indices such as plant height, compound number, length and width of main leaf, bulbil number per plant, bulbil diameter, fresh and dry weight, tuber fresh and dry weight, showed a similar rule of dynamic change, namely increased with the elongation of growth and development course. Their growth-development curves were all characterized as S type curve, in line with logistic equation, and top values all were attained during the first inverted seedling period. Especially, the highest tuber drying rate in the first inverted seedling period was significantly larger than that during the second one. As seen from the accumulated dynamics of total alkaloid, peak value was attained in the initial stage of the first inverted seedling period (Li, 2006, PP. 687-688). Therefore, in order to enhance the utilization value of *P. ternata* tuber, and quality and yield of *P. ternata*, the initial stage of the first inverted seedling period was the suitable period for harvest, and it would be optimum when leaf stem didn't wither to fall down.

Indices, such as fresh and dry weight of plant and tuber, showed a similar changing rule during the first and second inverted seedling period. They all decreased slightly or increased insignificantly during the period of total seedling compared to seedling period, and in light of this result we postulated that with the bourgeon of tuber after seedling, nutrition consumed, and photosynthetic accumulation was lower than decomposition; After total seedling, photosynthetic accumulation dominated, dry substance accumulation increased and attained the highest level till the centralized emergence of spathe; During the centralized emergence of spathe, *P. ternata* turned into reproductive growth phase from nutritional growth, and reproductive growth of *P. ternata* certainly would consume the nutritional substances deposited in the tubers. Nutritional provision center shifted to reproductive organs. Secondly, due to the inverted seedling period, high temperature and illumination in summer withered above-ground leaves to yellow discolouration or even falling off, leaf number and area decreased, photosynthesis attenuated, tubers tended to physiological dormancy, substances accumulation increased insignificantly or accumulation was lower than decomposition, and plant transferred from photosynthetic accumulation into consumption. Accordingly, fresh and dry weight of plant and tuber decreased slightly or kept stable during the period from centralized emergence of spathe to inverted seedling.

P. ternata growth-development was divided into 5 periods, viz. seedling, total seedling, centralized emergence of bulbil, centralized emergence of spathe, inverted seedling. In the practical produce, suitable cultivation measure should be undertaken, aimed at high quality and yield according to the growth characteristics of different growth-development period. During the period of seedling and total seedling, measures such as adding fertilizer, cultivation and so on should be undertaken, with attempt to promote the emergence of leaves, and built a base for the growth in the middle or later stage, especially for the formation of bulbil. Due to the association between compound leaf number and bulbil number, more compound leaves, more bulbil. According to the theoretical calculation, 3 leaves per plant would produce 3 bulbils. If inverted seedling occurred 3 times per annum, it would result in 64 different tubers and bulbils (Gu, 1994, PP. 44-48). During the period of centralized emergence of bulbil and spathe, field management should be strengthened, such as timely weeding and irrigating, prevention of plant diseases and insect pests and drought, proper sheltering, and suitable complement of fertilizer, with attempts to the produce of bulbils, enlargement of tubers, elongation of leave functional period, prevention of leaves from reaching inverted seedling period due to earlier wilt, and guarantees of nutritional substances bulbil formation and tuber augment required. Increment of tuber dry substances accumulation could also be obtained by removing spathes and attenuating the nutritional consumption due to reproductive growth, and thus high yield would be achieved.

References

Gu, D.X., Li, Y.X., & Xu, B.S. (1994). A study on the reproductive biology of *Pinellia ternata* (Thunb) Breit. *Journal of Plant Resources and Environment*, 3(4):44-48.

Li, X.W., Ma, X.J., & Song, J.Y., et al. (2006). Studies on the dynamic change of total alkaloid in *Pinellia ternata* (Thunb) Breit during the different growth-development period. *China Journal of Chinese Materia Medica*, 31(8):687-688.

National pharmacopoeia committee. (2005). *Pharmacopoeia of the people's republic of Chinese*. Beijing: Chemical industry press, 78-79.

Ren, D.Q., & Zhou, R.H. (2003).*Chinese medicinal materials quality control regulations*. Beijing: Chinese agriculture press, 9:12-17.
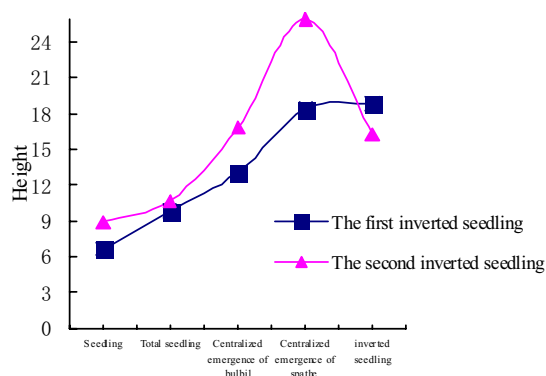

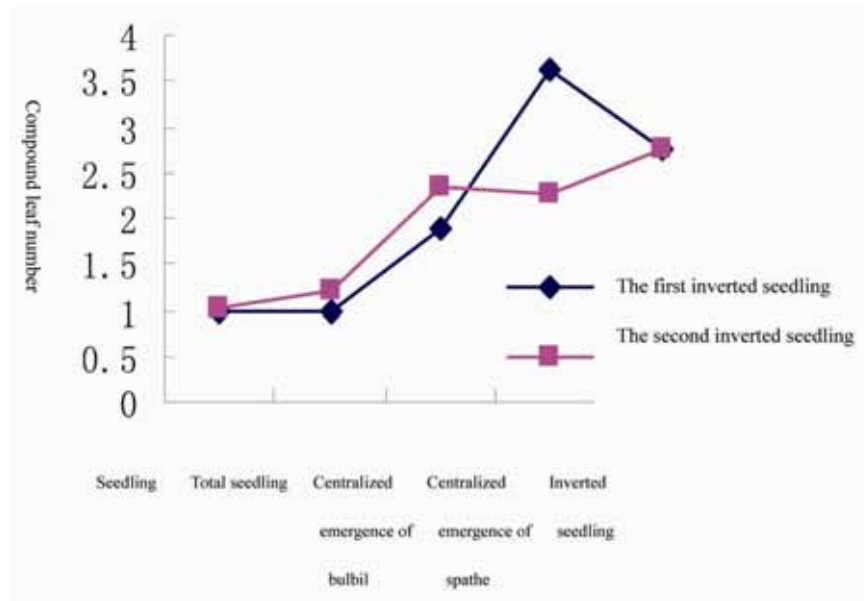Figure 1. Dynamic change of plant height in *P. ternata* during different growth-development period

Figure 2. Dynamic change of compound leaf number in Pinellia during different growth-development period
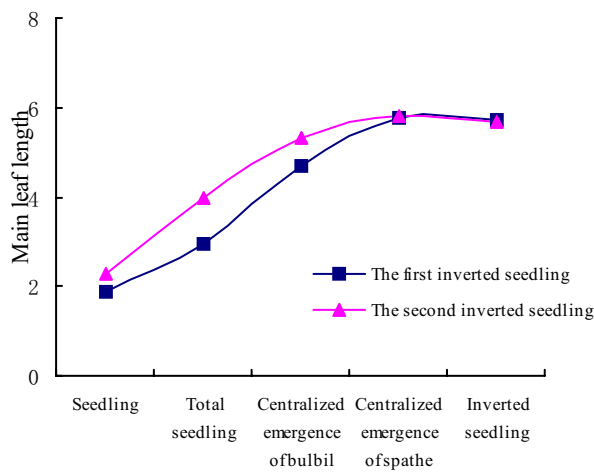

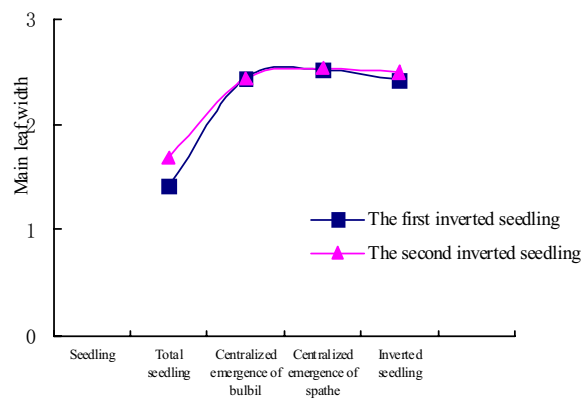Figure 3. Dynamic change of main leaf length in *P. ternata* during different growth-development period

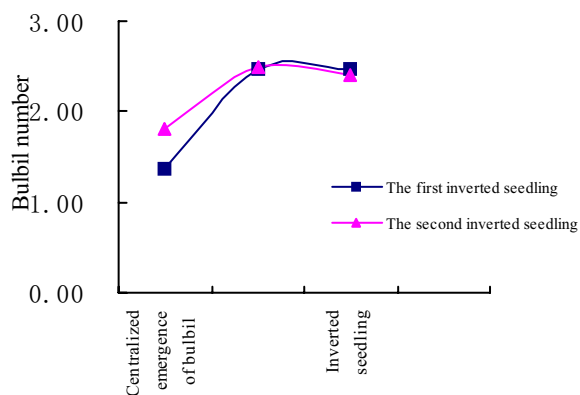Figure 4. Dynamic change of main leaf width in *P. ternate* during different growth-development period

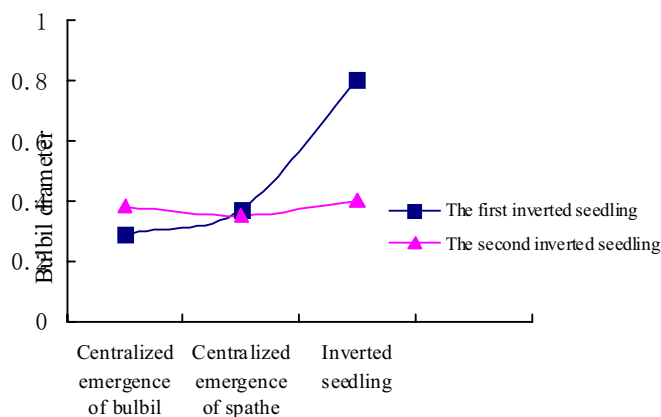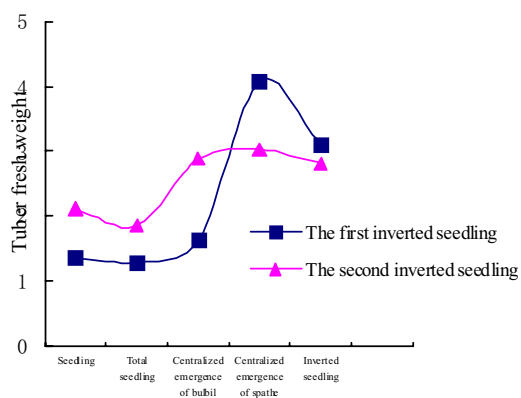Figure 5. Dynamic change of bulbil number in *P. ternata* during different growth-development period

Figure 6. Dynamic change of bulbil diameter in *P. ternata* during different growth-development period

Figure 7. Dynamic change of tuber fresh weight in *P. ternata* during different growth-development period
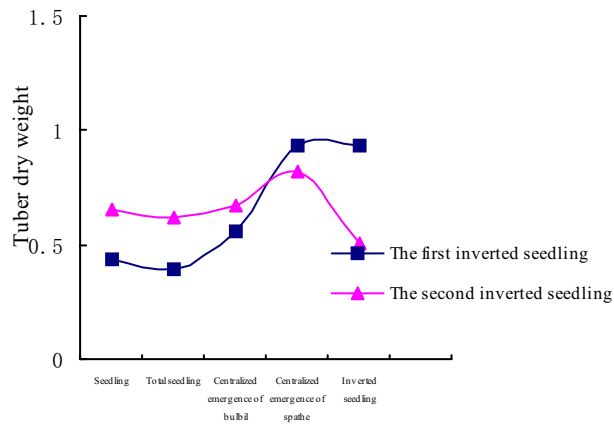
Figure 8. Dynamic change of tuber dry weight in *P. ternata* during different growth-development period

Powering a 35W DC Metal Halide High Intensity Discharge (HID) Lamp Based UCC3305

Hui Zhang & Fuzhong Wang

School of Science, Tianjin Polytechnic University

Tianjin 300160, China

E-mail: huizhang_09@163.com

Abstract

A high power density integrated electronic ballast for low wattage high intensity discharge (HID) lamp based UCC3305 is proposed in this paper. The UCC3305 integrates all of the functions required to control and drive one HID lamp. A highly efficient and reliable topology full-bridge inverter is adopted to supply the lamp with low-frequency square-wave positive and negative current avoiding acoustic resonances. The UCC3305 special microcontroller for HID ballast is used to raise control performance and simplify the control circuit. The presented circuit was verified with experimental results. A prototype of 35W digital HID ballast has been built with the whole efficiency of 89%.

Keywords: HID, UCC3305, Acoustic resonances

1. Introduction

Recent developments in HID (High Intensity Discharge) lamps have increasingly been attracting interest due to their highly efficient operation and excellent color characteristics. These lamps are suitable for various applications, including automobile headlamps and back lamps for projectors (Groot,1986, p.5) The HID lamps exhibit negative-resistance characteristics in the operating regions and lamp resistance varies as the time proceeds. Consequently, ballasts are needed, and high voltage is also needed to start lighting for the HID lamps. Many studies have been conducted on the ballasts for the HID lamps(Marchesan, 2006, pp.1107-1112).

Conventional ballasts, which operate at low frequencies(50 or 60 Hz)are not the most practical solution, due to their large size, excessive weight and low efficiency. In order to reduce the size of the ballasts for HID lamps, high frequency (1k-10kHz) ballasts have been developed. However, the arc in the HID lamp tends to be unstable at specific frequency operation ranges. This phenomenon is called acoustic resonance. One method for avoiding this phenomenon is to operate at a frequency above this range. However, the switching loss on the ballast becomes greater as the operating frequency increases (Lin, 2006, pp.254-262). In order to reduce the switching loss, a resonant-type inverter is used to provide electronic ballast for a fluorescent lamp (Ribas, 2005, pp.1118-1123).

The authors have investigated several types of electronic ballasts for the HID lamps(Azcondo, 2007, pp.815-823). This paper presents a novel type ballast for the HID lamps without the acoustic resonance. The key feature of the circuit topology is the application of an integrated controller. The proposed circuit is very simple. It is composed of a DC/DC converter, DC/AC converter and an ignition circuit. The design and performance of this proposed circuit are described in this paper. Based on the theoretical and experimental results, the proposed ballast circuit is suitable for the HID lamps.

2. Description of the UCC3305 controller

Block diagram of the UCC3305 is showed in Figure.1.It has 28 pins.

The features of the UCC3305 HID controller are outlined below(Caceres, 1999, pp.134-141):

. OV input protection

. Output fault protection/timing

. Power regulation vs. lamp voltage

. Lamp start-up/cool down simulation

. Current-mode control

. Fixed frequency operation

. DC or AC lamp drive capability

. High current drive capability

. On board charge pump to provide gate drive down to 6VDC

. Adjustable start-up to steady-state current Ratio

3. Function Description

The UCC3305 integrates all of the functions required to control and drive one HID lamp. It contains a complete current mode pulse width modulator, a lamp power regulator, lamp temperature compensation, and total fault protection.

3.1 Slope Compensation

In addition to a complete current mode PWM, the UCC3305 HID Controller contains internal slope compensation (Cimador, 1990, pp.597-604), a valuable function which improves current loop stability for high duty cycles. Slope compensation is accomplished with an on-chip current ramp and an off-chip resistor RSL. Larger values of RSL give more slope compensation and a more stable feedback loop.

3.2 Lamp Temperature Compensation

Lamp temperature compensation is critical for automobile headlamps, because without compensation, light output varies dramatically from a cold lamp to one that is fully warmed up. The UCC3305 HID Controller contains sophisticated internal circuitry to anticipate lamp temperature and also to compensate for lamp temperature. The circuits anticipate lamp temperature by monitoring charge on capacitors which charge when the lamp is on and discharge when the lamp is off.

The UCC3305 HID Controller compensates for lamp temperature by driving the lamp with a higher lamp power when the lamp is cold and reducing the power to a normal operating level when the lamp is warmed up. The capacitors which set these time constants are external film capacitors CS and CW, and are connected to SLOPEC and WARMUPC. CS and CW are critical capacitors and must be selected to match the time-temperature relationship of the lamp. In addition to changing the power regulation point, the WARMUPC capacitor voltage also changes the short circuit lamp current. The ratio of cold short circuit current to warmed-up short circuit current is set by the resistor from ADJ to ground.

3.3 Power Regulation

Regulate power to the lamp within +10% over a lamp voltage variation of 60 to100VDC. The LPOWER output of the UCC3305 is a voltage roughly proportional to lamp input power. The UCC3305 regulates constant lamp power over a wide range of lamp voltages.. For inputs to VOUTSENSE below 0.5V, such as would occur with a shorted lamp, the loop regulates constant load current. For inputs to VOUTSENSE greater than 0.82V, as might occur with a lamp that is open or not yet lit, the loop also regulates constant load current, but at a lower current than for a shorted lamp. In between those two voltages, the amplifier driving the LPOWER pin will sum the load current and load voltage and produce a signal roughly proportional to load power. The summing amplifier approximates power well enough to hold power within   10% over a factor of two in lamp voltage.

4. Ballast Design

HID Ballast Schematic is showed in Figure 2.This circuit shows the UCC3305 HID Lamp Controller IC in a fly back converter. The output of the converter is regulated at constant power, so that lamp intensity is relatively constant regardless of small lamp manufacturing variations.

4.1 Full Bridge Output Stage

The output of the flyback converter is directed to the AC lamp through a full bridge inverter. The full bridge is switched at a low frequency (typically 195Hz), so that the average lamp voltage is zero.

The low frequency switching is derived from the PWM oscillator. It is desirable to switch lamp polarity when running, but switching lamp polarity can interfere with clean starting. The UCC3305 has a logic output called NOTON which is high when the lamp is not running (Not On) and low when the lamp is running. This output is connected to the DIVPAUSE input so that the low frequency switching stops until the lamp is fully lit.

4.2 Powering The UCC3305

Conventional power MOSFETs require at least 8V of gate drive to ensure high efficiency and low on resistance.Despite this requirement, the UCC3305 HID Controller can be used to build a ballast that will drive power MOSFETs well with input supplies as low as 5V! The UCC3305 does this using a charge pump.

In this typical application, power for the UCC3305 HID Controller IC is derived from a 6.8V zener supply. This zener regulated supply gives the application overvoltage protection, reverse battery protection, low parts count,and low cost.

The output of the 6.8V zener supply drives the VCC pin of the UCC3305. VCC is the input to the UCC3305 charge pump. The charge pump generates a regulated 10V supply on the BOOST output. This 10V supply drives all other functions on the UCC3305.

4.3 Protection From Over Voltage

The most significant stresses in an automotive environment are the overvoltage conditions which can occur during load dump and double-battery jump start. At these times, the voltage into the ballast can go so high that even the most overdesigned power stage will be damaged.

5. Experimental Results

A prototype ballast with function of constant lamp-power control is built for HID lamp. The constant lamp-power control scheme is fulfilled by simultaneously sensing lamp voltage and lamp current to achieve constant instantaneous power for the HID lamp. The measured lamp voltage *Vlamp* and lamp current *Ilamp* are shown in Figure.3. The rms values of lamp voltage and current at steady state are 86.2V and 0.4A, respectively. Due to low-frequency square-wave-driven ballast, no acoustic resonances occur: Moreover, the efficiency of the presented ballast is approximately 89%.

6. Conclusion

The proposed ballast, combining a buck-boost-flyback converter with a Full-bridge inverter, is negative-voltage-driven and offers constant lamp-power control. A prototype ballast circuit has been built in order to drive a 35W HID lamp from 12V battery DC voltage. The experimental results have demonstrated a low-frequency, square-wave-driven electronic ballast that is free of acoustic resonance, thus verifying the functionality of the proposed circuit.

References

Azcondo, F. J. and Zane, R. C. Branas. (2007). Design of resonant inverters for optimal efficiency over lamp life in electronic ballast with phase control[J]. *Power Electron*, 2007, 22(3):815–823.

Cáceres, R. O and Barbi, I. (1999). "A Boost DC-AC converter: analysis, design, and experimentation", *IEEE Transactions on Power Electronics*, vol.14, pp. 134-141.

Cimador, G. and Prestifilippo, P. (1990). "An attractive new converter topology for AC/DC, DC/DC & DC/AC power conversion", *12th International Telecommunications Energy Conference*, pp. 597-604.

Groot, J. de and Vliet, J. V. (1986). The High-Pressure Sodium Lamp. 1st ed., *Mac Millan Educational*, London.

Lin, R. L. and Chen, Y. T.. (2006). Electronic ballast for fluorescent lamps with phase-locked loop control scheme[J]. *Power Electron*, 2006, 21(1): 254–262.

Marchesan, T. B., Cervi, M, Campos A.and Prado, R.N (2006). "A Family of Electronic Ballasts Integrating Power Factor Correction and Power Control Stages to Supply HPS Lamps", *Conference Records of the IEEE IAS Annual Meeting*, vol.3, pp. 1107-1112.

Ribas, J., Garcia, J., Cardesin, J., Dalla-Costa, M. A., Calleja ,J. A. and Corominas, E. L. (2005). "High Frequency Electronic Ballast for Metal Halide Lamps Based on a PLL Controlled Class E Resonant Inverter", *Conference on IEEE Power Electronics Specialists*, pp.1118-1123.
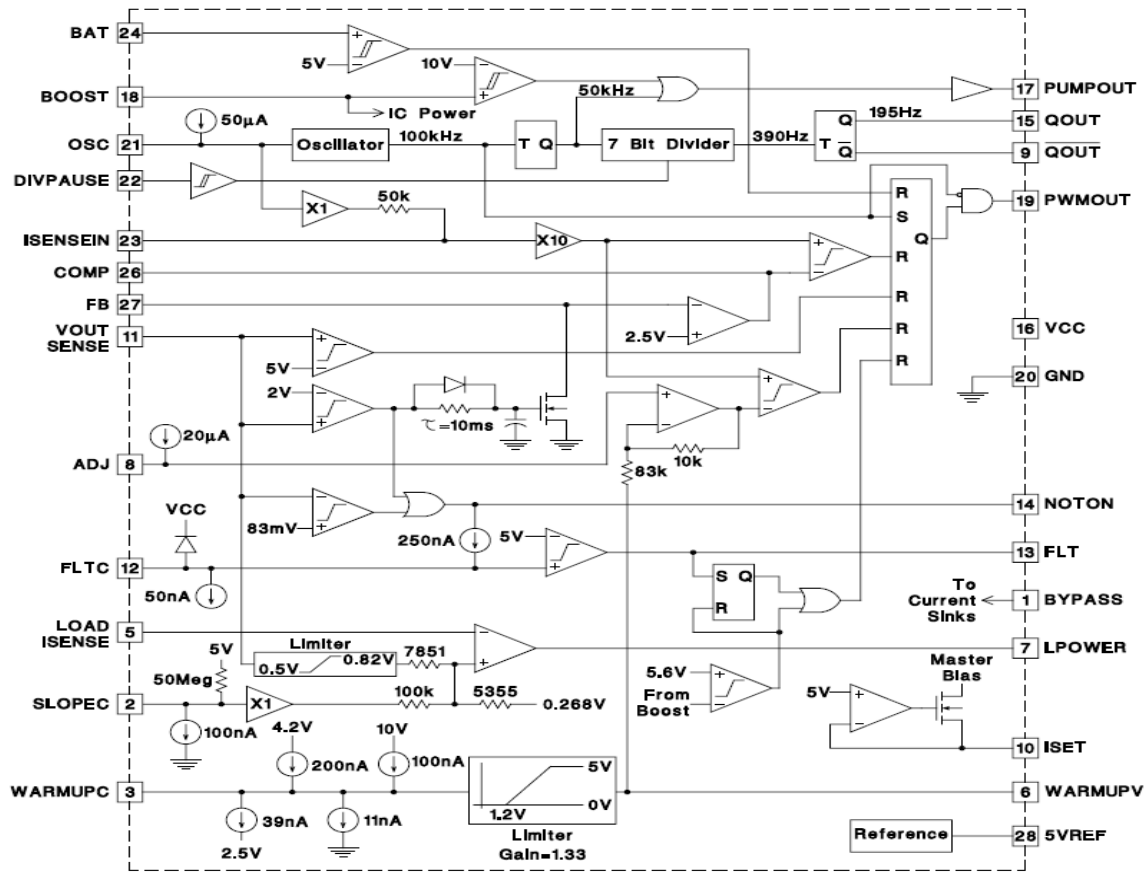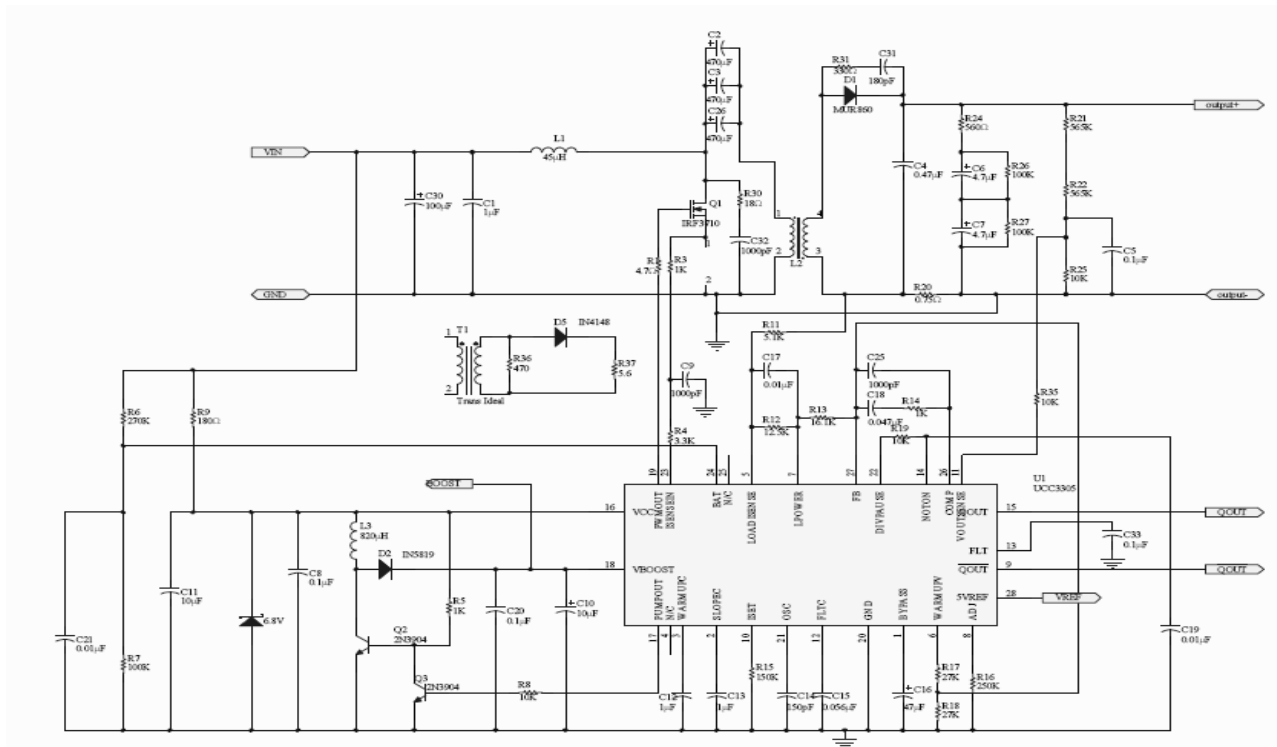
Figure 1. Block diagram of the UCC3305
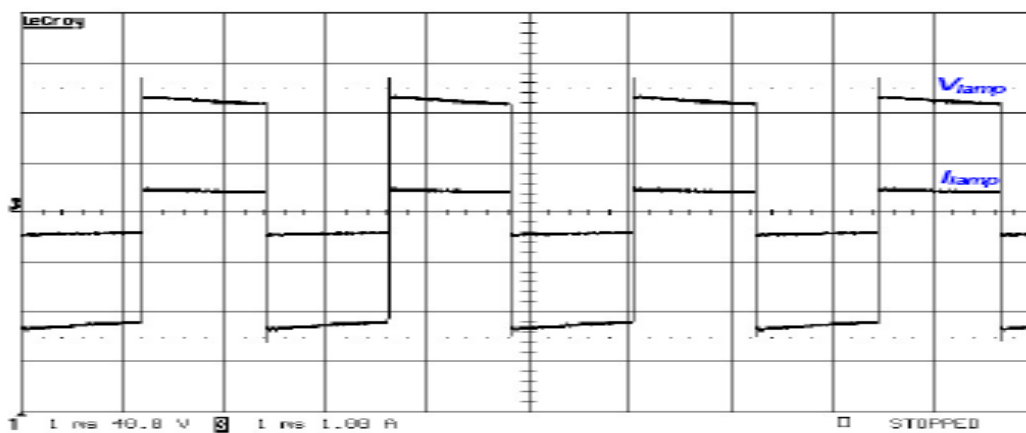

Figure 2. 35W DC HID Ballast Schematic

Figure 3. Lamp voltage *Vlamp* (40V/div) and current *Ilamp* (1A/div); time scale:1ms/div

Satellite Mapping of CO₂ Emission from Forest Fires in Indonesia Using AIRS Measurements

J. M. Rajab, M. Z. MatJafri, H. S. Lim & K. Abdullah

School of Physics, Universiti Sains Malaysia

11800 Penang, Malaysia

Tel: 60-4-653-3888   Fax: 60-4-657-9150

E-mail: jasim_rijab@yahoo.com, mjafri@usm.my, hslim@usm.my, khirudd@usm.my

Abstract

Results from the analysis of the retrieved carbon dioxide (CO_2) columns in the free troposphere are presented for one year (2005) obtained by the Atmospheric Infrared Sounder (AIRS) included on the EOS Aqua satellite launched on May 4, 2002. Providing information for several greenhouse gases, CO_2, CH_4, CO and O_3 is one goal of the AIRS instrument as well as to improve weather prediction and study the water and energy cycle. Carbon Dioxide (CO_2) is the most prominent Greenhouse gas in Earth's atmosphere and plays a key role in earth's climate. It is emitted into the air as humans exhale, from burning fossil fuels for energy and deforestation of the planet. The aim of this study is to generate Monthly CO_2 Distribution maps and to investigate the effects of Indonesia forest fires on CO_2 distributions over Peninsular Malaysia, north Sumatra and Singapore for 2005. The CO_2 concentration map of the study area was generated by using mole_fraction of CO_2 in free troposphere, obtained from AIRS/Aqua Level 3 monthly CO_2 retrieval product (AIRS+AMSU) V005 (AIRX3C2M) at GES DISC. Considerable variations were demonstrated in the annual changes of rainfall and drought patterns in various seasons (dry & wet season). Variations in the biomass burning and CO_2 emissions where noted over the study area, while the highest CO_2 occurred over industrial and congested urban zones and a greater draw down of CO_2 occurred in the pristine marine environment over northeast coasts of Sumatra during 2005. In particular, we observe a quasi-biennial variation in CO_2 emissions from study area with two peaks, the natural peak occurring at the end of each dry season (February to April), when biomass burning occurs, and the second peak at wet season (July to September), because of the influence of Indonesia forest fire. Examining satellite measurements, the results showed that the enhanced CO_2 emission correlates with occasions of less rainfall during dry season.

Keywords: Carbon dioxide, AIRS, Forest fire

1. Introduction

Many gases occur naturally in the atmosphere, while other from industrial wastes emissions. Our reliance on fossil fuels has been firmly solidified and global population has increased, so emissions of these gases have risen. While gases such as carbon dioxide occur naturally through the terrestrial biosphere and the ocean, humankind has altered the natural carbon cycle (through burning forest lands, oil, natural gas, wood or mining and burning coal) artificially move carbon from solid storage to its gaseous state, eighty-five percent of all human-produced carbon dioxide emissions come from the burning of fossil fuels like coal, natural gas and oil, including gasoline,(33% power plants, 22%cars and trucks, 12%major transportation, 33% Factories, home and heating systems). In the last half of the 20th century, carbon dioxide was the first greenhouse gas manifested to be increasing in atmospheric concentration with the first eventual measurements. Before the industrial revolution, concentrations were fairly stable at 280ppm, today they are around 370ppm and keeps rising steadily (Y. K. Tiwari et al., 2005). The increase from developing countries was three times that from developed countries, Middle East carbon dioxide emissions from burning of fossil fuels increased 35%, Africa increased 12% and Eastern Europe increased by 75% from 1990-1995(Greenhouse Gases, 2008). CO_2 is toxic in higher concentrations: 1% (10,000 ppm) will make some people feel drowsy, concentrations of 7% to 10% cause dizziness, headache, visual and hearing dysfunction and unconsciousness within a few minutes to an hour. Carbon dioxide in the

atmosphere is low, it is not practical to obtain the gas by extracting it from air. Some CO_2 is obtained from the combustion of coke or other carbon-containing fuels, while most commercial carbon dioxide is recovered as a by-product of other processes, such as the production of ethanol by fermentation and the manufacture of ammonia (Shakhashiri, 2008).

Southeast Asia is experiencing a similar rapid economic growth to that in Northeast Asia, is also a large source of several air pollutants, and may make an important contribution to regional and global pollution because of increasing anthropogenic emissions associated with biogenic emissions from large tropical forests, the greater oxidizing capacity in tropical regions due to higher UV intensity and humidity and rapid development and industrialization (Kato N, Akimoto H 1992; Streets DG et al., 2001). Because satellite remote sensing allows consistent, frequent, and comprehensive data collection at global scales and regional, scientists have discovered its use in studying biomass burning and fire emission. Parameters concluded from satellite data include fire counts, aerosol index, biomass density and burnt area (Cooke et al., 1996; Delmas et al., 1991). Carbon dioxide is the important trace gas retrieved by AIRS for the study of anthropogenic effects on climate.

The AIRS instrument is included among several instruments mounted on the (EOS) Earth Observing System Aqua satellite launched on 4 May 2002, with its partner, Advanced Microwave Sounding Unit (AMSU-A) are characterizing and observing the entire atmospheric column from the surface to the top of the atmosphere in terms of temperature and surface emissivity, cloud amount and height, atmospheric temperature and humidity profiles, and the spectral outgoing infrared radiation. The infrared brightness values coming up from Earth's surface and from the atmosphere are measure by AIRS. Their scan mirrors revolve around a central axis along the line of flight and directs infrared energy from the Earth into the instrument (Parkinson et al., 2003). The AIRS instrument, includes 2378 infrared spectral channels provides spectral coverage in the 3.74–4.61μm, 6.20–8.22μm, and 8.8–15.4μm infrared wavebands at a nominal spectral resolution of $\lambda/\Delta\lambda$=1200. The term "sounder" in the instrument's name refers to the fact that water vapor and temperature are measured as functions of height. The AIRS instrument also contains four visible/near-IR (Vis/NIR) channels between 0.40 and 0.94 μm, with a 2.3-km FOV. The most important minor gases retrieved from AIRS spectral radiances is CO_2 in the 712-750 cm^{-1} region (mid troposphere at a nadir resolution of 90km×90km), uses an analytical method based on the properties of partial derivatives. The AIRS broad swath makes it able to map the global distribution of CO_2 every day (Hartmut H. Aumann et al., 2003).

Immense plumes of the gas emitted from forest and grassland burning in Indonesia forest fires in 2005 caused a serious air pollution in Malaysia, northern Sumatra and Singapore. In Malaysia, air pollution reached extremely hazardous levels and forced schools and an airport to close. NOAA recorded 5420 hotspots from satellite images over the area between mid-July and mid-August. This study is based on CO_2 retrievals from a research version of the current AIRS operational physical algorithm, used Standard Level-3 Monthly gridded product (AIRX3C2M) Version 5 data, in the free troposphere using AIRS IR and AMSU, without-HSB, to investigate Monthly Distribution map of carbon dioxide over Peninsular Malaysia, north Sumatra and Singapore for 2005. By using Retrieved mole fraction of CO_2 in free troposphere, the CO_2 concentration map of the Peninsular Malaysia, northern Sumatra and Singapore was conducted, map were generated by using Photoshop CS & SigmaPlot 11.0 software to assess the effect of Indonesia forest fire August on atmosphere Carbon dioxide distribution.

2. Study Area and Data

An area covering 1.045X10^6 km^2, which includes the Peninsular Malaysia and northern Sumatra (centered at 100.755º E and 2.429º N) was selected for this study because a series of large fires occurred in the area in the August of 2005 (Figure 1).The extent of the domain was chosen so that it was sufficiently large to contain CO_2 plumes from severe fires in the area. The dimensions of the study domain are 1100 km E-W and 950 km N-S. The time period from January to December 2005 was selected to study the CO_2 distribution for 2005 covered the development of fires in the area since mid-June to mid-August was the period with the most number of detected fires.

The data used for this study include CO_2 data from AIRS and ground Station data; they were extracted for the study area and were processed to match in terms of space and time. The CO_2 data were derived from Atmospheric Infrared Sounder (AIRS) Version 5 Level 3 data available from the Goddard Earth Science Data and Information Services Center (GES DISC), as well as auxiliary data including the corresponding location and time along the satellite track in a HDF (Hierarchical Data Format) format on monthly basis. Using the location information, CO_2 data were gridded monthly at Geospatial Resolution of 2 degrees x 2.5 degrees (lat x lon). Twelve months composites (mean) were made to obtain near-complete coverage for 2005. The above-ground data were obtained from the Bukit Koto Tabang station is roughly 120 km north of Padang which the capital of the province West Sumatra, Indonesia, which is located in the tropical zone on the ridge of a high Barisan plateau at an altitude of 864.5 m a.s.l., latitude 0.20Sº longitude 100.32º and 40 km off the western coastline.

3. Acquisition and Specification

The AIRS spectrometer is devised to operate in synchronism with the microwave instruments AMSU-A1, AMSU-A2 and HSB. Its science objectives is to determinate the factors that control the global energy and water cycles, inquisition

of atmosphere-surface interactions, improving numerical weather prediction, evaluating climate variations and feedbacks and diagnosis of the effects of increased carbon dioxide, methane, ozone and other greenhouse gases (J.Le Marshall et al., 2006). Table 1 describes AIRS Technology – Specifications. We use Standard (Version 5.0), from Collection (AIRX3C2M), AIRS/Aqua Level 3 monthly CO_2 retrieval product in the free troposphere (AIRS+AMSU), and available at website: http://disc.sci.gsfc.nasa.gov/AIRS/data-holdings/by-data product/airsL3_C2M.shtml. AIRS radiance data in the regions at 700 cm-¹ and 2400 cm-¹ were used to retrieve mid-tropospheric CO_2 at nadir resolution of 90 km X 90 km (Hartmut H. Aumann et al., 2003). The AIRS mid-tropospheric CO_2 Level 3 Monthly Gridded Retrieval Product contains standard retrieval means, standard deviations and input counts as well as the latitude and longitude arrays giving the centers of the grid boxes.   Each file covers a calendar month. The mid-tropospheric CO_2 retrievals have been averaged and binned into 2.5°x 2° grid cells, from -180.0° to +180.0° longitude and from -60.0° to +90.0° latitude. The global, space-based measurements of the column-averaged CO_2 dry-air mole fraction, CO_2 could dramatically improve our understanding of the environmental processes that control the atmospheric CO_2 budget (Olsen, S. C., and Randerson, 2004).

4. Results and Discussion

The six maps in Figures 2a-2f illustrate the extent of AIRS monthly coverage of mid-tropospheric CO_2, the nominal peak of AIRS vertical sensitivity and the magnitude of the seasonal variations in free tropospheric CO_2 over Peninsular Malaysia, north Sumatra and Singapore for a dry season (November - April) of 2005.   A quasi-biennial variation in CO_2 over study area, plainly evident the high values of CO_2 occurred over northern and middle of Malaysia and southern and northern Sumatra, higher value was (380.751 ppm) at (lat. 4º - long. 100º) for December, and 381.636 ppm lat. 4º - long. 95º for November. During the late dry season between January and April, Observed CO_2 are lowest in the southern regions, increase at the northern regions, and are highest in the north of peninsular Malaysia 379.285 ppm at lat. 8º - long. 97.5º in April. The ~3.5 ppm higher CO_2 in the north than at south regions has mainly been attributed to long-range of air masses blowing from northeast and from western Asia (Pochanart P et al., 2003) as well as biomass burning occurs during these months, especially in the late dry season (boreal springtime), biomass burning pollution is raised by deep convection to the upper and middle troposphere, hold by the sub-tropical westerly flow, and transported to the central western Pacific.

Figures 3a-3f illustrate the extent of AIRS monthly coverage of mid-tropospheric CO_2 during wet season (May – October) of 2005 with high values in the north and low value in the middle region. Maximum values of CO_2 over the northern regions of Malaysia occur in June to September with minima over Sumatra occurs in May and October. Plainly evident in the monthly average, differences in the CO_2 spatial patterns for each of the peak months, variations are visible for other geographic regions, but none is as regular during these six months as those over Malaysia. During the wet season, Southeast Asia is governed by marine air masses from the middle and low latitudes of the Southern Hemispheric Indian Ocean, which   bring small to moderate amounts of air pollution to continental Southeast Asia. There is no clear observational evidence that the transport of air pollution exerts a strong effect on Southeast Asia, partly due to the transport attributes of the region, as well as previous study and investigations showed minima values of CO_2 occurred in wet season (Tsusumi Y and Matsueda H, 2000). In contrast, the highest observed CO_2 occurs during wet season 2005, plainly evident the impact of Indonesia forest fire August 2005 making opposite process at July to September. The ~2.5 ppm higher CO_2 in wet season than at dry season, especially occurs precisely in a region that experienced extensive fires over northern area of Peninsular Malaysia, nearest to the fire regions, this extended from Riau to north Sumatra in Indonesia affected 10,000 hectares of peat forest, smoke from forest fires were badly affected and caused serious air pollution, as the wind carried the smoke further afield from July to September 2005.

From Figures 2 & 3, with the exception of the forest fire influence, the highest CO_2 occurs over Industrial and congested urban zones and a greater draw down of CO_2 occurs in the pristine marine environment over southwest coasts of Peninsular Malaysia and west coasts of Sumatra. Variation in CO_2 over study area in 2005 is plainly evident in the monthly average AIRS CO_2 maps presented. Maximum values of CO_2 occur in June, July, and August, with minima in January, February, and March.

Figure 4 illustrates the annual average distribution of CO_2 for 2005. The local CO_2 maximum just north of Peninsular Malaysia (379.044ppm ,area above latitude 4) occurs precisely in a region that experienced extensive Indonesia forest fire in 2005. In contrast, the figure shows much less CO_2 over the middle and northeast coasts of Sumatra. Looking carefully at Figure 4, two very localized areas stand out against their backgrounds: Kuala Lumpur (3N, 101.5E) and Johor (1.5N, 103.5E), although, considered as broad Industrial and congested urban zones but their CO_2 values are still less than over the northern regions and this situation was mainly attributed to Indonesia forest fires. Also note, that a greater draw down of CO_2 occurs in the pristine marine environment over northeast coasts of Sumatra during 2005.

The variation of CO_2 values for 2005 from Bukit Koto Tabang station measurement (near the surface, Sumatra, Indonesia) and the corresponding mid-tropospheric column from AIRS are illustrated in Figure 5.Note two peaks for CO_2 measurement from ground station, the natural peak occurred in the late dry season between February and April,

and the CO_2 values had a large seasonal cycle whose amplitude was largert near the surface than column average CO_2 as a result of biomass burning. The second peak which occurred between July to September was caused by the Indonesia forest fire and thus phenomenon did not appear in the previous years. While from May to January the column average CO_2 was greater than surface CO_2, exceeded ~2-6 ppm.

5. Conclusion

As demonstrated here, AIRS' monthly observations of free troposphere CO_2 above the study area enable detailed analyses of both the spatial and temporal variations in emissions and the visualization of subsequent transport, AIRS can successfully detect CO_2 emission from large forest fires. We have investigated the wealth of information contained in the one year of AIRS data. AIRS/Aqua Level 3 monthly CO_2 retrieval Standard (Version 5.0) products in the free troposphere (AIRS+AMSU) V005 (AIRX3C2M), were used to evaluate the monthly CO_2 distributions, caused by the Indonesia forest fires.

Plainly evident the highest values of CO2 occurred during biomass burning in the late dry season and over Industrial and congested urban zones. The local CO2 maximum in the north Peninsular Malaysia occurs in a region that experienced the effect of extensive Indonesia forest fire in 2005. A greater draw down of CO2 occurs in the pristine marine environment over northeast coasts of Sumatra during 2005.

This paper has provided evidence for the impact of remote biomass burning and forest fire on Carbon Dioxide pollution levels above study area and enhanced our knowledge on AIRS detection of CO_2 emission from forest fire, the accuracy of remotely sensed tropospheric CO_2 columns and abundances from AIRS.

References

Christopher SA, Chou J, Welch RM, Kliche DV, Connors VS. (1998). Satellite investigations of fire, smoke and carbon monoxide during April 1994 MAPS mission: Case studies over tropical Asia. *J Geophys Res.*, 103, 19327-19336.

Cooke, W. F., Koffi, B., & Gre´goire, J. -M. (1996). Seasonality of vegetation fires in Africa from remote sensing data and application to a global chemistry model. *Journal of Geophysical Research*, 101, 21001– 21065.

Delmas, R. A., Loudjani, P., Podaire, A., & Menaut, J. -C. (1991). Biomass burning in Africa: An assessment of annually burned biomass. In Joel S. Levine (Ed.), The Global Impact of Biomass Burning: Atmospheric, Climatic, and Biospheric Implications (pp. 126–132). Cambridge7 The MIT Press.

Greenhouse Gases. (2008). http://www.umich.edu/~gs265/society/greenhouse.htm

Hartmut h. Aumann, Moustafa T. Chahine, Catherine Gautier, Mitchell D. Goldberg, Eugenia Kalnay,Larry M. Mcmillin, Hank Revercomb, Philip W. Rosenkranz, William L. Smith,David H. Staelin, L. Larrabee Strow, and Joel Susskind. (2003). AIRS/AMSU/HSB on the Aqua Mission: Design, Science Objectives, Data Products, And Processing Systems., *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, no. 2, February 2003.

J.L Marshall, J. Jung, et al. (2006). "Improving global analysis and forecasting with airs ", *American meteorological society, doi*: 10.1175/bams-87-7-891.

Kato, N. and H. Akimoto. (1992). Anthropogenic Emissions of SO_2 and NO_x in Asia: Emissions Inventories (plus errata), Atmos. Environ., 26A, 2997-3017, 1992.

Parkinson, C. L. (2003). Aqua: An earth-observing satellite mission to examine water and other climate variables, *IEEE Trans. Geosci. Remote Sens.*, 41, 173– 183.

Pochanart P, Akimoto H, Kajii Y, Sukasem P. (2003). Carbon monoxide, regional-scale transport, and biomass burning in tropical continental Southeast Asia: Observations in rural Thailand. *J.Geophys. Res.*, 108(D17), 4542, doi: 10.1029/2002JD003360 (2003).

Prof. Shakhashiri. (2008). General Chemistry, www.scifun.org

Streets DG, Tsai NY, Akimoto H, Oka K. (2001). Trends in emissions of acidifying species in Asia, 1987–1997. *Water, Air, and Soil Pollution* 130:187–192

Tsutsumi. Y and H. Matsueda. (2000). Relationship of ozone and CO at the summit of Mt. Fuji (35.35°N,138.73°E, 3776m above sea level)in summer 1997. *Atmospheric Environment*, 34, 553-561, 2000

Y. K. Tiwari, M. Gloor, R. Engelen, C. Rodenbeck, M. Heimann. (2005). Comparing model predicted atmospheric CO_2 with satellite retrievals and in-situ observations -Implications for the use of upcoming satellite data in atmospheric inversions, *Geophysical Research Abstracts*, Vol. 7, 09823, 2005.

Table 1. AIRS Technology – Specifications

Instantaneous Field of View (IFOV)	1.1°
Spectral Resolution	13.5x13.5 km in the nadir
coverage	pole-to-pole & covers the globe two times a day
orbit	438 miles (705.3km) polar, sun synchronous, 98.2+/- .1 degrees inclination, ascending node 1:30pm +/- 15 minutes, period 98.8 minutes
Ground Coverage	+/- 49.5 degrees around nadir
Ground Footprint	90 per scan, 22.4 ms footprint
Temporal Coverage	Global, twice daily swath (daytime and nighttime)
Radiometric Calibration	+/- 3% absolute error
Spatial Coverage	Scan Angle: +/- 49.5 around nadir IFOV: 0.185
Power / Mass	256 W / 166 kg
Swath Width	1650 km

Figure 1. The Study area
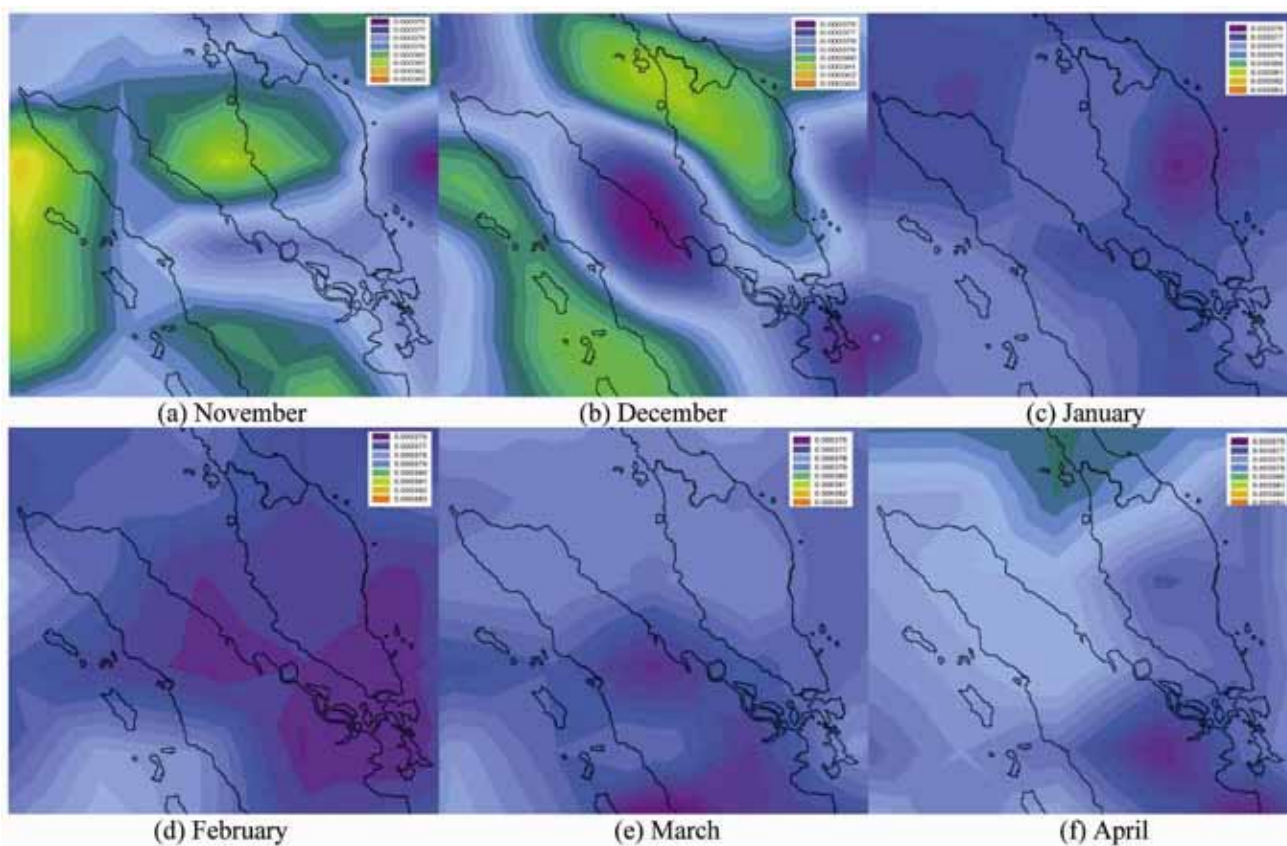
Figure 2. AIRS monthly coverage of mid-tropospheric CO_2, from November to April 2005
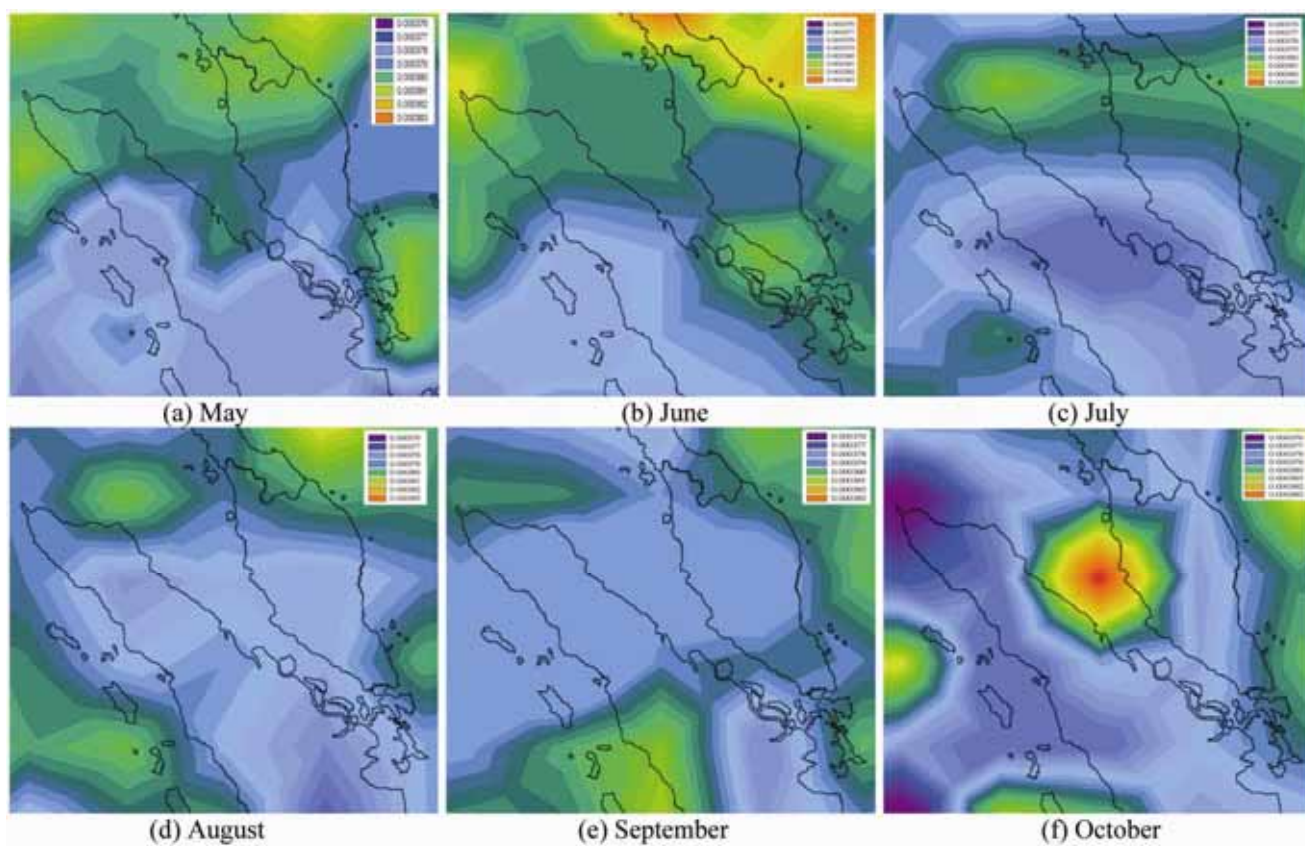
Figure 3. AIRS monthly coverage of mid-tropospheric CO_2, from May to October 2005
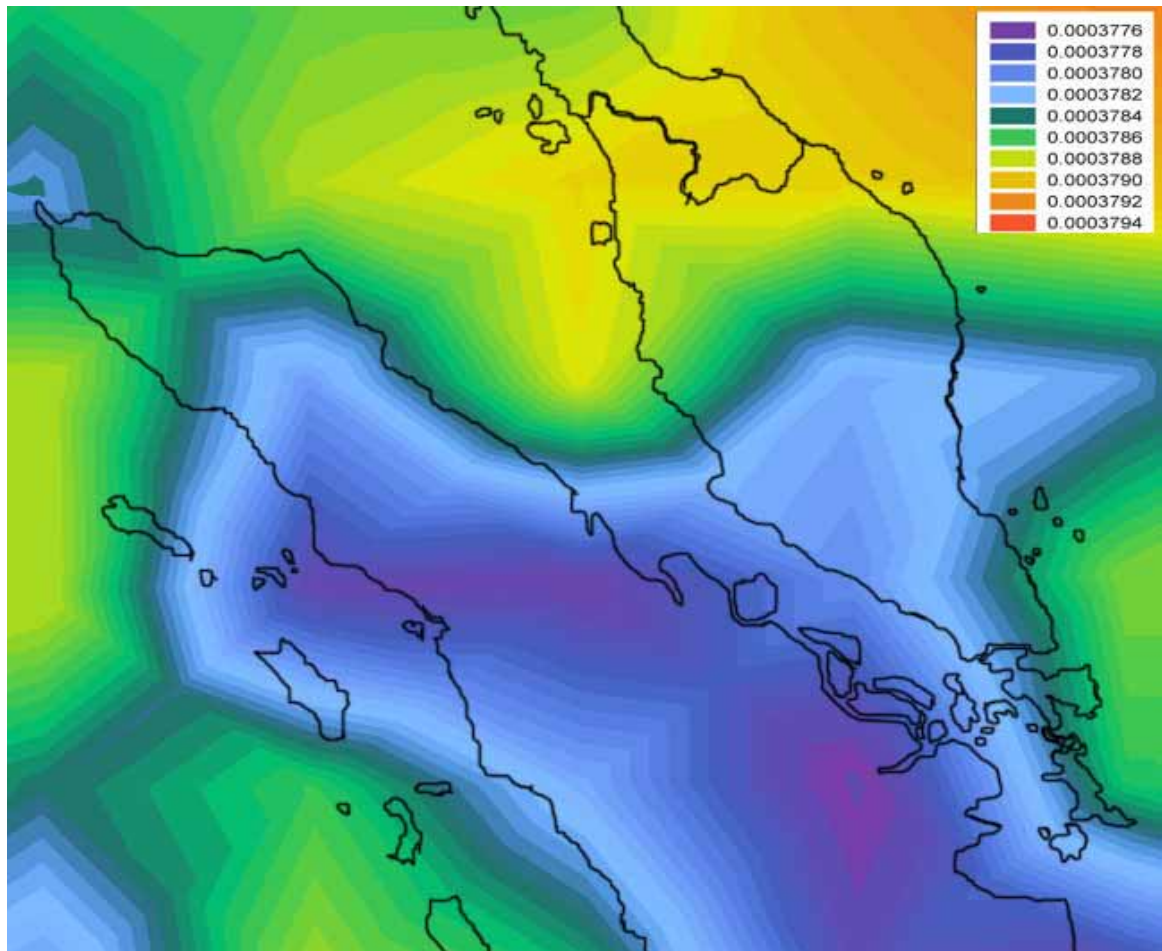
Figure 4. Presents the 2005 average distribution of CO_2
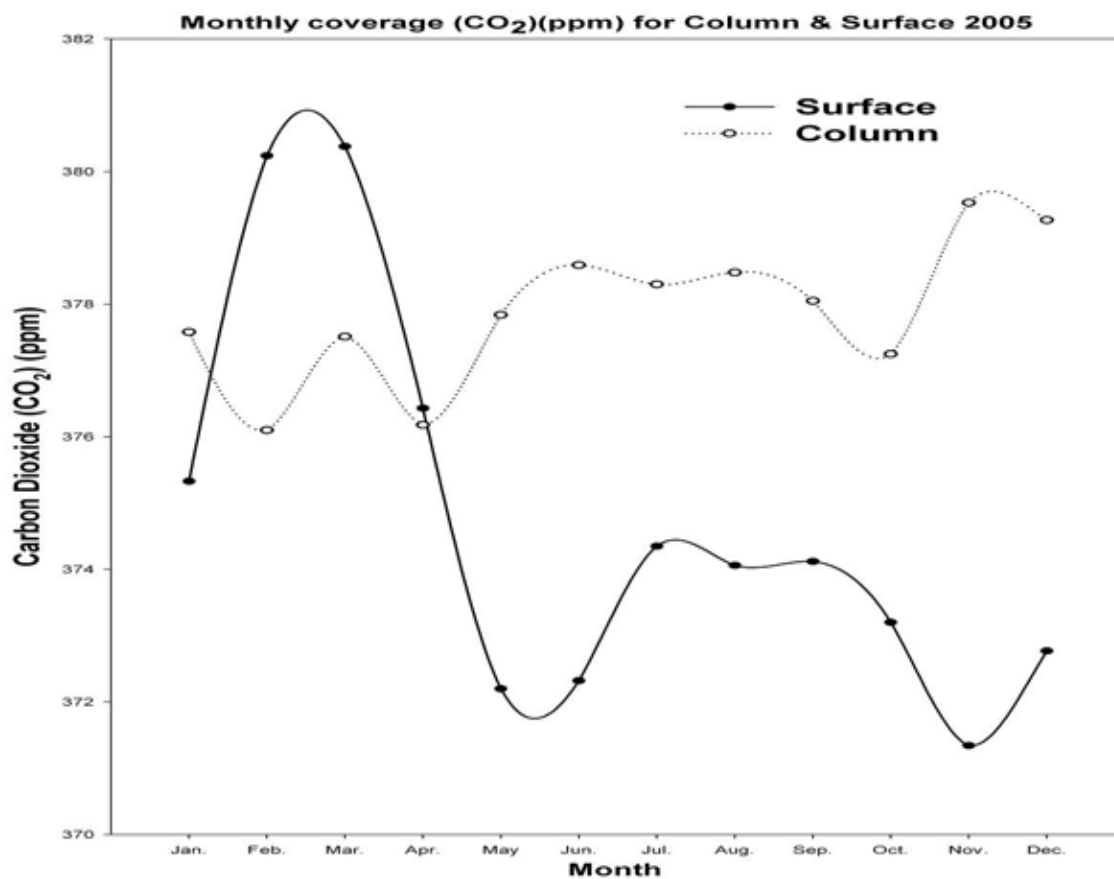
Figure 5. Monthly coverage carbon dioxide (CO2) (ppm) at (100.32 E, 0.20 S) for column
from AIRS (dashed line) and Surface (solid line)

Experimental Study on Continuous Degradation of Simulated Wastewater Containing Complex Black WAN with Nano TiO$_2$ Photocatalyst Prepared at Low Temperature

Wei Feng (Corresponding author)

Tianjin University, Tianjin 300072, China

Kehua Zou

State Key Laboratory of Odor Pollution Control

Tianjin Academy of Environmental Sciences, Tianjin 300191, China

Wentao Jin

College of Chemical Engineering

Tianjin University, Tianjin 300072, China

Zengxiu Zhai

College of Material Science and Chemical Engineering

Tianjin Polytechnic University, Tianjin 300160, China

E-mail: zhaizengxiu1983@sina.com

Jinhong Zou

Jiangxi Ganjiang Vocational College, Jiangxi 330000, China

Abstract

Nano TiO$_2$ photocatalyst has been prepared by the two-step microwave-assisted process at low temperature. The characterization results reveal that the product is titanium dioxide with a pure anatase crystal structure, and has a large specific surface area. In addition, a small size circulated photocatalytic reactor has been designed and developed. This paper investigates the activity of the above-mentioned photocatalyst by the orthogonal experiment on the degradation of simulated wastewater containing complex black, and studies the operation process of the self-made reactor.

Keywords: Microwave, Preparation at low temperature, Nano titanium dioxide, Circulated photocatalytic reactor

Preface

As is well known, TiO$_2$ is the best photocatalyst applied in the current photocatalysis technology (Gao et al, 2002). However, high temperature of 400 ℃ at the last stage of the traditional process for preparing TiO$_2$ photocatalyst, which converts TiO$_2$ into anatase phases of photocatalytic activity, not only leads to the particle agglomeration which has an adverse effect on the dispersion of TiO$_2$, but also results in the dehydroxylation of the crystal surface which causes a remarkable decline of the specific surface area and the hydroxylation degree, and makes the electron-hole pair separation more difficult, in turn has an adverse effect on the absorption of pollutants on the surface of catalyst, then leads to the decline of the catalytic activity. Moreover, high temperature always causes side reactions which lead to the decrease of the production efficiency, decline of the selectivity of carrier, and increase of the production cost. Therefore,

to develop a nano anatase photocatalyst and compounds therefrom plays an important role in the study of photocatalysis technology. In recent years, there has been a big progress in the research and development of nano anatase photocatalyst, such as hydrothermal reaction at low temperature (Langlet, M et al, 2003, p. 3945-3953), sol-gel process (Kotani, Y et al, 2001, p. 2144-2149), and hydrolytic precipitation at low temperature (Yamabi, S H et al, 2003, p. 86-93).

Microwave-assisted chemical synthesis is a newly developed method for preparing materials. Its process, as well as microstructure and quality of the obtained product, have incomparable advantages. Its reaction mechanism is very different from that of the traditional method. It is not clear if there is non-thermal effect in microwave-assisted reaction. In 1988, Baghurst et al (1988, p. 829-830) prepared a superconducting ceramics by microwave-assisted process. At present, many research achievements indicate that microwave-assisted process not only avoids the high temperature calcinations and saves energy, but also controls the crystal size and growth orientation (Wang, J et al, 2005, p. 1405-1408; Yang, H M et al, 2007, p. 1357-1363; Newalkar, B L et al, 2008, p. 271-276), and favors the preparation of nano TiO_2 photocatalyst at low temperature. There have been quite a few reports on the preparation of nano TiO_2 photocatalyst at low temperature (Yang, Shenghong et al, 2000, p. 354; Chen, Wenxin et al, 2004, p. 661-663; Zhou, Xiaoming et al, 2005, p. 277-279; Peiro, A M et al, 2002, p. 185-191).

Photocatalytic reactor is very different from traditional reactor. In photocatalytic reactor, light irradiation should be strong enough to activate the photocatalyst, photocatalyst can be easily separated from the reaction system and reused, light source is barrier-free, light utilization efficiency is high, processing capacity is large, and etc. Based on the modification of traditional reactor by the two-phase separation and precipitation process, as well as the wastewater recycling process, nano TiO_2 photocatalyst has been prepared by the microwave-assisted process at low temperature, which in turn accelerates the progress in industrialization of energy-saving photocatalysis oxidation technology.

1. Nano TiO_2 photocatalyst prepared by microwave-assisted process at low temperature

1.1 Materials, equipments and methods

Nano TiO_2 photocatalyst has been prepared by microwave-assisted hydrolysis process. Microwave initiates the vibration of O—H in H_2O, and results in a large amount of heat, which, in one hand, accelerates the hydrolysis of Ti^{4+} into a local supersaturated solution of TiO_2, and in other hand, increases the temperature of this supersaturated solution rapidly in a very short time, and results in a large amount of crystal nucleus. Nano-scale TiO_2 can be obtained only when the nucleus-forming speed is greater than the crystal-forming speed.

(1) Methods

Weigh a certain amount of $Ti(SO_4)_2$ into a premixed H_2SO_4 (HCl) solution with pH of 2 to obtain a 100 g/L acid solution of $Ti(SO_4)_2$. Regulate pH to 8 by adding concentrated ammonia drop-by-drop into the solution under the condition of strongly stirring. Irradiate the obtained solution for a certain period in an experimental microwave oven under different microwave conditions. Age, wash, remove part of water, and then irradiate for the second time in the microwave oven. The temperature is lower than 200 ℃ throughout the experiment.

(2) Equipments

ML08S-1 experimental microwave oven is from Nanjing Huiyan Microwave Co., Ltd.; DHS-25 pH meter is from Shanghai REX Instrument Factory. The crystal structure of the obtained TiO_2 is identified with X'pert PRO X-Ray Diffractormeter (Co target, Kα radiation, 30 mA tube current, 30 kV tube voltage, scan range of 10~90 °, step width of 0.0334 °) which is from PANalytical of Holland. The internal structure of crystal of the obtained TiO_2 is imaged with G^2F20 HRTEM from PHILIPS TECNAI.

(3) Materials

Main reagents: absolute ethanol, ammonia water, sulfuric acid, hydrochloric acid, titanium sulfate, barium chloride, potassium chloride, methyl orange solution and etc. All of them are analytical reagents.

1.2 Analysis on experimental results

Based on the previous investigations, it is determined to improve the crystal size and crystallinity of nano TiO_2 powder by increasing irradiation power and lengthening irradiation time under the pulsed microwave irradiation. The titled nano TiO_2 photocatalyst has been prepared under the conditions of using sulfuric acid, continuous microwave irradiation for the 1st time, pulsed microwave irradiation for the 2nd time, and power of 700 W. Figure 1 exhibits the XRD spectrum of nano TiO_2 obtained after 15 minutes of two-step microwave irradiation.

Figure 1 indicates that the self-made sample is a pure anatase of low crystallinity, which has an average diameter of 4.72 nm (calculated with Scherrer formula), and specific surface area of 225 m^2/g. It proves that the self-made sample has a small particle size and large specific area, which are favorable for the formation of hole/electron pair through electron transition, and form much more active reaction centers for the absorptions of pollutants during photocatalysis reaction.

Figure 2 exhibits the HRTEM images of nano TiO_2 obtained after 15 minutes of two-step microwave radiation. It reveals that the self-made sample is composed of 6~9 nm stout prismatic crystals which have irregular edges, obvious lattice defects, and severe agglomeration. The characterization results indicate that the reaction system releases a large amount of heat rapidly in a very short time under the action of microwave irradiation. As a result, the nucleus-forming speed is much greater than the crystal-forming speed, which results in evenly distributed small crystal particles. What's more, the crystallization degree is further improved by the two-step microwave-assisted process which makes the concentrated solution supersaturated. Therefore, as far as the crystal-forming effect is concerned, the two-step microwave-assisted process works the same way as the high temperature calcinations. However, high frequency vibration of reactant molecule initiated by the microwave irradiation results in much more irregular edges and lattice defects, which are favorable for the formation of hole/electron pair through electron transition, and accelerate the photocatalysis reaction.

2. Continuous degradation equipment of simulated wastewater containing complex black WAN

2.1 Continuous degradation equipment

Figure 3 exhibits the continuous degradation equipment of simulated wastewater.

TiO_2-contained wastewater flows from the raw solution tank into reactor, and then the three-phase fluidized photocatalysis reaction occurs in the reactor with gas flowing from gas distributor. The obtained liquid flows into the settling pond, and the supernatant outflows into the inner circulation self-dilution system. The continuous inner circulation reactor has the following features: (1) the treated wastewater can be reused as the circulation diluents; (2) the inner circulation self-dilution technology ensures the effective reutilization of nano TiO_2 photocatalyst; (3) based on the above-mentioned process, the continuous multi-level reaction system can be designed to increase the output and the comprehensive indexes of the treated wastewater.

2.2 Experimental method and result analysis

Formulate the complex black WAN-contained simulated wastewaters of different initial concentrations based on the practical industrial applications. Design a 3-factor-3-level orthogonal table (see table 1).

Measure the absorption (A) of the formulated complex black WAN-contained simulated wastewaters with UV-752 UV-Vis Spectrophotometer at wavelength of 740 nm, and calculate the decoloration rate (X) of above solutions. Measure the total organic carbon (TOC) of the degraded simulated wastewaters with TOC-5000A TOC detector.

$L_9(3^4)$ is used to schedule 9 operation conditions of orthogonal experiment, with the three factors being initial concentration of wastewater, feed flow rate and circulation flow rate. The effects of these factors on the chromaticity and TOC of the simulated wastewaters are investigated by the orthogonal experiment (see table 2).

Table 2 reveals that nano TiO_2 has a remarkable effect on the degradation of complex black WAN-contained simulated wastewater, which indicates that nano TiO_2 photocatalyst prepared by the microwave-assisted process at low temperature has good photocatalytic activity.

Figure 4 exhibits the changes of average value of the above-investigated factors at different levels. From figure 4, we can see that both COD_{Cr} removal rate and TOC removal rate change a little at the level of C_2 and circulation flow rate, however, decrease with the rise of feed flow rate. It indicates that the circulation photocatalytic reactor can achieve a better photocatalytic effect by regulating the feed flow rate. However, the circulation flow rate can be set in an easy-to-handle range since its effect on the degradation is little. The process should be further studied to develop a multi-level recycling processing equipment, which can increase the tolerance of equipment for processing wastewater. Therefore, the method studied here is valuable for broader applications.

3. Conclusion

(1) Nano TiO_2 photocatalyst has been prepared by the two-step microwave-assisted process at low temperature. The characterization results reveal that the obtained product has a larger specific surface area and much more crystal defects which favor the photocatalytic reaction.

(2) A small size circulated photocatalytic reactor has been designed and developed. The effects of the obtained nano TiO_2 photocatalyst on the degradation of simulated wastewater are investigated by the orthogonal experiment on the initial conditions of wastewater and operation process. The results indicate that the low feed flow rate favors the degradation of wastewater.

(3) The major task in the future study is to optimize the conditions for preparing nano TiO_2 photocatalyst by two-step microwave-assisted process at low temperature, increase its photocatalytic activity, develop the multi-level continuous processing equipment based on the method studied here, and figure out the way of reutilization of catalyst.

References

Bharat L Newalkar, T Chiranjeevi, N V Choudary & S Komarneni. (2008). Microwave-hydrothermal synthesis and

characterization of Co-VSB-5 microporous framework. *Journal of Porous Materials*, 15(3), 271-276.

Chen, Wenxin, Huang, Langhuan, Liu, Yingliang, Wang, Peng, Zhang, Jingxian, Rong, Jianhua & Tang, Lina. (2004). Nano-titania prepared by microwave heating hydrolysis. *Chemical Research and Application*, 16(5), 661-663.

D R Baghurst & D M P Mingos. (1988). Application of microwave-heating techniques for the synthesis of solid-state inorganic-compounds. *Journal of the Chemical Society-Chemical Communications*, 12, 829-830.

Gao, lian, Zheng, Shan & Zhang, Qinghong. (2002). *Nano TiO$_2$ photocatalyst and its application*. Beijing: Chemical Industry Press.

Huaming Yang, Xiaohui Su & Aidong Tang. (2007). Microwave synthesis of nanocrystalline Sb$_2$S$_3$ and its electrochemical properties. *Materials Research Bulletin*, 42(7), 1357-1363.

Jun Wang, Jieming Cao, Baoqing Fang, Peng Lu, Shaogao Deng & Haiyan Wang. (2005). Synthesis and characterization of multipod, flower-like, and shuttle-like ZnO frameworks in ionic liquids. *Materials Letters*, 59(11), 1405-1408.

Kotani Y, Matsuda A, Kogure T, Tatsumisago M & Minami T. (2001). Effects of addition of poly (ethylene glycol) on the formation of anatase nanocrystals in SiO$_2$-TiO$_2$ gel films with hot water treatment. *Chemistry of Materials*, 13(6), 2144-2149.

M Langlet, A Kim, M Audier, C Guillard & J M Herrmann. (2003). Liquid phase processing and thin film deposition of titania nanocrystallites for photocatalytic applications on thermally sensitive substrates. *Journal of Materials Science*, 38(19), 3945-3953.

Peiro A M, Vigil E, Peral J, Domingo C, Domenech X & Ayllon J A. (2002). Titanium (IV) oxide thin films obtained by a two-step soft-solution method. *Thin Solid Films*, 411(2), 185-191.

Satoshi Yamabi & Hiroaki Imai. (2003). Synthesis of rutile and anatase films with high surface areas in aqueous solutions containing urea. *Thin Solid Films*, 434(1-2), 86-93.

Yang, Shenghong, Zhang, Xiaoming, Zhang, Tingjie, Wang, Keguang & Zhu, Yubin. (2000). TiO$_2$ nano powders prepared by means of microwave radiation. *Rare Metal Materials and Engineering*, 29(5), 354.

Zhou, Xiaoming, Liu, Jianhong, Zhang, Peixin, Yang, Biming, Zhang, Qianling & Zhang, Xueli. (2005). Preparation of nano-particle TiO$_2$ at low temperature. *Journal of Shenzhen University (Science & Engineering)*, 22(3), 277-279.

Table 1. Factor/level of orthogonal experiment

Levels	Factors		
	A (Initial concentrations of simulated wastewater)	B (Feed flow rate)	C (Circulation flow rate)
1	C$_1$	Q$_1$ (0.01625 mL/s)	Q$_1$ (1.5015 mL/s)
2	C$_2$	Q$_2$ (0.03250 mL/s)	Q$_2$ (1.7745 mL/s)
3	C$_3$	Q$_3$ (0.04875 mL/s)	Q$_3$ (2.0475 mL/s)

- C$_1$ represents the wastewater containing 100 mg/L complex black WAN, 50 mg/L α-naphthol and 15 mg/L NaCl, which has COD$_{Cr}$ of 176 and TOC of 52.87.

- C$_2$ represents the wastewater containing 200 mg/L complex black WAN, 100 mg/L α-naphthol and 30 mg/L NaCl, which has COD$_{Cr}$ of 326 and TOC of 104.1.

- C$_3$ represents the wastewater containing 250 mg/L complex black WAN, 150 mg/L α-naphthol and 40 mg/L NaCl, which has COD$_{Cr}$ of 456 and TOC of 132.5.

- The amount of photocatalyst is 1 g/L, and UV radiation time is 2 h.

Table 2. L$_9$(3^4) orthogonal experiment

Samples	After 2 hours of reaction		
	Decoloration rate	COD removal rate	TOC removal rate
1	1	0.913	0.922
2	1	0.855	0.917
3	1	0.835	0.850
4	1	0.938	0.905
5	0.97	0.904	0.893
6	1	0.966	0.952
7	0.96	0.922	0.907
8	0.971	0.956	0.929
9	0.97	0.938	0.911
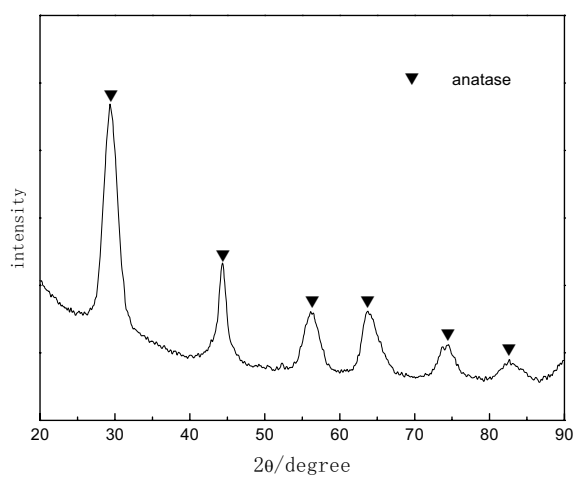
Figure 1. XRD spectrum of nano TiO$_2$ obtained after 15 minutes of two-step microwave radiation at different conditions

(a) 20 nm scale   (b) 10 nm scale
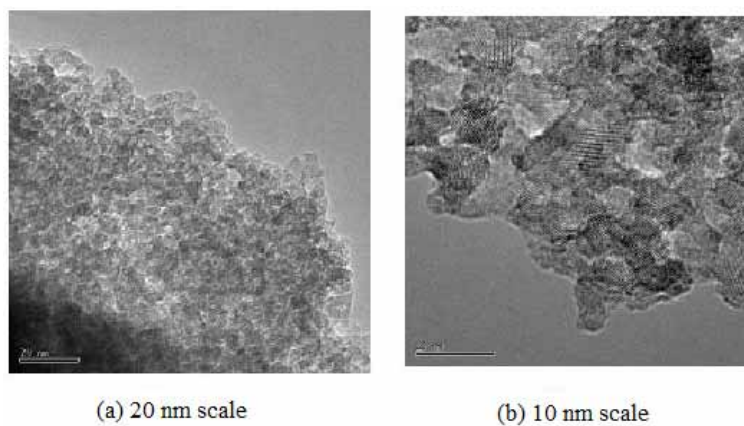
Figure 2. HRTEM images of nano TiO$_2$ obtained after 15 minutes of two-step microwave radiation
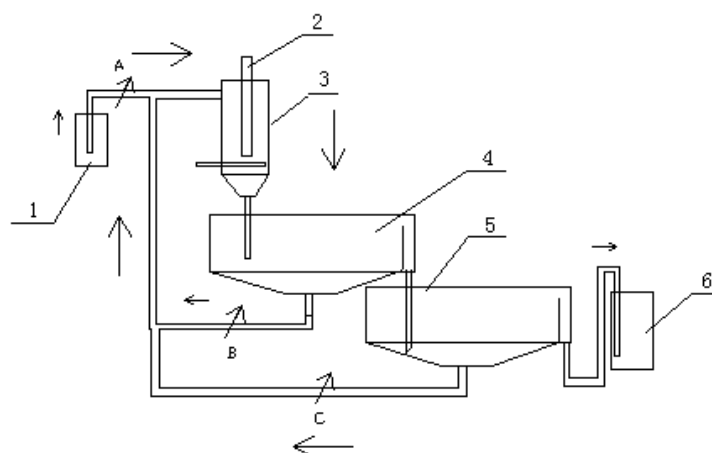
1– raw solution tank; 2– gas distributor; 3– photocatalytic reactor;

4– first solid/liquid separator; 5– second solid/liquid separator; 6– effluent tank;

A,B,C– peristaltic pump

Figure 3. The continuous degradation equipment of simulated wastewater
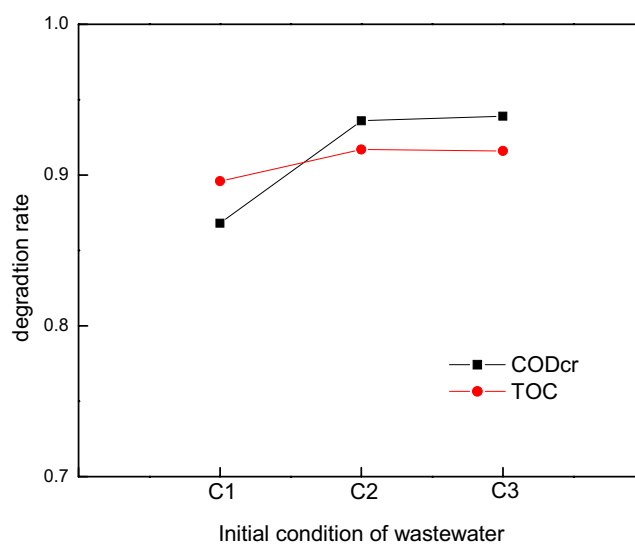
(a)
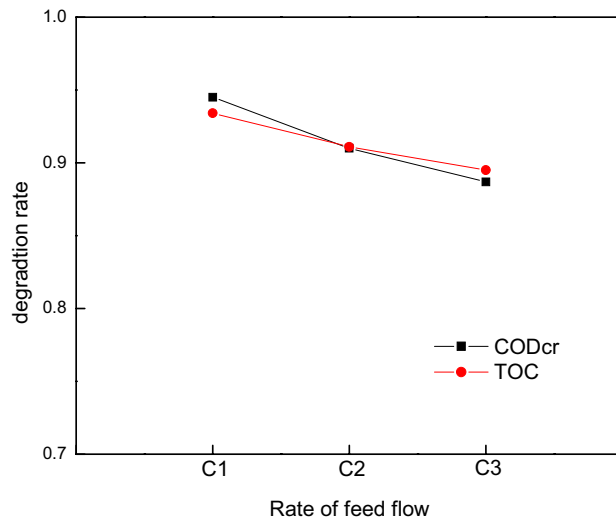
(b)
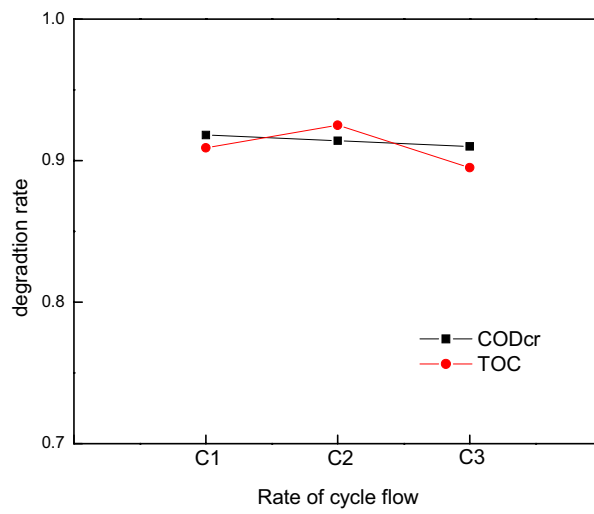
(c)

Figure 4. The evaluation of the factors at different levels

Ergonomics Analysis in the Industrial Design of Medical Instrument Products

Jun Yao

School of arts and design, China University of Mining and Technology, JiangSu, China

E-mail: cumtid@163.com

Yuerong Liang (Corresponding author)

School of Art and Design, Chinese Mining University

E-mail: chickenrongrong@163.com

Abstract

With the continuous development of economy and science, medical instrument entered the digital and informational age. Conflicts between the urgent requirement and severe lack for ergonomics in the industrial design of medical instrument products became increasingly fierce. In the present paper, we attempted to use the research achievements in the area such as product form design theory, ergonomics design theory and so on, with reference to the studies of other industrial products, brought forward relative design principle and design method for ergonomics design of medical instrument product, and aimed to accomplish the design work of medical instrument product perfectly.

Keywords: Medical instrument, Industrial design, Ergonomics, Humanization, User

1. Introduction

With the progress of society and rapid development of economics, human being has entered an age with considerably advanced science and technology. Rapid development of high-tech had brought human being biggish progress in medical field, and their average life increased greatly, to which invention and application of medical instrument product contributed considerably. Progress of science and technology brought people not only new function, but also new issues, namely the convenience of medical instrument products. The inconvenience of product application could decentralize people's energy for accomplishing assignment, and brought an uncontrollable sense imposed by machines. The convenience of product application was related to not only product function, but also the interface of all man-machine conversation, which was the right object that ergonomics research required. However, modern medical instrument design was quite short of consideration on ergonomics. In the design and production of some medical instrument production enterprises, no further investigation on ergonomics was undertaken, which hampered the development of China medical instrument. Accordingly, ergonomics consideration in industrial design of medical instrument was of great significance, and also the issues that many product designers of medical instrument should pay attention to.

2. Ergonomics and industrial design of medical instrument product

Ergonomics origin could trace back to the early 20th century. Modern ergonomics has been developed ever since 1960s and has been an independent systematic subject. Ergonomics was a subject to investigate the interaction among people, machine and work environment (Ding, 2005). First, it was an idea where users "people" were used as the starting point of product design, coordinative relations among people, machine and environment were investigated by all means, and some design methods were outlined. It required that the shape, color, performance and so on of products should be designed centering on the physiological and psychological characteristic of human being, in order to make those products more comfortable, safe and efficient for users' work and life.

Medical instruments were products serviced for life quality and health of people, and its ergonomic quality and application safety was related to people's life security. With the progress and development of science and technology, internal structure and operating control of medical instrument tended to be more and more accurate, which required

more and more harmony between people and machines during application, and thus ergonomics design should be undertaken in the production of medical instrument.

Due to various reasons, medical instrument design has long been accomplished by medical scientific and technical workers or mechanical electronic research department according to the requirement of certain medical application area, which made lots of medical instruments lack modern design idea in the design factors such as form, color, performance interface and so on, especially in ergonomics and humanization. All these led to their only purpose for treatment, however, very important factors that all such medical instruments should consider for users' humanization concern were severely ignored.

Users of the medical equipment were people who performed. It had two implications as seen from literal explanations (Wang, 2007):

1) Users were part of human being. User possessed the common characteristics of human being, and could reflect those characteristics when using any products. People behavior was not only affected by perception ability such as vision, hearing and so on, but also the ability of analyzing and solving problems, basic ability such as reaction ability fro stimulation as well. Meanwhile, it was restricted by factors such as psychological and character orientation, physical and cultural environment, education level, experiences and so on.

2) Users were people who made use of medical instrument. Its usable design focused on the special public related to product application. They could be the present users of product, and might also be the future users, or even potential users. Their behavior during the period of using products was closely associated with some characteristics related to products, such as knowledge for target product, expectant function of target product, and basic skill for using target product, time and frequency of using target product aftertime and so on. Medical instrument, as a special product, was obviously different from general consumable. A majority of medical instruments possessed double users, viz. doctor and patient. Doctors were positive users, while patients were passive users. Both belonged to "human being" category of ergonomics relationship.

Central idea of product design was "human being", "human-oriented". Design for people was one of the keys of product design. Medical instrument, as a product, was far different from general products, because it was associated with human health and life security. Therefore, properties of such product, such as usage, comfort, stability, amenity, security and so on, were paramount, among which ergonomics issues were very crucial, and humanization design was also a focus of medical instrument design.

3. Ergonomics analysis in medical instrument industrial design-a case of infrared breast scanner

Infrared breast scanner, according to the principle that different tissues had different infrared absorption, scanned breast tissue by infrared transillumination, obtained images by special infrared CCD camera, displayed all lesions of breast tissue on the screen after PC processing, and according to grayscale of image, and relation between vascular change and grayscale, diagnosed various breast diseases. In the present paper, analysis object referred was infrared breast scanner (SW-3003 common type) produced by Xuzhou Sanwei Medical equipment Co., Ltd in Jiangsu province (Figure 1). When using it, users were divided into two sorts. One was who detected, other was who was detected. Usually, people who were detected lay on the left side of the instrument for detection, and people who detected sat in the front of the instrument to detect patients; after data acquisition of the service conditions, we analyzed the product.

3.1 Dimension data analysis

Human body data collection was principal work for ergonomics analysis, and through consulting China human body dimension data, useful data was obtained, among which GB/T 10000-1988 was China important technological ergonomics standard. Data of this standard was frequently applied. Application task belonged to vision information working position. According to the design principle of vision information position, terminal position man-machine interface was depicted in Figure 2. Such task position design had the following points: A, human-chair interface; B, eye-visual screen interface; C, hand-keyboard interface; D, foot-floor interface.

When designers did their product design, human body dimension was one of the important references to determine his product size, and each percentage of human body dimensions was cardinal content of application (GB/T 12985-1991). Through data acquisition, consultation and analysis, external basic size, suitable for infrared breast scanner was ascertained which was depicted in Figure 3 in detail, namely human body dimension figure of normal sitting during work state. After data analysis, suitable height from typing operation area of infrared breast scanner to ground ranged from 66.0cm to 71.1cm; The least leg space was 63.5cm; The most reading distance from human body axis to screen ranged from 71.0cm to 76.0cm, in addition, if screen was rotatable and floating, adjustable height of screen was about 180mm, and adjustable angle was -5°~+15°; Optimum angle between the upper and front arm of hand control keyboard ranged from 70°~90°, and 100mm wrist rest space should remained between wrist joint and keyboard. The infrared breast scanner used in the present study basically accord with man-machine size requirement, and man-machine issues occurred in leg space. It failed to obligate enough space for leg, which resulted in leg disamenity and easy fatigue when

people operated it. It should be redesigned and ameliorated.

3.2 Color analysis

Color played a critical role in man-machine relation, and had a great impact on user's mentality. Product color matching in industrial design was primarily taken into consideration that color should assort with function and environment of equipment. It should not only meet the requirement of function, but also satisfied the requirement of human's taste. In the industrial design of medical instrument, effects of color were tremendous, and color was first applied in the medical treatment design was in 1925 in history. In surgery hospital of USA, during the operation, gleaming blood visual afterimage appeared in the white wall, which made vision in a tired state. In order to change such state, light green grey-complementary color of red was applied in the white wall, and thus vision fatigue of operative doctors was alleviated efficaciously.

As for infrared breast scanner, for objects were all female, some very simple but elegant and clean female color should be considered as subject color, such as quite light pink, pure blue and pure green. The infrared breast scanner applied in the present analysis used white as its subject color. Though it met the basic requirement, namely very simple but elegant and clean, it was short of sweet sense, and couldn't give the patients a sense of solicitude, but a sense of rather cold, which could not meet the requirement of caring for patients. Besides, medical instrument should give people a sense of tranquility and safety, and it was not suitable for the presence of vivid ornamental color; if there was no ornamental color at all, it would give people kind of calmness, cold and even lifeless sense, and no life vigor was recognized. The infrared breast scanner applied in the present analysis only had subject color-white, with no ornamental color, and it gave us an excessively rational and lifeless sense, which could imposed an passive psychological impact on users. It should be redesigned and ameliorated. Besides, in different countries and regions, due to the difference in historical tradition, national habits and custom, people had different and even contrary color preference and taboo. For instance, people advocated red, and usually used red to represent happiness and luck. But in some other country, such as Benin, red was considered as passive color... In Argentina, purple was avoided to use. While in Peru, it was one of the most popular colors (Wang, 1996).

3.3 Visual illusion in man-machine relation

When observing objects, for retina was stimulated by light, light not only promoted nerve system to react, but also result in increasing impact in transverse direction, which made the visual impression differ from the full size and shape of objects. This phenomenon was called visual illusion. Familiar shape illusions had length illusion, direction illusion, contrast illusion, size illusion, distance illusion, perspective illusion and so on; Color illusion had contrast illusion, size illusion, temperature illusion, distance illusion, fatigue illusion and so on. Some quite typical visual illusions were illustrated in Figure 4 in detail. In the industrial design of medical instrument, visual illusion could gain better utilization, for instance, as for very huge medical instrument, according to size visual illusion principle of visual illusion, medical instrument was designed using shape or color visual illusion, with attempts to reduce bulk in vision, reduce the sense volume of medical instrument, and psychological oppression imposed by huge medical instrument, and thus accomplish humanization design. The infrared breast scanner applied in the present analysis didn't apply visual illusion principle at all, which resulted in huge, stiffness and insipidity in appearance shape and brought passive emotion during operation. Therefore, it should be redesigned and ameliorated.

3.4 Barrier-free design analysis

"Design purpose is to meet the requirement of majority, not to service for minority. Especially for the oblivious majority, they should be paid more attention by designers" (Yin, 2003, P. 11). With the popularity of humanization design, more and more attention was paid to barrier-free design. As for "paths" to perceiving external information-vision, hearing and feeling, a majority of the disabled lacked one or some certain "paths" among them, but other "paths" were useable. For people with sensory disturbance, selection for the paths of information acceptance, namely selection of perception means, was the key to determine whether the product could realize the barrier-free design for information communication. Aimed at special cases and body limitation of specific population, selective application of various perception means to convey information could solve such problem.

4. Conclusions

In the present paper, we discussed the developmental deficiency of modern medical instrument products in ergonomics, and the deficiency seriously hampered the development and progress of medical instrument. With the development of social economy and scientific technology, further research on this area was necessary. Accordingly, as a producer and designer of medical instrument, issues of this aspect should be taken into great consideration, and in this way, product quality could be greatly ameliorated, which also played a critical role in the subsequent survival and development of medical instrument produce enterprises. Good and easy usage products, as well as humanization products certainly appealed to more consumers. This embodied "human-oriented" idea, accorded with the tide of social development, and was worthwhile for us to consider profoundly.

References

Ding, Y.L. (2005). *Ergonomics.* Beijing: Beijing Technology University Press.

National bureau of technical supervision. (1991). China national standard GB/T 12985-1991-General rules of using percentiles of the body dimensions for products design. Beijing: Chinese Standard Press.

Yin, D.B. (2003). *Design science introduction.* Changsha: Hunan Science and Technology Press, 11.

Wang, J.C. (1996). *Technology and art of modern industrial design.* Shanghai: China Textile University Press.

Wang, Y.L. (2007). Studies on the usable design of medical instrument man-machine interface. Nanjing: Master Degree Dissertations of Nanjing Forestry University.

Figure 1. Infrared breast scanner (SW-3003 common type)
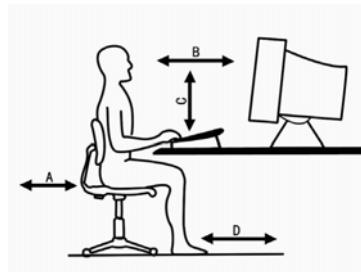
Figure 2. Human-machine interface of vision displayer station
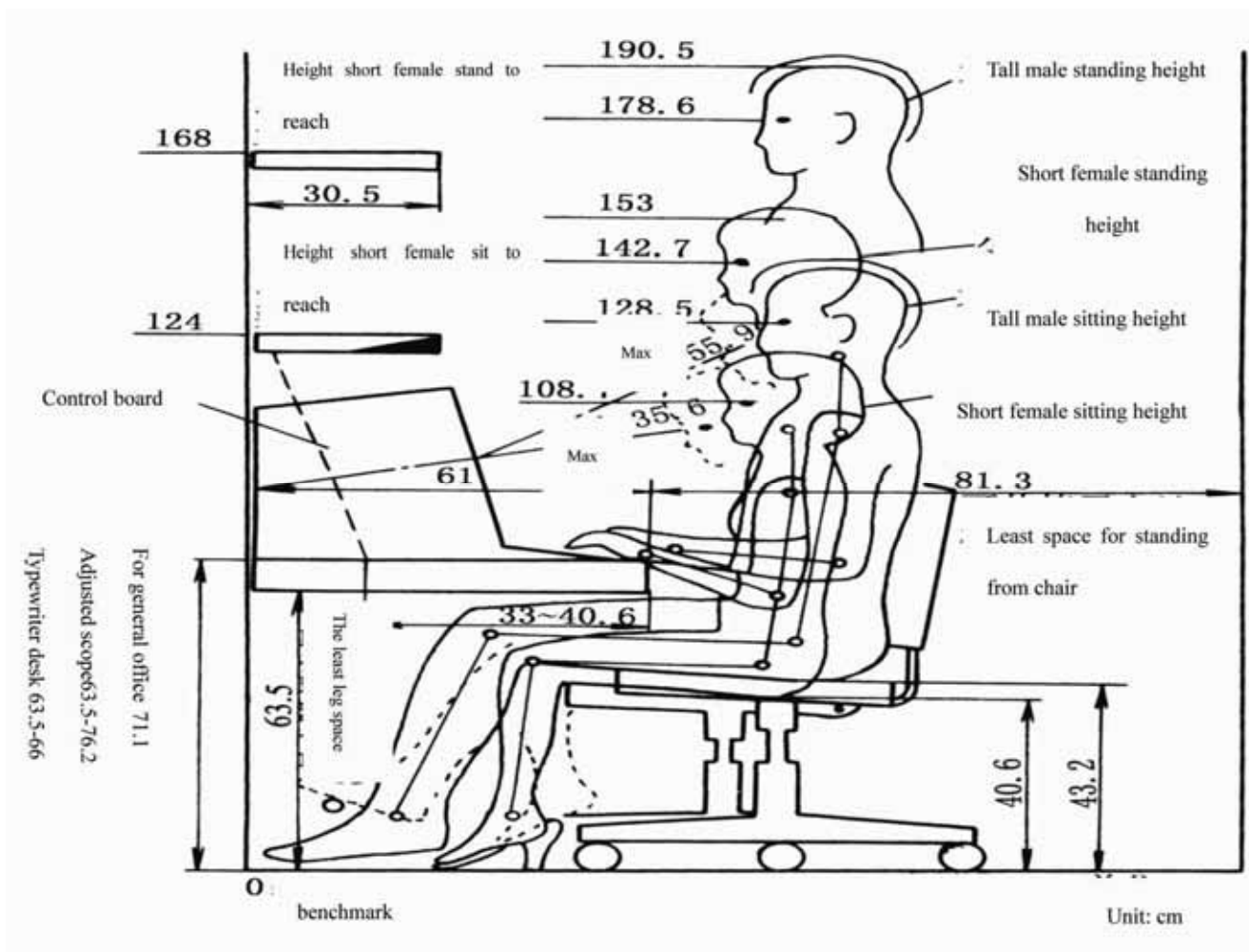
Figure 3. Human body dimension figure of normal sitting during work state
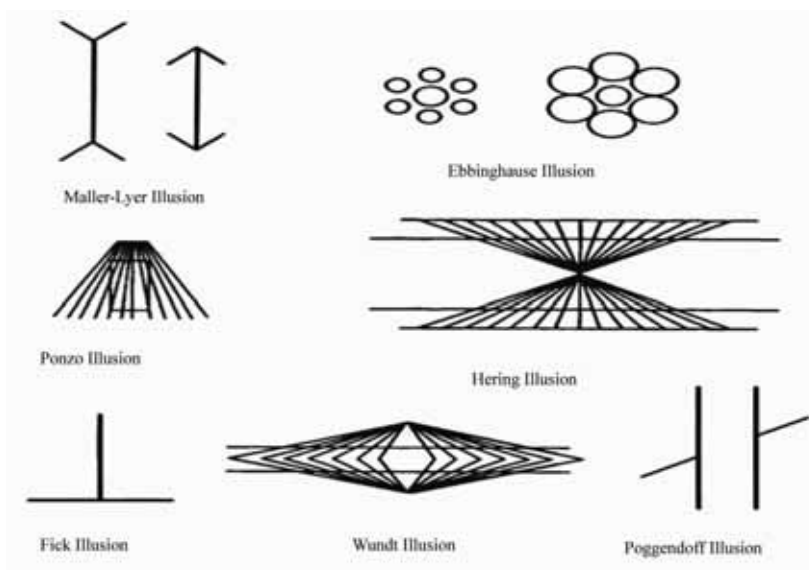
Figure 4. Typical visual illusions

Co(II), Ni(II), Cu(II), Zn(II) and Cd(II) Mixed Ligand Complexes of Theophylline and Cyanate: Synthesis and Spectroscopic Characterization

Shayma A. Shaker (Corresponding author) & Yang Farina

School of Chemical Science and Technology, University Kebangsaan Malaysia

43600 UKM Bangi, Selangor, Malaysia

E-mail: drshaimaa611@yahoo.com

Sadia Mahmmod & Mohean Eskender

Chemistry Department, Ibn-Alhaitham College of Education

University of Baghdad, Baghdad, Iraq

Abstract

New mixed ligand complexes were synthesized between theophylline and cyanate to give complexes with general formula $[M(Tp)_4Y_2]$ where Tp= theophylline, Y= cyanate ion and M=Co(II), Ni(II),Cu(II), Zn(II) and Cd(II). The resulting products were found to be crystallined which have been characterized using UV-Visible spectroscopic properties and Infrared spectra. Elemental analyses were performed using (C, H, N) and atomic absorption technique. The magnetic susceptibility and the conductivity for the complexes were also measured. The present results suggested the octahedral configuration for the metal complexes.

Keywords: Theophylline complexes, Cyanate complexes, Mixed ligand complexes

1. Introduction

The purines includes theophylline, theobromine and caffeine constitute an important class of anti-inflammatory agents (Satwinder, et al., 1995). Theophylline has biological importance as it is structurally related to nucleic acids components (Shohreh, et al., 2003). Thus it can be used as a drug in therapy for respiratory such as COPD or asthma under a variety of brand names and anticancer drugs. These drugs were extensively studied in terms of their inhibition and structural mode of binding against a family 18 chitinase from the opportunistic fungal pathogen *Aspergillus fumigates* (Francesco, et al., 2005; Jacek, et al., 2005; Kiriaki, et al, 2007). as, it can cause nausea, diarrhea, increasing of heart rate, arrhythmias and CNS excitation (Kiriaki, et al., 2007).

A few metal-theophylline complexes have shown significant antitumor activity (David, et al., 1999; Satwinder, et al., 1995). The anion derived from theophylline has often been used as a model ligand in studying the interaction with metal ions (Neville, et al., 1979). Therefore, it was of our interest to synthesis the complexation of metal ions like Co(II), Ni(II),Cu(II), Zn(II) and Cd(II) with theophylline and cyanate ion. The FTIR and UV-visible spectroscopic methods were used to examine the coordination site of these ligands with the metal ions.

2. Experimental

2.1 Material and Measurements

All chemicals were obtained from commercial sources and were used without further purifications ($CoCl_2.6H_2O$, $NiCl_2.6H_2O$, $CuCl_2.2H_2O$, $ZnCl_2$, $Cd(NO_3)_2.4H_2O$ and KOCN) from Riedeal–Dehaenage. Theophylline was used from BDH, methanol, ethanol and dimethylsulfoxide from Fluka.

The I.R spectra in the range of 4000-400 cm^{-1} were recorded as potassium bromid disc on Shimadzu FTIR-8300 Fourier transform infrared spectrophotometer. UV-Visible spectra were measured in DMSO using Shimadzu UV-Visible recorder spectrophotometer UV-160. Elemental analyses (C, H, N) were performed by the micro analytical unit on Berkin Elemer B-240 Elemental Analyzer. Determinations of metals were carried out using laboratory methods. Conductivity measurements were performed at 25°C in DMSO using Philips Pw-9526 digital conductivity meter. Melting points were determined using Stuart-Melting Point Apparatus, the magnetic susceptibility measurements were obtained using Balance Magnetic Susceptibility Model MsB-MK1.

2.2 General Procedure for Synthesis

An ethanolic solution 10 ml of theophylline (Tp) 1.06-2.96 g and 8 ml an aqueous solution of KOCN 0.24-0.48 g were respectively added to an aqueous solution of the metal salts. After constant stirring using appropriate amounts of materials needed as decided by the molar ratio (1:4:2) (M:Tp:Y), the resulting precipitates were filtered off, washed several times and recrystallized with 1:3 ethanol:water. Then, it dried in an oven at 65°C.

3. Results and Discussion

The prepared complexes were found to be solids, insoluble in water but they were soluble in some organic solvents like dimethylformamide and dimethylsulfoxide. The lower value observed of molar conductivities in DMSO indicates the non electrolyte behavior of the complexes (Kettle, 1975; Quaglian, et al., 1961). Elemental analysis (C, H, N) and metal determination were in good agreement with general formula given for the complexes. Table 1 gives in details the physical properties of the complexes.

3.1 Magnetic Susceptibility

The magnetic moment for the complex of cobalt ions should be around 3.872 B.M while the measured value of μ_{eff} was found to be higher than the expected value; this could be related to orbital-spin coupling (David, 1984; Ehssan, 1988). However the values of μ_{eff} of all Ni^{+2} and Cu^{+2} complexes were found to be closer to that of spinning only (Clyde, 1983; David, 1984). The complexes of Zn^{2+}, Cd^{2+} were diamagnetic as expected from their electron configuration. All these data are also included in Table1.

3.2 Electronic Spectral Studies

The electronic spectra for free ligands Tp and OCN^- ion shows absorption bands in the UV- region expressed as $\pi\rightarrow\pi^*$ and $n\rightarrow\pi^*$ transitions (Satwinder, 1995). The Co(II) complex exhibited bands at 282 and 585 nm which is attributed to the electronic transition of $^4T_1g_{(F)} \rightarrow ^4A_2g_{(F)}$ corresponding with charge transfer and the electronic transition of $^4T_1g_{(F)} \rightarrow ^4T_1g_{(P)}$ respectively.

The electronic spectra of Ni(II) complex shows electronic transition $^3A_2g_{(F)} \rightarrow ^3T_1g_{(P)}$ and $^3A_2g_{(F)} \rightarrow ^3T_2g_{(F)}$ at 284 and 993 nm respectively. These assignments correspond to Ni(II) octahedral complex.

The spectrum of Cu(II) complex shows bands in the visible region which are attributed to the electronic transition of $^2b_2g_{(D)} \rightarrow ^2b_1g_{(D)}$ and $^2e_2g_{(D)} \rightarrow ^2b_1g_{(D)}$ at 540 and 745 nm respectively.

Finally, the electronic configuration of Zn(II) and Cd(II) complexes were (d^{10}) which confirms the absence of any (d-d) transitions (Lever, 1968; Monther, 1983). All the absorption bands were fully assigned in Table 2.

3.3 Infrared Spectra Studies:

3.3.1 Infrared Spectra of Free Ligands

The spectrum of theophylline shows weak band at 3350 cm^{-1} which belongs to ν(NH). While the strong band attributed to ν(C=O) was found at 1720 cm^{-1} and the other strong band exhibited at 1668 cm^{-1} which is due to ν(N=C).

Here we turn our attention to cyanato compounds. Cyanate, as a terminal ligand, is potentially ambidentate and it can coordinate to metal ions through the nitrogen or the oxygen atoms. Information on the coordination mode of the cyanate ion may be obtained from infrared spectroscopy (Jose, et al., 2006). Moreover, the spectra of potassium cyanate shows very strong bands at 2165cm^{-1} and 1207cm^{-1} which belongs to ν(OCN) asymmetric and ν(OCN) symmetric (Kazuo, 1997; Socrates, 1980).

3.3.2 Infrared Spectra of the Complexes

Further information about the coordination behavior of the free ligands with metal ions were carried out by the comparison of the infrared spectra of the free ligands and their complexes. The absorption band appeared at 1668 cm^{-1} was caused by ν (-N=C) bond in the structure of Tp (Cross & Alan, 1969; John, 1965; Socrates, 1980; William & Flemming, 1973). In the spectra of their complexes, this absorption band was around 1635-1655 cm^{-1} which has shifted to lower frequencies by 33-13 cm^{-1}. Moreover, the spectra of the complexes exhibited weak bands between 460-490 cm^{-1} due to ν(M-N). This indicates the coordination of Tp with the metal ions through the nitrogen atom N9 in their structures (Elizabeth, & Elmer, 1979; Katsuyuki, & Hiroshi, 1980; Milan, 1981; Roger, et al., 1981).

Although the coordination chemistry of cyanate anion is well known, but surprisingly the adsorption of cyanate has been discussed only in a few papers (Jose, et al., 2006; Michael, & Rudolf, 1995). The asymmetrical stretching vibration was found to be in the region of 2200-2240 cm^{-1} and it was shifted to higher frequencies by 65-75cm^{-1}. The absorption band in the case of symmetrical stretching vibration was shifted to higher frequencies by 3-13 cm^{-1}. The usual observed infrared band is due to the CN stretching vibration, which was shifted to higher wave numbers with more positive potential. It was no metal cyanate complexes observed yet with such a strong shift of the frequencies of the CN stretching vibration to lower frequencies (Michael, & Rudolf, 1995). Therefore, we suggested that the

coordination of cyanate ion with metal ions through its oxygen atom. Similar absorption was found in the I.R spectra of the complexes $[Re(OCN)_6]^{2-}$, $[Re(OCN)_6]^-$ and $NBu_4[ReBr_4(OCN)(DMF)]$ where NBu_4= tetrabutylammonium cation and DMF=N,N-dimethylformamide (Bailey, & Kozak, 1969; Jose, et al., 2006; Kazuo, 1997; Socrates, 1980).

Another significant point, that is the effect of cyanate on the trans position relative to the cis position depends on qualitative evidence. Even this highly specific trans effect still does not have any theoretical justification but some studies have performed arguments suggested that the trans effect must be greatest on highly electronegative and poor donor ligands (John, & Lorraine, 1978). Thus, the spectra of the complexes exhibited lower frequency absorptions bands in the range of 565-605cm^{-1} have been assigned as δ(M-OCN). Table 3 gives the diagnostic absorptions and their assignments.

4. Conclusions

As part of our studies on the chemistry of transition metals, we have been able to synthesis and characterize a new mixed ligand complexes containing theophylline and cyanate ion coordinated with Co(II), Ni(II), Cu(II), Zn(II) and Cd(II). The molar conductivity of the complexes in DMSO solution was non electrolyte. The configuration of the complexes were performed that the theophylline coordinated with metal ions through the nitrogen atom N9 and cyanate ion coordinated with metal ions through the oxygen atom, Therefore, from the results presented the metal complexes have octahedral configuration, see Fig 1.

References

Bailey, R. A., & Kozak, S.L. (1969). Complex hexacyanates of rhenium and molybdenum, *J. Inorg. Nucl. Chem.*, 31, 689-693.

Clyde, M., J.R. Day & Joel Selbin. (1983). *Theoretical inorganic chemistry*.

Cross, A.D., & Alan Jones, R. (1969). *An introduction to partical infrared spectroscopy*.

David, L., Cozar, O., Forizs, E., Craciun, C., Ristoiu, D., & Balan, C. (1999). Local structure analysis of some Cu(II) theophylline complexes, *Spectrochimica Acta Part A*, 55, 2559–2564.

David Nicholls. (1984). *Complexes and first row transition elements.*

Ehssan, A. (1988). *Inorganic and coordination chemistry*, Mousel university, Iraq.

Elizabeth H. Griffith & Elmer L. Amma. (1979). Reaction of $PtCl_4^{2-}$ with theophylline: *X*-ray crystal structures of bis(theophyllinium) tetrachloroplatinate(II) and theophyllinium trichlorotheophyllineplatinate(II), *J.C.S. Chem. Comm*, 322-324.

Francesco V. Rao, Ole A. Andersen, Kalpit A. Vora, Julie A. Demartino & Daan M. F. Van Aalten. (2005). Methylxanthine drugs brief commu are chitinase inhibitors: Investigation of inhibition and binding modes, *Chemistry & biology*, ISSN 1074-5521, 12, 973-980.

Jacek Piosik, Anna Gwizdek-Wiśniewska, Katarzyna Ulanowska, Jakub Ochociński, Agata Czyż & Grzegorz Węgrzyn. (2005). Methylxanthines (caffeine, pentoxifylline and theophylline) decrease the mutagenic effect of daunomycin, doxorubicin and mitoxantrone, *Acta Biochemica Polonica*, 52 (4), 923–926.

John Dyer. (1965). *Application of absorption spectroscopy of organic compounds.*

Jose´ Martı´nez-Lillo, Donatella Armentano, Giovanni De Munno, Francesc Lloret, Miguel Julve & Juan Faus. (2006). Rhenium(IV) cyanate complexes: Synthesis, crystal structures and magnetic properties of NBu4[ReBr4(OCN)(DMF)] and (NBu4)2[ReBr(OCN)2(NCO)3], *Inorganica Chimica Acta*, 359,   4343–4349.

John K. Yandell & Lorraine A. Tomlins. (1987). The *trans* effect in Cobalt(III) complexes kinetics and mechanism of substitution of Cobalt(III) sulphito Complexes.   *Aust. J. Chem.,* 31, 561-71.

Kiriaki M. S. Rodrigues, Duclerc F. Parra, Maria José A. Oliveira, Oscar V. Bustillos & Ademar B. Lugão. (2007). Study of theophilline stability on polymer matrix. *International Nuclear Atlantic Conference* - INAC ,Santos, SP, Brazil, September 30 to October 5, *Associação Brasileira De Energia Nuclear – ABEN ISBN*: 978-85-99141-02-1.

Kettle, S.F.A. (1975). *Coordination compounds*. Thomas Nelson and Sons, London, 165.

Katsuyuki Aoki & Hiroshi Yamazaki. (1980). Interactions of tetrakis(-µ-carboxylato)dirhodium(II), an antitumour agent, with nucleic acid bases. *X*-Ray crystal structures of [Rh$_2$(acetato)$_4$(theophylline)$_2$] and [Rh$_2$(acetato)$_4$(caffeine)$_2$], *J. Chem. Soc. Chem. Commun.*, 186-188, DOI: 10.1039/C39800000186.

Kazuo Nakamoto. (1997). *Infrared and raman spectra of inorganic and coordination compounds*. (5th edn).

Lever, A. B. P. (1968). *Inorganic electronic spectroscopy*.

Monther.Y. Al. Janabi. (1983). *The physical methods in inorganic chemistry*.

Milan Melnik. (1981). Binuclear caffeine adducts of Cu(II)acetate and Cu(II)chloracetates with unusually high antiferromagnetic interaction. *J. inorg. Nucl. Chem.*, 43, 3035-3038.

Michael Bron & Rudolf Holze. (1995). Cyanate and thiocyanate adsorption at copper and gold electrodes as probed by in situ infrared and surface-enhanced raman spectroscopy, *Journal of Electroanalytical Chemistry,* 385, Issue 1, 105-113.

Neville H. A., Trevor G. A., John R. H., Gregory F. K. & Ian J. M. (1979). A new mode of theophyllinate anion–metal bonding: N(7), N(9)-bridging in theophyllinate complexes of trimethylplatinum(IV), *J.C.S. chem. Comm.*, 324-326, DOI: 10.1039/C39790000324.

Quaglian, J.V., Fujita, J., & Franz, G. (1961). *J. Am. Chem. Soc.*, 81, 3770.

Roger E. Cramer, Douglas M.Ho, William Van Doorn, James A., Ibers, Ted Norton & Midori Kashiwagi. (1981). Triphenylmethylphosphoniumtrichloro(caffeine)platinum(II). *Inorg. Chem,* 20, 2457-2461.

Satwinder S. Marwaha, Jasjest Kaur & Gurvinder S. Sodhi. (1995). Structure determination and anti-inflammatory activity of some purine complexes. *Met Based Drugs*, 2(1), 13–17, doi: 10.1155/MBD.

Shohreh Nafisi, Abolfazl Seyed Sadjadi,Sakine Shokrollah Zadeh & Maryam Damerchelli. (2003). Interaction of metal ions with caffeine and theophylline: stability and structural features.   *Journal of Biomolecular Structure & Dynamics*, ISSN 0739-1102, Volume 21, Issue Number 2.

Socrates, G. (1980). *Infrared characteristic group frequencies*.

William & Fleming. (1973). *Spectroscopic methods in organic chemistry*. (2nd edn).

Table 1. The physical properties of the metal complexes [M(Tp)$_4$Y$_2$]

Compound	Color	M.P ^{o}C	Elemental analysis: Calc. (Found)%				Magnetic moment Calc. (Found)%	Molar conductivity Ohm^{-1}cm^{2}mol^{-1}
			C	H	N	M		
[Co(Tp)$_4$Y$_2$]	Pink	260-262	41.722 (41.65)	3.731 (3.90)	29.191 (30.25)	6.66 (7)	3.872 (4.037)	18.75
[Ni(Tp)$_4$Y$_2$]	Light blue	>300	41.732 (40.83)	3.732 (3.67)	29.198 (29.27)	6.66 (7.33)	2.828 (2.766)	19.23
[Cu(Tp)$_4$Y$_2$]	Green blue	>300	41.500 (41.62)	3.711 (4.68)	29.036 (28.16)	7.3 (7.9)	1.732 (1.678)	15.78
[Zn(Tp)$_4$Y$_2$]	white	262-264	41.412 (41.35)	3.703 (3.59)	28.974 (28.91)	7.5 (6.16)	zero	15
[Cd(Tp)$_4$Y$_2$]	white	260-261	39.289 (39.37)	3.513 (3.63)	27.489 (26.54)	12.23 (12.6)	zero	15.78

Table 2. U.V-visible spectra of free ligands and their complexes $10^{-3}M$ in DMSO

Compound	λ_{max} n.m	ABS	Wave number cm^{-1}	ε_{max} $Lmol^{-1}cm^{-1}$	Transitions
KOCN	272	0.614	36764.705	614	$\pi \to \pi^*$
	310	0.718	32258.064	718	$n \to \pi^*$
TP	274	1.728	36496.35	1728	$\pi \to \pi^*$
	340	0.554	29411.764	544	$n \to \pi^*$
$[Co(TP)_4Y_2]$	282	1.939	35460.992	1939	$^4T_1g_{(F)} \to {}^4A_2g_{(F)}$
	585	0.132	17094.017	132	$^4T_1g_{(F)} \to {}^4T_1g_{(P)}$
$[Ni(TP)_4Y_2]$	284	1.935	35211.267	1935	$^3A_2g_{(F)} \to {}^3T_1g_{(P)}$
	993	0.012	10070.493	12	$^3A_2g_{(F)} \to {}^3T_2g_{(F)}$
$[Cu(TP)_4Y_2]$	287	2.090	34843.205	2090	Charge transfere
	540	0.024	18518.518	24	$^2b_2g_{(D)} \to {}^2b_1g_{(D)}$
	745	0.060	13422.818	60	$^2e_2g_{(D)} \to {}^2b_1g_{(D)}$
$[Zn(TP)_4Y_2]$	277	1.781	36101.083	1781	Charge transfere $M \to L$
$[Cd(TP)_4Y_2]$	279	1.858	35842.293	1858	Charge transfere $M \to L$

Table 3. The characteristic infrared absorptions of free ligands and their complexes

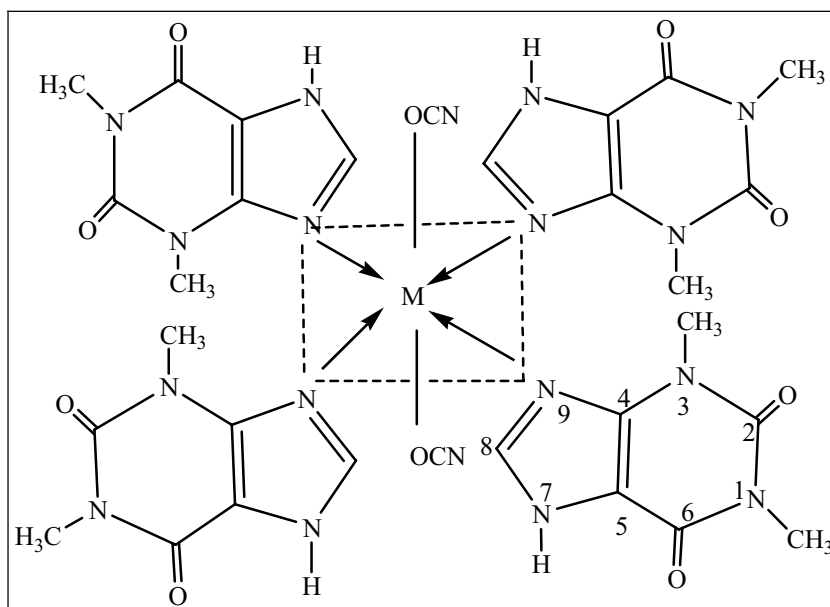| Compound	$\nu(NH)$	$\nu(OCN)_{asy}$	$\nu(C=O)$	$\nu(-N=C)$	$\nu(OCN)_{Sy}$	$\delta(M-OCN)$	$\nu(M-N)$
KOCN	-	2165	-	-	1207	-	-
TP	3350	-	1720	1668	-	-	-
$[Co(TP)_4Y_2]$	3360	2240	1725	1650	1210	605 575	460
$[Ni(TP)_4Y_2]$	3360	2240	1730	1650	1215	605 580	480
$[Cu(TP)_4Y_2]$	3320	2200	1700	1635	1210	600 565	460
$[Zn(TP)_4Y_2]$	3330	2220	1730	1655	1215	580	490
$[Cd(TP)_4Y_2]$	3320	2200	1725	1650	1220	590	480

Figure 1.

The Method of Spherical Surface Roughness Measurement

Xiaohui Xu & Yan Cui

School of physical electronics, Tianjin Polytechnic University, Tianjin 300160, China

E-mail:xiaohuixu2009@live.cn

Abstract

The spherical surface is measured by means of Atomic Force Microscope (AFM) on nanometer scale. In order to evaluate the surface quality of this kind of three dimensional measurement, the following parameters are suggested: R_a, R_q, R_y and T_p. Some questions related to this choice are discussed in this paper.

Keywords: Surface roughness measurement, Spherical surface, Atomic force microscope

1. Introduction

The spherical surface plays an important role in precision engineering. They are used in precision machines and instruments. People usually use interferometers and contact (stylus) instruments to measure the spherical surface roughness. The interferometer is non contact, but when the corrugation of the surface is less than 0.05um, it is difficult to evaluate it because the interface fringes are almost straight lines. (L C Leonard ,1998,p.433). The stylus can get several parameters by means of a computer. But it is a contact method. The limit of this method mainly depends on the radius of the scanning tip. For example, the instrument Tallest has the minimum radius of 0.1um.Besides,it is almost impossible for these two methods to operate with small balls due to their reduced diameter(3mm).Both methods are limited because they give line profiles.(Takahiko,2003,p,452-455).

Since Scanning Tunneling Microscope (STM) and Atomic Force Microscope (AFM) were invented, they have shown great potential in the measurement of the surface structure on nm scale. They have soon applied to precision engineering. Nigeria et al. used STM to measure the surface roughness of a "0" class block gauge. (Garcia, 1985, p.566).

Here we report on the results we have obtained by using an AFM to measure the spherical surface. We operated the AFM in the repulsive mode with a small force of $<5 \times 10^{-8}$N,which is more than two orders of magnitude lower than the force applied by the stylus instrument.AFM gives three-dimensional images and it can get nm resolution.

An important point is how to specify quantitatively the surface structure. For one dimensional measurement, there are international standards and various national standards. But for 3-dimensional measurement there are no formal documents that can be used. Here we propose four parameters: the arithmetical mean deviation of the profiles R_a, the root-mean-square deviation of the surface R_q, the maximum height of the surface R_y, and bearing area ratio T_p.

2. Experiments

We installed an AFM (Schnenberger, 1989,p.3131. Huang, 1991, p.1323-1328) based on a polarized laser interferometer. A tungsten tip was glued onto a commercial silicon wafer. The tip was made by direct-current (dc) etching (3V) of 0.1mm polycrystalline W wire. We measured the tips by Scanning Electron Micron Microscope (SEM).The radii of these tips were about 10nm.The force constant of the cantilever was about 12N/m. The frequency of 1.2 kHz.

We used the precision spherical surface which its diameter is 3mm.Because the system was sensitive to optical, mechanical, air and thermal fluctuation; we designed our system with cylindrical symmetry structure. We used materials which had small thermal expansion. The optical parts were fixed directly on a granite plate, and also the acoustic noise. The granite plate was suspended with soft springs. Finally we got an output of the signal sensitivity of 20mV/λ, the noise was about 5mV, and the thermal drift 0.05 λ/sec.

Usually we use the repulsive force mode. The constant force is about 5×10^{-8}N.The scanning frequency is 1-5Hz.Here we show two images of different measuring range.Fig.1 shows the three dimensional and top view representations of a 3μm×3μm area.Fig.2 shows the same data for size 1μm×1μm taken on a different area of the same sphere surface.

3. Parameters of 3-diamensional profiles

The data obtained by AFM can be processed to give a 3-D representation of the surface. Although this representation

can be sometimes self-evident, it is important to give quantitative information from the image. For that purpose and to give the non periodic character of the surface, we suggest that four parameters be used: the arithmetical mean deviation R_a, the root-mean-square deviation R_q, the maximum height R_y and the profile bearing area ratio T_p. The reference plane is the least squares mean plane. We use the same terminologies and the same symbols as those in ISO 4287/1-1984 and ISO 468-1982, in order that people can understand and use them easily conveniently.

The arithmetical mean deviation Ra id defined as

$$R_a = \frac{1}{N_x} \bullet \frac{1}{N_y} \sum_i^{N_x} \sum_j^{N_y} |Z_{ij}| \tag{1}$$

Here N_x: sampling points in X direction

N_y: sampling points in Y direction

Z_{ij}: height with respect to the least squares mean plane

The root-mean-square value Rq is defined as

$$R_q = \left\{ \frac{1}{N_x} \bullet \frac{1}{N_y} \sum_i^{N_x} \sum_j^{N_y} Z_{ij}^2 \right\}^{1/2} \tag{2}$$

Here N_x, N_y, Z_{ij} are the same as in(1).

The maximum height Ry is defined as

$$R_y = R_p + R_m \tag{3}$$

Here R_p is the distance between the highest point of the profiles and the reference planer is the distance between the lowest point of the profiles and the reference plane.

The profile bearing area ratio T_p is defined as

$$Tp = \frac{\eta_p}{S} \tag{4}$$

Here η_p is the sum of the section areas obtained by cutting the profile peaks with a plane parallel to the reference plane. S is the area of scanning. From this we can get the curve of the profile bearing area ratio representing the relationship between the value of the profile bearing area ratio and the section level.

In our experiment, in 3μm×3μm area ,Ra=4.1nm,Rq=5.2nm,Ry=34nm.The values for the 1μm×1μm image are Ra=4.1nm,Rq=4.9nm,Ry=25nm,The curves of bearing area ratio are shown in Fig.3.

4. Discussions

From our experiment, it is clear the AFM is a potential tool to measure an ultra-precise surface. Not only can conductiv3 sample be measured but also the non conductive one. Along with the development of science and technology, the requirements of specifying 3-dimensional high quality surface with nanometer scale are placed on the order of the day. The standardization of the parameters and values for the specification of 3-dimensional surface is absolutely necessary.

From the point of the view of metrology, the important and more difficulty thing is to calibrate the instrument and to estimate the uncertainty of the measuring method. There is a lot of work to be done. Some people (Li Y Z,1989,p.15) suggest using diffraction grating to calibrate μm scale range and using lattice constant to calibrate nanometer scale image. This is feasible at the present time. Maybe it is necessary to find some standard material which is very stable and suitable for calibrating the 3-dimensional measurements.

In this paper, we have not mentioned the terminology "sampling area". It is not the question that the values of R_a, R_q, R_y and T_p are not dependent on the "sampling area". In fact, like the "sampling length" in ISO 4287/1-1984 and ISO 468-1982,this is very important. The question is that because we performed few experiments we can not determine how big the "sampling area" is reasonable. At the moment, we suggest considering the "sampling area" as being equal to the scanning area. So it is clear that when someone mention the value of R_a, R_q, R_y or T_p of 3-dimensional profiles, he must point out the value of the scanning area.

References

Garcia N, Baro A M, Miranda R, Rohrer H, et al. (1985). *Metrologia*, 21, 566.

Huang Wen Hao, Baro A M and Sanez J J, J. (1991). *vac,sci.tech*, B9.No.2,1323-1328.

L C Leonard and Toal V, (1998). Roughness Measurement of Metallic Surfaces Based on the Laser Speckle Contrast

Method .Optics and Lasers in Engineering, 30(5), 433~440.

Li Y Z,Vazquez L et al. (1989). *Appl. Phys*, Lett,. 54, 15.

Schnenberger C and Alvarado S F, Rev. (1989). *Sci. Instrum.*50, 3131.

Takahiko Inari and Nobuya Aoki, (2003). Influences of Dark Output of a Video Camera on the Pattern Projection Method for Surface Roughness Measurement SICE Annual Conference in Fukui, August 4-6,452-455.
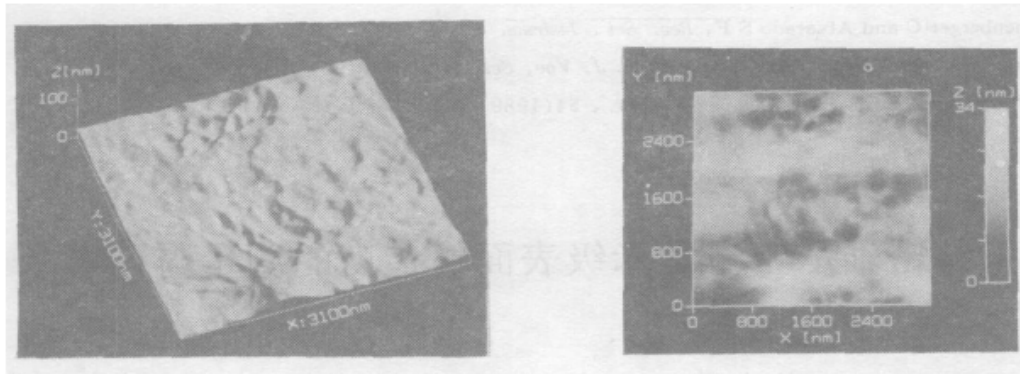

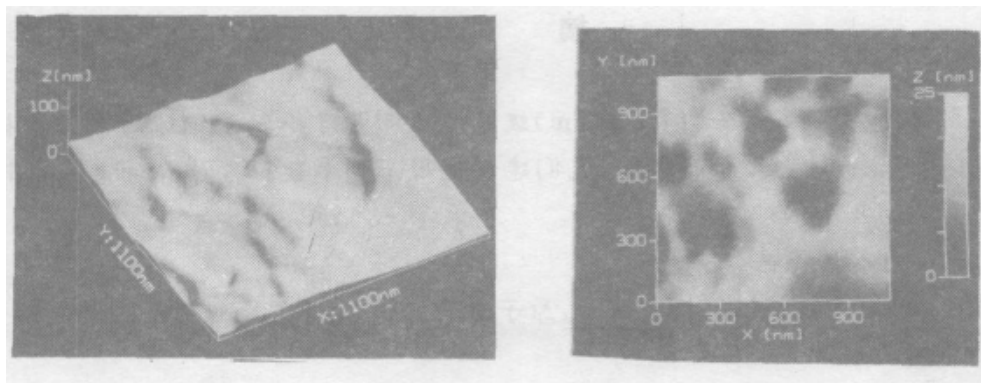Figure 1a shows a 3-dimensional representation of a sphere surface of 3mm .

Fig 1b is the top view.

Figure 2a and 2b show a smaller area of a different place in the same sphere surface as in Figure 1.
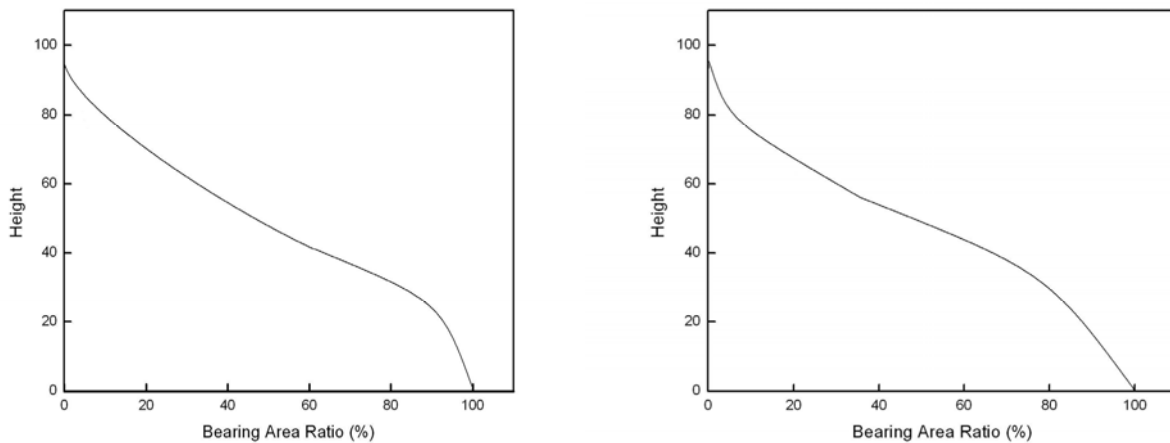
Figure 3a and 3b show the curves of bearing area ratio corresponding respectively to the regions depicted in Figures 1 and 2.

Modern Applied Science